Freedom and Voting Power

Itai Sher*

December 8, 2018

Abstract

This paper develops the *symmetric power order*, a measure of voting power for multicandidate elections. The measure generalizes standard pivotality-based voting power measures for binary elections, such as Banzhaf power. At the same time, the measure is not based on pivotality, but rather on a measure of freedom of choice in individual decisions. Indeed, I use the symmetric power order to show that pivotality only measures voting power in monotonic elections, and is not a good measure in multicandidate elections. Pivotality only provides an *upper bound* on voting power. This result establishes a relation between voting power and strategyproofness.

Keywords: Voting power, Banzhaf power, multicandidate elections, pivotality, freedom of choice.

*Department of Economics, University of Massachusetts Amherst, Crotty Hall, 412 North Pleasant Street, Amherst, MA, 01002, USA. Email: itaisher@gmail.com.

1 Introduction

What makes for a good voting system? Should we focus on the objective quality of outcomes or the match between voter preferences and voting outcomes? What about voting systems that encourage sincere voting, so that the inputs to voting at least reflect true preferences? Fairness is another important criterion.

There are large literatures in political science and economics that evaluate voting systems from many perspectives and using many different methods. One basic value that has been used to evaluate voting institutions is *agreement*, broadly understood as the idea that voting institutions should maximize agreement between the views of voters and electoral outcomes. This value can be elaborated in many ways. The impossibility theorem in Arrow's (1951) *Social Choice and Individual Values*, the work that launched the modern social choice literature, can be interpreted in terms of agreement. Arrow's theorem established that there does not exist a voting system that satisfies certain intuitive normative axioms. In the second edition of Arrow's monograph (1963), one of these axioms is the Pareto efficiency axiom, which says that if all voters prefer one alternative to another, the voting rule should rank the preferred alternative higher. This can be viewed as a minimal requirement of *agreement* between voter's views and electoral outcomes. Arrow's theorem says that, given some other conditions, we cannot have such agreement.

Another basic value is *control*: Citizens should collectively have as much control over voting outcomes as possible, and that control should be shared fairly. The notion of popular control is central to democracy.

The *voting power* literature can be interpreted as assessing voting institutions in terms of control. In *binary* elections, voting power is measured by *pivotality*, that is, the probability that a voter determines the outcome. Under a "normative" probability distribution according to which (i) each voter is equally likely to vote for either candidate, (ii) votes are independent, and (iii) there is no abstention, pivotality is known as *Banzhaf power*.[1]

[1] This is referred to as a "normative distribution" because, while these assumptions

Banzhaf power was discovered by Penrose (1946); it was rediscovered and made famous by Banzhaf (1964, 1966, 1968). The main competing index is the Shapley-Shubik power index (Shapley and Shubik 1954), which is an application of the Shapley value (Shapley 1953) to voting games. While it is not the most common interpretation, the Shapley-Shubik index can also be interpreted in terms of pivotality (Straffin 1977).[2,3] Gelman, Katz and Bafumi (2004) criticize the probability model underlying Banzhaf power; I discuss this in Section 2.1.4.

Through the lens of control, the evaluation of voting institutions can be given an attractive economic interpretation. We can think of control as a scarce resource: If we are faced with an indivisible public decision (e.g., shall we go to war?) not everyone can fully control the decision. We can share control, but if I share control with you, then my control must of necessity be less than if it were all mine. This gives rise to an allocation problem: the *allocation of control*. We seek voting institutions that allocate control *efficiently* and *fairly*.

I now illustrate this interpretation in terms of well-known results about voting power. Assume two candidates and n voters who vote according to the normative binomial distribution above. I compare two voting institutions that allocate power equally among voters: majority voting and random dictatorship. In a random dictatorship, each voter is selected with probability $\frac{1}{n}$ to be a "dictator" who can unilaterally determine the outcome. Under majority voting, the probability of being pivotal is $\frac{1}{2^{n-1}}\binom{n-1}{\frac{n-1}{2}}$. Because under random dictatorship, a voter is pivotal with probability $\frac{1}{n}$ and $\frac{1}{2^{n-1}}\binom{n-1}{\frac{n-1}{2}} > \frac{1}{n}$, we can declare that random dictatorship is *inefficient* in its allocation of control. Random dictatorship is dominated by majority voting. Indeed, majority vot-

are never satisfied in practice, some authors have argued that they provide a natural benchmark for evaluating voting institutions (see Felsenthal and Machover (1998) for a discussion).

[2]For the history of voting power, see Felsenthal and Machover (1998, 2005).

[3]Banzahf power has a relation to agreement and not just control. As already pointed out by Penrose (1946), under the normative distribution, the probability that a voter i agrees with the outcome is $\frac{1+\beta_i}{2}$, where β_i is i's (absolute) Banzhaf power. However this relationship is not robust to other assumptions about the distribution of voting behavior (Laruelle and Valenciano 2005), and becomes even more tenuous in mutlicandidate elections.

ing is the binary voting mechanism that maximizes average voter pivotality given the normative distribution (Dubey and Shapley 1979).[4] We can also use pivotality to assess the *fairness* of the allocation of control. Pivotality can be used to argue that the National Popular Vote is more fair than the Electoral College. Under the normative distribution, national popular vote gives each voter the same voting power, whereas the electoral college exaggerates the relative voting power of voters in large states (Penrose 1946, Banzhaf 1968). The national popular vote also leads to a higher average voter pivotality than the electoral college. It follows that if the scarce resource to be allocated is control, then national popular vote leads both to a larger pie, and to a more equally distributed one. An appealing feature of this way of evaluating elections is that it side-steps the problem of impossibility. If we can convert the evaluation of elections into an optimization problem – such as that of maximizing aggregate control with a penalty for unequal allocations – then we do not need to worry about whether election mechanisms satisfy a set of discrete conditions, such as Arrow's axioms.

One shortcoming of the above analysis is that it is limited to binary elections.[5] The distinction between two and more than two alternatives is fundamental to voting. For example, Arrow's impossibility result requires three alternatives: With two alternatives, majority voting satisfies all of Arrow's conditions, and there is no impossibility. If dealing with two alternatives were sufficient, then Arrow's theorem would never have achieved the influence it has had. In reality, there are almost never only two atlernatives, and if only two alternatives appear on the ballot, that means that the agenda setting procedure has done much of the work in determining the social choice. The purpose of this paper is to present a general theory for measuring and evaluating shared control in elections, including multicandidate elections.

There have been attempts to ageneralize Banzhaf power to multi-candidate elections, most notably by Bolger (1983, 1986, 1990, 2002). Closely related is work on voting power with abstention and multiple levels of approval

[4]Majority voting is optimal among *monotonic* binary voting mechanisms. See Definition 1.

[5]Another shortcoming is the assumption of independent votes (Gelman et al. 2004). This is discussed in Section 2.1.4.

(Felsenthal and Machover 1997, Freixas and Zwicker 2003). My work differs from these works in several respects.

The first and most fundamental difference is that my voting power measure – the *symmetric power order* – is founded on a ranking of individual decisions in terms of *freedom of choice*. This latter freedom ranking was axiomatized in Sher (2018*b*) and relates to the decision theoretic literature on preference for flexibility (Koopmans 1964, Kreps 1979, Nehring 1999, Dekel, Lipman and Rustichini 2001). Intuitively, freedom and power are closely allied concepts. Dowding and van Hees (2009) conclude their survey on freedom in social choice theory:

> ... it would be illuminating to examine the differences and similarities between the measurement of freedom and the measurement of power. Within cooperative game theory there is an extensive formal literature on the measurement of power. Given the close relationship between the concepts of freedom and power, it would be interesting to explore the extent to which the two types of analysis can profit from each other, and in particular insure that they are not measuring the same quantity.

The current paper carries out the program suggested by Dowding and van Hees, and concludes positively that it is possible to derive a measure of voting power as a special case of a measure of freedom.

The second difference is that whereas Bolger and others view themselves as generalizing the notion of pivotality to multicandidate elections, I use my ranking to *assess* pivotality as the guiding idea for measuring voting power. I use the symmetric power order to show that pivotality only measures voting power in *monotonic* elections. Since monotonicity is hard to achieve in multicandidate elections, it follows that when multiple candidates are genuinely competitive, pivotality does *not* measure voting power. I show however that pivotality does always provide an *upper bound* on voting power. My negative result relies on the Gibbard Satterthwaite theorem (Gibbard 1973, Satterthwaite 1975), a close cousin of Arrow's theorem, and thus implicitly relates power to incentives.[6]

The voting power literature has focused primarily on *monotonic* binary elections. My voting power measure coincides with pivotality for monotonic elections, and hence it *is* a generalization of pivotality-based measures of voting power, such as Banzhaf power, to multicandidate elections. At the same time, I claim that for multicandidate elections, which are generally not monotonic, pivotality and voting power come apart. To establish this result requires generalizing the concept of pivotality to multicandidate elections. The generalization of pivotality is a second distinct contribution of this paper. Thus one can view this paper as generalizing standard voting power measures to multicandidate elections in two different ways, and arguing that one way is correct and the other incorrect.

The symmetric power order only provides *partial* ranking of elections. In and of itself, this is not an advantage. After all, it would be desirable to be able to rank every pair of elections in terms of voting power. However, it can also be a disadvantage to rank elections that shouldn't be ranked. In Section 5.4, I show that when the symmetric power order does *not* rank election σ above election σ', one can always find a purpose that σ' allows a voter to achieve better than σ. This can provide a justification for leaving some pairs of elections unranked. Nevertheless, in Appendix A, I discuss refinements of the symmetric power order that can compare more pairs of elections.

The outline of the paper is as follows. Section 2 defines the symmetric power order. Section 3 analyzes binary elections and Section 4 analyzes multicandidate elections. Section 5 provides illustrations. An appendix presents proofs and technical details.

2 Defining voting power

This section presents the basic voting framework, the measure of freedom of choice, and the definition of voting power in terms of freedom of choice

[6]My result also relates closely to the Muller Satterthwaite theorem (Muller and Satterthwaite 1977). It is interesting to note here that positive association – a version of monotonicity – was the axiom in the original version of Arrow's theorem (1951) that was replaced by Pareto efficiency in the second edition (1963).

2.1 Framework

2.1.1 Candidates, voters, and votes

Let $C = \{1, \ldots, m\}$ be the set of **candidates**, $I = \{1, \ldots, n\}$ the set of **voters**, and V the set of **votes**. Votes could be of many forms, such as, "I vote for candidate c", or "I rank candidate c_1 first, candidate c_2 second, candidate c_3 third .... candidate c_m last", or "I approve of candidates c_1 and c_2 but of no others," or something else. I assume that V is finite.[7]

In addition to the voters in I, there there is also an automated agent, "voter 0", who is the **tie-breaker**. Intuitively, the tie-breaker is intended as a randomization device used to break ties, but in principle the tie-breaker could be used to introduce randomness into the voting mechanism in a more robust way. For example, under random dictatorship, the so-called tie-breaker may select the voter who is to be dictator. Let $I_0 = \{0, 1, \ldots, n\}$ be the set of voters, including the tie-breaker. V_0 is the finite set of "votes" for the tie-breaker. It is natural to assume that $V_0 \neq V$. For example, V_0 might be the set of possible sequences of 0's and 1's of a fixed length k, interpreted as the possible set of outcomes of k coin flips. I sometimes also refer to the tie-breaker's vote v_0 as the tie-breaker, where this will not cause any confusion.

A **vote profile** is a list $v = (v_0, v_1, \ldots, v_i, \ldots, v_n)$ where $v_0 \in V_0$ is the tie-breaker's vote, and, for $i \in I$, v_i is voter i's vote. v_{-i} is a list with i's vote removed: $v_{-i} = (v_0, v_1, \ldots, \; v_{i-1}, v_{i+1}, \ldots, v_n)$. If we put i's vote back in, we write $v = (v_i, v_{-i}) = (v_0, v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n)$. $\bar{V} = V_0 \times V^I$ is the set of all vote profiles, and for each $i \in I$, $V_{-i} = V_0 \times V^{I \setminus i}$ is the set of all vote profiles, excluding i's vote.

A **voting rule** is a function $f : \bar{V} \to C$ that maps a vote profile into a winning candidate. Many voting systems fit into this framework, including majority voting between two candidates, plurality voting among three candidates, the electoral college, instant runoff voting, the Borda rule, and the Condorcet rule.

[7]Nothing significant would change if different voters had different vote sets V_i.

2.1.2 Probabilities

Let μ_i be a probability distribution over V for voter i. For example in a binary election μ_i might specify that voter i votes for Clinton with probability .55, and votes for Trump with probability .45. These probabilities may be different for different voters. μ_0 is the tie-breaker's distribution. For example, if $V_0 = \{\text{Heads}, \text{Tails}\}$, μ_0 might specify that a coin, which could be used to break a tie, lands on either Heads or Tails with an equal probability of .5.[8] I assume that voting behavior is independent so that the probability distribution over voter profiles is the product distribution $\mu = \prod_{i=0}^{n} \mu_i$. $\mu_{-i} = \prod_{j \neq i} \mu_j$ is the joint distribution of votes other than the vote of voter i.

2.1.3 Elections

A **voting situation** or **election** is a pair $\sigma = (f, \mu)$, where f is a voting rule, and μ is a vote profile distribution. I am after a ranking of elections $\sigma = (f, \mu)$ in terms of voting power for each voter i. It is important to emphasize that my measure of voting power depends on *both* the voting rule *and* voting behavior. It is not sufficient to know only the voting rule.[9]

2.1.4 Comments

Independence

The assumption that different voters' vote probabilities are independent is unrealistic. Think, for example, about James Comey's announcement days before the 2016 US presidential election, which was plausibly a common cause of a number of voters' voting decisions. There are many factors that may correlate the votes of different voters. The votes of voters in the same social network, the same geographic area, or who watch the same television station may be correlated. Knowing how Ohio voted might cause you to update your beliefs about how Michigan voted. Gelman et al. (2004) argue that standard

[8] Assume for simplicity that the coin is flipped simultaneously with the voting, regardless of whether there is a tie, but we only appeal to the outcome of the flip if there is a tie.

[9] Laruelle and Valenciano (2005) also emphasize the importance of both ingredients.

voting indices – in particular the Banzhaf power index – are not empirically adequate because they ignore correlation of votes.

Independence is unrealistic, but my purpose here is largely orthogonal to the question of whether votes are correlated: It is to provide a new foundation for voting power measures and to challenge pivotality as a basis for measuring voting power. I do this in the simplest and most extensively studied context: that of independent votes. Considering correlated votes introduces some subtle conceptual issues that I wish to avoid.[10] Future work will explore correlated votes.

Preferences

I treat the distribution over votes as a primitive. However, one would think that votes are mediated by preferences: So that preferences are (randomly) determined, and then preferences causally influence votes. So much of the randomness in votes might be explained by randomness in preferences.

This is true, but I will ignore preferences in this paper. This paper aims to develop a measure of voting power as a function of voting rules f and voting behavior μ. To measure power, at least on the conception I present here, it is not necessary to know preferences if one knows voting behavior.

If one has a theory of how behavior arises from preferences (e.g., Bayesian Nash equilibrium with an equilibrium selection rule) then one can use my measure to measure voting power on the basis of voting rules and the distributions of *preferences*. However one does not need a *theory* of voting behavior to *measure* voting power.

Causal influence of the voting rule on behavior

The choice of voting rule f doubtless causally influences the distribution of votes μ. This influence may be mediated through preferences and equilibrium, as discussed above. In light of this, it is important to specify what the theory presented here can and cannot do. The theory can answer counterfactual questions of the form "How would voting power change if the voting

[10]See Machover (2007) and Bovens and Beisbart (2011).

rule were to change from f to f' *and* behavior changed from μ to μ'?”, but without further assumptions, it cannot answer questions of the form “How would voting power change if the voting rule changed from f to f'?” However, given any theory of how voting rules determine voting behavior as input, the theory developed here can answer the second question.

Current follow-up research considers the question of multicandidate voting rules that are *optimal* with regard to voting power. This does require additional assumptions – at least weak assumptions – about how voting rules determine voting behavior. However, without a way of measuring voting power in multicandidate elections, it is impossible to even ask the question of which voting rules are optimal with respect to voting power. So, by providing a measure of voting power, this paper does make a contribution to answering to such questions.

2.2 General strategy: Rank menus to rank elections

I am after a ranking of voting situations $\sigma = (f, \mu)$ in terms of voting power. This section lays out my approach. To this end, imagine that instead of selecting candidates by election, society selected candidates by lottery. Recall that $C = \{1, \ldots, m\}$ is the set of candidates. A **lottery** ℓ is a probability distribution over candidates. ℓ_c is the probability ℓ assigns to candidate c. Formally, a lottery is a vector $\ell = (\ell_c : c \in C)$ in $\mathbb{R}^C$ satisfying $\ell_c \geq 0$ for all $c \in C$ and $\sum_{c \in C} \ell_c = 1$. $\Delta(C)$ is the set of all candidate lotteries.

Now imagine that instead of a candidate being selected by a predetermined lottery, a voter may select a lottery from a menu of lotteries $M = \{\ell, \ell', \ell'', \ldots\}$. In fact, an election essentially gives each voter a menu of lotteries. Consider the election between Clinton and Trump. If Ann votes for Clinton, then this leads to a lottery ℓ^C in which Clinton wins with some probability and Trump wins with some probability. If instead, Ann votes for Trump, this leads to a different lottery ℓ^T in which Trump wins with a slightly higher probability. In a large election, the different lotteries available to a voter are typically only very slightly different, but they are different. The lotteries ℓ^C and ℓ^T that Ann faces are determined by the voting rule f (e.g., electoral college vs. national popular vote) and by the probability distribution of others' votes μ_{-i}, where $i =$ Ann. The menu of votes $\{\text{Clinton}, \text{Trump}\}$ corresponds to a menu of lotteries $\left\{\ell^C, \ell^T\right\}$, and thus, with regard to Ann's voting decision, the election between Clinton and Trump also corresponds to the menu $\left\{\ell^C, \ell^T\right\}$.

In any election, if voter i selects a vote v_i, this will lead a lottery over candidates in the manner described above. The candidates may no longer be just two. Let $\ell_i(v_i|\sigma)$ be the lottery that voter i selects by selecting vote v_i in voting situation $\sigma = (f, \mu)$. $[\ell_i(v_i|\sigma)]_c = \mu_{-i}(\{v_{-i} : f(v_i, v_{-i}) = c\})$ is the probability that this lottery assigns to candidate c.[11] Define i's **electoral menu** in σ to be

$$M_i(\sigma) = \{\ell_i(v_i|\sigma) : v_i \in V\}. \tag{1}$$

Voter i's electoral menu is the menu of lotteries that i can select by selecting some vote. This leads to a key idea, represented schematically as:

Key Idea: *Rank Menus* $\Rightarrow$ *Rank Elections.*

The idea is that if we can rank menus in terms of the power they allow, this will lead to a ranking of voting institutions. In particular

- *Voter i's power in election σ is given by i's power in the menu $M_i(\sigma)$.*

To rank menus, I employ a ranking of menus in terms of *freedom of choice*.

2.3 Measuring freedom

This section presents my ranking of menus in terms of freedom, which was introduced and studied in Sher (2018*b*). The ranking is justified axiomatically. While I briefly justify the axioms, there is not space for detailed arguments in their favor here. Such detailed arguments, discussion of limitations of the axioms, and additional results can be found in Sher (2018*b*).

[11]Observe that the validity of the expression $[\ell_i(v_i|\sigma)]_c = \mu_{-i}(\{v_{-i} : f(v_i, v_{-i}) = c\})$ depends on the assumption that different voters vote independently.

Let $Z = \{1, \ldots, m\}$ be a finite set of outcomes. In the electoral context, $Z = C$. However, to emphasize that the ranking applies to individual choice more generally, not just to elections, I distinguish Z from C. Z is any set of outcomes that an individual might care about. For example, the outcomes in Z might correspond to the possibilities of winning or losing a contest. Different actions that the agent may take may lead to different winning probabilities, or, in other words, different lotteries over winning and losing. We can then identify the actions with the lotteries to which they lead. We want to know how much freedom the agent has in virtue of the lotteries that she can bring about through her actions. The set of such lotteries is her menu. I assume that the agent's menu contains lotteries rather than outcomes because I want to be able to measure freedom when the agent has *imperfect control*.

Let $\Delta(Z)$ be the set of all lotteries over Z. A generic lottery is denoted by $\ell = (\ell_z : z \in Z)$. A menu of lotteries is a closed subset of $\Delta(Z)$. $\mathscr{M}$ is the set of all menus of lotteries.

2.3.1 Axioms

I now present the axioms for a freedom ranking $\precsim$ on the set of menus $\mathscr{M}$.

Axiom 1 ***(Quasiorder)*** *$\precsim$ is transitive and reflexive.*

Observe that I do not assume that $\precsim$ is complete; that is, there may be menus that are incomparable in terms of freedom.

The next three axioms are due to Dekel, Lipman, and Rustichini (2001) who axiomatize preference for flexibility, rather than freedom.[12] Sher (2018*b*) discusses the relationship between the interpretation of the axioms in terms of preference for flexibility and in terms of freedom of choice in more detail.

Axiom 2 ***(Continuity)*** *For any convergent sequences $\{M_i\}$,$\{N_i\}$ in $\mathscr{M}$, and $M, N \in \mathscr{M}$,* $(M_i \to M$ and $N_i \to N$ and $[\forall i, M_i \precsim N_i]) \Rightarrow M \precsim N$.[13]

[12] Dekel et al. (2001) employ a weak order axiom rather than a quasiorder axiom. That is, they assume that $\precsim$ is complete. Kochov (2007) and Galaabaatar (2010) characterize a version of preference for flexibility with the quasiorder axiom. The version of the continuity axiom used here is also due to them.

[13] Convergence is in the metric topology generated by the Hausdorff distance.

Axiom 3 ***(Opportunity Monotonicity)***
$\forall M, N \in \mathscr{M}, M \subseteq N \Rightarrow M \precsim N$.

It is intuitive that a larger menu – larger in the sense that more alternatives have been added – provides more freedom. That is what opportunity monotonicity says.

The next axiom requires a little machinery. For lotteries $\ell, \ell' \in \Delta(Z)$ and $\lambda \in [0,1]$, the mixture $\lambda\ell + (1-\lambda)\ell'$ is the lottery that assigns to each outcome z the probability $\lambda\ell_z + (1-\lambda)\ell'_z$. The *mixture of menus* $\lambda M + (1-\lambda) M'$ is defined by taking the mixtures of all selections of lotteries in those menus,

$$\lambda M + (1-\lambda)M' := \{\lambda\ell + (1-\lambda)\ell' : \ell \in M, \ell' \in M'\}.$$

For future reference, it is useful to generalize this operation to any number of menus. Let $(p_1, \ldots, p_h) \in \mathbb{R}^h$ be a probability vector. That is, $p_k \geq 0$ for $k = 1, \ldots, h$ and $\sum_{k=1}^h p_i = 1$. For any set of h lotteries $\left(\ell^1, \ldots, \ell^h\right)$, define $\sum_{k=1}^h p_k \ell^k$ to be the mixed lottery that puts probability $\sum_{k=1}^h p_k \ell_z^k$ on each outcome z. For any collection $(M_1, \ldots, M_h)$ of menus of lotteries, define the **Minkowski average**

$$\sum_{k=1}^h p_k M_k = \left\{ \sum_{k=1}^h p_k \ell^k : \ell^k \in M_k \text{ for } k = 1, \ldots, h \right\}. \tag{2}$$

Thus, $\sum_{k=1}^h p_k M_k$ is the set of all weighted averages of selections from the menus $(M_1, \ldots, M_h)$, where the weight on the selection from M_k is p_k.

$\sum_{k=1}^h p_k M_k$ is a genuine menu of lotteries – i.e., an element of $\mathscr{M}$. However, we can also regard $\sum_{k=1}^h p_k M_k$ as a lottery over menus (of lotteries) because $\sum_{k=1}^h p_k M_k$ is the set of lotteries that an agent could generate (ex ante) by selecting a (pure) strategy for selecting a lottery from each menu M_k given that she faces each menu M_k with probability p_k. For this reason, I also refer to $\sum_{k=1}^h p_k M_k$ as a **menu lottery**. As a special case, $\lambda M + (1-\lambda) M'$ can be regarded as a lottery over the two menus M and M'. We are now in a position to present the next axiom.

Axiom 4 ***(Independence)***
$\forall \lambda \in (0,1), \forall L, M, N \in \mathscr{M}, \quad M \precsim N \Leftrightarrow \lambda M + (1-\lambda)L \precsim \lambda N + (1-\lambda)L.$

To understand why this axiom is compelling for freedom, I present an example which was previously presented in Sher (2018*a*) and Sher (2018*b*). Consider an agent who faces the prospect of going to prison outside of North America with probability $1-\lambda$. If the agent does not go to prison, she will either become a US citizen or a Canadian citizen. Let US be the choice set that the agent would have in virtue of becoming a US citizen and C the choice set that she would have in virtue of being a Canadian citizen. P is the choice set the agent would have in virtue of being in prison. Scenario A, in which the agent becomes a US citizen if she does not go to prison, is represented by the menu $\lambda US + (1-\lambda)P$ and Scenario B, in which she becomes a Canadian citizen if she does not go to prison, is represented by $\lambda C + (1-\lambda)P$. The independence axiom applied to this case says that

$$C \precsim US \Leftrightarrow \lambda C + (1-\lambda)P \precsim \lambda US + (1-\lambda)P.$$

That is, the agent has more freedom in Scenario A than in Scenario B if and only if she has more freedom in the US than in Canada. This is intuitively correct: the possibility of going to prison, which occurs with the same probability in both scenarios, cancels out. This is very similar to the intuition for the independence axiom in standard expected utility theory, except that here it applies to freedom rather than to preference.

The final axiom is due to Sher (2018*b*). Some machinery is required. A **permutation** π of Z is a bijection $\pi : Z \to Z$. Π is the set of permutations on Z. For each $\ell \in \Delta(Z)$, and $\pi \in \Pi$, define $\ell^{\pi} \in \Delta(Z)$ by:

$$\ell^{\pi}_z = \ell_{\pi(z)}, \qquad \forall \pi \in \Pi, \forall z \in Z.$$

So ℓ^{π} is the lottery over outcomes that results from permuting the probability of outcomes in ℓ according to π: that is, ℓ^{π} assigns to z the same probability that ℓ assigns to $\pi(z)$. The menu M^{π} is formed by permuting each lottery in M according to π. That is:

$$M^{\pi} = \{\ell^{\pi} : \ell \in M\}. \tag{3}$$

Axiom 5 ***(Neutrality)*** $M \sim M^{\pi}, \quad \forall M \in \mathscr{M}, \forall \pi \in \Pi.$

The axiom asserts indifference between a menu and any permutation of that menu. In other words, the evaluation of the menu is blind to the identity of the particular items that feature in the menu's lotteries, and depends only on the *structure* of the menu. This axiom would be too strong as an axiom for preference for flexibility, but it can be reasonable as a normative axiom about freedom if we want the evaluation of freedom to capture only the scope for choice and to be neutral about different alternatives. This is particularly natural in the context of *voting power*: It is natural to define a voter's power in a way that is independent of the identity of the candidates. The most commonly used voting power indices, including the Banzhaf and Shapley-Shubik indices exhibit this kind of neutrality towards candidates. However, I discuss relaxations of the neutrality axiom in Section 2.4 and footnote 14.

I refer to quasiorder, continuity, order monotonicity, independence, and neutrality collectively as the **freedom axioms**.

2.3.2 The symmetry order

This section presents the symmetry order, which serves as the basis for my proposed voting power order, and relates it to the freedom axioms. For any $M \in \mathscr{M}$, let $\mathrm{co}(M)$ be the convex hull of M. For any menu M, define the **symmetrization** of M by

$$S(M) := \mathrm{co}\left[\frac{1}{m!}\sum_{\pi\in\Pi} M^{\pi}\right]. \tag{4}$$

$S(M)$ is the Minkowski average (see (2)) of all $m!$ permutations of M (see (3)). This definition involves a lot of averaging: First, we take an (equal weight) Minkowski average of all $m!$ permutations of M; then we take a convex hull of the resulting menu, which amounts to taking all weighted averages of lotteries in $\frac{1}{m!}\sum_{\pi\in\Pi} M^{\pi}$.

The **symmetry order** $\precsim^{*}$ on $\mathscr{M}$ is defined by:

$$M \precsim^{*} M' \Leftrightarrow S\left(M\right) \subseteq S\left(M'\right), \quad \forall M, M' \in \mathscr{M}. \tag{5}$$

That is, the symmetry order ranks menu M' above menu M if the symmetrization of M' contains the symmetrization of M. $\prec^{*}$ is the asymmetric part of the symmetry order. Sher (2018*b*) presents some diagrams that illustrate the symmetrization of a menu (4) and the symmetry order (5) in simple cases.

Say that one quasiorder $\precsim$ is **coarser** than another order $\precsim'$ if $\forall M, M' \in \mathscr{M}, M \precsim M' \Rightarrow M \precsim' M'$. The symmetry order is justified by the following proposition, due to Sher (2018*b*):

Theorem 1 *The symmetry order is the coarsest quasiorder satisfying the freedom axioms.*

This means that $M \precsim^{*} M'$ exactly when it is a *consequence* of the freedom axioms that M' is ranked above M – that is, whenever *all* quasiorders satisfying the axioms rank M' above M. However, if it is merely *consistent* with the axioms that M is ranked above M', but not *implied* by the axioms, then $M \not\precsim^{*} M'$. In this sense, $\precsim^{*}$ is not like a subjective preference, but is rather principle-based: It encodes what is implied by a certain set of principles.

2.4 From freedom to voting power

Recall that the electoral menu $M_i\left(\sigma\right)$ is the set of candidate lotteries that the voter can achieve by varying her vote in σ (see (1)). Let $\sigma = (f, \mu)$ and $\sigma' = (f', \mu')$ be two elections. I define the **symmetric power order** $\precsim_i^{\circ}$ on elections in terms of the symmetry order $\precsim^{*}$ on menus by a condition, which I call the **freedom-voting power translation**:

$$\sigma \precsim_i^{\circ} \sigma' \Leftrightarrow M_i\left(\sigma\right) \precsim^{*} M_i\left(\sigma'\right). \tag{6}$$

The symmetric power order $\precsim_i^\circ$, unlike the symmetry order $\precsim^*$, is indexed by a voter i. It is a voting power order for a single voter. One can also define the **average voting power order** $\precsim^\circ$ on elections:

$$\sigma \precsim^\circ \sigma' \Leftrightarrow \frac{1}{n}\sum_{i=1}^{n} M_i(\sigma) \precsim^* \frac{1}{n}\sum_{i=1}^{n} M_i(\sigma'), \tag{7}$$

where (7) appeals to the notion of a Minkowski average (see (2)). The average voting power in σ is the voting power that a voter would have if she were selected to occupy the role of each voter i in σ with probability $\frac{1}{n}$.

My proposal bundles together two aspects: (i) that voting power is defined by $\sigma \precsim_i^\circ \sigma' \Leftrightarrow M_i(\sigma) \precsim M_i(\sigma')$ for *some* menu order $\precsim$, and (ii) that the menu order in (i) is the symmetry order: $\precsim = \precsim^*$. These two aspects can be separated. Appendix A provides alternative orders $\precsim$ that can be used to define voting power in (i). Some of these alternatives are *coarsenings* of $\precsim^*$ and others are *refinements*. One might prefer a coarsening if one thinks the freedom axioms are too strong. Conversely, one might prefer a refinement if one is concerned that the symmetry order, and hence the symmetric power order, is incomplete and leaves too many pairs of elections unranked. Indeed one may refine $\precsim^*$ to the extent that the resulting voting power order is complete. A virtue of refinements of $\precsim^*$ is that they can be constructed to satisfy all of the freedom axioms, which is desirable if one finds these axioms to be normatively compelling.

Generalizing in another way, if one does not like the neutrality axiom (Axiom 5), one could substitute for $\precsim^*$ in (6) an arbitrary menu preference $\precsim$ satisfying the Dekel et al. (2001) axioms (the axioms in Section 2.3.1 other than neutrality, with the addition of a completeness axiom) or quasiorder version of the Dekel Lipman Rustichini axioms (i.e., a version without the completeness axiom) (Kochov 2007, Galaabaatar 2010). This would produce a voting power order that is similar to the one studied here, but treats the option to choose some candidates as more significant to voting power than the option to choose others.[14] A discussion of when the neutrality axiom is

[14]One might even explore a scheme whereby the preference $\precsim_i$ used in (6) in place of $\precsim^*$ is personalized to i so that, e.g., options more likely to be chosen by μ_i would

and is not reasonable (in settings more general than elections) can be found in Section 3.3 of Sher (2018*b*).

I conclude this section by mentioning two types of elections under which the symmetric power order simplifies. An election σ is **anonymous** if $M_i(\sigma) = M_{i'}(\sigma)$ for all $i, i' \in I$. An election σ is **neutral for** i if for all $\pi \in \Pi$, $[M_i(\sigma)]^\pi = M_i(\sigma)$ (see (3)). An election σ is **neutral** if if is neutral for all i. Intuitively, an election is anonymous if it treats all voters in the same way, and it is neutral if it treats all candidates in the same way. Anonymity and neutrality, as I have defined them, depend not only on f but also on μ. Examples of neutral and anonymous elections are given in Section 5. For anonymous elections, $\precsim^\circ = \precsim_i^\circ, \forall i \in I$. That is, we only have to look at the perspective of a single "representative" voter to gauge average voting power.

Proposition 1 *If σ and σ' are neutral, then $\sigma \precsim_i^\circ \sigma' \Leftrightarrow M_i(\sigma) \subseteq \text{co}[M_i(\sigma')]$.*[15]

This proposition says that for neutral elections, to check whether $\sigma \precsim_i^\circ \sigma'$, we have to check only whether every lottery in the electoral menu $M_i(\sigma)$ can be formed as a weighted average of lotteries in $M_i(\sigma')$.

3 Binary elections

This section applies the symmetric power order to binary elections, the most extensively studied elections in the voting power literature. I show that in "ordinary" elections – the sorts of elections that have been studied – the symmetric power order coincides with the standard pivotality approach. This provides a novel foundation for pivotality measures for binary elections: If one accepts the freedom axioms (Section 2.3) and the freedom-voting power translation (Section 2.4), then one should accept pivotality as a measure of voting power. If one asks: "Why should I accept pivotality as the measure of voting power?", the freedom axioms provide the answer.

be more likely to considered valuable in the additive expected utility representation of $\precsim_i$ (see Definition 4 and Theorem 4 of Dekel et al. (2001)). Working this out properly would require establishing a connection between voting behavior and evaluation of voting institutions in terms of voting power that goes beyond the current paper, but would be an interesting direction for future work.

[15]Observe that $M_i(\sigma) \subseteq \text{co}[M_i(\sigma')]$ is equivalent to $\text{co}[M_i(\sigma)] \subseteq \text{co}[M_i(\sigma')]$.

However, problems with pivotality are already inherent in the binary case. When one considers "strange" elections, the symmetric power order deviates from pivotality, and, on an intuitive level, the symmetry order seems to give the right answer and pivotality the wrong answer. One might think that these strange elections are just curiosities, and so carry no larger meaning. However, Section 4 will show that the reasons that pivotality in general fails to be a good measure in multicandidate elections are the same as the reasons that pivotality fails to be a good measure for strange binary elections.

3.1 Monotonic elections

3.1.1 Monotonicity

Define a **binary election** to be an election with two candidates, $C = \{1, 2\}$, *and* for which the vote set is $V = \{1, 2\}$; that is, a vote is either a vote for 1 or for 2. Note that $V = \{1, 2\}$ is part of the definition of a binary election.[16] Recall that v_{-i} is the vote profile consisting of the votes of all voters other than i (including the tie-breaker). $(1, v_{-i})$ is the vote profile in which i votes for 1 and all other voters vote as in v_{-i}. Similarly, $(2, v_{-i})$ is the vote profile in which i votes for 2, and all other voters vote as in v_{-i}.

Definition 1 *A voting rule f is **monotonic** if for all voters i in I and vote profiles $v_{-i} \in V_{-i}$, $f(1, v_{-i}) = 2 \Rightarrow f(2, v_{-i}) = 2$.*

That is, if candidate 2 wins when i votes for 1, and the only change is that i changes her vote to 2, then 2 still wins. The definition implies symmetrically that switching one's vote from 2 to 1 can only shift the election to 1, if it has any effect on the election at all.[17] Monotonicity is a very natural condition. Voting for a candidate should not cause the other candidate to win. It would be strange if this happened. A notational caveat: monotonicity applied to elections should not be confused with *opportunity monotonicity* axiom (Section 2.3). This terminological clash arises from merging two literatures.

[16] An alternative definition of binary elections would assume that $C = \{1, 2\}$ but would allow V to be arbitrary. The definition I choose simplifies the presentation. Section 4 covers $C = \{1, 2\}$ and V arbitrary as a special case.

[17] By contraposition, the conditional in Definition 1 is equivalent to $f(2, v_{-i}) = 1 \Rightarrow f(1, v_{-i}) = 1$.

3.1.2 Pivotality

I now define pivotality in binary elections. I begin by defining the notion of a **pivotal event**:

$$\text{Pivotal Event}_i^f = \{v_{-i} \in V_{-i} : f(1, v_{-i}) \neq f(2, v_{-i})\}. \quad (8)$$

That is, the pivotal event Pivotal Event_i^f is the set of vote profiles v_{-i} for voters other than i (including the tie-breaker) such that if i changes her vote, the outcome of the election changes. That is, the pivotal event is the event in which i's vote makes a difference.

Define i's pivotality probability $\text{Piv}_i(\sigma)$ in election σ by

$$\text{Piv}_i(\sigma) = \mu_{-i}\left(\text{Pivotal Event}_i^f\right) \quad (9)$$

$\text{Piv}_i(\sigma)$ is the probability that i's vote makes a difference σ. Whereas the pivotal event (8) depends only on the voting rule f and not on voting behavior μ, the pivotality probability (9) depends on both f and μ_{-i}.

3.1.3 Equivalence of pivotality and the symmetric power order

I now relate pivotality to the symmetric power order:

Theorem 2 *For monotonic binary elections,* $\sigma \precsim_i^\circ \sigma' \Leftrightarrow \text{Piv}_i(\sigma) \leq \text{Piv}_i(\sigma')$.

This result shows that in monotonic binary voting situations, voting power – as measured by the symmetry order – coincides with pivotality. When voters vote independently, and each voter is equally likely to vote for either candidate, pivotality is known as **Banzhaf power**, which is the most prominent voting power index. So the approach of the current paper coincides with the most prominent voting power index for binary elections.

Now I explain why Theorem 2 is true. When there are only two candidates, we can identify any lottery over candidates ℓ with the probability ℓ_2 that the candidate gives to candidate 2. Thus a menu of lotteries can be regarded as a set of numbers, or in other words, a subset of the interval $[0, 1]$. In particular, for binary elections, we can write $M_i(\sigma) =$

$\{[\ell_i(1|\sigma)]_2, [\ell_i(2|\sigma)]_2\}$, where, recall, $[\ell_i(1|\sigma)]_2$ is the probability that 2 wins if i votes for 1, and $[\ell_i(2|\sigma)]_2$ is the probability that 2 wins if i votes for 2. Define $\eta_i^\sigma = [\ell_i(2|\sigma)]_2 - [\ell_i(1|\sigma)]_2$. Thus η_i^σ is the increased probability of a victory for 2 that results from i voting for 2 rather than 1. It is easy to see that for binary monotonic elections $\mathrm{Piv}_i(\sigma) = \eta_i^\sigma$. Moreover, regarding sets of binary lotteries a subsets of $[0,1]$ as above, and using (4), a simple calculation shows that for binary monotonic elections $S(M_i(\sigma)) = \left[\frac{1}{2} - \frac{\eta_i^\sigma}{2}, \frac{1}{2} + \frac{\eta_i^\sigma}{2}\right]$. That is, $S(M_i(\sigma))$ is an interval of length η_i^σ, centered on $\frac{1}{2}$. Using (6) and (5) it follows that

$$\sigma \precsim_i^\circ \sigma' \Leftrightarrow M_i(\sigma) \precsim_i^* M_i(\sigma') \Leftrightarrow S(M_i(\sigma)) \subseteq S(M_i(\sigma'))$$

$$\Leftrightarrow \left[\frac{1}{2} - \frac{\eta_i^\sigma}{2}, \frac{1}{2} + \frac{\eta_i^\sigma}{2}\right] \subseteq \left[\frac{1}{2} - \frac{\eta_i^{\sigma'}}{2}, \frac{1}{2} + \frac{\eta_i^{\sigma'}}{2}\right] \Leftrightarrow \eta_i^\sigma \leq \eta_i^{\sigma'} \Leftrightarrow \mathrm{Piv}_i(\sigma) \leq \mathrm{Piv}_i(\sigma').$$

This establishes Theorem 2.[18]

A notable feature of Theorem 2 is that whereas in general the symmetry order, and hence also the symmetric power order, is only a quasiorder – it is not complete – the symmetric power order is complete on the set of monotonic binary voting situations. This is explained by the fact that when there are only two outcomes (or candidates), $\precsim^*$ is the *unique* order satisfying all the freedom axioms. Note that Theorem 2 relies on *all* of the freedom axioms. In the appendix, I display a weak order $\precsim$ that satisfies all of the axioms except independence and a weak order $\precsim'$ that satisfies all of the axioms except neutrality such that if $\precsim$ or $\precsim'$ were substituted for $\precsim^*$ in the definition (6) of $\precsim_i^\circ$, Theorem 2 would fail.

3.2 Nonmonotonic elections

The voting power literature has focussed almost exclusively on monotonic elections. This is justified because nonmonotonic binary elections are strange. However, it is interesting and – as we shall see – important to note that in nonmonotonic elections the connection between pivotality and voting power breaks down.

[18] I am grateful to an anonymous referee for suggesting that I explain the argument along these lines.

Proposition 2 *There exist nonmonotonic binary elections σ in which 1. Voter i's pivotality attains its* maximum *possible value:* $\mathrm{Piv}_i(\sigma) = 1$, *while 2. Voter i has the* minimum *possible voting power:* $\sigma \precsim^{\circ} \sigma'$ *for all* σ'.

This result shows not only that without monotonicity the relationship between pivotality and the symmetry order breaks down, but that the two can be diametrically opposed. This result is established by the following example.

The matching pennies election

Suppose there are two candidates, 1 and 2, and two voters, Ann and Bob. Each voter may vote for either 1 or 2. The voting rule is:

- If both voters vote for the same candidate, then candidate 2 wins.
- If both voters vote for different candidates, then candidate 1 wins.

This election is nonmonotonic: If, starting from a situation in which both voters vote for 1, one voter switches her vote to 2, this change will cause 1 to win. Voting behavior is as follows: Each voter votes for each candidate with probability 1/2 and the votes of the two voters are independent.[19] Refer to this situation as the **matching pennies election**, denoted σ^{mp}.

In σ^{mp}, each voter is pivotal with probability 1: It is always true that if either voter changes her vote, the winner changes. This is of course the maximum possible pivotality; it is not possible to be pivotal with a probability greater than 1. However, despite this high level of pivotality, neither voter has any meaningful control. Look at the situation from Ann's point of view: If Ann were to vote for candidate 1, then Ann knows only that candidate 1 will win with probability 1/2 (depending on whether Bob votes for 1 or 2). If Ann were to vote for candidate 2, then again candidate 1 will win with probability 1/2. No matter how Ann votes, this will lead to each candidate winning with probability 1/2. Indeed, in this situation each voter has the minimum possible voting power according to $\precsim_i^{\circ}$. To see this, observe

[19] The model does not formally assume that voters have preferences, but if voters did have preferences, the behavior assumed here would be in equilibrium, whatever those preferences might be.

that the electoral menu for each voter contains only the single lottery according to which each candidate wins with probability 1/2. Treating menus of binary lotteries as subsets of $[0,1]$ as in Section 3.1.3, $M_i(\sigma^{mp}) = \left\{\frac{1}{2}\right\}$ and $S(M_i(\sigma^{mp})) = \left\{\frac{1}{2}\right\} \subseteq \left[\frac{1}{2} - \frac{|\eta_i^\sigma|}{2}, \frac{1}{2} + \frac{|\eta_i^\sigma|}{2}\right] = S(M_i(\sigma))$ for all binary elections σ. (The absolute value $|\eta_i^\sigma|$ is used here because for nonmonotonic elections, it is possible that $\eta_i^\sigma < 0$.) It follows that $\sigma^{mp} \precsim_i^\circ \sigma$ for all binary elections σ.[20] So while each voter has maximum pivotality, each voter has minimum control. Thus the symmetry order gives the intuitive conclusion that neither voter has any power in this election. In light of this example, it is natural to conclude that in general voting power and pivotality are not the same.

3.3 Pivotality vs influence

There is another way of looking at the results of the previous sections. Define a voter's **influence** by

$$\mathrm{Inf}_i(\sigma) = |\Pr_\sigma(2 \text{ wins} \,|\, i \text{ votes for } 2) - \Pr_\sigma(2 \text{ wins} \,|\, i \text{ votes for } 1)| \qquad (10)$$

where $\Pr_\sigma$ is the probability measure induced by σ.[21] In the voting power literature, pivotality and influence are often conflated,[22] and for good reason: For monotonic elections, influence and pivotality are the same. The following proposition – a variant of the preceding results – shows that for nonmonotonic elections, pivotality and influence diverge.

[20] In fact, $\sigma^{mp} \precsim_i^\circ \sigma$ for any election σ with any number of candidates (and any number of voters). We can view σ^{mp} as a multicandidate election with n voters and m candidates in which only Ann and Bobs' votes have an impact and only candidates 1 and 2 can win. Viewed in this way, $M_i(\sigma^{mp})$ contains only the lottery in which 1 wins with probability 1/2, 2 wins with probability 1/2, and all other candidates win with probability 0. Sher (2018*b*)) establishes that singleton menus have minimal freedom according to the symmetry order: $\{\ell\} \precsim^* M$ for all lotteries ℓ and menus M. Hence viewing σ^{mp} as a multicandidate election, $\sigma^{mp} \precsim_i^\circ \sigma$ for all multicandidate elections σ.

[21] An equivalent way of defining $\mathrm{Inf}_i(\sigma)$ is as $\mathrm{Inf}_i(\sigma) = |\eta_i^\sigma|$.

[22] Some writers, such as Felsenthal and Machover (1998), distinguish these notions. Their Theorem 3.2.12 is equivalent to Part 1 of Proposition 3 under the (stronger) assumption that each voter is equally likely to vote for either candidate.

Proposition 3 *1. For monotonic binary elections,* $\mathrm{Inf}_i(\sigma) = \mathrm{Piv}_i(\sigma)$.
2. For (possibly nonmonotonic) binary voting situations, $\mathrm{Inf}_i(\sigma) \leq \mathrm{Piv}_i(\sigma)$. *Moreover, there exist nonmonotonic binary* σ *for which* $\mathrm{Piv}_i(\sigma) = 1$ *but* $\mathrm{Inf}_i(\sigma) = 0$. *3. In all binary elections – monotonic or nonmonotonic – influence represents the symmetric power order:* $\sigma \precsim_i^{\circ} \sigma' \Leftrightarrow \mathrm{Inf}_i(\sigma) \leq \mathrm{Inf}_i(\sigma')$.

When the two diverge, the symmetry order tracks influence, not pivotality.

4 Multicandidate elections

We have established that pivotality measures voting power for *monotonic* – but not *nonmonotonic* – binary elections. This section establishes a similar result for multicandidate elections. There is, however, an important difference. In the binary case, the most reasonable elections, such as majority voting, are monotonic, and nonmonotonic elections can be dismissed as curiosities. In contrast, in multicandidate elections, monotonicity becomes very demanding so that the most reasonable elections are not monotonic. Hence, it is in the multicandidate case that the symmetric power order – embodying an idea distinct from pivotality – really comes into its own.

4.1 Defining pivotality in multicandidate elections

To proceed, we must generalize pivotality to multicandidate elections. This is nontrivial: With multiple candidates, not all instances of pivotality are the same. A voter may be pivotal over 2, 3 or more candidates. For example, under plurality rule with k candidates tied for first place, each voter is pivotal over k candidates. In a random dictatorship with m candidates, each voter is pivotal over m candidates with probability $\frac{1}{n}$.

4.1.1 Pivotal events

We now define pivotal events – but unlike in the binary definition (8) – there is not just one pivotal event, but a *family* of pivotal events, indexed by the number of candidates over which the voter is pivotal:

$$\text{Pivotal Event}_i^f(k) = \{v_{-i} \in V_{-i} : |\{f(v_i, v_{-i}) : v_i \in V\}| = k\}. \tag{11}$$

To understand the notation: For any finite set X, $|X|$ is the number of elements in X. Pivotal Event$_i^f(k)$ is set of votes v_{-i} such that if the vote profile of others is v_{-i}, by varying her vote, i can bring it about that one of exactly k candidates can win. In other words, Pivotal Event$_i^f(k)$ is the event that i is pivotal over k candidates. Here k ranges between 1 and m, where m is the number of candidates. Observe that Pivotal Event$_i^f(1)$ is the event in which voter i is *not* pivotal; that is, Pivotal Event$_i^f(1)$ is the event such that every vote by i will lead the same candidate to win the election. To be pivotal over only a single candidate is not to be pivotal at all.

In principle, one could individuate pivotal events more finely: For example, one could split Pivotal Event$_i^f(2)$ into $\binom{m}{2}$ pivotal events, one for each distinct pair of candidates: the event that i is pivotal exactly between candidates 1 and 2, the event that i is pivotal between 1 and 3, and so on. In Section 4.1.4, I will explain why it is not necessary to individuate pivotal vents more finely in this way.

4.1.2 Pivotality probabilities

As there is not just one pivotal event but a collection of pivotal events, one for each number between 1 and m, there is not just one pivotality probability but a collection of pivotality probabilities:

$$\text{Piv}_i(k; \sigma) = \mu_{-i}\left(\text{Pivotal Event}_i^f(k)\right). \tag{12}$$

$\text{Piv}_i(k; \sigma)$ is the probability that voter i is pivotal over k candidates.

Recalling that Pivotal Event$_i^f(1)$ is the event that voter i is not pivotal, $\text{Piv}_i(1; \sigma)$ is the probability that i is not pivotal. Observe that $\text{Piv}_i(k; \sigma) \geq 0$ for all k, and $\sum_{k=1}^{m} \text{Piv}_i(k; \sigma) = 1$, and define i's **pivotality distribution** as $\overline{\text{Piv}}_i(\sigma) = (\text{Piv}_i(1; \sigma), \ldots, \text{Piv}_i(k; \sigma), \ldots, \text{Piv}_i(m; \sigma))$. $\overline{\text{Piv}}_i(\sigma)$ is the probability distribution of the number of candidates over which i is pivotal in σ.

4.1.3 The Pivotality Order

I now use the concepts of the preceding sections to (partially) rank multicandidate elections in terms of pivotality. For k between 1 and m, define the **deterministic menu**

$$\mathbf{k} = \{\delta_1, \delta_2, \ldots, \delta_k\},$$

where δ_c is the degenerate lottery that puts probability 1 on candidate c. Thus $\mathbf{k}$ is the menu that effectively allows you to choose among the first k candidates. For any election σ, define the **pivotality lottery** $\mathscr{L}_i^\sigma$ for i in σ:

$$\mathscr{L}_i^\sigma = \sum_{k=1}^{m} \text{Piv}_i\,(k;\sigma)\,\mathbf{k}. \tag{13}$$

$\mathscr{L}_i^\sigma$ is the special case of the Minkowski average (2) in which $p_k = \text{Piv}_i\,(k;\sigma)$ and $M_k = \mathbf{k}$. $\mathscr{L}_i^\sigma$ is the lottery according to which, with probability $\text{Piv}_i\,(1;\sigma)$, the probability with which i would not have been pivotal in σ, the voter faces menu $\mathbf{1} = \{\delta_1\}$, or, in other words, the voter has no choice; with probability $\text{Piv}_i\,(2;\sigma)$, the probability with which i would have been pivotal between two candidates in σ, i faces menu $\mathbf{2} = \{\delta_1, \delta_2\}$, so that i has a choice between two candidates; . . . ; and with probability $\text{Piv}_i\,(m;\sigma)$, i faces the menu $\mathbf{m} = \{\delta_1, \ldots, \delta_m\}$. Define the **pivotality order** $\precsim_i^{\text{piv}}$ on elections by

$$\sigma \precsim_i^{\text{piv}} \sigma' \Leftrightarrow \mathscr{L}_i^\sigma \precsim^* \mathscr{L}_i^{\sigma'}. \tag{14}$$

I claim that the pivotality order is a natural generalization of pivotality to multicandidate elections. That is, one can say that i is more pivotal in σ' than in σ if $\sigma \precsim_i^{\text{piv}} \sigma'$. That this is the correct way of generalizing pivotality is not obvious, so in the next section, I will provide some justification.

4.1.4 Justifying the Pivotality Order

First observe that the pivotality order is indeed a generalization:

Proposition 4 *For binary elections,* $\sigma \precsim_i^{\text{piv}} \sigma' \Leftrightarrow \text{Piv}_i\,(2;\sigma) \le \text{Piv}_i\,(2;\sigma')$.

But there are many ways to generalize. Why not in a different way? Wouldn't it be a more pure approach to individuate pivotal events more finely, as suggested in the end of Section 4.1.1. We could have defined a pivotal event and pivotal probability for each set S of candidates: $\text{Pivotal Event}_i^f(S) = \{v_{-i} : \{f(v_i, v_{-i}) : v_i \in V\} = S\}$ and $\text{Piv}_i(S;\sigma) = \mu_{-i}\left(\text{Pivotal Event}_i^f(S)\right)$. For each $S \subseteq C$, let $\mathbf{S} = \{\delta_c : c \in S\}$. Then we can re-define the notion of a pivotality lottery (13) so that it does not collapse all sets of candidate S of size k into the single set $\mathbf{k}$. In particular, define $\hat{\mathscr{L}}_i^\sigma := \sum_{S\subseteq C, S\neq\emptyset} \text{Piv}_i(S;\sigma)\,\mathbf{S}$.

Proposition 5 $\hat{\mathscr{L}}_i^\sigma \sim^* \mathscr{L}_i^\sigma$. *Consequently,* $\sigma \precsim_i^{\text{piv}} \sigma' \Leftrightarrow \hat{\mathscr{L}}^\sigma \precsim^* \hat{\mathscr{L}}^{\sigma'}$.

Thus, to define the pivotality order, it makes no difference whether we individuate pivotality events finely to be of the form $\mathbf{S}$ or coarsely to be of the form $\mathbf{k}$; either way, we arrive at the same pivotality order. This is a consequence of the independence and neutrality axioms.

An opposite idea is that because in large elections, at least ordinary large elections, a voter is much more likely to be pivotal over exactly two candidates than over any larger number of candidates, we should collapse all levels pivotality, and represent pivotality just by the probability of being pivotal over two candidates $\text{Piv}_i(2;\sigma)$ or the probability $1 - \text{Piv}_i(1;\sigma)$ of being pivotal over any number of candidates. The next proposition shows that the pivotality order does indeed capture this idea, under appropriate conditions:

Proposition 6
If $|\text{Piv}_i(2;\sigma') - \text{Piv}_i(2;\sigma)| > (m-1)\sum_{k=3}^m |\text{Piv}_i(k;\sigma') - \text{Piv}_i(k;\sigma)|$, *then* $\sigma \precsim_i^{\text{piv}} \sigma' \Leftrightarrow 1 - \text{Piv}_i(1;\sigma) \leq 1 - \text{Piv}_i(1;\sigma')$.[23]

Notice that what really matters for $1 - \text{Piv}_i(1;\sigma)$ to be a good measure of pivotality is that the magnitude of the *difference* between higher level pivotalities is small relative to that of 2-pivotalities. In general, however, pivotality over more than two candidates may matter. For example, does a voter have more power if she participates in majority voting over 2 candidates or if she participates in a random dictatorship over m candidates? How does

[23] It is interesting to note that $1 - \text{Piv}_i(1;\sigma) \leq 1 - \text{Piv}_i(1;\sigma')$ is *always* a necessary condition for $\sigma \precsim_i^{\text{piv}} \sigma'$, but to be a sufficient condition, the inequality in the proposition is required.

the answer depend on m and the number of voters n? A further point is that even when only pivotality over two candidates matters, we cannot in general use the single number $1 - \mathrm{Piv}_i(1;\sigma)$ to pose the question of whether pivotality overestimates or underestimates voting power. In contrast, as will be shown in Theorems 3 and 4, the pivotality lotteries $\sum_k \mathrm{Piv}(k;\sigma)\,\mathbf{k}$ used to define $\precsim_i^{\mathrm{piv}}$ can also be used to determine whether pivotality correctly measures voting power.

A final thought might be that we should measure pivotality by comparing pivotality distributions $\overline{\mathrm{Piv}}_i(\sigma)$ via first order stochastic dominance (FOSD). This would be justified by the observation that being pivotal over a greater number of candidates is superior to being pivotal over a small number. So, in general, increasing the probability that one is pivotal over at least k candidates intuitively seems to increase one's pivotality. That is precisely the criterion that FOSD captures. The following proposition bears on this.

Proposition 7 *1.*

$$\sigma \precsim_i^{\mathrm{piv}} \sigma' \Leftrightarrow \left(\sum_{k=1}^{m-j} \binom{m-k}{j} \left(\mathrm{Piv}_i(k;\sigma') - \mathrm{Piv}_i(k;\sigma)\right) \leq 0, \textit{for } j = 1, \ldots, m-1 \right).$$

2. If $\overline{\mathrm{Piv}}_i(\sigma')$ first order stochastically dominates $\overline{\mathrm{Piv}}_i(\sigma)$, then $\sigma \precsim_i^{\mathrm{piv}} \sigma'$.
3. If $\overline{\mathrm{Piv}}_i(\sigma')$ second order stochastically dominates $\overline{\mathrm{Piv}}_i(\sigma)$, then $\sigma \precsim_i^{\mathrm{piv}} \sigma'$. Neither the converse of 2 nor of 3 hold in general, so that $\precsim_i^{\mathrm{piv}}$ is strictly stronger than both first and second order stochastic dominance.

Part 1 provides a characterization of $\precsim_i^{\mathrm{piv}}$ in terms of a set of linear inequalities. It follows from the fact that the set contains *several* inequalities that $\precsim_i^{\mathrm{piv}}$ is only a quasiorder (i.e., it is not complete). Nevertheless, $\precsim_i^{\mathrm{piv}}$ provides a stronger ranking than first order stochastic dominance, or even second order stochastic dominance.[24] It is a virtue of $\precsim_i^{\mathrm{piv}}$ that it captures FOSD, but it is also a virtue that $\precsim_i^{\mathrm{piv}}$ is stronger than FOSD, because FOSD is too weak of a criterion. To see this, suppose that in situation A (denoted σ^A), voter i is a dictator who can simply decide whether candidate 1 or 2 wins. Voter

[24] Second order strochastic dominance is stronger than first order stochastic dominance. For this reason, part 3 of the proposition actually implies part 2.

i is pivotal over this pair with probability 1. In situation B (denoted σ_ε^B), with probability $1-\varepsilon$, voter i has no influence. With probability ε, voter i decides whether candidate 1, 2, or 3 wins. FOSD is unable to compare these two situations, but, intuitively, if ε is sufficiently small, voter i is more pivotal in σ^A. In contrast, Proposition 6 implies that if ε is sufficiently small, $\sigma_\varepsilon^B \precsim_i^{\text{piv}} \sigma^A$. So $\precsim_i^{\text{piv}}$ does not suffer from the same problem.

I hope that the above analysis has persuaded the reader that $\precsim_i^{\text{piv}}$ is the right way of generalizing pivotality to multicandidate elections; below, I will argue that pivotality is *not* the right way to measure *voting power* in multicandidate elections.

4.2 Monotonicity

Next I generalize monotonicity to multicandidate elections in the standard way, although with an accommodation for the possibility that votes are not rankings. I start with the case in which votes are rankings. Let $\mathscr{R}$ be the set of rankings or total orders: That is, the set of antisymmetric, transitive, and complete relations R on C.[25] A total order R ranks all candidates in C from best to worst. A total order does not allow ties. For two candidates c and c', cRc' means that c is ranked above c'. Call a voting rule f **preferential** if its vote set V is such that $V=\mathscr{R}$. The term "preferential" is used because the ranking R is typically interpreted as the voter's preference. The term **ranked voting** is also often used to refer to preferential voting systems. Examples of preferential voting rules include instant runoff voting, the Borda rule, and the Condorcet rule. For preferential rules, we express a vote profile as $\bar{R}=(v_0, R_1, \ldots, R_i, \ldots, R_n)$. Observe the $\bar{R}$ includes the tie-breaker v_0, which is not assumed to be a ranking; that is, $V_0 \neq \mathscr{R}$. We also write $\bar{R}=(R_i, R_{-i})$ where R_i is i's preferential vote in the vote profile and $R_{-i}=(v_0, R_1, \ldots, R_{i-1}, R_{i+1}, \ldots, R_n)$ is the vote profile that results from removing i's vote from the vote profile. Let $\mathscr{R}_{-i}$ be the set of all such profiles R_{-i}.

[25]Recall that a relation R on C is **antisymmetric** if and only if for all $c, c' \in C, (cRc' \text{ and } c'Rc) \Rightarrow c=c'$.

Definition 2 *Let i be a voter in I. A preferential voting rule f is **monotonic for** i if* $\forall c \in C, \forall R_i, R_i' \in \mathscr{R}, \forall R_{-i} \in \mathscr{R}_{-i}$,

$$[f(R_i, R_{-i}) = c \text{ and } \forall c' \in C, (cR_ic' \Rightarrow cR_i'c')] \Rightarrow f(R_i', R_{-i}) = c.$$

*A voting rule is **monotonic** if it is monotonic for all voters i in I.*

This sort of monotonicity is known as **Maskin monotonicity** or **strong positive association** (Muller and Satterthwaite 1977). It says that when a candidate c moves up in a voter's ranking, this cannot *cause* c to lose when c would have won otherwise.

Intuitively, monotonicity in Definition 2 generalizes monotonicity for binary elections (Definition 1) to multicandidate elections. To see this formally, in a binary election with candidates 1 and 2, reinterpret a vote for 1 as the ranking R^1 such that $1R^12$, and a vote for 2 as the ranking R^2 such that $2R^21$. Thus, we can think of vote for 1 as the claim, "I prefer candidate 1 to candidate 2". Under this translation, we can write $\mathscr{R} = V$, and then it is easy to verify that Definition 1 coincides with Definition 2.

Not all mutlicandidate elections are preferential. For example, under plurality voting, a vote just names a single candidate rather than providing a ranking, under approval voting (Weber 1977, Brams and Fishburn 1978, Brams and Fishburn 2007), a voter submits a set of candidates of which the voter approves, and Balinski and Laraki (2011) have proposed a voting system called majority judgement, under which voters assign grades to candidates, akin to letter grades that students receive in schools. For my results to apply also to nonpreferential voting rules, I generalize Definition 2. Say that a subset V' of V is ***i*-sufficient** if $\forall v_i \in V, \forall v_{-i} \in V_{-i}, \exists v_i' \in V', f(v_i, v_{-i}) = f(v_i', v_{-i})$.

Definition 3 *A voting rule f with vote set V has a **monotonic interpretation for i** if there exists an i-sufficient subset V' of V and a surjective function $\phi_i : \mathscr{R} \to V'$ such that* $\forall c \in C, \forall R_i, R_i' \in \mathscr{R}, \forall v_{-i} \in V_{-i}$,

$$[f(\phi_i(R_i), v_{-i}) = c \text{ and } \forall c' \in C, (cR_ic' \Rightarrow cR_i'c')] \Rightarrow f(\phi_i(R_i'), v_{-i}) = c.$$

Refer to ϕ_i as ***the interpretation for i****. A voting rule is* ***has a monotonic interpretation*** *if it has a monotonic interpretation for all voters i in I.*

A voting rule has a monotonic interpretation if we can reinterpret the votes as rankings in such a way that the voting rule becomes monotonic. In a voting rule with a monotonic interpretation, we can interpret the vote $\phi_i(R_i)$ *as if* it were a claim that that the voter's preference is R_i. Notice that the definition allows for the possibility $R_i \neq R'_i$ but $\phi_i(R_i) = \phi_i(R'_i)$. In other words, a single vote may have multiple interpretations. Some votes v_i may have no interpretations (if v_i is not in the range of ϕ_i). However, the i-sufficiency of the range of ϕ_i implies that the set of votes that do have an interpretation is sufficiently rich.

4.3 A basic result on pivotality and voting power

We are now in a position to begin to assess whether and when pivotality measures voting power in multicandidate elections.

Theorem 3 *Let $i \in I$ and $\sigma = (f, \mu)$. 1. For all voting situations σ,*

$$M_i(\sigma) \precsim^* \sum_k \mathrm{Piv}_i(k; \sigma)\,\mathbf{k}.$$

2. Suppose that $\mu_j(v_j) > 0$ for all $j \in I_0 \setminus i$ and $v_j \in V$. Then

$$M_i(\sigma) \sim^* \sum_k \mathrm{Piv}_i(k; \sigma)\,\mathbf{k}$$

if and only if f has a monotonic interpretation for i.

Part 2 of the theorem says that in monotonic elections, the pivotality lottery gives a voter the same voting power as participating in the election. Part 2 is the multicandidate analog of Theorem 2, which says that for monotonic binary voting mechanisms, pivotality measures voting power. The following is an immediate corollary of part 2.

Corollary 1 *Let $\sigma = (f, \mu)$ and $\sigma' = (f', \mu')$ be such that f and f' have monotonic interpretations for i. Then $\sigma \precsim_i^{\circ} \sigma' \Leftrightarrow \sigma \precsim_i^{\text{piv}} \sigma'$.*

In other words, the pivotality order and symmetric power order coincide on elections with monotonic interpretations.

What can be said about elections that do not have monotonic interpretations? Part 1 says that pivotality is always weakly an overestimate of voting power, and (by part 2) in the absence of monotonicity, a *strict* overestimate. We'll see why pivotality overestimates voting power below.

4.4 Some lemmas from social choice

An important question is whether part 2 of Theorem 3 is a positive or negative result: Does it say that pivotality is usually or a good measure of voting power or rarely a good measure? Those familiar with the social choice literature may immediately see the answer. I now review a few basic results from social choice that are relevant.

Definition 4 *A preferential voting rule is **strategyproof** if*

$$\forall i \in I, \forall R_i, R_i' \in \mathscr{R}, \forall R_{-i} \in \mathscr{R}_{-i}, f\left(R_i, R_{-i}\right) R_i f\left(R_i', R_{-i}\right).$$

In other words, a preferential voting rule is strategyproof if and only if no voter ever has an incentive to lie to the rule (i.e., misrepresent her preferences) even if she knows how others have voted.[26]

Voting rule f is **deterministic** if it does not depend on the tie-breaker v_0.[27]

Proposition 8 *(Muller and Satterthwaite 1977) A deterministic preferential voting rule is monotonic if and only if it is strategyproof.*

I refer to this as the Muller Satterthwaite theorem, although that term is often used to refer to a corollary of Proposition 8 instead.[28] The analog of

[26] This version of the definition is slightly nonstandard because my definition of R_{-i} is such that R_{-i} includes the tie-breaker v_0.

[27] Formally, a voting rule is deterministic if for all v_0, v_0', v_{-0}, $f\left(v_0, v_{-0}\right) = f\left(v_0', v_{-0}\right)$.

[28] Often the Muller Satterthwaite theorem is used to refer to the proposition that *If there are three candidates, then every preferential voting rule that is monotonic and onto*

the Muller Satterthwaithe result for binary voting rules is that voting for one's favorite candidate is a dominant strategy for all voters if and only if the voting rule is monotonic. Proposition 8 has a useful corollary.

Definition 5 *Voting rule f has a* ***strategyproof interpretation for i*** *if and only if* $\forall R_i \in \mathscr{R}, \exists v_i^{R_i} \in V, \forall v_i \in V, \forall v_{-i} \in V_{-i}, f\left(v_i^{R_i}, v_{-i}\right) R_i f\left(v_i, v_{-i}\right).$

Corollary 2 *Voting rule f has a monotonic interpretation for i if and only if f has a strategyproof interpretation for i.*

Next for any nonempty subset D of the set of candidates C and total order $R \in \mathscr{R}$, define $\hat{c}(D, R)$ to be the candidate $c \in D$ that is top ranked by R among candidates in D. Call $\hat{c}$ the **choice function**.[29] Define the **range** of a voting rule in the standard way: $\text{range}(f) = \left\{c \in C : \exists v \in \bar{V}, f(v) = c\right\}.$ The range of f is the set of candidates who can possibly win the election. A preferential voting rule is **dictatorial on its range** if there exists $i \in I$ such that $\forall R_i \in \mathscr{R}, \forall R_{-i} \in \mathscr{R}_{-i}, f(R_i, R_{-i}) = \hat{c}(\text{range}(f), R_i)$. In other words, a preferential voting rule is dictatorial on its range if one of the voters is a "dictator" such that the voting rule always chooses the dictator's favorite candidate among those that are ever chosen by the voting rule; a dictatorial voting rule simply ignores the votes of voters other than the dictator. I now state the famous Gibbard Statterthwaite theorem.

Proposition 9 *(Gibbard 1973, Satterthwaite 1975) Let f be a deterministic preferential voting rule with at least three candidates in its range. f is strategyproof if and only if it is dictatorial on its range.*

(i.e., its range is C) *is dictatorial.* In fact, in Muller and Satterthwaite (1977), this latter result is treated as a corollary, and the main result is Proposition 8. Indeed, in this section I combine Proposition 8 with the Gibbard Satterthwaite theorem to derive (a variant of) this corollary. I do this instead of working directly with the corollary because I want to emphasize the connection to strategyproofness, which I believe will be important for future work.

[29]Formally, $\hat{c}(D, R) = c$ if and only if (i) $c \in D$ and (ii) $\forall c' \in D, cRc'$.

4.5 Pivotality fails to measures voting power

4.5.1 The pivotality gap

This section uses the preceding results to show that pivotality is rarely a good measure of voting power in multicandidate elections. I start with a definition. For any voting rule f and $v_0 \in V$, define the function $f_{v_0} : V^I \to C$, the **section of f at** v_0, by $f_{v_0}(v_1, \ldots, v_n) = f(v_0, v_1, \ldots, v_n), \forall (v_1, \ldots, v_n) \in V^I$. Then observe that $\text{range}(f_{v_0}) = \{f(v_0, v_{-0}) : v_{-0} \in V^I\}$.

Definition 6 *A voting rule f is* ***dictatorial when not binary*** *if there exist functions $\chi : V_0 \to I$ and $\psi : I \times \mathscr{R} \to V$ such that for all $v = (v_0, v_1, \ldots, v_i, \ldots, v_n) \in \bar{V}$, $i \in I$, and $R \in \mathscr{R}$,*

1. $\left(|\text{range}(f_{v_0})| = 2 \text{ and } \psi(i, R) = v_i \text{ and } v \in \text{Pivotal Event}_i^f(2)\right) \Rightarrow$ $f(v) = \hat{c}(\text{range}(f_{v_0}), R)$.

2. $(|\text{range}(f_{v_0})| > 2 \text{ and } \psi(i, R) = v_i \text{ and } \chi(v_0) = i) \Rightarrow$ $f(v) = \hat{c}(\text{range}(f_{v_0}), R)$.

To understand this definition, observe first that ψ is similar to the interpretations ϕ_i in Definition 3. $\psi(i, R) = v_i$ can be interpreted to mean that vote v_i is interpreted as a claim that i's preference is R. The definition says that when the "tie-breaker" selects a vote v_0, there are two possibilities: either the election is essentially binary conditional on v_0 (only two winning candidates are possible conditional on v_0), in which case the election is essentially a monotonic binary election conditional on v_0; or, conditional on v_0, the election is essentially a dictatorship, with the dictator conditional on v_0 being $\chi(v_0)$. There are various subtleties associated with this definition, which I discuss in a footnote.[30] Using the Gibbard Satterthwaite theorem, we obtain the following result.

[30] This footnote discusses a few subtleties associated with Definition 6. First, note that $\psi(R, i)$ does not depend on v_0, so that votes must have the same interpretations independently of the tie breaker v_0. Observe next that as in the case of Definition 3, not all votes v_i need have an interpretation as a preference R_i; on top of the votes with interpretations, there can be some additional votes without interpretations. This means that the dictator $i = \chi(v_0)$, while she can get her preferred outcome by using vote $\psi(i, R_i)$ when her preference is R_i, she may also have the option to cede her power by using a vote

Proposition 10 *It f has a monotonic interpretation, then f is dictatorial when not binary.*

Interestingly the converse of this result does not hold. That is, there exist voting rules f that are dictatorial when not binary, but that are not monotonic. The appendix provides such a counterexample.

I now provide a theorem that highlights the failure of pivotality to measure voting power in multicandidate elections.

Theorem 4 *Let $\sigma = (f, \mu)$, and assume that for all $v \in \bar{V}$, $\mu(v) > 0$.*

1. *If f is not dictatorial when not binary, then there exists $i \in I$ such that*

$$M_i(\sigma) \prec^* \sum_k \mathrm{Piv}_i(k; \sigma)\, \mathbf{k}.$$

2. *If f does not have a monotonic interpretation for i, then there exists a voting rule f' such that for all μ', $(f', \mu') \prec_i^{\mathrm{piv}} (f, \mu)$ but $(f, \mu) \prec_i^{\circ} (f', \mu')$.*

Part 1 of the theorem says that whenever an election is *non-trivially* non-binary – in the sense that there is a positive probability that the rule's random input v_0 is such that conditional on v_0, three candidates may win and no voter has the dictatorial power to determine the winner – there will be a **pivotality gap**: The pivotality lottery will be a strict overestimate of voting power for some voter i. Pivotality and voting power are distinct. As the property of being dictatorial when not binary is very restrictive, this shows that for reasonable multicandidate elections, pivotality will never equal voting power.

---

v_i with no interpretation. For example, consider a voting rule such that voter 1 is the dictator who may choose any candidate she wishes, but she also has a vote v_i that cedes here power and allows voter 2 to become the dictator. This is compatible with the formal definition of the voting rule f being dictatorial when not binary.

A second subtlety is that for f to satisfy the formal definition of being dictatorial when not binary, it must be monotonic when binary. Finally, if f is dictatorial when not binary, it is also possible that conditional on some v_0, the winning candidate is determined independently of the voters' votes so that $|\mathrm{range}(f_{v_0})| = 1$.

Part 2 of the theorem shows that not only does pivotality overestimate voting power, but it does not even provide a good *ordinal* measure of voting power: σ may supply more pivotality to i than σ' while σ' gives i more voting power.[31] This part is essentially a converse of Corollary 1.

The next section presents an example that illustrates why the theorem holds.[32] It also illustrates the connection to the Muller Satterthwaite and Gibbard Satterthwaite theorems. It is striking that while the model assumes nothing about incentives – indeed voters are not even formally assumed to have preferences – limits due to incentives – the Muller Satterthwaite and Gibbard Satterthwaite theorems – translate into limits on voting power.

4.5.2 An Illustration of the Theorem

The illustration that pivotality overestimates voting power compares two situations.

Situation A: An election

Ann participates in a plurality voting election in which

- With probability .1, in votes of other voters, candidates 1 and 2 are tied.
- With probability .1, in votes of other voters, candidates 1 and 3 are tied.
- With probability .1, in votes of other voters, candidates 2 and 3 are tied.
- With probability .7, the winner does not depend on Ann's vote.

In this election, Ann must submit a vote for one of the three candidates. When she submits her vote, Ann does not know which candidates the election will come down to.

Situation B: A pivotality lottery

Ann participates in the pivotality lottery corresponding to Situation A:

- With probability .1, Ann is offered a choice between candidates 1 and 2.
- With probability .1, Ann is offered a choice between candidates 1 and 3.
- With probability .1, Ann is offered a choice between candidates 2 and 3.
- With probability .7, a candidate is installed and Ann has no choice.

This formulation of the pivotality lottery distinguishes the events of being pivotal over different pairs of candidates, but as shown by Proposition

[31] Observe that while the the antecedent in part 2 of the theorem is that f does not have a monotonic interpretation, it follows from Proposition 10 that part 2 would continue to hold if the antecedent were instead that f is not dictatorial when not binary.

[32] The section illustrates part 1 of the theorem, but as the proof of the theorem shows, the two parts are closely related.

5, Situation B is equivalent to the pivotality lottery $\sum_k \text{Piv}_i\left(k;\sigma^A\right)\mathbf{k} = (.7 \times \mathbf{1}) + (.3 \times \mathbf{2})$ corresponding to σ^A, where σ^A is the election in Situation A above.

The key difference between Situation A and Situation B is that in Situation A, Ann must vote before knowing who the election comes down to, whereas in Situation B, Ann has the opportunity to make her choice *after* it is determined who the election comes down to. In Situation A, Ann votes simultaneously with all other voters. In Situation B, it is *as if* Ann votes last, knowing how all other votes have been cast.

Intuitively, Ann has more power in Situation B than in Situation A. If Ann votes for candidate 1, then Ann's vote will make the difference if the election comes down to 1 and 2 or to 1 and 3, but Ann will effectively have no influence if the election comes down to 2 and 3. Still, Ann's vote is pivotal if there is a tie between 2 and 3 for first place (so that the tie-breaker settles the election): Had Ann voted for 2 or 3, whichever of the two she had voted for would have won, but, at the same time, the outcome in this case is the same as if Ann had not participated in the election. So while Ann is pivotal when the election comes down to 2 or 3, her actual vote for 1 has no influence. And this circumstance is unavoidable, given the rules of the election: Whichever candidate c Ann votes for, she will have no influence if the election comes down to the other two candidates.

The above discussion clearly shows why Ann's pivotality overstates her voting power in the election. In contrast, in situation B, the pivotality lottery corresponding to the election, Ann can always make full use of her pivotality. Whenever Ann is pivotal among a collection of candidates, she is told who she is pivotal among, and she may then simply choose among these candidates.

This allows us to see the relevance of the Gibbard Satterthwaite theorem for Theorem 4. Suppose that instead of participating in the election, Ann were allowed to submit a preference R_i, and whenever Ann would have been pivotal over a collection S of candidates, Ann's top choice in S according to R_i is determined to be the winner (and if Ann would not have been pivotal, the election proceeds as if Ann had submitted an ordinary vote to the election).

This would be very much like having the opportunity to vote last.[33]

The ability to effectively submit a preference that determines the outcome as described above is essentially equivalent to the voting rule's having a strategyproof interpretation, which by the Muller Satterthwaite theorem, is equivalent to the rule's having a monotonic interpretation. It is also equivalent to fully realizing one's pivotality. But by the Gibbard Satterthwaite theorem, it is in general impossible for the voting rule be such that it has a strategyproof interpretation for every voter at once. In other words, it is impossible for everyone to effectively move last, knowing how others have moved. But because not everyone's opportunities can be given a strategyproof interpretation – not everyone can effectively vote last – not everyone can fully realize their pivotality. This explains Theorem 4.

Observe that in the above example, the probability of being pivotal over more than two candidates is zero: $\mathrm{Piv}_i\left(k;\sigma^A\right) = 0$ for $k \geq 3$. Among such elections, $\mathrm{Piv}_i\left(2;\sigma\right)$ measures pivotality. Yet if σ^B is an election corresponding to Situation B in which Ann effectively moves last by submitting a preference R_i as described in the paragraph before the preceding one,[34] then $\mathrm{Piv}_i\left(2;\sigma^A\right) = \mathrm{Piv}_i\left(2;\sigma^B\right)$ but Ann has more voting power in σ^B than in σ^A. So in mutlicandidate elections, the problem for pivotality as a measure does not arise because a voter may be pivotal over more than two candidates; it arises because a voter may not *know* over which candidates she will be

[33]There is a subtle point here: Submitting a preference that determines the outcome as described above is not exactly like being able to choose last. Consider Situation B above. Suppose that if one were to move last, one would adopt the following strategy: if the choice comes down to 1 and 2, choose 1; if the choice comes down to 2 and 3, choose 2; if the choice comes down to 1 and 3, choose 3. There is no preference R_i such that choosing the top ranked element from $\{1,2\},\{2,3\}$, and $\{1,3\}$ leads respectively to choices of 1, 2, and 3. So in some sense, one would have more power if one were to choose last than if one were to submit a preference to determine the choice on one's behalf. However, the symmetry order does not treat this additional freedom as significant: The symmetry order treats the situation in which one submits a preference that "chooses last" on one's behalf and in which one really moves last as equivalent in terms of freedom. This is related to the fact that the symmetry order satisfies the axioms for preference for flexibility (Dekel et al. 2001) and so has a representation in terms of a collection of agents who value flexibility (Sher 2018*b*); to value flexibility is not to value inconsistency.

[34]It is straightforward to extend this paper's framework to dynamic elections so that what I say here about σ^B also applies to the election in which Ann literally moves last.

pivotal, and hence cannot take full advantage of her pivotality.[35]

5 Illustrations

This section illustrates the symmetric power order, focusing on techniques for calculating whether one election provides more voting power than another. In this section, it will be convenient to assume that there are $n+1$ voters instead of n, that n is divisible by 6 and that $1 \leq n \leq 10,000$.

5.1 Plurality voting vs random dictatorship

The introduction explained that with two candidates, assuming that voters independently vote for each candidate with probability $\frac{1}{2}$, voters achieve greater pivotality under majority voting than under random dictatorship. Since both majority voting and random dictatorship are monotone binary elections, pivotality measures voting power (Theorem 2). So, as in the traditional analysis, according to the symmetric power order, majority voting provides voters with more voting power than random dictatorship.

Does the above result extend to more than two candidates? In this case pivotality no longer measures voting power. Suppose there are three candidates. Under random dictatorship, each voter is selected with probability $\frac{1}{n+1}$ and then may select any candidate she wishes. Ordinarily, to fully specify an election, we must specify voting behavior. However, in random dictatorship, each voter's electoral menu is independent of others' voting behavior, so that to compare random dictatorship to other elections, we can leave voting behavior unspecified. Under plurality voting, the candidate receiving the most votes wins, and if there is a tie, it is broken randomly. Assume that each voter votes for each candidate with probability $\frac{1}{3}$ and votes are independent. Let σ_3^r and σ_3 be, respectively, random dictatorship and plurality voting with three candidates.

[35] This is closely related to the problem of "spoiler candidates" where a voter would like to express a preference for a noncompetitive candidate but only if the election is not competitive; her problem is that she does not know whether the election will be competitive.

Observe that both random dictatorship and plurality voting are anonymous and neutral elections in the sense of Section 2.4. Because of anonymity, individual i's voting power order $\precsim_i^\circ$ coincides with the average voting power order $\precsim^\circ$. Because of neutrality, to establish that voters have more voting power under plurality voting than under random dictatorship, it is necessary and sufficient to establish that every lottery in $M_i(\sigma_3^r)$ is a convex combination of lotteries in $M_i(\sigma_3)$ (Proposition 1). It is worth noting that this criterion would still be sufficient – but not necessary – if either of the elections were not neutral. As I show in the appendix:

$$\begin{aligned}
\mathrm{Piv}_i(2;\sigma_3) &= \frac{1}{3^n}\left(3\sum_{k=\frac{n}{3}+1}^{\frac{n}{2}}\binom{n}{k,k,n-2k} + 3\sum_{k=\frac{n}{3}+1}^{\frac{n}{2}}\binom{n}{k,k-1,n-2k+1}\right)\\
\mathrm{Piv}_i(3;\sigma_3) &= \frac{1}{3^n}\binom{n}{\frac{n}{3},\frac{n}{3},\frac{n}{3}}
\end{aligned} \tag{15}$$

Let $\Phi_n := \frac{1}{2}\mathrm{Piv}_i(2;\sigma_3) + \mathrm{Piv}_i(3;\sigma_3)$. Representing a lottery by a column $\ell = \begin{pmatrix}\ell_1\\ \ell_2\\ \ell_3\end{pmatrix}$, where ℓ_c is the probability of candidate c for $c = 1,2,3$, we have

$$\begin{aligned}
M_i(\sigma_3^r) &= \left\{\begin{pmatrix}\frac{n+3}{3n+3}\\ \frac{n}{3n+3}\\ \frac{n}{3n+3}\end{pmatrix}, \begin{pmatrix}\frac{n}{3n+3}\\ \frac{n+3}{3n+3}\\ \frac{n}{3n+3}\end{pmatrix}, \begin{pmatrix}\frac{n}{3n+3}\\ \frac{n}{3n+3}\\ \frac{n+3}{3n+3}\end{pmatrix}\right\}\\
M_i(\sigma_3) &= \left\{\begin{pmatrix}\frac{1+2\Phi_n}{3}\\ \frac{1-\Phi_n}{3}\\ \frac{1-\Phi_n}{3}\end{pmatrix}, \begin{pmatrix}\frac{1-\Phi_n}{3}\\ \frac{1+2\Phi_n}{3}\\ \frac{1-\Phi_n}{3}\end{pmatrix}, \begin{pmatrix}\frac{1-\Phi_n}{3}\\ \frac{1-\Phi_n}{3}\\ \frac{1+2\Phi_n}{3}\end{pmatrix}\right\}
\end{aligned} \tag{16}$$

These menus are derived in the appendix. A simple calculation shows that

$$\alpha_n\begin{pmatrix}\frac{1+2\Phi_n}{3}\\ \frac{1-\Phi_n}{3}\\ \frac{1-\Phi_n}{3}\end{pmatrix} + \alpha_n\begin{pmatrix}\frac{1-\Phi_n}{3}\\ \frac{1+2\Phi_n}{3}\\ \frac{1-\Phi_n}{3}\end{pmatrix} + (1-2\alpha_n)\begin{pmatrix}\frac{1-\Phi_n}{3}\\ \frac{1-\Phi_n}{3}\\ \frac{1+2\Phi_n}{3}\end{pmatrix} = \begin{pmatrix}\frac{n}{3n+3}\\ \frac{n}{3n+3}\\ \frac{n+3}{3n+3}\end{pmatrix}, \tag{17}$$

where $\alpha_n = \frac{(1+n)\Phi_n - 1}{3(1+n)\Phi_n}$. $\alpha_n \geq 0 \Leftrightarrow \Phi_n \geq \frac{1}{n+1}$. Direct calculation shows that $\Phi_n \geq \frac{1}{n+1}$ for all n in the range we are considering. Moreover, it is immediate that $\alpha_n \leq \frac{1}{3}$. So $1 - 2\alpha_n \geq 0$. It follows that the left hand side of (17) is a convex combination of lotteries in $M_i(\sigma_3)$. The right hand side of (17) is a lottery in $M_i(\sigma_3^r)$. Similar equations hold for the other two lotteries in

$M_i(\sigma^r)$. So $\sigma^r \precsim^\circ \sigma$; in other words, with three candidates, voters have more voting power under plurality voting than under random dictatorship.

5.2 Majority voting vs random majority voting

Let random majority voting be the preferential voting system according to which voters submit rankings of candidates (with no ties). The tie-breaker selects each pair of candidates with equal probability. If the tie-breaker selects candidates c_1 and c_2, then the submitted rankings are used to determine votes for c_1 or c_2: A vote is submitted on voter i's behalf for c_1 if i ranked c_1 above c_2 and for c_2 if i ranked c_2 above c_1. The voter who receives the most votes wins, with ties broken randomly.

We would like to compare random majority voting to standard majority voting with two *fixed* candidates. We consider this comparison formally here in part because it will be useful for the comparison of Section 5.3 below. To make the two elections comparable, let us model standard majority voting as an election with m candidates (like random majority voting), but in which there are only two votes, a vote for 1 and a vote for 2. We can assume formally that $C = \{1, \ldots, m\}$ while $V = \{1, 2\}$. Whichever of candidates 1 and 2 receives more votes wins.

Suppose that under random majority voting, each voter submits each possible ranking with the same probability, namely, $\frac{1}{m!}$, whereas under standard majority voting, each candidate submits each of the two votes, 1 and 2, with probability $\frac{1}{2}$. In both cases, votes are independent. Let $\hat{\sigma}_m$ be random majority voting and σ_2 be standard majority voting.

There are several ways of seeing that majority voting and random majorty voting are equivalent in terms of voting power. Here is one way. Both standard majority voting and random majority voting are monotonic. Therefore by Theorem 3, pivotality measures voting power in both of these elections.[36] In both elections it is only possible to be pivotal over two candidates. So, for both of these elections, voting power is measured by $\mathrm{Piv}_i(2; \sigma)$. In standard

[36] Observe that Theorem 4 does not claim that in $\hat{\sigma}_m$, pivotality is a strict overestimate of voting power because conditional on any "tie-breaker" v_0 – which is a pair of candidates c_1 and c_2 – only two candidates can win.

majority voting, each voter is pivotal with probability $\frac{1}{2^n}\binom{n}{\frac{n}{2}}$. In random majority voting, given any pair of candidates c_1 and c_2, conditional on the tie-breaker selecting c_1 and c_2, voter i is pivotal with probability $\frac{1}{2^n}\binom{n}{\frac{n}{2}}$. So overall, voter i is pivotal with probability $\frac{1}{2^n}\binom{n}{\frac{n}{2}}$. So random standard majority voting and random majority voting are equivalent in voting power.

5.3 Majority voting vs plurality voting

I now compare majority voting with two candidates – the election σ_2 defined in Section 5.2 – to plurality voting with three candidates – σ_3 defined in Section 5.1. Recall that σ_2 and σ_3 assume symmetric behavior toward the candidates on the ballot. The substantive question is: Does adding candidates increase voting power?

Recall that $\text{Piv}_i(2;\sigma_2) = \frac{1}{2^n}\binom{n}{\frac{n}{2}}$ and that $\text{Piv}_i(2;\sigma_3)$ and $\text{Piv}_i(3;\sigma_3)$ are given by (15). Direct calculation establishes that $\text{Piv}_i(2;\sigma_2) \leq \text{Piv}_i(2;\sigma_3) + \text{Piv}_i(3;\sigma_3)$. Moreover, under σ_3, it is possible to be pivotal for either two or three candidates, whereas under σ_2, it is only possible to be pivotal over two candidates. So one is both more likely to be pivotal under σ_3, and when one is pivotal under σ_3, one may be pivotal over a greater number of candidates. It follows that $\sigma_2 \prec_i^{\text{piv}} \sigma_3$. So if pivotality measured voting power, σ_3 would provide more voting power than σ_2. However, because σ_3 does not have a monotonic interpretation, pivotality does not measure voting power with regard to these elections (Theorem 3), so we must appeal to the symmetric power order.

Proposition 1 tells us that the comparisons of voting power are especially simple for neutral elections. Unfortunately, when $m = 3$, σ_2 is not neutral, because candidates 1 and 2 are treated differently from 3: 3 is not on the ballot, and so cannot win. However $\hat{\sigma}_3$ – random majority voting – is neutral when $m = 3$, and we have seen in the previous section that $\sigma_2 \sim^{\circ} \hat{\sigma}_3$. So comparing σ_2 and σ_3 is equivalent to comparing $\hat{\sigma}_3$ and σ_3. Proposition 1 now tells us that comparing $\hat{\sigma}_3$ and σ_3 is equivalent to checking whether one of $\text{co}(M_i(\hat{\sigma}_3))$ and $\text{co}(M_i(\sigma_3))$ is contained in the other, and if neither is contained in the other, the two are incomparable.

Let $U = \mathbb{R}^C$ be the set of utility functions on candidates $u = (u(c) : c \in C)$. For any utility function u and lottery ℓ, let $u \cdot \ell = \sum_{c \in C} u(c) \ell_c$ be the expected utility of ℓ for utility function u. The following corollary follows immediately from Proposition 1 using the separating hyperplane theorem.

Corollary 3 *Let σ and σ' be* neutral *elections. Then,* $\sigma \precsim_i^\circ \sigma' \Leftrightarrow \forall u \in U, \max_{\ell \in M_i(\sigma)} u \cdot \ell \leq \max_{\ell \in M_i(\sigma')} u \cdot \ell$.

This corollary is remarkable because it gives a **utilitarian** interpretation to voting power for neutral elections.[37] To apply the corollary to our question, consider two agents, Ann and Bob. Ann has a utility function on candidates u_a given by $u_a(1) = 1, u_a(2) = u_a(3) = 0$. Ann wants to *promote* 1's victory, and if one does not win, she does not care who does. Bob's utility function is u_b given by $u_b(1) = u_b(2) = 1, u_b(3) = 0$. Bob wants to *prevent* 3's victory. He does not care who wins if 3 loses.

We now calculate Ann's **indirect expected utility** $V_a(\hat{\sigma}_3)$ from having access to the menu $M_i(\hat{\sigma}_3)$, defined by $V_a(\hat{\sigma}_3) = \max_{\ell \in M_i(\hat{\sigma}_3)} u_a \cdot \ell$. $V_a(\hat{\sigma}_3)$ is what Ann's expected utility would be if she were to participate in $\hat{\sigma}_3$, taking voter i's role. Note that we are not assuming that Ann literally participates in the election. Ann is a *hypothetical* agent; she is not agent i, and her behavior is not governed by μ_i. We simply imagine as a *thought experiment* that Ann may vote however she wishes in place of i, and everyone else's behavior is held fixed at μ_{-i}. If Ann were to participate in $\hat{\sigma}_3$, she would submit a ranking with 1 on top. If the tie-breaker selects candidates 2 and 3 to participate in the election, Ann's utility is automatically 0. If the tie-breaker selects either 1 and 2 or 1 and 3 to participate in the election, then Ann's expected utility is slightly greater than $\frac{1}{2}$ because – if say 1 and 2 are selected – each wins with probability $\frac{1}{2}$ when Ann is not pivotal and when Ann is pivotal, which happens with probability $\frac{1}{2^n}\binom{n}{\frac{n}{2}}$, 1 wins, and Ann gets a utility of 1. Using these facts, $V_a(\hat{\sigma}_3) = \frac{1}{3} + \frac{1}{3}\text{Piv}_i(2; \sigma_2)$ (see the appendix). Making similar calculations for both Ann and Bob for both elections, we

[37] This connects the analysis to the literature on utilitarian assessment of voting rules (Beisbart, Bovens and Hartmann 2005, Barbera and Jackson 2006, Azrieli and Kim 2014). As explained in Section 5.4, *purely* utilitarian considerations cannot capture voting power.

attain $V_a(\sigma_3) = \frac{1}{3} + \frac{1}{3}\text{Piv}_i(2;\sigma_3) + \frac{2}{3}\text{Piv}_i(3;\sigma_3)$, $V_b(\hat{\sigma}_3) = \frac{2}{3} + \frac{1}{3}\text{Piv}_i(2;\sigma_2)$, and $V_b(\sigma_3) = \frac{2}{3} + \frac{1}{2}\left[\frac{1}{3}\text{Piv}_i(2;\sigma_3) + \frac{2}{3}\text{Piv}_i(3;\sigma_3)\right]$. It follows that

$$V_a(\hat{\sigma}_3) \leq V_a(\sigma_3) \Leftrightarrow \text{Piv}_i(2;\sigma_2) \leq \text{Piv}_i(2;\sigma_3) + 2\text{Piv}_i(3;\sigma_3)$$

$$V_b(\hat{\sigma}_3) \leq V_b(\sigma_3) \Leftrightarrow \text{Piv}_i(2;\sigma_2) \leq \frac{1}{2}\left[\text{Piv}_i(2;\sigma_3) + 2\text{Piv}_i(3;\sigma_3)\right]$$

Direct calculation shows that $V_a(\sigma_3) > V_a(\hat{\sigma}_3)$ and $V_b(\sigma_3) < V_b(\hat{\sigma}_3)$ for all n in the range we are considering. In other words, if your goal were to promote a single candidate, or in other words, bring it about that a single candidate wins, you would prefer plurality to random majority voting, and if your goal were to prevent a single candidate from winning, you would prefer random majority voting. It follows from Corollary 3 that random majority voting and plurality voting are unordered by the symmetric power order, and for good reason: plurality voting gives you more power to cause someone to win and random majority voting gives you more power to prevent someone from winning. Since $\sigma_2 \sim^* \hat{\sigma}_3$, it follows that majority voting and plurality voting are also unordered. So while pivotality orders majority voting with two candidates and plurality voting with three candidates, and prefers plurality voting, the two elections are unordered in voting power.

5.4 A final lesson

Let us generalize a little from the preceding example. Call a utility function $u \in U$ a **justification** for σ' over σ if $\max_{\ell \in M_i(\sigma)} u \cdot \ell \leq \max_{\ell \in M_i(\sigma')} u \cdot \ell$. For any permutation $\pi \in \Pi$ of candidates in C (see Section 2.3), let $u^\pi \in U$ be the utility function defined by $u^\pi(c) = u(\pi(c))$ for all $c \in C$. u^π is a utility function that, relative to u, interchanges the utilities of the candidates according to π. Call a justification u for σ' over σ **impartial** if for all permutations π, u^π is also a justification for σ' over σ. We refer to such a justification as impartial because it does not treat one candidate differently than any other candidate. A justification is **partial** if it is not impartial.

Observation 1 *For neutral elections, every justification is an impartial justification.*

For example, when comparing σ_3 and $\hat{\sigma}_3$ both of the justifications of promoting and of preventing a candidate were impartial. Viewed in this light, Corollary 3 says that whenever there is an impartial justification for σ' over σ, $\sigma' \not\precsim_i^\circ \sigma$. An impartial justification for σ' over σ – one in terms of a goal that is not specific to a given candidate but to what one can achieve were one to favor any candidate – is a good reason not to rank σ over σ' in voting power, and so provides an argument against refining the symmetric power order to make more elections comparable.

In contrast to Observation 1, when considering non-neutral elections, a partial justification for σ' over σ is consistent with the possibility that σ provides more voting power than σ'. A partial justification refers to a particular candidate and so may prefer an election because it favors that candidate rather than because it grants a certain power to the voter. For example, if u ranks candidate 1 highest, then u is a justification for the rigged election according to which 1 wins independently of voting over any other election. Clearly the rigged election gives each voter no voting power.[38] Another way of stating the point is that if the word "neutral" were taken out of Corollary 3, the corollary would cease to be true.

The question then arises as to whether it is possible to extend the notion of an impartial justification to non-neutral elections and derive an analog of Corollary 3 for all elections, neutral or non-neutral? The answer is yes.

Proposition 11 *For any elections σ and σ',*

$$\sigma \precsim_i^\circ \sigma' \Leftrightarrow \frac{1}{m!}\sum_{\pi\in\Pi} \max_{\ell\in M_i(\sigma)} u^\pi \cdot \ell \leq \frac{1}{m!}\sum_{\pi\in\Pi} \max_{\ell\in M_i(\sigma')} u^\pi \cdot \ell, \quad \forall u \in U \tag{18}$$

The expression $\sum_{\pi\in\Pi} \max_{\ell\in M_i(\sigma)} u^\pi \cdot \ell$ gives your expected utility from participating in election σ behind a veil of ignorance in which you forget which utility you assign to which candidate, assuming that when you vote you will no longer be behind the veil. This means that the symmetric power order links voting power to expected utility with a twist. It is not the expected utility of an actual voter, but of a hypothetical voter, who asks, "having

[38]This also shows why utilitarian evaluation of elections, without the qualification of impartiality, cannot capture voting power.

not yet formed my opinions on candidates, which election would I prefer to participate in, assuming I form those opinions before I vote?" Election σ' allows more voting power than election σ precisely when this sort of impartial reflection unambiguously favors σ'.

6 Conclusion

This paper has presented a new approach to voting power, founded in freedom of choice. The ideal of popular control is fundamental to democracy. However, it is difficult to formalize popular control and to evaluate competing voting institutions in terms of this ideal. Popular control is distinct from agreement. We can agree with the outcome of an election even if we had no influence in shaping it. The voting power order presented here measures the control that a voter has over electoral outcomes and it allows us to assess whether one election is better than another in terms of the control it provides.

References

Arrow, K. J. (1951), *Social Choice and Individual Values*, John Wiley & Sons.

Arrow, K. J. (1963), *Social Choice and Individual Values*, 2nd edn, Yale University Press.

Azrieli, Y. and Kim, S. (2014), 'Pareto efficiency and weighted majority rules', *International Economic Review* **55**(4), 1067–1088.

Balinski, M. and Laraki, R. (2011), *Majority judgment: measuring, ranking, and electing*, MIT press.

Banzhaf, J. F. (1964), 'Weighted voting doesn't work: A mathematical analysis', *Rutgers Law Review* **19**, 317–343.

Banzhaf, J. F. (1966), 'Multi-member electoral districts. do they violate the "one man, one vote" principle', *The Yale Law Journal* **75**(8), 1309–1338.

Banzhaf, J. F. (1968), 'One man, 3.312 votes: a mathematical analysis of the electoral college', *Villanova Law Review* **13**, 304–332.

Barbera, S. and Jackson, M. O. (2006), 'On the weights of nations: Assigning voting weights in a heterogeneous union', *Journal of Political Economy* **114**(2), 317–339.

Beisbart, C., Bovens, L. and Hartmann, S. (2005), 'A utilitarian assessment of alternative decision rules in the Council of Ministers', *European Union Politics* **6**(4), 395–418.

Bolger, E. M. (1983), 'The Banzhaf index for multicandidate presidential elections', *SIAM Journal on Algebraic Discrete Methods* **4**(4), 442–458.

Bolger, E. M. (1986), 'Power indices for multicandidate voting games', *International Journal of Game Theory* **15**(3), 175–186.

Bolger, E. M. (1990), 'A characterization of an extension of the Banzhaf value to multicandidate voting games', *SIAM Journal on Discrete Mathematics* **3**(4), 466–477.

Bolger, E. M. (2002), 'Characterizations of two power indices for voting games with r alternatives', *Social Choice and Welfare* **19**(4), 709–721.

Bovens, L. and Beisbart, C. (2011), 'Measuring voting power for dependent voters through causal models', *Synthese* **179**(1), 35–56.

Brams, S. and Fishburn, P. C. (2007), *Approval voting*, Springer.

Brams, S. J. and Fishburn, P. C. (1978), 'Approval voting', *American Political Science Review* **72**(3), 831–847.

Dekel, E., Lipman, B. L. and Rustichini, A. (2001), 'Representing preferences with a unique subjective state space', *Econometrica* **69**(4), 891–934.

Dowding, K. and van Hees, M. (2009), Freedom of choice, *in* P. Anand, P. K. Pattanaik and C. Puppe, eds, 'The Handbook of Rational and Social Choice: An Overview of New Foundations and applications', Oxford University Press, pp. 374–392.

Dubey, P. and Shapley, L. S. (1979), 'Mathematical properties of the Banzhaf power index', *Mathematics of Operations Research* **4**(2), 99–131.

Felsenthal, D. and Machover, M. (1998), *The Measurement of Voting Power*, Edward Elgar.

Felsenthal, D. S. and Machover, M. (1997), 'Ternary voting games', *International journal of game theory* **26**(3), 335–351.

Felsenthal, D. S. and Machover, M. (2005), 'Voting power measurement: a story of misreinvention', *Social choice and welfare* **25**(2-3), 485–506.

Freixas, J. and Zwicker, W. S. (2003), 'Weighted voting, abstention, and multiple levels of approval', *Social Choice and Welfare* **21**(3), 399–431.

Galaabaatar, T. (2010), 'Menu choice without completeness'. Unpublished manuscript.

Gelman, A., Katz, J. N. and Bafumi, J. (2004), 'Standard voting power indexes do not work: an empirical analysis', *British Journal of Political Science* **34**(4), 657–674.

Gibbard, A. (1973), 'Manipulation of voting schemes: a general result', *Econometrica* **41**(4), 587–601.

Kochov, A. (2007), 'Subjective states without the completeness axiom'. Unpublished Manuscript, University of Rochester.

Koopmans, T. C. (1964), On the flexibility of future preferences, *in* M. W. Shelly and G. L. Bryan, eds, 'Human Judgment and Optimality', John Wiley and Sons.

Kreps, D. M. (1979), 'A representation theorem for "preference for flexibility"', *Econometrica* **47**, 565–577.

Laruelle, A. and Valenciano, F. (2005), 'Assessing success and decisiveness in voting situations', *Social Choice and Welfare* **24**(1), 171–197.

Machover, M. (2007), 'Discussion topic: Voting power when voters' independence is not assumed'.

Muller, E. and Satterthwaite, M. A. (1977), 'The equivalence of strong positive association and strategy-proofness', *Journal of Economic Theory* **14**(2), 412–418.

Nehring, K. (1999), 'Preference for flexibility in a Savage framework', *Econometrica* **67**(1), 101–119.

Penrose, L. S. (1946), 'The elementary statistics of majority voting', *Journal of the Royal Statistical Society* pp. 53–57.

Satterthwaite, M. A. (1975), 'Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions', *Journal of Economic Theory* **10**(2), 187–217.

Shapley, L. S. (1953), A value for n-person games, *in* H. W. Kuhn and A. W. Tucker, eds, 'Contributions to the Theory of Games II', Vol. 28 of *Annals of Mathematics Studies*, Princeton University Press, pp. 307–317.

Shapley, L. S. and Shubik, M. (1954), 'A method for evaluating the distribution of power in a committee system', *American Political Science Review* **48**(3), 787–792.

Sher, I. (2018*a*), 'Evaluating allocations of freedom', *The Economic Journal* **128**(612), F65–F94.

Sher, I. (2018*b*), 'Neutral freedom and freedom as control'. Available at SSRN `https://ssrn.com/abstract=2652913`.

Straffin, P. D. (1977), 'Homogeneity, independence, and power indices', *Public Choice* **30**(1), 107–118.

Weber, R. J. (1977), 'Comparison of voting systems', *Cowles Foundation Discussion paper No. 498. Yale University* .

Appendix

Appendix A contains refinements of the symmetric power order and Appendix B contains proofs of theorems and technical details.

A Alternative voting power orders

This section presents a coarsening and some refinements of the symmetry order $\precsim^*$. The motivation for introducing these orders was discussed in Section 2.4. Each one of these alternative orders $\precsim$ is an ordering on $\mathscr{M}$, the set of menus, or in other words, the set of closed subsets of $\Delta(C)$. Each of these orders can be substituted for $\precsim^*$ in (6) to define an alternative voting power order.

First define the **convex hull order** $\precsim^{\text{co}}$ by $M \precsim^{\text{co}} M' \Leftrightarrow \text{co}(M) \subseteq \text{co}(M')$. $\precsim^{\text{co}}$ is a coarsening of $\precsim^*$. $\precsim^{\text{co}}$ fails to satisfy the neutrality axiom. However, Proposition 1 implies that the voting power order induced by the convex hull order agrees with the symmetric power order on all neutral elections. For non-neutral elections, the convex hull order compares fewer elections than the symmetric power order.

Next I parameterize the set of *all* refinements of $\precsim^*$ that satisfy the freedom axioms. Let $U = \mathbb{R}^C$ be the set of all utility functions $u = (u(c) : c \in C)$ on C. For any utility function $u \in U$ and permutation $\pi \in \Pi$ of candidates in C, let $u^\pi \in U$ be the utility function defined by $u^\pi(c) = u(\pi(c))$ for all $c \in C$. For any subset E of U and permutation $\pi \in \Pi$, define $E^\pi = \{u^\pi : u \in E\}$. A probability measure p on U is **symmetric** if for all measurable $E \subseteq U$ and all $\pi \in \Pi$, $p(E) = p(E^\pi)$. Let $\Delta^*_{\text{sym}}(U)$ be the set of symmetric probability measures on U with compact support. Recall that for any $u \in U$ and $\ell = (\ell_c : c \in C) \in \Delta(C)$, $u \cdot \ell = \sum_{c \in C} u(c)\,\ell_c$. For any $P \subseteq \Delta^*_{\text{sym}}(U)$, define $\precsim^P$ by

$$M \precsim^P M' \Leftrightarrow \left(\int_U \max_{\ell \in M} (u \cdot \ell)\, dp \leq \int_U \max_{\ell \in M'} (u \cdot \ell)\, dp, \forall p \in P \right).$$

Such rankings have an interpretation in terms of an agent who has ex ante uncertainty about which candidates she will prefer when she votes and hence values flexibility (Dekel, Lipman and Rustichini 2001) and has symmetric uncertainty about which candidates will be desirable at the ex ante stage (Sher 2018). See the preceding papers for an elaboration of this interpretation. It follows from results in Sher (2018) that $\left\{\precsim^P : P \subseteq \Delta^*_{\text{sym}}(U), P \neq \emptyset\right\}$ is precisely the set of refinements of $\precsim^*$ that satisfy the freedom axioms. Because orders $\precsim^P$ satisfy the freedom axioms, many of the properties of the symmetric power order are shared by voting power orders induced by orders of the form $\precsim^P$. For example, Theorem 2 holds for any voting power order defined by a refinement of this form.

An advantage of the orders $\precsim^P$ is that they can compare more elections than $\precsim^*$. However, there is a question of how these additional comparisons are to be justified. Section 5.4 suggests a barrier – not necessarily insurmountable – to justifying them. At one extreme, we may want to our voting power order to be a complete order and hence representable by a real valued voting power index. Such an order could compare every pair of elections. We accomplish this by setting $P = \{p\}$. We denote such orders by $\precsim^p$. A few special cases are worth discussing. Consider $\precsim^p$ for $p \in \Delta^*_{\text{sym}}(U)$ such that for some $u \in U$, p puts probability $\frac{1}{m!}$ on u^π for each $\pi \in \Pi$. This leads to an ordering, which we may denote $\precsim^u$, that is closely related to the generalized notion of an impartial justification associated with Proposition 11. One interesting case arises for the utility function u defined by $u(c) = c$ for $c = 1, \ldots, m$. This utility function ranks the candidates and treats successive candidates in the ranking as being equidistant. Because each u^π is equiprobable under

p, this ranking embodies implicit ex ante uncertainty about how the candidates will be ranked, but once ranked, each pair of successive candidates will be separated by the same distance. Another interesting class of cases arises by selecting the utility function u to be u^k defined by $u^k(c) = 1$ for $c = 1, \ldots, k$ and $u^k(c) = 0$ for $c = k+1, \ldots, m$. We can denote the corresponding order by $\precsim^k$. Elections that rank highly with respect to $\precsim^k$ are elections that are desirable if one's goal is to bring it about that some candidate in a set S of size k wins (when there are m total candidates). Special cases occur when $k = 1$, in which case highly ranked elections are those that are effective with regard to the goal of *promoting* a singe candidate, and when $k = m - 1$, for which highly ranked elections are those are effective with regard to *preventing* a singe candidate from winning. The goals of promoting and preventing were discussed in Sections 5.3-5.4.

B Proofs of theorems

Theorem 1

The proof is in Sher (2018), which is available at SRRN at `https://ssrn.com/abstract=2652913`.

Proposition 1

If σ is neutral, then $M_i(\sigma) = [M_i(\sigma)]^\pi$. So $\frac{1}{m!}\sum_{\pi\in\Pi}[M_i(\sigma)]^\pi = \frac{1}{m!}\sum_{\pi\in\Pi} M_i(\sigma) \subseteq \text{co}[M_i(\sigma)]$. So, using (4), $S(M) = \text{co}\left[\frac{1}{m!}\sum_{\pi\in\Pi}[M_i(\sigma)]^\pi\right] = \text{co}[M_i(\sigma)]$. It follows that if both σ and σ' are neutral, then by (6) and (5), $\sigma \precsim_i^\circ \sigma' \Leftrightarrow M_i(\sigma) \precsim^* M_i(\sigma') \Leftrightarrow S(M_i(\sigma)) \subseteq S(M_i(\sigma')) \Leftrightarrow \text{co}[M_i(\sigma)] \subseteq \text{co}[M_i(\sigma')]$. Observe finally that $\text{co}[M_i(\sigma)] \subseteq \text{co}[M_i(\sigma')] \Leftrightarrow M_i(\sigma) \subseteq \text{co}[M_i(\sigma')]$. This completes the proof.

Observe that the result only requires the assumption that σ and σ' are neutral for i.

Theorem 2

Here I present a different proof than the one that was sketched in the text. The reason is that the proof below appeals directly to the freedom axioms and thus applies not only to the symmetric power order but to any voting power order that is induced by a refinement of $\precsim^*$ that satisfies the freedom axioms. It thus also establishes that any refinement of the symmetry order that satisfies the freedom axioms coincides with pivotality on monotone binary elections.

Recall that in a monotone binary election the vote set is $V = \{1, 2\}$ and that δ_j is the degenerate lottery that puts probability 1 on candidate j. We have

$$\begin{aligned} M_i(\sigma) &= \{\ell_i(1|\sigma), \ell_i(2|\sigma)\} = \{[\ell_i(1|\sigma)]_1\delta_1 + [\ell_i(1|\sigma)]_2\delta_2, [\ell_i(2|\sigma)]_1\delta_1 + [\ell_i(2|\sigma)]_2\delta_2\} \\ &= [\ell_i(2|\sigma)]_1\{\delta_1\} + [\ell_i(1|\sigma)]_2\{\delta_2\} + ([\ell_i(2|\sigma)]_2 - [\ell_i(1|\sigma)]_2)\{\delta_1, \delta_2\} \\ &\sim^* ([\ell_i(2|\sigma)]_1 + [\ell_i(1|\sigma)]_2)\{\delta_1\} + ([\ell_i(2|\sigma)]_2 - [\ell_i(1|\sigma)]_2)\{\delta_1, \delta_2\} \\ &= [1 - ([\ell_i(2|\sigma)]_2 - [\ell_i(1|\sigma)]_2)]\{\delta_1\} + ([\ell_i(2|\sigma)]_2 - [\ell_i(1|\sigma)]_2)\{\delta_1, \delta_2\} \end{aligned} \tag{B.1}$$

For the third equality, observe that monotonicity of f and the fact that vote probabilities are independent across voters imply that $[\ell_i(2|\sigma)]_2 - [\ell_i(1|\sigma)]_2 \geq 0$. The indifference follows from the fact that all singleton menus are indifferent according to the symmetry order (see Proposition 4 of Sher (2018)) and the independence axiom. Let $1[\cdot]$ be the indicator function such that for any condition φ, $1[\varphi] = 1$ if φ is true, and $1[\varphi] = 0$ if φ is false. Then

$$\text{Piv}_i(\sigma) = \sum_{v_{-i}\in V_{-i}} 1[f(1, v_{-i}) \neq f(2, v_{-i})]\,\mu_{-i}(v_{-i})$$

$$= \sum_{v_{-i} \in V_{-i}} \left(1\left[f\left(2, v_{-i}\right)=2\right]-1\left[f\left(1, v_{-i}\right)=1\right]\right) \mu_{-i}\left(v_{-i}\right)$$
$$= \left[\ell\left(2 \left| \sigma_{-i}\right.\right)\right]_2-\left[\ell_2\left(1 \left| \sigma_{-i}\right.\right)\right]_2,$$

where again the second equality uses the monotonicity of f. It follows that

$$\begin{aligned}
\operatorname{Piv}_i(\sigma) \leq \operatorname{Piv}_i(\sigma') \Leftrightarrow\; & \left[\ell_i\left(2 \left| \sigma\right.\right)\right]_2-\left[\ell_i\left(1 \left| \sigma\right.\right)\right]_2 \leq \left[\ell_i\left(2 \left| \sigma\right.\right)\right]_2-\left[\ell_i\left(1 \left| \sigma\right.\right)\right]_2 \\
\Leftrightarrow\; & \left[1-\left(\left[\ell_i\left(2 \left| \sigma\right.\right)\right]_2-\left[\ell_i\left(1 \left| \sigma\right.\right)\right]_2\right)\right]\{\delta_1\}+\left(\left[\ell_i\left(2 \left| \sigma\right.\right)\right]_2-\left[\ell_i\left(1 \left| \sigma\right.\right)\right]_2\right)\{\delta_1, \delta_2\} \\
& \precsim^* \left[1-\left(\left[\ell_i\left(2 \left| \sigma'\right.\right)\right]_2-\left[\ell_i\left(1 \left| \sigma'\right.\right)\right]_2\right)\right]\{\delta_1\} \\
& \quad +\left(\left[\ell_i\left(2 \left| \sigma'\right.\right)\right]_2-\left[\ell_i\left(1 \left| \sigma'\right.\right)\right]_2\right)\{\delta_1, \delta_2\} \\
\Leftrightarrow\; & M_i(\sigma) \precsim^* M_i(\sigma') \Leftrightarrow \sigma \precsim_i^{\circ} \sigma'.
\end{aligned} \tag{B.2}$$

where the second equivalence follows from the fact that the symmetry order strictly prefers all menus that are not singletons to all singleton menus (Proposition 4 of Sher (2018)) and the independence axiom (Section 2.3.1), and the last equivalence follows from (B.1). □

Finally, I present two orders $\precsim$ and $\precsim'$ on $\mathscr{M}$, each of which violate only one of the freedom axioms but which induce voting power orders for which Theorem 2 fails. As we consider the binary case, we can represent a lottery by the probability it gives to candidate 2. Thus $\mathscr{M}$ reduces to the set of all closed subsets of $[0,1]$. First consider $\precsim$ represented by the function

$$g\left(M\right)=\max\left\{\max\left\{\ell, 1-\ell\right\}: \ell \in M\right\}, \quad \forall M \subseteq [0,1], M \text{ closed}.$$

where, above ℓ is the probability assigned to candidate 2. $\precsim$ satisfies all of the freedom axioms except independence. $\precsim$ prefers a monotonic election with electoral menu $\{.5, .9\}$ to a monotonic election with electoral menu $\{.2, .7\}$, even though the latter yields a higher pivotality.

Next consider the preference $\precsim'$ represented by

$$h\left(M\right)=\max\left\{\ell: \ell \in M\right\}, \quad \forall M \subseteq [0,1], M \text{ closed}.$$

This preference satisfies all freedom axioms except neutrality. It prefers an election in which candidate 2 wins for sure – so that pivotality is zero – to majority voting.

Proposition 3

Recall that $1[\cdot]$ is the indicator function such that for any condition φ, $1[\varphi]=1$ if φ is true, and $1[\varphi]=0$ if φ is false.

First, I prove part 2. We have

$$\begin{aligned}
\operatorname{Piv}_i^{\text{ea}}\left(\sigma\right) &= \left| \sum_{v_{-i} \in V_{-i}} 1\left[f\left(2, v_{-i}\right)=2\right] \mu_{-i}\left(v_{-i}\right) - \sum_{v_{-i} \in V_{-i}} 1\left[f\left(1, v_{-i}\right)=2\right] \mu_{-i}\left(v_{-i}\right)\right| \\
&= \left| \sum_{v_{-i} \in V_{-i}} \left(1\left[f\left(2, v_{-i}\right)=2\right]-1\left[f\left(1, v_{-i}\right)=2\right]\right) \mu_{-i}\left(v_{-i}\right)\right| \\
&\leq \sum_{v_{-i} \in V_{-i}} \left|1\left[f\left(2, v_{-i}\right)=2\right]-1\left[f\left(1, v_{-i}\right)=2\right]\right| \mu_{-i}\left(v_{-i}\right) \\
&= \sum_{v_{-i} \in V_{-i}} 1\left[f\left(1, v_{-i}\right) \neq f\left(2, v_{-i}\right)\right] \mu_{-i}\left(v_{-i}\right) \\
&= \operatorname{Piv}_i\left(\sigma\right),
\end{aligned} \tag{B.3}$$

where the inequality is an instance of the triangle inequality. The election σ for which $\text{Piv}_i(\sigma) = 1$ but $\text{Piv}_i^{\text{ea}}(\sigma) = 1$ is the matching pennies election (see Section 3.2). This establishes part 2.

Part 1 follows from the fact that for monotone voting rules,

$$1[f(2, v_{-i}) = 2] - 1[f(1, v_{-i}) = 2] \geq 0,$$

so that the inequality in (B.3) becomes an equality.

I now prove part 3. A derivation similar to (B.1) shows that for all σ,

$$\begin{aligned} M_i(\sigma) \sim^* &\left[1 - \left(\max_{v_i \in \{1,2\}} [\ell_i(v_i|\sigma)]_2 - \min_{v_i \in \{1,2\}} [\ell_i(v_i|\sigma)]_2\right)\right] \{\delta_1\} \\ &+ \left(\max_{v_i \in \{1,2\}} [\ell_i(v_i|\sigma)]_2 - \min_{v_i \in \{1,2\}} [\ell_i(v_i|\sigma)]_2\right) \{\delta_1, \delta_2\} \end{aligned} \tag{B.4}$$

Next, observe that

$$\begin{aligned} \text{Piv}_i^{\text{ea}}(\sigma) &= \max_{v_i \in \{1,2\}} \sum_{v_{-i} \in V_{-i}} 1[f(v_i, v_{-i}) = 2] \mu_{-i}(v_{-i}) - \min_{v_i \in \{1,2\}} \sum_{v_{-i} \in V_{-i}} 1[f(v_i, v_{-i}) = 2] \mu_{-i}(v_{-i}) \\ &= \max_{v_i \in \{1,2\}} [\ell_i(v_i|\sigma)]_2 - \min_{v_i \in \{1,2\}} [\ell_i(v_i|\sigma)]_2 . \end{aligned} \tag{B.5}$$

Part 3 of the proposition now follow from (B.4) and (B.5), using a derivation similar to (B.2). □

Proposition 4

First assume that $\text{Piv}_i(2;\sigma) < \text{Piv}_i(2;\sigma')$. It then follows, using the independence axiom and the fact that $\mathbf{1} \prec^* \mathbf{2}$ (which follows from Proposition 4 of Sher (2018)) that

$$\begin{aligned} &\text{Piv}_i(1;\sigma')\,\mathbf{1} + \text{Piv}_i(2;\sigma)\,\mathbf{2} + \left(\text{Piv}_i(2;\sigma') - \text{Piv}_i(2;\sigma)\right)\mathbf{1} \\ \prec^* &\text{Piv}_i(1;\sigma')\,\mathbf{1} + \text{Piv}_i(2;\sigma)\,\mathbf{2} + \left(\text{Piv}_i(2;\sigma') - \text{Piv}_i(2;\sigma)\right)\mathbf{2}. \end{aligned}$$

Regrouping the probabilities, the preceding inequality is equivalent to

$$\text{Piv}_i(1;\sigma)\,\mathbf{1} + \text{Piv}_i(2;\sigma)\,\mathbf{2} \prec^* \text{Piv}_i(\sigma')\,\mathbf{1} + \text{Piv}_i(2;\sigma')\,\mathbf{2},$$

which, in turn, is equivalent to $\sigma \precsim_i^{\text{piv}} \sigma'$ in the binary case. A similar derivation shows that $\text{Piv}_i(2;\sigma') \leq \text{Piv}_i(2;\sigma)$ implies $\sigma' \precsim_i^{\text{piv}} \sigma$. □

Proposition 5

Observe that for any nonempty subset S of C with $|S| = k$, there exists $\pi \in \Pi$ such that $\mathbf{S}^\pi = \mathbf{k}$. The neutrality axiom therefore implies $\mathbf{S} \sim^* \mathbf{k}$. Next observe that

$$\begin{aligned} \hat{\mathscr{L}}_i^\sigma &= \sum_{S \subseteq C, S \neq \emptyset} \text{Piv}_i(S;\sigma)\,\mathbf{S} = \sum_{k=1}^m \sum_{S \subseteq C, |S|=k} \text{Piv}_i(S;\sigma)\,\mathbf{S} \sim^* \sum_{k=1}^m \sum_{S \subseteq C, |S|=k} \text{Piv}_i(S;\sigma)\,\mathbf{k} \\ &\sim^* \sum_{k=1}^m \sum_{S \subseteq C, |S|=k} \text{Piv}_i(S;\sigma)\,\text{co}(\mathbf{k}) = \sum_{k=1}^m \left[\sum_{S \subseteq C, |S|=k} \text{Piv}_i(S;\sigma)\right] \text{co}(\mathbf{k}) \end{aligned}$$

$$=\sum_{k=1}^{m}\mathrm{Piv}_i\left(k;\sigma\right)\mathrm{co}\left(\mathbf{k}\right)\sim^{*}\sum_{k=1}^{m}\mathrm{Piv}_i\left(k;\sigma\right)\mathbf{k}=\mathscr{L}_i^{\sigma},$$

where the first indifference follows from the neutrality axiom, and the second and third indifferences follow from the independence axiom and the fact that for all menus M, $M\sim^{*}\mathrm{co}\left(M\right)$. This latter fact follows from definition of $\precsim^{*}$ and the properties of the Minkowski sum. □

Proposition 6

I assume that $m\geq 3$, because otherwise the result follows immediately from the freedom axioms. Assume that

$$\left|\mathrm{Piv}_i\left(2;\sigma'\right)-\mathrm{Piv}_i\left(2;\sigma\right)\right|>\left(m-1\right)\sum_{k=3}^{m}\left|\mathrm{Piv}_i\left(k;\sigma'\right)-\mathrm{Piv}_i\left(k;\sigma\right)\right| \tag{B.6}$$

Let $j\in\{1,\ldots,m-2\}$. Some algebra and the fact that $\sum_{k=1}^{m}\mathrm{Piv}_i\left(k;\sigma''\right)=1$ for all elections σ'', imply that

$$\sum_{k=1}^{m-j}\binom{m-k}{j}\left(\mathrm{Piv}_i\left(k;\sigma'\right)-\mathrm{Piv}_i\left(k;\sigma\right)\right)\leq 0 \tag{B.7}$$

$$\Leftrightarrow\left(\mathrm{Piv}_i\left(2;\sigma'\right)-\mathrm{Piv}_i\left(2;\sigma\right)\right)\geq\sum_{k=3}^{m}\alpha_k^j\left(\mathrm{Piv}_i\left(k;\sigma\right)-\mathrm{Piv}_i\left(k;\sigma'\right)\right), \tag{B.8}$$

where $\alpha_k^j=\begin{cases}\frac{\binom{m-1}{j}-\binom{m-k}{j}}{\binom{m-1}{j}-\binom{m-2}{j}} & \text{if } 3\leq k\leq m-j,\\ \frac{\binom{m-1}{j}}{\binom{m-1}{j}-\binom{m-2}{j}} & \text{if } m-j+1\leq k\leq m,\end{cases}$ for $j\in\{1,\ldots,m-2\}$. Observe that for k such that $3\leq k\leq m-j$, $0<\frac{\binom{m-1}{j}-\binom{m-k}{j}}{\binom{m-1}{j}-\binom{m-2}{j}}<\frac{\binom{m-1}{j}}{\binom{m-1}{j}-\binom{m-2}{j}}=\frac{m-1}{j}$. It follows that

$$0\leq\alpha_k^j\leq m-1,\quad\forall j\in\{1,\ldots,m-2\},\forall k\in\{3,\ldots,m\}. \tag{B.9}$$

To complete the proof, it is sufficient to establish the following three facts

1. If $\mathrm{Piv}_i\left(2;\sigma'\right)>\mathrm{Piv}_i\left(2;\sigma\right)$, then $\mathrm{Piv}_i\left(1;\sigma\right)>\mathrm{Piv}_i\left(1;\sigma'\right)$ and $\sigma\prec_i^{\mathrm{piv}}\sigma'$.
2. If $\mathrm{Piv}_i\left(2;\sigma'\right)=\mathrm{Piv}_i\left(2;\sigma\right)$, then $\mathrm{Piv}_i\left(1;\sigma'\right)=\mathrm{Piv}_i\left(1;\sigma\right)$ and $\sigma\sim_i^{\mathrm{piv}}\sigma'$.
3. If $\mathrm{Piv}_i\left(2;\sigma'\right)<\mathrm{Piv}_i\left(2;\sigma\right)$, then $\mathrm{Piv}_i\left(1;\sigma\right)>\mathrm{Piv}_i\left(1;\sigma'\right)$ and $\sigma'\prec_i^{\mathrm{piv}}\sigma$.

So first assume

$$\mathrm{Piv}_i\left(2;\sigma'\right)>\mathrm{Piv}_i\left(2;\sigma\right). \tag{B.10}$$

Because for all σ'', $\sum_{k=1}^{m}\mathrm{Piv}_i\left(k;\sigma''\right)=1$, we have

$$\mathrm{Piv}_1\left(1;\sigma'\right)<\mathrm{Piv}_1\left(1;\sigma\right)\Leftrightarrow\ \mathrm{Piv}_i\left(2;\sigma'\right)-\mathrm{Piv}_i\left(2;\sigma\right)>\sum_{k=3}^{m}\left(\mathrm{Piv}_i\left(k;\sigma\right)-\mathrm{Piv}_i\left(k;\sigma'\right)\right) \tag{B.11}$$

(B.6) and (B.10) imply that the inequality on the left hand side of biconditional (B.11) holds. It follows that $\text{Piv}_i(1;\sigma) > \text{Piv}_i(1;\sigma')$.[1] Next observe that for $j \in \{1, \ldots, m-2\}$,

$$\begin{aligned}\text{Piv}_i\left(2;\sigma'\right) - \text{Piv}_i\left(2;\sigma\right) &= \left|\text{Piv}_i\left(2;\sigma'\right) - \text{Piv}_i\left(2;\sigma\right)\right| > (m-1)\sum_{k=3}^{m}\left|\text{Piv}_i\left(k;\sigma'\right) - \text{Piv}_i\left(k;\sigma\right)\right| \\ &\geq \sum_{k=3}^{m}\alpha_k^j\left|\text{Piv}_i\left(k;\sigma'\right) - \text{Piv}_i\left(k;\sigma\right)\right| \geq \sum_{k=3}^{m}\alpha_k^j\left(\text{Piv}_i\left(k;\sigma'\right) - \text{Piv}_i\left(k;\sigma\right)\right),\end{aligned}$$

where the equality follows from (B.10), the first inequality follows from (B.6),and the second and third inequalities follow from (B.9). This establishes the inequality in (B.8). Using the biconditional in (B.7)-(B.8), we have now established the established the inequalities (B.7) for $j \in \{1, \ldots, m-2\}$. Note that if $j = m-1$, the inequality (B.7) is equivalent to $\text{Piv}_i(1;\sigma) \geq \text{Piv}_i(1;\sigma')$, which we have established above. It now follows from the inequalities (B.7) and part 1 of Proposition 7 that $\sigma \precsim_i^{\text{piv}} \sigma'$. We have now established statement 1 above. The proof of statement 3 is similar. For statement 2, observe that if $\text{Piv}_i(2;\sigma') = \text{Piv}_i(2;\sigma)$, then (B.6) implies that $\text{Piv}_i(k;\sigma') = \text{Piv}_i(k;\sigma)$ for $k \in \{3, \ldots, m\}$. It follows from the fact that for any election, the pivotality probabilities for different k sum to 1, that $\text{Piv}_i(1;\sigma') = \text{Piv}_i(1;\sigma)$ as well. This establishes 2, and completes the proof. □

Proposition 7

I first prove part 1.I start with some preliminaries. Recall that $\mathscr{M}$, the set of **menus**, is the set of all closed subsets of the set $\Delta(C)$ of lotteries on C. Let $U = \mathbb{R}^C$ be the set of all utility functions $u = (u(c) : c \in C)$ on C. Recall that for any $u \in U$ and $\ell = (\ell_c : c \in C) \in \Delta(C)$, $u \cdot \ell = \sum_{c \in C} u(c)\,\ell_c$. For any utility function $u \in U$ and permutation of candidates $\pi \in \Pi$ of candidates on C, let $u^\pi \in U$ be the utility function defined by $u^\pi(c) = u(\pi(c))$ for all $c \in C$. For any subset E of U and permutation $\pi \in \Pi$, define $E^\pi = \{u^\pi : u \in E\}$. A probability measure p on U is **symmetric** if for all measurable $E \subseteq U$ and all $\pi \in \Pi$, $p(E) = p(E^\pi)$. Let $\Delta^*_{\text{sym}}(U)$ be the set of symmetric probability measures on U with compact support. Let $D = \{\delta_c : c \in C\}$, where recall that δ_c is the degenerate lottery that puts probability 1 on candidate c. Observe that for any $u \in U$, $u \cdot \delta_c = u(c)$. Let $\mathscr{M}_d$ be the set of all nonempty subsets of D. $\mathscr{M}_d$ is the set of **deterministic menus**. Recall that for $k = 1, \ldots, m$, $\mathbf{k} = \{\delta_1, \ldots, \delta_k\}$. So $\mathbf{k} \in \mathscr{M}_d$. Observe that $\mathscr{M}_d \subseteq \mathscr{M}$.

Lemma B.1 *Let $W : \mathscr{M}_d \to \mathbb{R}$. Define $w : \{1, \ldots, m\} \to \mathbb{R}$ by*

$$w(k) := W(\mathbf{k}) \text{ for } k = 1, \ldots, m \tag{B.12}$$

Define $\Delta w(k) := w(k) - w(k-1)$ for $k = 2, \ldots, m$ and $\Delta w(1) := w(1)$. Then

$$\exists \mu \in \Delta^*_{\text{sym}}(U), \forall M \in \mathcal{M}_d, \quad W(M) = \int_U \left(\max_{\delta_c \in M} u \cdot \delta_c\right)\mu(du) \tag{B.13}$$

if and only if

$$\sum_{j=0}^{k}(-1)^{k-j}\binom{k}{j}\Delta w(m-j) \geq 0, \qquad \text{for } k = 0, \ldots, m-2. \tag{B.14}$$

[1] Recall that we are assuming that $m \geq 3$.

Proof. Consider $W : \mathcal{M}_d \to \mathbb{R}$. A slight modification of Proposition 1 in Nehring (1999) shows that[2]

$$\exists \mu \in \Delta^* (U), \forall M \in \mathcal{M}_d, \quad W(M) = \int_U \left(\max_{\delta_c \in M} u_c \right) \mu(du) \tag{B.15}$$

if and only if there exists $\lambda = (\lambda_S : S \in \mathscr{M}_d) \in \mathbb{R}^{\mathscr{M}_d}$ such that:

$$\begin{aligned} \forall M \in \mathcal{M}_d, \quad W(M) = \sum_{S \in \mathscr{M}_d : S \cap M \neq \emptyset} \lambda_S, \text{ and} \\ \forall S \in \mathscr{M}_d \setminus \{D\}, \quad \lambda_S \geq 0. \end{aligned} \tag{B.16}$$

Observe that the difference between (B.13) and (B.15) is that the former assumes that $\mu \in \Delta^*_{\text{sym}}(U)$, whereas the latter only assumes that $\mu \in \Delta^*(U)$. It is not difficult to modify Nehring's proof to show that W satisfies (B.13) if and only if there exists a $\lambda \in \mathbb{R}^{\mathscr{M}_d}$ that satisfies (B.16) and

$$\forall S, T \in \mathscr{M}_d, \quad |S| = |T| \Rightarrow \lambda_S = \lambda_T. \tag{B.17}$$

Define the function $\eta : \{1, \ldots, m\} \times \{1, \ldots, m\} \to \mathbb{N}$ by $\eta(k, j) := |\{S \subseteq D : |S| = j, \mathbf{k} \cap S \neq \emptyset\}|$. In other words, $\eta(k, j)$ is the number of sets of cardinality j that intersect $\mathbf{k} = \{\delta_1, \ldots, \delta_k\}$. It is easy to see that there exists $\lambda \in \mathbb{R}^{\mathscr{M}_d}$ satisfying (B.16) and (B.17) if and only if there exists $\lambda = (\lambda_j : j = 1, \ldots, m) \in \mathbb{R}^m$ satisfying:

$$\begin{aligned} w(k) = \sum_{j=1}^{n} \eta(k, j) \lambda_j \quad \forall k \in \{1, \ldots, m\}, \\ \lambda_j \geq 0, \quad \forall j \in \{1, \ldots, m-1\}. \end{aligned} \tag{B.18}$$

If for each $k = 2, \ldots, m$, we subtract the equation for $k-1$ from the equation for k in (B.18), we arrive at the system:

$$\Delta w(m - (k-1)) = \sum_{j=1}^{k} \binom{k-1}{j-1} \lambda_j, \quad \forall k \in \{1, \ldots, m\} \tag{B.19}$$

$$\lambda_j \geq 0, \quad \forall j \in \{1, \ldots, m-1\}, \tag{B.20}$$

To see this derivation, observe that $\eta(m - (k-1), j) - \eta((m - (k-1)) - 1, j)$ is equal to $\binom{k-1}{j-1}$ if $j \leq k$, and is equal to 0 otherwise. This is because $\{\delta_1, \ldots, \delta_{(m-(k-1))-1}\} \subseteq \{\delta_1, \ldots \delta_{(m-(k-1))}\}$, so any set that intersects $\{\delta_1, \ldots, \delta_{(m-(k-1))-1}\}$ also intersects $\{\delta_1, \ldots \delta_{(m-(k-1))}\}$, and, when $j \leq k$, the sets of cardinality j that intersect $\{\delta_1, \ldots, \delta_{(m-(k-1))-1}\}$ but not $\{\delta_1, \ldots, \delta_{(m-(k-1))-1}\}$ are the sets that contain $\delta_{m-(k-1)}$ and $j-1$ elements of the set $\{\delta_{(m-(k-1))+1}, \delta_{(m-(k-1))+2}, \ldots, \delta_m\}$, which has cardinality $k-1$; if $j > k$, there are no sets of cardinality j that intersect $\{\delta_1, \ldots \delta_{(m-(k-1))}\}$ but not $\{\delta_1, \ldots, \delta_{(m-(k-1))-1}\}$.[3] Likewise, adding the equalities corresponding to $\Delta w(1)$ through $\Delta v(k)$ in (B.19), we recover the equality corresponding to $w(k)$ in (B.18). So the systems (B.18) and (B.19-B.20) are equivalent.

[2] A Nehring proved the equivalence for probability measures μ with finite support. A slight modification of the proof shows that the equivalence continues to hold when we consider measures with compact support (i.e., measures in $\Delta^*(U)$).

[3] Given that we define $\Delta w(1) = w(1)$, it is easy to see that the equation for $w(1)$ in (B.18) is the same as the equation for $\Delta v(m - (m-1))$ in (B.19).

Inverting the system (B.19) amounts to inverting a Pascal's matrix (see Call and Velleman (1993)). So we can invert the system invert the system (B.19) and attain

$$\lambda_k = \sum_{j=1}^{k} (-1)^{k-j} \binom{k-1}{j-1} \Delta w(m-(j-1)), \qquad k = 1, \ldots, m.$$

It follows that there is a solution to (B.19)-(B.20) if and only if:

$$\sum_{j=1}^{k} (-1)^{k-j} \binom{k-1}{j-1} \Delta w(m-(j-1)) \geq 0, \qquad \text{for } k = 1, \ldots, m-1, \tag{B.21}$$

However (B.21) is equivalent to (B.14). This completes the proof of the lemma. □

Recall that $C = \{1, \ldots, m\}$. We can represent a function $w : C \to \mathbb{R}$ by a point $w = (w(k) : k \in C) \in \mathbb{R}^m$. Let $\mathscr{W}$ be the set of functions $w \in \mathbb{R}^m$ that satisfy (B.14).

Lemma B.2 *Let $w \in \mathbb{R}^m$ be such that $w(m) = 0$. w satisfies (B.14) if and only if w satisfies*

$$\sum_{j=m-k}^{m-1} (-1)^{k-(m-j)+1} \binom{k}{m-j} w(j) \geq 0, \text{ for } k = 1, \ldots, m-1. \tag{B.22}$$

Proof. We have

$$\sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} \Delta w(m-j) = \sum_{j=1}^{k} (-1)^{k-j} \binom{k}{j} (w(m-j) - w(m-j-1))$$
$$= \left[\sum_{j=1}^{k} \left[(-1)^{k-k} \binom{k}{j} - (-1)^{k-j+1} \binom{k}{j-1} \right] w(m-j) \right] - (-1)^{k-k} \binom{k}{k} w(m-(k+1))$$
$$= \left[\sum_{j=1}^{k} (-1)^{k-j} \left[\binom{k}{j-1} + \binom{k}{j} \right] w(m-j) \right] - (-1)^{k-k} \binom{k}{k} w(m-(k+1))$$
$$= \left[\sum_{j=1}^{k} (-1)^{k-j} \binom{k+1}{j} \right] w(m-j) + (-1)^{k-(k+1)} \binom{k+1}{k+1} w(m-(k+1))$$
$$= \sum_{j=1}^{k+1} (-1)^{k-j+1} \binom{k+1}{j} w(m-j) = \sum_{i=m-(k+1)}^{m-1} (-1)^{k-(m-i)} \binom{k+1}{m-i} w(i),$$

where the second equality used $w(m) = 0$. It follows that w satisfies (B.14) if and only if w satisfies

$$\sum_{i=m-(k+1)}^{m-1} (-1)^{k-(m-i)} \binom{k+1}{m-i} w(i) \text{ for } k = 0, 1, \ldots, m-2$$

or equivalently if and only if w satisfies (B.22). □

A vector $p = (p_c : c = 1, \ldots, m) \in \mathbb{R}^m$ is a **probability vector** if $p_c \geq 0$ for all $c \in \{1, \ldots, m\}$ and $\sum_{c=1}^{m} p_c = 1$. Let $\mathscr{P}$ be the set of all probability vectors in $\mathbb{R}^m$. (Formally, $\mathscr{P}$ is the same as $\Delta(C)$, but the interpretation is different.) For any $p \in \mathscr{P}$, define $M^p = \sum_{k=1}^{n} p_k \mathbf{k}$.

Lemma B.3 $\forall p, q \in \mathscr{P}, (M^p \precsim^* M^q \Leftrightarrow \forall w \in \mathscr{W}, \sum_{k=1}^{m} (q_k - p_k) w(k) \geq 0)$.

For any and $M \in \mathscr{M}$ and $\mu \in \Delta^*(U)$, define $\tilde{W} : \mathscr{M} \times \Delta^*(U) \to \mathbb{R}$ by

$$\tilde{W}(M;\mu) = \int_U \max_{\ell \in \mathscr{M}} (u \cdot \ell)\, d\mu(u).$$

It is east to see that

$$\tilde{W}(M^p;\mu) = \sum_{k=1}^{m} \tilde{W}(\mathbf{k};\mu)\, p_k, \quad \forall p \in \mathscr{P}, \forall \mu \in \Delta^*(U) \tag{B.23}$$

It follows that

$$\begin{aligned} M^p \precsim^* M^q &\Leftrightarrow \tilde{W}(M^p;\mu) \le \tilde{W}(M^q;\mu), \quad \forall \mu \in \Delta^*_{\text{sym}}(U) \\ &\Leftrightarrow \sum_{k=1}^{m} (q_k - p_k)\, \tilde{W}(\mathbf{k};\mu), \qquad \forall \mu \in \Delta^*_{\text{sym}}(U) \\ &\Leftrightarrow \sum_{k=1}^{m} (q_k - p_k)\, w(k) \ge 0, \qquad \forall w \in \mathscr{W}, \end{aligned}$$

where the first biconditional follows from Theorem 1 of Sher (2018), the second follows from (B.23), and the third from Lemma B.1. □

Lemma B.4 $\forall p, q \in \mathscr{P}, \left(M^p \precsim^* M^q \Leftrightarrow \left(\sum_{k=1}^{m-j} \binom{m-k}{j} (q_k - p_k) \le 0, \textit{for } j = 1, \ldots, m-1 \right) \right).$

By Lemma B.3, $M^p \precsim^* M^q$ is equivalent to

$$\forall w \in \mathbb{R}^m, \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} \Delta w(m-j) \ge 0, \;\; \text{for } k = 0, \ldots, m-2 \Rightarrow \sum_{j=1}^{m} (p_j - q_j)\, w(j) \ge 0 \tag{B.24}$$

Suppose that $w = (w(k) : k = 1, \ldots, m)$ is a solution to the set of inequalities in the antecedent of (B.24). Then $w' = (w'(k) : k = 1, \ldots, m) = (w(k) - w(m) : k = 1, \ldots, m)$ is also a solution to these inequalities. Likewise w is a solution to the set of inequality in the consequent of (B.24) if and only if w' is a solution to this inequality. Therefore, in evaluating the above conditional, we may restrict attention to $w \in \mathbb{R}^m$ with $w(m) = 0$ and (B.24) is equivalent to:

$$\begin{aligned} \forall w \in \mathbb{R}^m, \sum_{j=1}^{k} (-1)^{k-j} \binom{k}{j} \Delta w(m-j) \ge 0, \;\; &\text{for } k = 1, \ldots, m-2 \text{ and } w(m) = 0 \\ &\Rightarrow \sum_{j=1}^{m-1} (p_j - q_j)\, w(j) \ge 0. \end{aligned} \tag{B.25}$$

By Lemma B.22, this is equivalent to

$$\forall w \in \mathbb{R}^{m-1}, \sum_{j=m-k}^{m-1} (-1)^{k-(m-j)+1} \binom{k}{m-j} w(j) \ge 0, \;\; \text{for } j = 1, \ldots, m-1 \Rightarrow \sum_{j=1}^{m-1} (p_j - q_j)\, w(j) \ge 0. \tag{B.26}$$

Note that we did not have to include $w(m) = 0$ in the other inequalities in the antecedent contain $w(m)$, and the inequality in the consequent does not contain $w(m)$ either. By the theorem of the alternative (B.22) holds if and only if there exists a solution $x = (x_k : k = 1, \ldots, m-1)$ to

$$\sum_{k=m-j}^{m-1} (-1)^{k-(m-j)+1} \binom{k}{m-j} x_k = (p_j - q_j), \qquad \text{for } j = 1, \ldots, m-1, \tag{B.27}$$

$$x_k \geq 0, \qquad \text{for } k = 1, \ldots, m-1. \tag{B.28}$$

Using calculations very similar to the proof of Theorem 1 in Call and Velleman (1993), we invert (B.27), we attain:

$$x_k = \sum_{j=1}^{m-k} \binom{m-j}{k} (p_j - q_j), \quad \text{for } k = 1, \ldots, m-1$$

Thus, (B.27)-(B.27) has a solution if and only if $\sum_{j=1}^{m-k} \binom{m-j}{k} (p_j - q_j) \geq 0$ for $k = 1, \ldots, m-1$. We have now established that this condition is equivalent to $M^p \precsim^* M^q$, which establishes the result. □

Observe finally that when we set $q = \overline{\text{Piv}}_i(\sigma')$ and $p = \overline{\text{Piv}}_i(\sigma)$, $M^p \precsim^* M^q$ is equivalent to $\sigma \precsim_i \sigma'$, and Lemma B.4 establishes part 1 of Proposition 7.

I now prove part 2. Let $q = (q_1, \ldots, q_m) = \overline{\text{Piv}}_i(\sigma')$ and $p = (q_1, \ldots, p_m) = \overline{\text{Piv}}_i(\sigma)$. Suppose that q first order stochastically dominates p. Then

$$\sum_{i=1}^{j} (q_i - p_i) \leq 0, \qquad \text{for } j = 1, \ldots, m-1. \tag{B.29}$$

By part 1, to establish that $\sigma \precsim^* \sigma'$, it is sufficient to establish

$$.\sum_{i=1}^{m-j} \binom{m-i}{j} (q_i - p_i) \leq 0, \quad \text{for } j = 1, \ldots, m-1. \tag{B.30}$$

Choose $j \in \{1, \ldots, m-1\}$. Using the fact that $\binom{j-1}{j} = 0$, it follows that

$$\sum_{i=1}^{m-j} \binom{m-i}{j} (q_i - p_i) = \sum_{i=1}^{m-j} \left[\binom{m-i}{j} - \binom{j-1}{j} \right] (q_i - p_i)$$
$$= \sum_{i=1}^{m-j} \sum_{k=i}^{m-j} \left[\binom{m-k}{j} - \binom{m-k-1}{j} \right] (q_i - p_i) = \sum_{k=1}^{m-j} \sum_{i=1}^{k} \left[\binom{m-k}{j} - \binom{m-k-1}{j} \right] (q_i - p_i)$$
$$= \sum_{k=1}^{m-j} \left[\binom{m-k}{j} - \binom{m-k-1}{j} \right] \sum_{i=1}^{k} (q_i - p_i) \leq 0$$

where the inequality follows from (B.29) and the fact that $\binom{m-k}{j} - \binom{m-k-1}{j} \geq 0$. This establishes (B.30), as was desired, and establishes part 2.

Now I prove part 3: Again let $q = \overline{\text{Piv}}_i(\sigma')$ and $p = \overline{\text{Piv}}_i(\sigma)$. Suppose that q second order stochastically dominates p. Then

$$\sum_{i=1}^{j} \sum_{k=1}^{i} (q_k - p_k) \leq 0, \quad \text{for } j = 1, \ldots, m-1. \tag{B.31}$$

Observe that $\sum_{i=1}^{j}\sum_{k=1}^{i}(q_k - p_k) = \sum_{k=1}^{j}\sum_{i=k}^{j}(q_k - p_k) = \sum_{k=1}^{j}(j-k+1)(q_k - p_k)$. It follows that (B.31) can equivalently be written as

$$\sum_{i=1}^{j}(j-i+1)(q_i - p_i) \leq 0, \quad \text{for } j = 1, \ldots, n-1. \tag{B.32}$$

Again, by part 1, to establish that $\sigma \precsim^{\circ} \sigma'$, it is sufficient to establish (B.30).

First observe that

$$\sum_{i=1}^{m-1}\binom{m-i}{1}(q_i - p_i) = \sum_{i=1}^{m-1}(m-i)(q_i - p_i) \leq 0,$$

because the inequality coincides with (B.32) with $j = m-1$.

Next consider $j \in \{2, \ldots, m-1\}$. We appeal to the following combinatorial equality:[4]

$$\sum_{k=j-1}^{m-i-1}(m-i-k)\binom{k-1}{j-2} = \binom{m-i}{j} \text{ for } j = 2, \ldots, m-i. \tag{B.33}$$

Using (B.33), we get

$$\sum_{i=1}^{m-j}\binom{m-i}{j}(q_i - p_i) = \sum_{i=1}^{m-j}\sum_{k=j-1}^{m-i-1}(m-i-k)\binom{k-1}{j-2}(q_i - p_i)$$
$$= \sum_{k=j-1}^{m-2}\sum_{i=1}^{m-k-1}(m-i-k)\binom{k-1}{j-2}(q_i - p_i) = \sum_{k=j-1}^{m-2}\binom{k-1}{j-2}\sum_{i=1}^{m-k-1}(m-i-k)(q_i - p_i) \leq 0,$$

where the inequality follows from the fact that it is a weighted sum on inequalities of the form (B.32) with positive weights. This establishes (B.30), as was desired. This establishes part 3.

To proof that the converses of parts 2 and 3 fail is omitted, but is available upon request. □

Theorem 3

It follows from Proposition 5 that Theorem 3 is equivalent to the following proposition.

Proposition B.1 *Let $i \in I$.*

1. *For all voting situations σ, $M_i(\sigma) \precsim^* \sum_{S\in\mathscr{C}} \mathrm{Piv}_i(S;\sigma)\,\mathbf{S}$.*
2. *Let $\sigma = (f, \mu)$. Suppose that $\mu_j(v_j) > 0$ for all $j \in I_0 \setminus i$ and $v_j \in V_j$. Then, $M_i(\sigma) \sim^* \sum_{S\in\mathscr{C}} \mathrm{Piv}_i(S;\sigma)\,\mathbf{S}$ if and only if f has a monotonic interpretation for i.*

Given the above-stated equivalence, to establish Theorem 3, it is sufficient to establish Proposition B.1. Therefore, in this appendix, I prove Proposition B.1.

Let $\mathscr{C}$ be the set of all nonempty subsets of C. Throughout the proof, I fix election σ and voter i. Therefore to simplify notation, define $\lambda_S = \mathrm{Piv}_i(S;\sigma)$ for $S \in \mathscr{C}$. It follows from the definition

[4](B.33) can be derived from the following equality in Knuth (1997, p. 59, equation (25)): $\sum_{h=0}^{r}\binom{r-h}{t}\binom{s+h}{u} = \binom{r+s+1}{t+u+1}$ for all nonnegative integers t, r, s, and u. Setting $t = 1, r = m-i-1, s = 0$, and $u = j-2$, this equation becomes $\sum_{h=0}^{m-i-1}(m-i-1-h)\binom{h}{j-2} = \binom{m-i}{j}$. Using the change of variables $k-1 = h$, the equation becomes $\sum_{k=1}^{m-i}(m-i-k)\binom{k-1}{j-2} = \binom{m-i}{j}$. Observing finally that $(m-i-k)\binom{k-1}{j-2} = 0$ except when $j-1 \leq k \leq m-i-1$, we arrive at (B.33).

of $\text{Piv}_i(S;\sigma)$ (see Section 4.1.4) that $\lambda_S \geq 0$ for all $S \in \mathscr{C}$ and $\sum_{S\in\mathscr{C}} \lambda_S = 1$. Define a **menu strategy** to be a function $\zeta : \mathscr{C} \to \Delta(C)$ satisfying

$$[\zeta(S)]_c > 0 \Rightarrow c \in S, \quad \forall S \in \mathscr{C}, \forall c \in C, \tag{B.34}$$

where $[\zeta(S)]_c$ is the probability that the lottery $\zeta(S)$ assigns to c. We can think of ζ as function that assigns to each nonempty subset S a lottery over S. Condition (B.34) says that we can indeed think of $\zeta(S)$ as a lottery over S rather than over the larger set C. For each menu strategy ζ, define the **induced lottery** $\ell^\zeta \in \Delta(C)$ to be the lottery defined by $\ell^\zeta = \sum_{S\in\mathscr{C}} \lambda_S \zeta(S)$. Thus ℓ^ζ is the distribution of outcomes induced by the compound lottery in which, first, each set S is selected with probability λ_S, and then a candidate is selected from S according to the lottery $\zeta(S)$. Recall that for any $S \in \mathscr{C}, \mathbf{S} = \{\delta_c : c \in S\}$, where δ_c is the degenerate lottery that puts probability 1 on c.

Lemma B.5 *For any menu strategy ζ, $\ell^\zeta \in \text{co}\left[\sum_{S\in\mathscr{C}} \lambda_S \mathbf{S}\right]$.*

Proof. Consider a two player extensive form game with perfect information in which first player 1 selects a nonempty subset S of C, and then player 2 selects a candidate $c \in S$. The outcome of the game is then c. Suppose that player 1 selects each set S with probability λ_S. The set $\sum_{S\in\mathscr{C}} \lambda_S \mathbf{S}$ can be interpreted as the set of distributions over outcomes that player 2 can generate in this game by choosing a pure strategy. The set $\text{co}\left[\sum_{S\in\mathscr{C}} \lambda_S \mathbf{S}\right]$ can then be interpreted as the set of distributions over outcomes that player 2 can generate by selecting a mixed strategy. ζ can be interpreted as a behavior strategy in this game, and ℓ^ζ as the distribution of outcomes that ζ generates. The Lemma now follows from Kuhn's theorem (Kuhn 1953), which asserts the equivalence of behavior and mixed strategies in extensive form games with perfect recall. □

Lemma B.6 $M_i(\sigma) \subseteq \text{co}\left[\sum_{S\in\mathscr{C}} \lambda_S \mathbf{S}\right]$.

Proof. Choose $\ell \in M_i(\sigma)$. Then there exists $v_i \in V$ such that $\ell = \ell_i(v_i|\sigma)$ (see Section 2.2). For each $c \in C$, define the event $\text{Vote}_i^f(v_i, c) = \{v_{-i} \in V_{-i} : f(v_i, v_{-i}) = c\}$. Let $\mathscr{C}_+ := \{S \in \mathscr{C} : \lambda_S > 0\}$. Define menu strategy ζ by

$$[\zeta(S)]_c = \frac{\mu_{-i}\left(\text{Vote}_i^f(v_i, c) \cap \text{Pivotal Event}_i^f(S)\right)}{\mu_{-i}\left(\text{Pivotal Event}_i^f(S)\right)}, \quad \forall S \in \mathscr{C}_+, \forall c \in C.$$

For $S \in \mathscr{C} \setminus \mathscr{C}_+$, define $([\zeta(S)]_c : c \in C)$ arbitrarily so that (B.34) is satisfied. It is easy to see that ζ, defined in this way, satisfies (B.34). Observe that for any $c \in C$,

$$\begin{aligned}
\sum_{S\in\mathscr{C}} \lambda_S [\zeta(S)]_c &= \sum_{S\in\mathscr{C}_+} \lambda_S [\zeta(S)]_c = \sum_{S\in\mathscr{C}_+} \lambda_S \frac{\mu_{-i}\left(\text{Vote}_i^f(v_i, c) \cap \text{Pivotal Event}_i^f(S)\right)}{\lambda_S} \\
&= \sum_{S\in\mathscr{C}_+} \mu_{-i}\left(\text{Vote}_i^f(v_i, c) \cap \text{Pivotal Event}_i^f(S)\right) = \mu_{-i}\left(\text{Vote}_i^f(v_i, c)\right) = [\ell_i(v_i|\sigma)]_c = \ell_c
\end{aligned}$$

So $\ell = \sum_{S\in\mathscr{C}} \lambda_S [\zeta(S)]$. Lemma B.5 implies that $\ell \in \text{co}\left[\sum_{S\in\mathscr{C}} \lambda_S \mathbf{S}\right]$. □

Lemma B.7 *For all menus $M, N \in \mathscr{M}$, if $M \subseteq \text{co}(N)$, then $M \precsim^* N$.*

Proof. This follows from (4) and (5) and the properties of the Minkowski sum. □

Lemmas B.6 and B.7 imply that $M_i(\sigma) \precsim^* \sum_{S\in\mathscr{C}} \lambda_S \mathbf{S}$, establishing Part 1 of Proposition B.1.

I now prove part 2. To this end, assume that $\sigma = (f, \mu)$ is such that f has a monotonic interpretation for i. It will be useful to define

$$B(f, v_{-i}) = \{f(v_i, v_{-i}) : v_i \in V_i\}. \tag{B.35}$$

$B(f, v_{-i})$ is the set of candidates whose election i can bring about by varying her vote, holding fixed the votes v_{-i} of other voters (including the tie-breaker). Next observe that we can interpret $\mathbb{R}^C$ as the set of utility function $u = (u(c) : c \in C)$ on C. Consider any $u \in \mathbb{R}^C$. Let $R_i^u \in \mathscr{R}$ be such that

$$cR_i^u c' \Rightarrow u(c) \geq u(c'), \quad \forall c, c' \in C. \tag{B.36}$$

Since f has a monotonic interpretation for i, it follows from Corollary 2 that there exists $v_i^u \in V$ such that for all $v_{-i} \in V_{-i}$ and $c \in B(f, v_{-i})$, $f(v_i^u, v_{-i}) R_i^{u_i} c$.[5] Hence,

$$\forall v_{-i} \in V_{-i}, \forall c \in B(f, v_{-i}), u(f(v_i^u, v_{-i})) \geq u(c). \tag{B.37}$$

A few more preliminaries are required. A **selection function** is a function $h : \mathscr{C} \to C$ satisfying $h(S) \in S, \forall S \in \mathscr{C}$. For any $\ell \in \Delta(C)$, let $u \cdot \ell = \sum_{c\in C} u(c) \ell_c$.

Lemma B.8 *For all $\ell \in \sum_{S\in\mathscr{C}} \lambda_S \mathbf{S}$, $u \cdot \ell_i(v_i^u|\sigma) \geq u \cdot \ell$.*

Choose $\ell \in \sum_{S\in\mathscr{C}} \lambda_S \mathbf{S}$. There exists a selection function h such that $\ell = \sum_{S\in\mathscr{C}} \lambda_S \delta_{h(S)}$. Observe that

$$\begin{aligned}
&u \cdot \ell = u \cdot \sum_{S\in\mathscr{C}} \lambda_S \delta_{h(S)} = u \cdot \sum_{S\in\mathscr{C}} \mu_{-i}(\{v_{-i} \in V_{-i} : B(f, v_{-i}) = S\}) \delta_{h(S)} \\
&= u \cdot \sum_{S\in\mathscr{C}} \left(\sum_{\{v_{-i}\in V_{-i}:B(f,v_{-i})=S\}} \mu_{-i}(v_{-i}) \right) \delta_{h(S)} = u \cdot \sum_{v_{-i}\in V_{-i}} \mu_{-i}(v_{-i}) \delta_{h(B(f,v_{-i}))} \\
&= \sum_{v_{-i}\in V_{-i}} \mu_{-i}(v_{-i}) u(h(B(f, v_{-i}))) \\
&\leq \sum_{v_{-i}\in V_{-i}} \mu_{-i}(v_{-i}) u(f(v_i^u, v_{-i})) = \sum_{c\in C} \mu_{-i}(\{v_{-i} \in V_{-i} : f(v_i^u, v_{-i}) = c\}) u(c) = u \cdot \ell_i(v_i^u|\sigma).
\end{aligned}$$

In the above derivation, the inequality follows from (B.37). This establishes the lemma. □

The following is an immediate consequence of Lemma B.8.[6]

Corollary B.1 *If $\sigma = (f, \mu)$ is such that f has a monotonic interpretation for i, then for all $u \in \mathbb{R}^C$, $\max\{u \cdot \ell : \ell \in M_i(\sigma)\} \geq \max\{u \cdot \ell : \ell \in \sum_{S\in\mathscr{C}} \mathrm{Piv}_i(S;\sigma) \mathbf{S}\}$.*

It follows from the separating hyperplane theorem that when f has a monotonic interpretation for i, $\mathrm{co}\left(\sum_{S\in\mathscr{C}} \lambda_S \mathbf{S}\right) \subseteq \mathrm{co}(M_i(\sigma))$. So $\sum_{S\in\mathscr{C}} \lambda_S \mathbf{S} \precsim^* M_i(\sigma)$. It now follows from part 1 of Proposition B.1, which has been established above, that if f has a monotonic interpretation for i,

[5] I here use the notation v_i^u, rather than the more cumbersome notation $v_i^{R_i^u}$, which would be more parallel to the notation in Corollary 2.

[6] We can use max rather than sup in the inequality in Corollary B.1 because $\sum_{S\in\mathscr{C}} \mathrm{Piv}_i(S;\sigma) \mathbf{S}$ and $M_i(\sigma)$ are finite. $M_i(\sigma)$ is finite because V is finite.

then $M_i(\sigma) \sim^* \sum_{S\in\mathscr{C}} \text{Piv}_i(S;\sigma)\,\mathbf{S}$. This establishes the right-to-left direction in the biconditional in part 2.

To complete the proof, we must establish the other direction, that is, that if f does not have a monotonic interpretation for i, then $M_i(\sigma) \prec^* \sum_{k=1}^{m} \text{Piv}_i(S;\sigma)\,\mathbf{S}$. To this end, assume that f does not have a monotonic interpretation for i. Assume, moreover that $\mu_j(v_j) > 0$ for all $j \in I_0 \setminus i$. It follows that $\mu_{-i}(v_{-i}) > 0$ for all $v_{-i} \in V_{-i}$. Since f does not have a monotonic interpretation for i, Corollary 2 implies that there exists $R_i \in \mathscr{R}$ such that

$$\forall v_i \in V, \exists v_i' \in V, \exists v_{-i} \in V_{-i}, f\left(v_i', v_{-i}\right) P_i f\left(v_i, v_{-i}\right), \tag{B.38}$$

where P_i is the asymmetric part of R_i. Now consider a utility function $u \in \mathbb{R}^C$ that represents R_i in the sense that

$$u(c) \geq u\left(c'\right) \Leftrightarrow cR_ic', \quad \forall c, c' \in C. \tag{B.39}$$

Lemma B.9 *If f does not have a monotonic interpretation for i, then*

$$\max\{u\cdot\ell : \ell \in M_i(\sigma)\} < \max\left\{u\cdot\ell : \ell \in \sum_{S\in\mathscr{C}} \lambda_S\mathbf{S}\right\}.$$

Proof. Choose $\ell \in M_i(\sigma)$. There exists $v_i' \in V$ such that $\ell = \ell_i(v_i'|\sigma)$ (see Section 2.2). (B.38) and (B.39) imply that there exists v_{-i} such that

$$u\left(f\left(v_i', v_{-i}\right)\right) < \max_{c\in B(f,v_{-i})} u(c). \tag{B.40}$$

(For the definition of $B(f, v_{-i})$, see (B.35) above.) Observe that

$$\begin{aligned} u\cdot\ell = u\cdot\ell_i\left(v_i'|\sigma\right) &= \sum_{v_{-i}\in V_{-i}} u\left(f\left(v_i', v_{-i}\right)\right)\mu_{-i}(v_{-i}) \\ < \sum_{v_{-i}\in V_{-i}} \max_{c\in B(f,v_{-i})} u(c)\,\mu_{-i}(v_{-i}) &= \sum_{S\in\mathscr{C}} \sum_{\{v_{-i}\in V_{-i}: B(f,v_{-i})=S\}} \max_{c\in S} u(c)\,\mu_{-i}(v_{-i}) \\ = \sum_{S\in\mathscr{C}} \max_{c\in S} u(c)\,\text{Piv}_i(S;\sigma) &= \max\left\{u\cdot\ell : \ell \in \sum_{S\in\mathscr{C}} \lambda_S\mathbf{S}\right\}, \end{aligned}$$

where the inequality in the above derivation follows from (B.40). Because ℓ was an arbitrary element of $M_i(\sigma)$, this establishes the lemma. □

Recall that Π is the set of permutations on C. For any utility function $u \in \mathbb{R}^C$ and $\pi \in \Pi$, define the utility function $u^\pi \in \mathbb{R}^C$ by $u^\pi := u \circ \pi$ (where $\circ$ denotes function composition).

Lemma B.10 *For all $\pi \in \Pi$, $\max\{u^\pi\cdot\ell : \ell \in M_i(\sigma)\} \leq \max\left\{u^\pi\cdot\ell : \ell \in \sum_{S\in\mathscr{C}} \lambda_S\mathbf{S}\right\}$.*

Proof. This is an immediate consequence of Lemma B.6. □

The following proposition is proven in the Appendix of Sher (2018).

Proposition B.2 *For all $M, M' \in \mathscr{M}$, $M \preccurlyeq^* M'$ if and only if:*

$$\frac{1}{m!}\sum_{\pi\in\Pi} \max_{\ell\in M}(u^\pi\cdot\ell) \leq \frac{1}{m!}\sum_{\pi\in\Pi} \max_{\ell\in M'}(u^\pi\cdot\ell), \qquad \forall u \in \mathbb{R}^C.$$

Lemmas B.9 and B.10 and Proposition B.2 imply that if f does not have a monotonic interpretation for i, then $M_i(\sigma) \prec^* \sum_{S\in\mathscr{C}} \lambda_S\mathbf{S}$. This completes the proof of the theorem. □

Corollary 2

First, I generalize some definitions and results in the text. Consider a voting rule of the form $f' : \mathscr{R} \times V_{-i} \to C$. We can interpret this as a voting rule in which voter i submits a ranking R_i and all other voters i' submit votes $v_{i'} \in V$ (and the tie-breaker submits $v_0 \in V_0$). Call such a rule a **preferential for** i.[7] The main text only discusses voting rules for which all agents have the same vote set.[8] I now extend some definitions and results from the text to voting rules that are preferential for i, and in which i may have a different vote set than the others. First, the analogs of Definitions 2 and 4 for rules that are preferential for i is

Definition B.1 *Let i be a voter in I. Let f' be a voting rule that is preferential for i.*

1. *f' is **monotonic for** i if $\forall c \in C, \forall R_i, R_i' \in \mathscr{R}, \forall v_{-i} \in V_{-i}$,*

$$\left[f'\left(R_i, v_{-i}\right) = c \text{ and } \forall c' \in C, \left(cR_ic' \Rightarrow cR_i'c'\right)\right] \Rightarrow f'\left(R_i', v_{-i}\right) = c.$$

2. *f' is **strategyproof for** i if*

$$\forall R_i, R_i' \in \mathscr{R}, \forall v_{-i} \in V_{-i}, f'\left(R_i, v_{-i}\right) R_i f'\left(R_i', v_{-i}\right).$$

Next I state an analog of the Muller Satterthwaite theorem (Proposition 8) for voting rules that are preferential for i.

Proposition B.3 *Let f' be voting rule that is preferential for i. f' is monotonic for i if and only if f' is strategyproof for i.*

The proof is identical as that for the standard Muller Satterthwaite theorem.

With these preliminaries out of the way, first suppose that $f : V_0 \times V^I \to C$ has a monotonic interpretation for i. Then there exists a corresponding interpretation ϕ_i with i-sufficient range V' (see Definition 3). We want to show that f has a strategyproof interpretation for i. Now define $f' : \mathscr{R} \times V_{-i} \to C$ (a rule that is preferential for i) by

$$f'\left(R_i, v_{-i}\right) = f\left(\phi_i\left(R_i\right), v_{-i}\right), \quad \forall R_i \in \mathscr{R}, \forall v_{-i} \in V_{-i}.$$

Because ϕ_i is a monotonic interpretation for i, f' is monotonic for i. So by the Muller Satterthwaite theorem (Proposition B.3), f' is strategyproof for i. Now set $v^{R_i} = \phi_i\left(R_i\right), \forall R_i \in \mathscr{R}$. Choose $v_i' \in V$ and $v_{-i} \in V_{-i}$. Then because V' is i-sufficient, there exists $v_i^{R_i'} \in V'$ such that $f\left(v_i', v_{-i}\right) = f\left(v^{R_i'}, v_{-i}\right)$, and, because f' is strategyproof, we have

$$f\left(v_i^R, v_{-i}\right) = f'\left(R_i, v_{-i}\right) R_i f'\left(R_i', v_{-i}\right) = f\left(v_i^{R'}, v_{-i}\right) = f\left(v_i', v_{-i}\right).$$

So f has a strategyproof interpretation for i.

Going in the other direction, suppose that $f : V_0 \times V^I \to C$ has a strategyproof interpretation for i. Now define $f' : \mathscr{R} \times V_{-i} \to C$ by

$$f'\left(R_i, v_{-i}\right) = f\left(v^{R_i}, v_{-i}\right), \quad \forall R_i \in \mathscr{R}, \forall v_{-i} \in V_{-i}.$$

[7] If $V = \mathscr{R}$, then f' is not only preferential for i, but also preferential.

[8] Nothing essential depends on this.

Then f' is strategyproof for i. So by the Muller Satterthwaite theorem (Proposition B.3), f' is monotonic for i. Now let $\phi_i(R_i) = v^{R_i}, \forall R_i \in \mathscr{R}$, and let $V' = \phi_i(\mathscr{R})$, the range of ϕ_i. Then choose $R_i, R_i' \in \mathscr{R}$, and $c \in C$ such that $f(\phi_i(R_i), v_{-i}) = c$ and $\forall c', cR_ic' \Rightarrow cR_i'c'$. Then $f'(R_i, v_{-i}) = c$. So, since f is monotonic for i, $f'(R_i', v_{-i}) = c$. So $f'(\phi_i(R_i'), v_{-i}) = c$.

To complete the proof that f has a monotonic interpretation for i, it is now sufficient to show that V' is i-sufficient. So choose $v_{-i} \in V_{-i}$ and $c \in C$ such that there exists $v_i \in V$ with $f(v_i, v_{-i}) = c$. Now choose the total order $R_i \in \mathscr{R}$ with $cR_ic', \forall c' \in C$. Then since f has a strategyproof interpretation for i, $f\left(v^{R_i}, v_{-i}\right) R_i f(v_i, v_{-i}) = c$. It follows that $f\left(v^{R_i}, v_{-i}\right) = c$. This establishes that V' is i-sufficient, and completes the proof. □

Proposition 10

Assume that f has a monotonic interpretation. Then, by Corollary 2, for all i, f has a strategyproof interpretation for i.

For each $i \in I$ and $R_i \in \mathscr{R}$, $v_i^{R_i}$ is defined so as to satisfy the condition in Definition 4. Then define $\psi(i, R_i) = v_i^{R_i}$ for all $i \in I$ and $R_i \in \mathscr{R}$.

Next I define χ. To this end, for each $v_0 \in V_0$, define the deterministic preferential voting rule $\tilde{f}_{v_0} : \mathscr{R}^I \to C$ by

$$\tilde{f}_{v_0}(R_1, \ldots, R_n) = f\left(v_0, v_1^{R_1}, \ldots, v_n^{R_n}\right), \quad \forall (R_1, \ldots, R_n) \in \mathscr{R}^I. \tag{B.41}$$

Observe that, given our assumptions, $\tilde{f}_{v_0}$ is strategyproof.

Lemma B.11 range $\left(\tilde{f}_{v_0}\right)$ = range (f_{v_0}).

Proof. It is immediate that range $\left(\tilde{f}_{v_0}\right) \subseteq$ range (f_{v_0}). So we must only prove that range $\left(\tilde{f}_{v_0}\right) \supseteq$ range (f_{v_0}). To do so, choose $v_{-0} \in V^I$. For each $i \in I$, choose $R_i \in \mathscr{R}$ such that $f(v_0, v_{-0}) R_i c, \forall c \in C$. Since R_i is a total order, and f has a strategyproof interpretation for 1, we have

$$f\left(v_0, v_1^{R_1}, v_2, \ldots, v_n\right) = f(v_0, v_{-0}).$$

Now assume for the inductive hypothesis that

$$f\left(v_0, v_1^{R_1}, \ldots, v_{i-1}^{R_{i-1}}, v_i, v_{i+1}, \ldots, v_n\right) = f(v_0, v_{-0}).$$

Then by the same logic as in the base case,

$$f\left(v_0, v_1^{R_1}, \ldots, v_i^{R_i}, v_{i+1}, \ldots, v_n\right) = f(v_0, v_{-0}).$$

It follows that $f\left(v_0, v_1^{R_1}, \ldots, v_n^{R_n}\right) = f(v_0, v_{-0})$. Since $\left(v_1^{R_1}, \ldots, v_n^{R_n}\right) \in V'^I$, it follows that $f(v_0, v_{-0}) \in$ range $\left(\tilde{f}_{v_0}\right)$, and hence range $(f_{v_0}) \subseteq$ range $\left(\tilde{f}_{v_0}\right)$. □

It follows from the lemma that $\left|\text{range}\left(\tilde{f}_{v_0}\right)\right| > 2 \Leftrightarrow |\text{range}(f_{v_0})| > 2$. Choose any $v_0 \in V_0$ with $|\text{range}(f_{v_0})| > 2$. By the Gibbard Satterthwaite theorem (Proposition 8), $\tilde{f}_{v_0}$ is dictatorial on its range. Let i_{v_0} be the "dictator" in $\tilde{f}_{v_0}$, and define $\chi(v_0) = i_{v_0}$. If v_0 is such that $|\text{range}(f_{v_0})| = 2$, then let $\chi(v_0)$ be arbitrary.

Now choose $v^* = (v_0^*, v_1^*, \ldots, v_i^*, \ldots, v_n^*) \in \bar{V}$, $i \in I$, and $R_i \in \mathscr{R}$. First, suppose that the antecedent in condition 2 holds with v^* playing the role of v, that is $\left|\text{range}\left(f_{v_0^*}\right)\right| > 2$, $\chi(i, R_i) = v_i^*$, and $\chi(v_0) = i$.

Lemma B.12 *There exists $R_{-i} \in \mathscr{R}^{I \setminus i}$ such that $\tilde{f}_{v_0^*}(R_i, R_{-i}) = f(v^*)$.*

The proof is almost identical to the proof of Lemma B.11. It now follows that

$$f(v^*) = \tilde{f}_{v_0^*}(R_i, R_{-i}) \in \hat{c}\left(\text{range}\left(\tilde{f}_{v_0^*}\right), R_i\right) = \hat{c}\left(\text{range}\left(f_{v_0^*}\right), R_i\right),$$

where the first equality follows from Lemma B.12, the last equality from Lemma B.11, that the middle $\in$ relation follows from the fact that $i = \chi(v_0^*)$ is a dictator (in the range of $\tilde{f}_{v_0^*}$).

Finally suppose that the antecedent in condition 1 of the proposition holds, with v^* playing the role of v. Then $v_i^* = v_i^R$. It follows from the fact that $f\left(v_i^R, v_{-i}^*\right) Rf\left(v_i, v_{-i}^*\right), \forall v_i \in V_i$, that the consequent of 1 holds (with v^* playing the role of v). $\square$

Example of a voting rule that is dictatorial when not binary but not monotonic

This section informally sketches an example of a class of voting rules that are dictatorial when not binary but not monotonic. Suppose that there are 3 candidates. Voter 1 is the dictator, and may choose her favorite candidate through her vote. However, voter 1has a vote that allows her to abstain, effectively ceding his power. If voter 1 does this then the residual election among the other voters is not monotonic. Such an election is dictatorial when not binary (it satisfies the conditions in Definition 6), but it is not monotonic.

Theorem 4

Part 1 of the theorem is an immediate consequence of Theorem 3 and Proposition 10.

For part 2, let f be a voting rule that is not monotonic for some voter i. Choose μ satisfying $\mu(v) > 0$ for all $v \in \bar{V}$. Let $\sigma = (f, \mu)$. Now, for each $\varepsilon \in [0, 1]$, I construct a mechanism f_i^ε. f_i^ε is subscripted by i because f_i^ε treats voter i in a special way. f_i^ε has vote set $\mathscr{R}$. f_i^ε works as follows. For $k = 1, \ldots, m$, with probability $(1 - \varepsilon)\,\text{Piv}_i(k; \sigma)$, the winning candidate is the candidate that is maximal in the menu $\mathbf{k}$ with respect to voter i's vote R_i (i.e., with respect to i's submitted ranking). For $j = 1, \ldots, i-1, i+1, \ldots, n$, with probability $\frac{\varepsilon}{n-1}$, the winning candidate is the candidate who is maximal (among all candidates) with respect to voter j's vote R_j. Let μ' be an *arbitrary* vote profile distribution in $\Delta\left(V_0 \times \mathscr{R}^I\right)$ of the form $\mu' = \prod_j \mu'_j$, and let $\sigma^\varepsilon = (f^\varepsilon, \mu')$.

By construction,

$$M_i(\sigma^\varepsilon) \sim_i^* \varepsilon \mathbf{1} + (1 - \varepsilon) \sum_{k=1}^{m} \text{Piv}_i(k; \sigma)\,\mathbf{k} \tag{B.42}$$

(Observe that this does not depend on μ'.) We have[9]

$$M_i(\sigma^\varepsilon) \to \sum_{k=1}^{m} \text{Piv}_i(k; \sigma)\,\mathbf{k} \text{ as } \varepsilon \to 0.$$

Because f is not monotone, it follows from Theorem 3 that

$$M_i(\sigma) \prec^* \sum_{k=1}^{m} \text{Piv}_i(k; \sigma)\,\mathbf{k}.$$

[9] Convergence is in the metric topology generated by the Hausdorff distance.

By the continuity axiom (Axiom 2 in Section 2.3.1), if $\varepsilon > 0$ is sufficiently small,

$$M_i(\sigma) \prec^* \varepsilon \mathbf{1} + (1-\varepsilon) \sum_{k=1}^{m} \mathrm{Piv}_i(k;\sigma)\,\mathbf{k} \tag{B.43}$$

It follows from the fact that f is not monotonic for i, that $\sum_{k=1}^{m} \mathrm{Piv}_i(k;\sigma)\,\mathbf{k} \neq \mathbf{1}$. It follows from the independence axiom that (Axiom 4) that for $\varepsilon > 0$,

$$\begin{aligned} &\varepsilon \mathbf{1} + (1-\varepsilon) \sum_{k=1}^{m} \mathrm{Piv}_i(k;\sigma)\,\mathbf{k} \\ \prec_i^* \; &\varepsilon \sum_{k=1}^{m} \mathrm{Piv}_i(k;\sigma)\,\mathbf{k} + (1-\varepsilon) \sum_{k=1}^{m} \mathrm{Piv}_i(k;\sigma)\,\mathbf{k} \\ = &\sum_{k=1}^{m} \mathrm{Piv}_i(k;\sigma)\,\mathbf{k} \end{aligned} \tag{B.44}$$

If follows from (B.42) and (B.43) that for sufficiently small $\varepsilon > 0$, $f \prec_i^{\circ} f_i^{\varepsilon}$. It follows from (B.44) that $f_i^{\varepsilon} \prec_i^{\mathrm{piv}} f$. This completes the proof of part 2. □

Proposition 11

This proposition is an immediate consequence of (6) and Proposition B.2 (which is presented in the proof of Theorem 3). □

Facts from Section 5

I begin by specifying the tie-breaker in σ_3. I assume that $V_0 = \mathscr{R}$, where recall that $\mathscr{R}$ is the set of total orders R_0 over C, which in this case, is equal to $\{1,2,3\}$. The tie-breaker selects each R_0 in $\mathscr{R}$ with probability $\frac{1}{6}$. If there is a tie for first place, then the tied candidate that is ranked highest by the tie-breaker wins.

Recall that there are $n+1$ voters. We look at things from the perspective of voter i. In σ_3 a voter can be pivotal between two candidates in two different ways. First, within the two candidates can be tied for first place in the votes of voters other than i. This can occur in $3 \times \sum_{k=\frac{n}{3}+1}^{\frac{n}{2}} \binom{n}{k,k,n-2k}$ ways. We multiply by 3 because there are three sets of 2 candidates that might be tied for first. Note that because n is odd, the candidate not tied for first place must be losing by at least two votes. So regardless of the tie-breaker, i cannot bring it about that this candidate wins. The probability of each configuration of votes for voters other than i is $\frac{1}{3^n}$ because each voter can vote in three ways. So the probability that two candidates are tied for first place among voters other than i is $\frac{1}{3^n} \times 3 \times \sum_{k=\frac{n}{3}+1}^{\frac{n}{2}} \binom{n}{k,k,n-2k}$. There is a second way that i can be pivotal over two candidates: That is if the candidate in first place is one vote ahead of the candidate in second place in v_{-i} and the tie-breaker favors the candidate in second place. This occurs with probability $\frac{1}{2} \times \frac{1}{3^n} \times 6 \times \sum_{k=\frac{n}{3}+1}^{\frac{n}{2}} \binom{n}{k,k-1,n-2k+1}$, where we multiply by $\frac{1}{2}$ because the tie-breaker favors the candidate in second place over the candidate in first place with probability $\frac{1}{2}$.[10] We multiply by 6 because there are six sequences of candidates in first, second, and third place. The previous expression simplifies to $\frac{1}{3^n} \times 3 \times \sum_{k=\frac{n}{3}+1}^{\frac{n}{2}} \binom{n}{k,k-1,n-2k+1}$. The above analysis accounts for the expression for $\mathrm{Piv}_i(2;\sigma_2)$ in (15).

[10] Observe that because n is divisible by 6, $k-1 > n-2k+1$.

Voter i is pivotal among all three candidates only if all three are tied for first place. (Recall that because n is even, it is not possible that two voters are tied for first place in v_{-i}, and the third voter is one vote behind.) The probability that all three candidates are tied for first place is $\frac{1}{3^n}\binom{n}{\frac{n}{3},\frac{n}{3},\frac{n}{3}}$. This accounts for the expression for $\mathrm{Piv}\,(3;\sigma_3)$ in (15).

Next I explain the menus in (16). First consider random dictatorship with with three candidates, and consider the lottery induced by a vote for candidate 1. If voter i is selected to be dictator, then candidate 1 will win. This occurs with probability $\frac{1}{n+1}$. If i is not selected dictator, which happens with probability $\frac{n}{n+1}$, candidate 1 wins $\frac{1}{3}$ of the time. So the total probability that 1 wins if i votes for 1 is $\frac{1}{n+1}+\frac{1}{3}\frac{n}{n+1}=\frac{3+n}{3n+3}$. The remaining probability of $1-\frac{3+n}{3n+3}=\frac{2n}{3n+3}$ is evenly split among the other two candidates. This explains the first lottery in $M\left(\sigma_3^r\right)$. The other two lotteries are explained similarly. Next consider plurality voting among three candidates. Suppose that voter i votes for 1. Then when i is pivotal over all three candidates, 1 will win. This occurs with probability $\mathrm{Piv}_i\,(3;\sigma_3)$. Also, when i is pivotal over 1 and 2 or 1 and 3, then 1 will win. This occurs with probability $\frac{2}{3}\mathrm{Piv}_i\,(2;\sigma_3)$. Note, however that when i is pivotal over 2 and 3, then 1 will *lose*. This happens with probability $\frac{1}{3}\mathrm{Piv}_i\,(2;\sigma_3)$. When i is not pivotal at all, which occurs with probability $1-\mathrm{Piv}_i\,(2;\sigma_3)-\mathrm{Piv}_i\,(3;\sigma_3)$, 1 wins with probability $\frac{1}{3}$. Thus in total, if i votes for 1, 1 will win with probability $\mathrm{Piv}_i\,(3;\sigma_3)+\frac{2}{3}\mathrm{Piv}_i\,(2;\sigma_3)+\frac{1}{3}\left(1-\mathrm{Piv}_i\,(2;\sigma_3)-\mathrm{Piv}_i\,(3;\sigma_3)\right)=\frac{1}{3}+\frac{1}{3}\mathrm{Piv}_i\,(2;\sigma_3)+\frac{2}{3}\mathrm{Piv}_i\,(3;\sigma_3)=\frac{1+2\Phi_n}{3}$. The remaining probability of $2\left(\frac{1-\Phi_n}{3}\right)$ is evenly split among the other two candidates. This explains the first lottery in $M_i\left(\sigma_3\right)$. The other two lotteries are explained similarly.

Finally, I explain the expressions of the form $V_a\left(\sigma\right)$ and $V_b\left(\sigma\right)$ in Section 5.3. First, I explain $V_a\left(\hat{\sigma}_3\right)$. As explained in the text, if the tie-breaker selects candidates 2 and 3 to participate in the election, Ann's utility is 0. This happens with probability $\frac{1}{3}$. If the tie-breaker selects candidates 1 and 2 to participate in the election, which occurs with probability $\frac{1}{3}$, then if Ann is pivotal between 1 and 2, which happens with probability $\mathrm{Piv}_i\,(2;\sigma_2)$, then candidate 1 wins. When Ann is not pivotal, 1 wins with probability $\frac{1}{2}$. Hence conditional on the tie-breaker selecting 1 and 2, Ann's expected utility is $\mathrm{Piv}_i\,(2,\sigma_2)+\frac{1}{2}\left(1-\mathrm{Piv}_i\,(2,\sigma_2)\right)=\frac{1}{2}+\frac{1}{2}\mathrm{Piv}_i\,(2,\sigma_2)$. Ann gets the same expected utility conditional on the tie-breaker selecting 1 and 3. Hence Ann's overall expected utility is $V_a\left(\hat{\sigma}_3\right)=\frac{1}{3}\times 0+\frac{2}{3}\left[\frac{1}{2}+\frac{1}{2}\mathrm{Piv}_i\,(2;\sigma_2)\right]=\frac{1}{3}+\frac{1}{3}\mathrm{Piv}_i\,(2;\sigma_2)$.

Let us now calculate $V_b\left(\hat{\sigma}_3\right)$. If the tie-breaker selects 1 and 2 to participate in the election, which happens with probability $\frac{1}{3}$, then Bob gets a utility of 1. If the tie-breaker selects 1 and 3 to participate in the election, which also happens with probability $\frac{1}{3}$, then Bob gets a utility of $\frac{1}{2}+\frac{1}{2}\mathrm{Piv}_i\,(2;\sigma_2)$. The same applies if the tie-breaker selects 2 and 3 to participate in the election. So, overall, Bob gets a utility of $V_b\left(\hat{\sigma}_3\right)=\frac{1}{3}+\frac{2}{3}\left[\frac{1}{2}+\frac{1}{2}\mathrm{Piv}_i\,(2;\sigma_2)\right]=\frac{2}{3}+\frac{1}{3}\mathrm{Piv}_i\,(2;\sigma_2)$.

Next, I calculate $V_a\left(\sigma_3\right)$. Under plurality voting, Ann will vote for 1, and Ann's expected utility is equal to the probability that 1 wins, given that Ann votes for 1. We already calculated this probability when calculating the menu $M_i\left(\sigma_3\right)$. Using that calculation, Ann's expected utility is $V_a\left(\sigma_3\right)=\frac{1}{3}+\frac{1}{3}\mathrm{Piv}_i\,(2;\sigma_3)+\frac{2}{3}\mathrm{Piv}_i\,(3;\sigma_3)$.

Finally, I calculate $V_b\left(\sigma_3\right)$. Under plurality voting, Bob will vote for 1 or 2. He is indifferent between the two. So we may assume without loss of generality that he votes for 1. Then Bob's expected utility will be the probability that either 1 or 2 will win if he votes for 1. We already calculated each of these two probabilities above. So Bob's expected utility is $V_b\left(\sigma_3\right)=\frac{1}{3}+\frac{1}{3}\mathrm{Piv}_i\,(2;\sigma_3)+\frac{2}{3}\mathrm{Piv}_i\,(3;\sigma_3)+\frac{1}{3}\left[1-\frac{1}{2}\mathrm{Piv}_i\,(2;\sigma_3)-\mathrm{Piv}_i\,(3;\sigma_3)\right]=\frac{2}{3}+\frac{1}{6}\mathrm{Piv}_i\,(2;\sigma_3)+\frac{1}{3}\mathrm{Piv}_i\,(3;\sigma_3)$.

References for Appendix

Call, G. S. and Velleman, D. J. (1993), 'Pascal's matrices', *American Mathematical Monthly* pp. 372–376.

Dekel, E., Lipman, B. L. and Rustichini, A. (2001), 'Representing preferences with a unique subjective state space', *Econometrica* **69**(4), 891–934.

Knuth, D. E. (1997), *The art of computer programming: fundamental algorithms*, Addison Wesley Longman.

Kuhn, H. W. (1953), Extensive games and the problem of information, *in* H. W. Kuhn and A. W. Tucker, eds, 'Contributions to the Theory of Games, II', Vol. 28 of *Annals of Mathematics Studies*, pp. 193–216.

Nehring, K. (1999), 'Preference for flexibility in a Savage framework', *Econometrica* **67**(1), 101–119.

Sher, I. (2018), 'Neutral freedom and freedom as control'. Available at SSRN `https://ssrn.com/abstract=2652913`.