SALEM GAZETTE

VOLUME XXVII.] SALEM, FRIDAY, APRIL 2, 1812.

FEDERALISTS! FOLLOWERS OF WASHINGTON!

AGAIN behold and shudder at the exhibition of this terrific Dragon, brought forth to swallow and devour your Liberties and *equal* Rights. Unholy party spirit and inordinate love of power gave it birth,—your patriotism and hatred of tyranny must by one vigorous struggle strangle it in its infancy. The iniquitous Law, which cut up and severed this Commonwealth into Districts, is kindred to the arbitrary deeds of NAPOLEON when he partitioned the territories of innocent nations to suit his sovereign will. This Law inflicted a grievous wound on the Constitution,—it in fact subverts and *changes our Form of Government*, which ceases to be *Republican* as long as an *Aristocratic* HOUSE OF LORDS under the form of a Senate tyrannizes over the People, and silences and stifles the voice of the *Majority*.

When Tyranny and arbitrary Power thus make inroads upon the Rights of the People, what becomes the duty of the citizen? Shall he submit quietly and ignominiously to the decrees of Usurpers? Are the citizens of this Republic less jealous of their rights than their ancestors? Will you, then, permit a Party to disfranchise the People,—to convert the Senate Chamber into a Fortress in which ambitious office-seekers may entrench themselves and set at defiance the frowns of the People? No,—this usurping Faction must be dislodged from its strong-hold.

Arise, then, *injured Citizens!* Turn out! turn out! let Monday Next be the day of your Emancipation,—by one manful Struggle reclaim your usurped Rights—and frown into obscurity those audacious men who unblushingly boasted—"*We have secured the Senate for ten years, and should have been Fools if we had not done it.*" Prove on election day that the Folly of their men is equal to their want of honesty and contempt of the People. Elect patriots who will be loyal to the Constitution, and faithful to the interests of the State,—men who will harmonize with our illustrious STRONG,—and who will not sell our Drafted Soldiery, like Hessians, to any Foreign Power.

Salem,

FRIDAY, APRIL 2, 1812.

SOUTH with a Vengeance!

The Right of Suffrage.

Voting Theory Primer for Rationalists

1. A voting theory primer for rationalists
2. 5 general voting pathologies: lesser names of Moloch
3. Multi-winner Voting: a question of Alignment

A voting theory primer for rationalists

What is voting theory?

Voting theory, also called social choice theory, is the study of the design and evaulation of democratic voting methods (that's the activists' word; game theorists call them "voting mechanisms", engineers call them "electoral algorithms", and political scientists say "electoral formulas"). In other words, for a given list of candidates and voters, a voting method specifies a set of valid ways to fill out a ballot, and, given a valid ballot from each voter, produces an outcome.

(An "electoral system" includes a voting method, but also other implementation details, such as how the candidates and voters are validated, how often elections happen and for what offices, etc. "Voting system" is an ambiguous term that can refer to a full electoral system, just to the voting method, or even to the machinery for counting votes.)

Most voting theory limits itself to studying "democratic" voting methods. That typically has both empirical and normative implications. Empirically, "democratic" means:

- There are many voters
- There can be more than two candidates

In order to be considered "democratic", voting methods generally should meet various normative criteria as well. There are many possible such criteria, and on many of them theorists do not agree; but in general they do agree on this minimal set:

- Anonymity; permuting the ballots does not change the probability of any election outcome.
- Neutrality; permuting the candidates on all ballots does not change the probability of any election outcome.
- Unanimity: If voters universally vote a preference for a given outcome over all others, that outcome is selected. (This is a weak criterion, and is implied by many other stronger ones; but those stronger ones are often disputed, while this one rarely is.)
- Methods typically do not directly involve money changing hands or other enduring state-changes for individual voters. (There can be exceptions to this, but there are good reasons to want to understand "moneyless" elections.)

Why is voting theory important for rationalists?

First off, because democratic processes in the real world are important loci of power. That means that it's useful to understand the dynamics of the voting methods used in such real-world elections.

Second, because these real-world democratic processes have all been created and/or evolved in the past, and so there are likely to be opportunities to replace, reform, or

add to them in the future. If you want to make political change of any kind over a medium-to-long time horizon, these systemic reforms should probably be part of your agenda. The fact is that **FPTP, the voting method we use in most of the English-speaking world, is absolutely horrible**, and there is reason to believe that reforming it would substantially (though not of course completely) alleviate much political dysfunction and suffering.

Third, because understanding social choice theory helps clarify ideas about how it's possible and/or desirable to resolve value disputes between multiple agents. For instance, if you believe that superintelligences should perform a "values handshake" when meeting, replacing each of their individual value functions by some common one so as to avoid the dead weight loss of a conflict, then social choice theory suggests both questions and answers about what that might look like. (Note that the ethical and practical importance of such considerations is not at all limited to "post-singularity" examples like that one.)

In fact, on that third point: my own ideas of ethics and of fun theory are deeply informed by my decades of interest in voting theory. To simplify into a few words my complex thoughts on this, I believe that voting theory elucidates "ethical incompleteness" (that is, that it's possible to put world-states into ethical preference order partially but not fully) and that this incompleteness is a good thing because it leaves room for fun even in an ethically unsurpassed world.

What are the branches of voting theory?

Generally, you can divide voting methods up into "single-winner" and "multi-winner". Single-winner methods are useful for electing offices like president, governor, and mayor. Multi-winner methods are useful for dividing up some finite, but to some extent divisible, resource, such as voting power in a legislature, between various options. Multi-winner methods can be further subdivided into seat-based (where a set of similar "seats" are assigned one winner each) or weighted (where each candidate can be given a different fraction of the voting power).

What are the basics of single-winner voting theory?

(Note: Some readers may wish to skip to the summary below, or to read the later section on multi-winner theory and proportional representation first. Either is valid.)

Some of the earliest known work in voting theory was by Ramon Llull before his death in 1315, but most of that was lost until recently. Perhaps a better place to start would be in the French Academy in the late 1700s; this allows us to couch it as a debate (American Chopper meme?) between Jean-Charles de Borda and Nicolas de Condorcet.

> **Condorcet**: "Plurality (or 'FPTP', for First Past the Post) elections, where each voter votes for just one candidate and the candidate with the most votes wins, are often spoiled by vote-splitting."

Borda: "Better to have voters rank candidates, give candidates points for favorable rankings, and choose a winner based on points." (Borda Count)

Condorcet: "Ranking candidates, rather than voting for just one, is good. But your point system is subject to strategy. Everyone will rate some candidate they believe can't win in second place, to avoid giving points to a serious rival to their favorite. So somebody could win precisely because nobody takes them seriously!"

Borda: "My method is made for honest men!"

Condorcet: "Instead, you should use the rankings to see who would have a majority in every possible pairwise contest. If somebody wins all such contests, obviously they should be the overall winner."

In my view, Borda was the clear loser there. And most voting theorists today agree with me. The one exception is the mathematician Donald Saari, enamored with the mathematical symmetry of the Borda count. This is totally worth mentioning because his last name is a great source of puns.

But Condorcet soon realized there was a problem with his proposal too: it's possible for A to beat B pairwise, and B to beat C, while C still beats A. That is, pairwise victories can be cyclical, not transitive. Naturally speaking, this is rare; but if there's a decision between A and B, the voters who favor B might have the power to artificially create a "poison pill" amendment C which can beat A and then lose to B.

How would a Condorcet cycle occur? Imagine the following election:

1: A>B>C

1: B>C>A

1: C>A>B

(This notation means that there's 1 voter of each of three types, and that the first voter prefers A over B over C.) In this election, A beats B by 2 to 1, and similarly B beats C and C beats A.

Fast-forward to 1950, when theorists at the RAND corporation were inventing game theory in order to reason about the possibility of nuclear war. One such scientist, Kenneth Arrow, proved that the problem that Condorcet (and Llull) had seen was in fact a fundamental issue with any ranked voting method. He posed 3 basic "fairness criteria" and showed that no ranked method can meet all of them:

- Ranked unanimity: if every voter prefers X to Y, then the outcome has X above Y.
- Independence of irrelevant alternatives: If every voter's preferences between some subset of candidates remain the same, the order of those candidates in the outcome will remain the same, even if other candidates outside the set are added, dropped, or changed.
- Non-dictatorial: the outcome depends on more than one ballot.

Arrow's result was important in and of itself; intuitively, most people might have guessed that a ranked voting method could be fair in all those ways. But even more important than the specific result was the idea of an impossibility proof for voting.

Using this idea, it wasn't long until Gibbard and Satterthwaite independently came up with a follow-up theorem, showing that no voting system (ranked or otherwise) could possibly avoid creating strategic incentives for some voters in some situations. That is to say, there is no non-dictatorial voting system for more than two possible outcomes and more than two voters in which every voter has a single "honest" ballot that depends only on their own feelings about the candidates, such that they can't sometimes get a better result by casting a ballot that isn't their "honest" one.

There's another way that Arrow's theorem was an important foundation, particularly for rationalists. He was explicitly thinking about voting methods not just as real-world ways of electing politicians, but as theoretical possibilities for reconciling values. In this more philosophical sense, Arrow's theorem says something depressing about morality: if morality is to be based on (potentially revealed) preferences rather than interpersonal comparison of (subjective) utilities, it cannot simply be a democratic matter; "the greatest good for the greatest number" doesn't work without inherently-subjective comparisons of goodness. Amartya Sen continued exploring the philosophical implications of voting theory, showing that the idea of "private autonomy" is incompatible with Pareto efficiency.

Now, in discussing Arrow's theorem, I've said several times that it only applies to "ranked" voting systems. What does that mean? "Ranked" (also sometimes termed "ordinal" or "preferential") systems are those where valid ballots consist of nothing besides a transitive preferential ordering of the candidates (partial or complete). That is, you can say that you prefer A over B or B over A (or in some cases, that you like both of them equally), but you cannot say how strong each preference is, or provide other information that's used to choose a winner. In Arrow's view, the voting method is then responsible for ordering the candidates, picking not just a winner but a second place etc. Since neutrality wasn't one of Arrow's criteria, ties can be broken arbitrarily.

This excludes an important class of voting methods from consideration: those I'd call rated (or graded or evaluational), where you as a voter can give information about strength of preference. Arrow consciously excluded those methods because he believed (as Gibbard and Satterthwaite later confirmed) that they'd inevitably be subject to strategic voting. But since ranked voting systems are also inevitably subject to strategy, that isn't necessarily a good reason. In any case, Arrow's choice to ignore such systems set a trend; it wasn't until approval voting was reinvented around 1980 and score voting around 2000 that rated methods came into their own. Personally, for reasons I'll explain further below, I tend to prefer rated systems over purely ranked ones, so I think that Arrow's initial neglect of ranked methods got the field off on a bit of a wrong foot.

And there's another way I feel that Arrow set us off in the wrong direction. His idea of reasoning axiomatically about voting methods was brilliant, but ultimately, I think the field has been too focused on this axiomatic "Arrovian" paradigm, where the entire goal is to prove certain criteria can be met by some specific voting method, or cannot be met by any method. Since it's impossible to meet all desirable criteria in all cases, I'd rather look at things in a more probabilistic and quantitative way: how often and how badly does a given system fail desirable criteria.

The person I consider to be the founder of this latter, "statistical" paradigm for evaluating voting methods is Warren Smith. Now, where Kenneth Arrow won the Nobel Prize, Warren Smith has to my knowledge never managed to publish a paper in a peer-reviewed journal. He's a smart and creative mathematician, but... let's just say, not exemplary for his social graces. In particular, he's not reluctant to opine in varied

fields of politics where he lacks obvious credentials. So there's plenty in the academic world who'd just dismiss him as a crackpot, if they are even aware of his existence. This is unfortunate, because his work on voting theory is groundbreaking.

In his 2000 paper on "Range Voting" (what we'd now call Score Voting), he performed systematic utilitarian Monte-Carlo evaluation of a wide range of voting systems under a wide range of assumptions about how voters vote. In other words, in each of his simulations, he assumed certain numbers of candidates and of voters, as well as a statistical model for voter utilities and a strategy model for voters. Using the statistical model, he assigned each virtual voter a utility for each candidate; using the strategy model, he turned those utilities into a ballot in each voting method; and then he measured the total utility of the winning candidate, as compared to that of the highest-total-utility candidate in the race. Nowadays the name for the difference between these numbers, scaled so that the latter would be 100% and the average randomly-selected candidate would be 0%, is "Voter Satisfaction Efficiency" (VSE).

Smith wasn't the first to do something like this. But he was certainly the first to do it so systematically, across various voting methods, utility models, and strategic models. Because he did such a sensitivity analysis across utility and strategic models, he was able to see which voting methods consistently outperformed others, almost regardless of the specifics of the models he used. In particular, score voting, in which each voter gives each candidate a numerical score from a certain range (say, 0 to 100) and the highest total score wins, was almost always on top, while FPTP was almost always near the bottom.

More recently, I've done further work on VSE, using more-realistic voter and strategy models than what Smith had, and adding a variety of "media" models to allow varying the information on which the virtual voters base their strategizing. While this work confirmed many of Smith's results — for instance, I still consistently find that FPTP is lower than IRV is lower than approval is lower than score — it has unseated score voting as the undisputed highest-VSE method. Other methods with better strategy resistance can end up doing better than score.

Of course, something else happened in the year 2000 that was important to the field of single-winner voting theory: the Bush-Gore election, in which Bush won the state of Florida and thus the presidency of the USA by a microscopic margin of about 500 votes. Along with the many "electoral system" irregularities in the Florida election (a mass purge of the voter rolls of those with the same name as known felons, a confusing ballot design in Palm Beach, antiquated punch-card ballots with difficult-to-interpret "hanging chads", etc.) was one important "voting method" irregularity: the fact that Ralph Nader, a candidate whom most considered to be ideologically closer to Gore than to Bush, got far more votes than the margin between the two, leading many to argue that under almost any alternative voting method, Gore would have won. This, understandably, increased many people's interest in voting theory and voting reform. Like Smith, many other amateurs began to make worthwhile progress in various ways, progress which was often not well covered in the academic literature.

In the years since, substantial progress has been made. But we activists for voting reform still haven't managed to use our common hatred for FPTP to unite behind a common proposal. (The irony that our expertise in methods for reconciling different priorities into a common purpose hasn't let us do so in our own field is not lost on us.)

In my opinion, aside from the utilitarian perspective offered by VSE, the key to evaluating voting methods is an understanding of strategic voting; this is what I'd call

the "mechanism design" perspective. I'd say that there are 5 common "anti-patterns" that voting methods can fall into; either where voting strategy can lead to pathological results, or vice versa. I'd pose them as a series of 5 increasingly-difficult hurdles for a voting method to pass. Because the earlier hurdles deal with situations that are more common or more serious, I'd say that if a method trips on an earlier hurdle, it doesn't much matter that it could have passed a later hurdle. Here they are:

(0. Dark Horse. As in Condorcet's takedown of Borda above, this is where a candidate wins precisely because nobody expects them to. Very bad, but not a serious problem in most voting methods, except for the Borda Count.)

1. Vote-splitting / "spoiled" elections. Adding a minor candidate causes a similar major candidate to lose. Very bad because it leads to rampant strategic dishonesty and in extreme cases 2-party dominance, as in Duverger's Law. Problematic in FPTP, resolved by most other voting methods.

2. Center squeeze. A centrist candidate is eliminated because they have lost first-choice support to rivals on both sides, so that one of the rivals wins, even though the centrist could have beaten either one of them in a one-on-one (pairwise) election. Though the direct consequences of this pathology are much less severe than those of vote-splitting, the indirect consequences of voters strategizing to avoid the problem would be exactly the same: self-perpetuating two-party dominance. This problem is related to failures of the "favorite betrayal criterion" (FBC). Problematic in IRV, resolved by most other methods.

3. Chicken dilemma (aka Burr dilemma, for Hamilton fans). Two similar candidates must combine strength in order to beat a third rival. But whichever of the two cooperates less will be the winner, leading to a game of "chicken" where both can end up losing to the rival. This problem is related to failures of the "later-no-harm" (LNH) criterion. Because LNH is incompatible with FBC, it is impossible to completely avoid the chicken dilemma without creating a center squeeze vulnerability, but systems like STAR voting or 3-2-1 minimize it.

4. Condorcet cycle. As above, a situation where, with honest votes, A beats B beats C beats A. There is no "correct" winner in this case, and so no voting method can really do anything to avoid getting a "wrong" winner. Luckily, in natural elections (that is, where bad actors are not able to create artificial Condorcet cycles by strategically engineering "poison pills"), this probably happens less than 5% of the time.

Note that there's a general pattern in the pathologies above: the outcome of honest voting and that of strategic voting are in some sense polar opposites. For instance, under honest voting, vote-splitting destabilizes major parties; but under strategic voting, it makes their status unassailable. This is a common occurrence in voting theory. And it's a reason that naive attempts to "fix" a problem in a voting system by adding rules can actually make the original problem worse.

(I wrote a separate article with further discussion of these pathologies)

Here are a few of the various single-winner voting systems people favor, and a few (biased) words about the groups that favor them:

FPTP (aka plurality voting, or choose-one single-winner): Universally reviled by voting theorists, this is still favored by various groups who like the status quo in countries like

the US, Canada, and the UK. In particular, incumbent politicians and lobbyists tend to be at best skeptical and at worst outright reactionary in response to reformers.

IRV (Instant runoff voting), aka Alternative Vote or RCV (Ranked Choice Voting... I hate that name, which deliberately appropriates the entire "ranked" category for this one specific method): This is a ranked system where to start out with, only first-choice votes are tallied. To find the winner, you successively eliminate the last-place candidate, transferring those votes to their next surviving preference (if any), until some candidate has a majority of the votes remaining. It's supported by FairVote, the largest electoral reform nonprofit in the US, which grew out of the movement for STV proportional representation (see the multi-winner section below for more details). IRV supporters tend to think that discussing its theoretical characteristics is a waste of time, since it's so obvious that FPTP is bad and since IRV is the reform proposal with by far the longest track record and most well-developed movement behind it. Insofar as they do consider theory, they favor the "later-no-harm" criterion, and prefer to ignore things like the favorite betrayal criterion, summability, or spoiled ballots. They also don't talk about the failed Alternative Vote referendum in the UK.

Approval voting: This is the system where voters can approve (or not) each candidate, and the candidate approved by the most voters wins. Because of its simplicity, it's something of a "Schelling point" for reformers of various stripes; that is, a natural point of agreement as an initial reform for those who don't agree on which method would be an ideal end state. This method was used in Greek elections from about 1860-1920, but was not "invented" as a subject of voting theory until the late 70s by Brams and Fishburn. It can be seen as a simplistic special case of many other voting methods, in particular score voting, so it does well on Warren Smith's utilitarian measures, and fans of his work tend to support it. This is the system promoted by the Center for Election Science (electology.org), a voting reform nonprofit that was founded in 2012 by people frustrated with FairVote's anti-voting-theory tendencies. (Full disclosure: I'm on the board of the CES, which is growing substantially this year due to a significant grant by the Open Philanthropy Project. Thanks!)

Condorcet methods: These are methods that are guaranteed to elect a pairwise beats-all winner (Condorcet winner) if one exists. Supported by people like Erik Maskin (a Nobel prize winner in economics here at Harvard; brilliant, but seemingly out of touch with the non-academic work on voting methods), and Markus Schulze (a capable self-promoter who invented a specific Condorcet method and has gotten groups like Debian to use it in their internal voting). In my view, these methods give good outcomes, but the complications of resolving spoil their theoretical cleanness, while the difficulty of reading a matrix makes presenting results in an easy-to-grasp form basically impossible. So I personally wouldn't recommend these methods for real-world adoption in most cases. Recent work in "improved" Condorcet methods has showed that these methods can be made good at avoiding the chicken dilemma, but I would hate to try to explain that work to a layperson.

Bucklin methods (aka median-based methods; especially, Majority Judgment): Based on choosing a winner with the highest median rating, just as score voting is based on choosing one with the highest average rating. Because medians are more robust to outliers than averages, median methods are more robust to strategy than score. Supported by French researchers Balinski and Laraki, these methods have an interesting history in the progressive-era USA. Their VSE is not outstanding though; better than IRV, plurality, and Borda, but not as good as most other methods.

Delegation-based methods, especially SODA (simple optionally-delegated approval): It turns out that this kind of method can actually do the impossible and "avoid the Gibbard-Satterthwaite theorem in practice". The key words there are "in practice" — the proof relies on a domain restriction, in which voters honest preferences all agree with their favorite candidate, and these preference orders are non-cyclical, and voters mutually know each other to be rational. Still, this is the only voting system I know of that's 100% strategy free (including chicken dilemma) in even such a limited domain. (The proof of this is based on complicated arguments about convexity in high-dimensional space, so Saari, it doesn't fit here.) Due to its complexity, this is probably not a practical proposal, though.

Rated runoff methods (in particular STAR and 3-2-1): These are methods where rated ballots are used to reduce the field to two candidates, who are then compared pairwise using those same ballots. They combine the VSE advantages of score or approval with extra resistance to the chicken dilemma. These are currently my own favorites as ultimate goals for practical reform, though I still support approval as the first step.

Quadratic voting: Unlike all the methods above, this is based on the universal solvent of mechanism design: money (or other finite transferrable resources). Voters can buy votes, and the cost for n votes is proportional to n^2. This has some excellent characteristics with honest voters, and so I've seen that various rationalists think it's a good idea; but in my opinion, it's got irresolvable problems with coordinated strategies. I realize that there are responses to these objections, but as far as I can tell every problem you fix with this idea leads to two more.

TL; DR?

- Plurality voting is really bad. (Borda count is too.)
- Arrow's theorem shows no ranked voting method is perfect.
- Gibbard-Satterthwaite theorem shows that no voting method, ranked or not, is strategy-free in all cases.
- Rated voting methods such as approval or score can get around Arrow, but not Gibbard-Satterthwaite.
- Utilitarian measures, known as VSE, are one useful way to evaluate voting methods.
- Another way is mechanism design. There are (1+)4 voting pathologies to worry about. Starting from the most important and going down: (Dark horse rules out Borda;) vote-splitting rules out plurality; center squeeze would rule out IRV; chicken dilemma argues against approval or score and in favor of rated runoff methods; and Condorcet cycles mean that even the best voting methods will "fail" in a few percent of cases.

What are the basics of multi-winner voting theory?

Multi-winner voting theory originated under parliamentary systems, where theorists wanted a system to guarantee that seats in a legislature would be awarded in proportion to votes. This is known as proportional representation (PR, prop-rep, or #PropRep). Early theorists include Henry Droop and Charles Dodgson (Lewis Carroll).

We should also recognize Thomas Jefferson and Daniel Webster's work on the related problem of apportioning congressional seats across states.

Because there are a number of seats to allocate, it's generally easier to get a good answer to this problem than in the case of single-winner voting. It's especially easy in the case where we're allowed to give winners different voting weights; in that case, a simple chain of delegated voting weight guarantees perfect proportionality. (This idea has been known by many names: Dodgson's method, asset voting, delegated proxy, liquid democracy, etc. There are still some details to work out if there is to be a lower bound on final voting weight, but generally it's not hard to find ways to resolve those.)

When seats are constrained to be equally-weighted, there is inevitably an element of rounding error in proportionality. Generally, for each kind of method, there are two main versions: those that tend to round towards smaller parties (Sainte-Laguë, Webster, Hare, etc.) and those that tend to round towards larger ones (D'Hondt, Jefferson, Droop, etc.).

Most abstract proportional voting methods can be considered as greedy methods to optimize some outcome measure. Non-greedy methods exist, but algorithms for finding non-greedy optima are often considered too complex for use in public elections. (I believe that these problems are NP-complete in many cases, but fast algorithms to find provably-optimal outcomes in all practical cases usually exist. But most people don't want to trust voting to algorithms that nobody they know actually understands.)

Basically, the outcome measures being implicitly optimized are either "least remainder" (as in STV, single transferable vote), or "least squares" (not used by any real-world system, but proposed in Sweden in the 1890s by Thiele and Phragmen). STV's greedy algorithm is based on elimination, which can lead to problems, as with IRV's center-squeeze. A better solution, akin to Bucklin/median methods in the single-winner case, is BTV (Bucklin transferable vote). But the difference is probably not a big enough deal to overcome STV's advantage in terms of real-world track record.

Both STV and BTV are methods that rely on reweighting ballots when they help elect a winner. There are various reweighting formulas that each lead to proportionality in the case of pure partisan voting. This leads to an explosion of possible voting methods, all theoretically reasonable.

Because the theoretical pros and cons of various multi-winner methods are much smaller than those of single-winner ones, the debate tends to focus on practical aspects that are important politically but that a mathematician would consider trivial or ad hoc. Among these are:

- The role of parties. For instance, STV makes partisan labels formally irrelevant, while list proportional methods (widely used, but the best example system is Bavaria's MMP/mixed member proportional method) put parties at the center of the decision. STV's non-partisan nature helped it get some traction in the US in the 1920s-1960s, but the only remnant of that is Cambridge, MA (which happens to be where I'm sitting). (The other remnant is that former STV advocates were key in founding FairVote in the 1990s and pushing for IRV after the 2000 election.) Political scientist @jacksantucci is the expert on this history.
- Ballot simplicity and precinct summability. STV requires voters to rank candidates, and then requires keeping track of how many ballots of each type there are, with the number of possible types exceeding the factorial of the

number of candidates. In practice, that means that vote-counting must be centralized, rather than being performed at the precinct level and then summed. That creates logistical hurdles and fraud vulnerabilities. Traditionally, the way to resolve this has been list methods, including mixed methods with lists in one part. Recent proposals for delegated methods such as my PLACE voting (proportional, locally-accountable, candidate endorsement; here's an example) provide another way out of the bind.
- Locality. Voters who are used to FPTP (plurality in single-member districts) are used to having "their local representative", while pure proportional methods ignore geography. If you want both locality and proportionality, you can either use hybrid methods like MMP, or biproportional methods like LPR, DMP, or PLACE.
- Breadth of choice. Ideally, voters should be able to choose from as many viable options as possible, without overwhelming them with ballot complexity. My proposal of PLACE is designed to meet that ideal.

Prop-rep methods would solve the problem of gerrymandering in the US. I believe that PLACE is the most viable proposal in that regard: maintains the locality and ballot simplicity of the current system, is relatively non-disruptive to incumbents, and maximizes breadth of voter choice to help increase turnout.

Oh, I should also probably mention that I was the main designer, in collaboration with dozens of commenters on the website Making Light, of the proportional voting method E Pluribus Hugo, which is now used by the Hugo Awards to minimize the impact and incentives of bloc voting in the nominations phase.

Anticlimactic sign-off

OK, that's a long article, but it does a better job of brain-dumping my >20 years of interest in this topic than anything I've ever written. On the subject of single-winner methods, I'll be putting out a playable exploration version of all of this sometime this summer, based off the work of the invaluable nicky case (as well as other collaborators).

I've now added a third article on this topic, in which I included a paragraph at the end asking people to contact me if they're interested in activism on this. I believe this is a viable target for effective altruism.

5 general voting pathologies: lesser names of Moloch

Earlier, I wrote a primer on voting theory. Among the things I discussed were 5 types of pathologies suffered by different single-winner voting methods. I presented these as 5 sequential hurdles for voting method design. That is, since they are in what I view as decreasing importance and increasing difficulty, you should check your voting method against each hurdle in order, and stop as soon as it fails to pass.

Then I read Eliezer's book on Inadequate Equilibria, and Scott's "Meditations on Moloch". They argue that the point of civilization is to provide mechanisms to get out of pernicious equilibria, and the kakistotropic tendencies of civilization they characterize as "Moloch" are basically cases where pernicious incentives reinforce each other. I realized that the simple two-player games such as Prisoners' Dilemma that serve as intuition pumps for game theory lack some of the characteristics of my 5 voting pathologies. So I want to go back and explain those pathologies more carefully, to help build up intuition about how multi-player, single-outcome games differ from two-player ones.

A key point here is that I'm talking about single-winner voting methods; that is, "games" where the number of possible outcomes is far less than the number of players. In this case, it's not a matter of seeking an individual advantage for yourself; the only way for you to win is for your entire faction to win equally. This means that I will not be talking about the oldest and deepest name of Moloch, which is Malthus. All the Molochs in this essay can and should be killed or (mostly) tamed.

Also note that this essay is not the one I'd write if I were only trying to recruit the rationalist community to become electoral reform activists. As an activist, I think that the most important and short-term-viable electoral reforms are in the multi-winner space: solving the problem of coordinating public goods not directly through mechanism design, but indirectly through a combination of mechanism design and representation. Some of my reasons for thinking that are contingent and have no place here. The one that's not: I think that the problem of "ain't nobody got time for all that politics" is worse than the principal-agent problem of a well-designed representative mechanism. Regardless, I think that this community would rather hear first about these names for Moloch.

In order, my pathologies — hurdles for multi-agent shared-outcome mechanism design — are:

Dark Horse

Let's say that you have a 3-candidate election using the Borda count, and your electorate has the following true utilities:

49: A9.0 B1.0 D0.0

48: A1.0 B9.0 D0.0

3: A1.0 B0.0 D9.0

Under the Borda count, each voter must give the three candidates 2, 1, and 0 points in some order. If the B voters strategize, the election might look like:

49: A2 B1 D0

48: A0 B2 D1

3: A1 B0 D2

B wins with a total of 145. The A voters might try to retaliate with a similar strategy:

49: A2 B0 D1

48: A0 B2 D1

3: A1 B0 D2

But now D wins with a total of 103, even though D was honest last preference for 97% of voters.

This "Dark Horse 2" example becomes even harder to resolve if you make it "Dark Horse 3":

34: A9.0 B2.0 C1.0 D0.0

33: A2.0 B9.0 C1.0 D0.0

33: A2.0 B1.0 C9.0 D0.0

I'll let you work it out for yourself, but the upshot is that each group has an incentive to give D the second-most points; that if one or two groups are strategic, they can profit; but if all three are strategic, all of them lose. D can win in this situation with literally zero honest support — an epically pathological result.

What does it feel like in this situation:

To win honestly? "All is right with the world."

To weakly-lose when everyone's honest? "I am slightly tempted to strategize."

To weakly-lose when the opponents are strategic? "I need to stop being a sucker, and counter-strategize."

To win strategically? "I feel a little bit guilty, but at least I won."

To strongly-lose strategically? "WTF? This system sucks. If possible, I should change it. If not, maybe I should learn my lesson and not strategize. But regardless, those other evil sneaky strategizers against me MUST learn theirs."

This is the closest to a standard prisoners dilemma of all of the voting pathologies. As with the standard prisoners dilemma, "social glue" (that is, heuristics developed through successful cooperation in iterated scenarios) can generally avoid breakdown. But it's also the easiest to avoid using mechanism design: just don't use the Borda count (or any other strictly-ranked point-based method). That is to say, don't force people to dishonestly support D in merely in order to oppose some other candidate.

So "Dark Horse" is a name for a Moloch that's outstandingly evil but not particularly powerful.

Lesser evil

If you live in the US, UK, Canada, or India — or any other country that uses First Past the Post voting — you already know this Moloch well. In a system where you can only vote for one, you'd better not "waste your vote" on the option you most truly support; you must instead support the lesser evil, the least-bad of the viable options. The logical end-point is a world with only two options, each of which has far stronger incentives to make the other side look bad than to actually pursue the common good. If you're lucky, one or both of those two options will pursue the common good for the fun of it; if you're unlucky, they'll each be as corrupt as they can get away with without losing support to the other side; but either way, there's relatively little you can do about it.

Of course, I should point out that this game theory doesn't always play out exactly in real life. The US has only 2 parties that matter, but most other FPTP countries have a bit more than that, even if the top two matter more than they should. So if you want to continue to spar with the teeth of this Moloch instead of just cutting off its head, OK, you're not doomed to lose every time. Just most of the time.

In terms of election scenarios, this looks something like the following. Utilities are:

15: A9.0, B8.0, C0.0

36: A8.0, B9.0, C0.0

24: A0.0, B9.0, C8.0

25: A0.0, B1.0, C9.0

Votes are:

15+36=51: A

24+25=49: C

This is an equilibrium because, in most games where there are far more players than outcomes, almost everything is an equilibrium; no one voter could get a better outcome by changing their vote, even though the society as a whole would be far happier if they could elect B. Any A voter who moved to B would be helping C win; any C voter who moved to B would be making it easier for A to win, even if next election honest C>A voters are a majority.

I probably don't have to tell you what this one feels like, but here goes anyway:

On top of the winning coalition (15 A voters): "All is right with the world."

On the bottom of the winning coalition (36 B>A>C voters): Conflicted. On the one hand, "the lesser evil is still evil". On the other hand, "a vote for B is a vote for C". Both are true; this dilemma is inescapable without changing the voting method. Short-term incentives favor continuing to vote for A, and in fact actively suppressing

discussion of A's flaws and B's ideas; but human nature favors getting mad at A and exaggerating their flaws. Either way, mind-killing is likely.

On the bottom of the losing coalition (24 B>C>A voters): Enraged. Ripe for a demagogue.

On the top of the losing coalition (25 C voters): Must... try... harder. Next time, we'll win!

This is a lesser Moloch, in that we could easily kill it by changing the voting method. Note that proportional representation can (if it's done well) be just as good at killing this Moloch as the single-winner methods discussed below! But it's still strong enough to rule over most of you who are reading these words.

Center Squeeze

OK, you say; if the Lesser Evil is enabled by the existence of wasted votes, let's fix that by moving all the votes until they're not wasted. You've just invented Instant Runoff Voting (IRV). Each voter ranks the candidates; votes are piled up by which candidate they rank first; and then, iteratively, the smallest pile is eliminated and those votes are moved to whichever remaining pile they rank highest (if any). You can stop as soon as one pile has a majority of remaining votes, because that pile is guaranteed to win.

This would solve the spoiler problem of the 2000 Florida presidential election. Here's a simplified version of utilities in that scenario (B/G/N stand of course for Bush/Gore/Nader):

490: B9.0 G1.0 N0.0 (Bush>Gore)

100: B1.0 G9.0 N0.0 (Gore>Bush)

389: B0.0 G9.0 N1.0 (Gore>Nader)

10: B0.0 G1.0 N9.0 (Nader>Gore)

6: B0.0 G0.0 N9.0 (Nader>nobody)

5: B1.0 G0.0 N9.0 (Nader>Bush)

Under FPTP, honest voting would "spoil" the election and let Bush win. But under IRV, the Nader supporters can vote honestly; when Nader is eliminated, those votes will transfer, so Gore will beat Bush 499 to 495.

But what happens if Nader appeals to more voters, and 300 of the Gore>Nader voters shift to Nader>Gore? That would mean that Nader had 321 first-choice supporters, and Gore only 189. So Gore would be eliminated first, 100 of those votes would shift to Bush, and Bush would win! In this scenario, the centrist Gore was "squeezed" on both sides and prematurely eliminated, even though he could have beaten either of the others in a 1-on-1 race.

And the result is that, just like in the real election, Nader's supporters ended up helping cause the election of Bush, the candidate most of them like the least. That spoilage doesn't happen until after Nader passes 25%, but it still happens. And this

problem is real; it happened in the Burlington 2009 mayoral election (though in that case, the voters whose honesty worked against them were the Republicans).

Now, Center Squeeze is a much smaller problem than Lesser Evil. If you have a choice, you'd rather run a race with a minefield between 25% and 50% of the way, than one where the minefield stretches from the beginning up to 50%. If you're skillful, maybe you can build up enough speed in the first 25% to leap over the minefield. And parties that stay under 25% can at least get more attention than those who are stuck around 0% as in Lesser Evil.

What does this one feel like?

Win, not spoiled: "All is right with the world."

Small fringe party, vote honestly, still matter: "At least I tried."

Medium fringe party, vote honestly, spoil the election: Dilemma. Some will decide to be strategic; others will say "wasn't my fault. It was the fault of those treacherous centrists who ranked the greater evil as their second choice."

Centrist, lose due to spoilage: "Huh? What happened? We're the rightful Condorcet winners, how can we lose?"

Large fringe party, win due to spoilage on the other side: "Ha! My far-off enemies were so disgusting that some of my nearby former enemies joined my cause! I deserved that."

Large fringe party, don't win: "Hmm... how can I divide my enemies?"

This Moloch is a relatively benign one, who acts to protect incumbent winners but allows dissenting voices up to a certain point. Living under its reign (as, arguably, Australia now does) involves occasional craziness but is mostly OK. Still, it can be killed.

Chicken Dilemma

This scenario actually exists in two separate versions, depending on the voting method: slippery and non-slippery slope. Both share the same underlying voter utility scenario, with two similar candidates who must team up in order to beat a third one:

35: A9.0 B8.0 C0.0 (A>B)

25: A8.0 B9.0 C0.0 (B>A)

40: A0.0 B0.0 C9.0 (C)

For the slippery slope version, let's assume the election uses approval voting: voters can approve as many candidates as they want, and the most approvals wins. If voters approve any candidate with a utility above 5.0, the ballots will be:

35+25=60: AB

40: C

A and B end up in an exact tie for first place (as Burr and Jefferson did in 1800; thus, the chicken dilemma is sometimes called the Burr dilemma). C, the candidate whom the majority opposes, has been safely defeated; but the outcome between A and B is essentially random. Incentives are clearly high for the first two groups of voters to approve only their favorite candidate. If 1 of the A>B voters votes for only A, then A wins; but then, 2 of the B voters can get B to win by switching to only B; and next 2 more A voters defect; etc. It's a slippery slope until over 20 of each group defect, and then C wins, an outcome the majority hates.

In game theory terms, this is a "chicken" or "snowdrift" game, with 2 equilibria: either the A voters stably cooperate and the B voters stably defect, so that B wins, or vice versa. But in emotional terms, neither of these equilibria feel stable: both are arguably "unfair" cases where one group is exploiting the other's cooperation. It might be "fair" if the smaller group was reliably the one to cooperate, but that's hard to coordinate in practice in cases where the sizes are similar, both sides will probably bet that they are the larger group. So in practical terms, probably the more "stable" outcomes are "both enforce cooperation, and hope there's some odd C voters who care enough to swing the election one way or the other", or "both bicker and defect".

To improve matters, we can use a non-slippery-slope voting method such as 3-2-1 voting. In this method, voters rank each candidate "good", "OK", or "bad", and the winner is decided in 3 steps. First, choose 3 semifinalists, those with the most "good" ratings; then of those, choose 2 finalists, those with the fewest "bad" ratings; then of those, the winner is the one rated higher on more ballots (the pairwise winner).

(When choosing the third semifinalist, there are two additional rules. First, to avoid a clone-candidate incentive, they must not be from the same party as both of the first two or, in a nonpartisan race, do not count their "good" ratings on the same ballots as also rated the first semifinalist "good". Second, to avoid a dark horse issue, they must have at least 1/2 as many "good" ratings as the first semifinalist. If no candidate meets these criteria, then skip step 2.)

In this method, if each voter votes honestly, then all 3 will be semifinalists (eliminating any also-rans whom we left out of the scenario for simplicity); A and B will be finalists (eliminating the majority loser C); and A will win, as the honest pairwise winner between those two.

It's still possible, in this scenario, for 21 B voters to defect, rate A as "bad", and cause B to win. But if under 20 of them do so, it doesn't change the result. Thus, there's no "slippery slope". Even though "everyone cooperates" is not a strong Nash equilibrium in strict game theory terms, it is probably strong enough to endure in practical terms.

Is it possible to make a voting method without even a non-slippery chicken dilemma? Yes, we've already seen that: IRV. But since defectors in the chicken dilemma look exactly like fringe voters in center squeeze, it's impossible to fully solve the chicken dilemma like this without creating a center squeeze problem — one I'd argue is worse, at least as compared to the non-slippery CD.

What does a non-slippery CD feel like? If both sides cooperate, I'd argue that it feels basically fair to everyone involved. If the smaller side wins through strategic defection, that feels unfair, and technically it's an equilibrium; but I'd argue that human stubbornness is enough to counter-defect as a punishment, and thus iterate back to cooperation. 9 So non-slippery CD isn't really Moloch at all. And as for slippery CD... it's mean, but capricious, and can sometimes be distracted or overcome.

Condorcet Cycles

Here's the scenario. Instead of utilities, I'll just give preferences, because there's almost no way to make this one "realistic".

34: A>B>C

33: B>C>A

33: C>A>B

This scenario is so unavoidably strategic that it's at the heart of a proof of the Gibbard-Satterthwaite theorem that no (non-dictatorial) voting method can entirely avoid strategy. If one of the three groups preemptively throws their favorite under the bus and embraces their second choice, the ballots will show at least a 66% majority for that second choice, so any democratic voting method will elect that candidate. So to all three groups, this situation will feel like a dilemma between racing to signal they'll compromise first and most convincingly, or hoping that the group before them in the cycle makes the compromise.

In practice, Condorcet cycles probably happen only 1-5% of the time. This is true in the most sophisticated voter utility models I can create (hierarchical "crosscat" Dirichlet clusters in ideology/priority space), and also in empirical evidence (where cyclical preferences seem rare but not nonexistent). So this last lesser Moloch is one which can never be defeated, but which spends most of its time in the deep woods and only occasionally rampages out, doing surprisingly little damage in the process.

Conclusion

I set out to write this because I thought that multiplayer game theory has some fundamental differences from single-player game theory and specifically that we need to stop leaning so hard on the prisoners' dilemma. Having written it, I realize that though I touched on these issues, I spent most of the time going over more basic points of voting methods. So I'm not sure this essay is exactly what I wanted it to be, but I think what it is can still be at least somewhat useful; I hope you feel the same way.

I guess my larger point is that evolution has actually equipped us pretty well with social strategies for dealing with PD or CD, but that by that same token we humans are particularly subject to pernicious equilibria of the "lesser evil" variety. The feeling of "we all agree these aren't the best options but looking for better ones would waste energy we need to spend fighting against the worse one" (lesser evil) seems like at least as important a paradigm of Moloch as "if I weren't evil someone else would be" (tragedy of the commons/multiplayer prisoner's dilemma/dark horse). It's important to remind ourselves that mechanism design offers a way out of lesser evil (and thus also center squeeze); not just in politics, but wherever it occurs.

Multi-winner Voting: a question of Alignment

This is my third (and for now, last) essay about voting theory for rationalists. In the first two, I focused primarily on single-winner voting theory; that is, methods for aggregating group preferences into a final verdict on some choice. Ideally, single-winner methods would be used in cases where decisions are inherently collective, while other mechanisms such as markets are better for cases where decisions are more individual. (As I touched on in the earlier articles, Sen's theorem puts limits on how precisely that distinction can be made; but that's not the point here. I'm going to take it as given that there are some cases where collective action is called for and others where action should be left up to individuals, and I don't want to spend time here arguing about the relative frequency or importance of those two kinds of situations.)

Why isn't multi-winner voting theory just a generalization of the single-winner kind?

If we've covered the best means for collective decisions, and individual decisions are out of the scope of voting theory, then what's left? Governance. That is, cases of collective action that aren't a single decision point but an ongoing series of decisions.

Such cases probably aren't best served by a series of separate single-winner elections, for a few of reasons. To begin with, it's not cognitively efficient; it would be silly for every citizen to need the expertise in order to make decisions about the minutia of every policy area. In fact, direct democracy tends to favor negative-sum rent-seeking: small groups extracting concentrated benefits by imposing diffuse costs, merely because they're the only ones motivated enough to sweat the details. And finally, it's not predictable: in many cases, governance should be coherent even at the sacrifice of some responsiveness.

In cases of governance, voting is not the final step, but merely one step in a larger process of decision-making. Thus, traditional multi-winner voting theory would look at ways to resolve this by electing a set of representatives to take those decisions.

To the rationalist community, such a multi-step process immediately raises the question of alignment. Just as designers of artificial intelligence should worry about whether their initial goals will be warped into a contrary outcome through the process of design and improvement, so should people like me designing multi-step mechanisms of governance worry about how values are preserved, lest small misalignments in each step add up to major disconnects in outcome.

Proportionality

Of course, if you're worried about preserving some property over a multi-step process, the first thing to do is to define that property. In this case, the key property is the proportions of decision-makers with each given set of utilities. Proportional multi-winner voting methods are those that are designed to (roughly) preserve these

proportions. Thus, collective decisions can be made by smaller groups; the ugly dynamics of mass argument can replaced by the hopefully-healthier ones of a smaller group. (Though flawed, the concept of "Dunbar number" is relevant here.)

Note that voting theory itself has nothing to say about how to define the original group whose proportions should be preserved. That is, it doesn't answer questions of who should be able to vote or how many votes each voter should have, in defining the original proportion. I'd argue that the safest and ultimately best rule today is for each human above a certain age to be allowed to vote; but that's out of the main scope of this article.

In stating the goal of "proportionality", I've been deliberately a bit vague about defining it. If voters come pre-sorted into comprehensive and mutually-exclusive partisan sets, it's relatively easy to define "Droop proportionality", in which each party gets a minimum proportion of seats in the legislature. But what if divisions of opinion are more complicated than that — continuous and/or multidimensional? In that case, there are various desirable proportionality properties, and some degree of tradeoff between them.

As a statistician, I should mention that there is one democratic "voting" method which will satisfy every possible proportionality property, at least "asymptotically" as the legislature grows in size. I'm talking about random sampling, or, as it's called when used for governance, sortition. By the law of large numbers, a random sample will tend to resemble the underlying population in proportions as to any and all individual characteristics, at least if the sample is large enough. In practice, sortition is rarely used for governance, though advocates of "citizens' assemblies", "citizen juries", "deliberative polls", and the like are trying to change that.

If we require voting methods to be deterministic, there are still a number of methods that have been designed to ensure proportionality; all such methods are called "proportional representation". (Since that's a mouthful and PR has too many meanings already, the best abbreviation is prop-rep.) In general, though perfect proportionality is impossible, most prop-rep methods come close enough that their other, more-pragmatic differences are more important.

Values and beliefs

Of course, representation (proportional or otherwise) is a goal regarding values, but decisions are also based on beliefs; and when it comes to beliefs, the goal should be truth, not representation. The idea of futarchy is about creating a political system that separates values and beliefs, so that values are resolved using a voting method (presumably one where any sub-steps preserve proportionality), while beliefs are resolved using prediction markets. While I'm skeptical of the possibility of designing markets that are immune to bubbles, values-based manipulation, or other systematic distortions, I think that the idea of trying to design a system that respects the separate logics of both values and beliefs is a good one.

Note that the current US voting system actually does try to do this to some extent, it just does a really crappy job. If political parties were groups of people with perfectly homogeneous values, then party primaries would not be the worst way of selecting smart, knowledgeable people with those values and thus of getting a slightly extrapolated volition as compared to mere sortition. Of course, we know that in many

real-world cases, primaries are more about ideological litmus tests than qualifications like expertise or intelligence.

Still, that suggests that proportional voting methods should probably include mechanisms for both intra-party and inter-party selection of candidates. In particular, closed-list proportional methods, which offload intra-party selection to some partisan mechanism probably dominated by insiders, are a bad idea.

(A related dichotomy is that between instrumental rationality, which involves both values and beliefs, and epistemic rationality, which involves only beliefs. So this issue can be seen as about finding ways to decrease the misalignment between the incentives for an instrumental and an epistemic rationalist.)

Parties

Another important question about voting methods is the party system they encourage.

First question: should there be parties at all? Though some people would disagree, I'd suggest that parties play an inevitable, and in some regards a positive, role in a political process. Yes, they do have bad effects, such as mind-killing tribal thinking; but they also have good ones, such as serving as useful cognitive heuristics for voters, and possibly allowing intraparty sorting to have more of a focus on qualifications and ability rather than ideology. Furthermore, even if you do believe they are bad on net, getting rid of them is really hard. Metaphorically speaking, if you try to design a voting system that bars the door against parties, you may find that they just make a hole in a load-bearing wall as they force their way in anyway.

Second question: how many parties should there be? Too few, and you get a stagnant "monopolistic" or "duopolistic" system in which zero-sum thinking leads to negative-sum outcomes. (For a real-world example, look at the USA.) Too many, and you encourage politicians who make narrow, single-issue appeals. (For a real-world example, look at Israel.)

Political scientists often view the distinction of few or many parties as by considering the representative voting method as just one step in a larger process of forming a majority coalition to take a societal decision. In other words, they speak of systems which encourage few parties as encouraging pre-election coalition-building, and those that encourage many parties as encouraging post-election coalition-building. In my view, it's good to have a little of both.

A useful way to measure number of parties is "effective number of parties" (ENP). The formula is

, where s_i is the size of party i as a fraction of all voters. Intuitively, this is the reciprocal of the fraction of voters in the party of the average voter (thus naturally weighting larger parties more). In other words, if the average voter's party size is 1/3 of the electorate, then ENP is 3. I'd aim for something between 3 and 4 as ideal.

I'd argue that choosing a voting method that tends towards such a moderate ENP will also tend to encourage better rationality within the legislature. As I said above, in a two-party system, with only one ideological dimension, winning or losing the eternal battle against the other side is all that matters, and so norms of debate (including

rationality norms) go out the window. And a highly fragmented world of single-issue parties actually has exactly the same problem; since each party is focused on just one issue, they have no reason to subscribe to overarching norms. It's only when there are more than two parties which each care about more than one issue that norms become selfishly worthwhile to each; though the norms might work against them on any one issue, insofar as they're positive-sum norms they will tend to work for each party's interests more than they work against them.

Voter strategy: free riding and vote management

In essentially all proportional methods (except weighted/proxy systems), an individual voter has an incentive not to vote for a candidate whom they know will win anyway, in order to avoid having any of their voting power "used up" by that foregone conclusion. But even though this incentive exists to some extent across many methods, its strength varies. All else equal, it's better to look for methods where this incentive is relatively weak.

On a collective level, this incentive is somewhat self-limiting. That is, if nobody votes for a popular candidate just because they're a "sure thing", then that candidate won't win after all. So collectively this incentive isn't so much for "free riding" as for "vote management": giving each candidate exactly the minimum number of votes they need to win. For instance, a party might try to equalize the number of votes that favor each of their candidates by instructing voters to vote based on their birthday.

Pragmatics (1)

So from the above, we're looking for a voting method that's reasonably proportional; that allows voter input on both within-party and between-party choices; that encourages a moderate number of parties; and that has a relatively weak free-riding incentive. That is an underdetermined set of constraints; there are a number of methods which do all of those to (what I'd consider) a pretty good degree. To choose between those proposals, we can add in pragmatic questions. Which methods are easiest for voters? Which are easiest to count? Which are likely to be most politically viable (which includes being non-disruptive to incumbents, at least, when disruption doesn't serve a useful purpose for any of the values above)? Which have the best track record?

Proportional Method Lego

Most proportional methods can be thought of as combinations of a few basic building blocks:

- Greedy assignment and deweighting. Choose winners one at a time according to who has the "most votes", then reweight the ballots that helped them get elected so that some of their voting power is used up. There are various reweighting schemes that work. Say there are 40% of the ballots that all are among the strongest ballots helping elect the same 3 winners out of a total of 9 seats. They can be reweighted to 20, 10, 5; to 13.3, 8, 4.28; to 30, 20, 10; or to 28.89, 17.78, 6.67. All of these schemes, if applied to all groups, will end up with

a proportional result; they differ in whether they round leftovers towards larger or smaller parties and in the strength of their free riding incentives. Note that greedy algorithms are actually approximations of more-complex globally-maximizing algorithms. Mostly voting methods do not use global maximizers, simply because they're harder to explain.

- Elimination and transfer. Eliminate "losers" and transfer their votes based on some implicit or explicit preference order. Note that when combined with the above, this sequential elimination is an extra, unnecessary greedy approximation. In the single-winner case, it's what leads to the center squeeze problem.
- Descending threshold. Instead of elimination and transfer, you can progressively lower some threshold, and count ballots as supporting all candidates they rate above that threshold. Even though one ballot may count as supporting multiple candidates, it will still be deweighted if any of those candidates actually wins, so it does not get any additional voting power. This is theoretically-superior to elimination and transfer, but the difference is usually small in practice, and this has far less of a track record of real-world use.
- Districts (single- or multi-member): Simplify matters by dividing up into sub-elections. These may be entirely separate, or unified by mixed-member or biproportional mechanisms (below). Traditionally, the variable name used to denote district magnitude is "M".
- Mixed member. Some seats are assigned by a fully nonproportional system (such as FPTP by districts), while others are later assigned by a proportional system so as to adjust the proportions. This is often accompanied by a dual ballot; for instance in Bavaria, you may vote for one candidate in your own district and one candidate outside your district but in your region.
- Biproportionality. Results are constrained so that there is exactly a certain number of each kind. (This is akin to stratified sampling in survey design.) For instance, there could be a rule that there should be exactly 1 winner per equal-population district, or that there must be at least X% of winners of each gender, or that certain seats are reserved for a native ethnicity.
- Ranked ballots. Voters rank candidates in preference order.
- Delegation. Each candidate makes a (partial?) ranking or rating of the other candidates, and this is (optionally?) used to fill in preferences on ballots cast for that candidate. Most proposals have candidates pre-register preferences, to avoid corruption and so that voters can use this information when casting their ballots, but in theory it would be possible to allow candidate preferences to be set after the election.
- Pooling. Similar ballots (for instance, those that prefer a given candidate, or those that prefer a given party) are averaged and then counted together. This sacrifices some information about the details of each ballot, in order to make counting summable from the precinct level. Note that without delegation and/or pooling, proportional methods are not summable, which can present practical problems in vote-counting such as chain-of-custody.
- Open party lists. Essentially, this means that there are separate mechanisms for assigning each party an appropriate number of seats, and for choosing which of that party's candidates get those seats. This can allow for simpler ballots; for instance, a voter can choose a single candidate and that can be counted both as a party vote in a proportional system and as a vote for that candidate in a nonproportional within-party ordering. (Note that open party lists can be seen as just a special case of pooling, but since they're a common idea, I'm listing them separately.)
- Party thresholds. That is, parties with under a given percentage (such as 5%) are not given any seats. This is a mechanism to stop "fringe parties" from winning

proportional seats; in other words, to keep the ENP from growing too large. But it's very much a blunt instrument, especially if votes for sub-threshold can't transfer to other similar parties. In real-world elections, party thresholds and "divide and conquer" have let parties with as little as 38% of the popular vote get legislative majorities in supposedly "proportional" methods, with serious long-term consequences.
- Individual local thresholds. Individual candidates with under a given percentage (such as 25%) of votes from their local district are eliminated. Since this usually is used in combination with vote transfers, it's much less of a blunt instrument than party thresholds. For instance, a party with just 15% of the vote region-wide will probably have some candidates with over 25% of the vote in their district; these candidates will get transfers from their co-party-members and thus probably win seats. And even if the party gets no seats, their votes will be transferred to a similar party, not just be wasted.

Combining the above building blocks, we can build various voting methods:

- Regional open list: Open list (pooling by party). Districts, typically with 10-40 seats each. For the proportional backbone, because of pooling, there are many which give the same outcome, but can be seen as a greedy/deweighting method.
- STV: (Single Transferrable Vote) Districts typically M=5 or so. Ranked ballots, deweighting, and elimination. Used in Ireland and Malta, and at some levels in Australia.
- MMP: Mixed member: FPTP + open list. (Good example: Bavaria. Bad example: Wales.)
- DMP: (Dual Member Proportional) Mixed member: FPTP + biproportional open list, so that there are exactly 2 winners per district.
- LPR: (Local Proportional Representation) Biproportional + STV.
- PLACE: (Proportional, Locally-Accountable Candidate Endorsement) Preferences are set by a hybrid of delegation and individual pooling. There's an individual local threshold of 25%. Seats are biproportional, so that there's exactly one winner per district. The back-end method is STV.

Pragmatics (2): I think PLACE is awesome

I'm going to switch from just explaining multi-winner voting theory to advocating for a specific method, so I should start out by explaining where I'm coming from. I'm a US activist for voting reform; on the board of the Center for Election Science (electology.org). My object-level politics, and my social milieu, tend to be pretty much on the left of the spectrum, but I also have real meta-level politics in favor of democracy. Ask me about any given issue and I'll happily explain why my own views are smarter than those of the median voter; but across all of those issues I know that the crowd is probably wiser than I am as often as not.

I've been thinking seriously about voting theory for over 20 years, and it's the main reason I am now getting a doctorate in statistics. In that time, I've designed many voting methods. The ones I consider best (3-2-1, PLACE, EPH, and SODA) are designed to optimize on the characteristics I think are important. When I argue for these methods, of course I'm biased. But I'd suggest that when I argue "My method is best normatively because it optimizes characteristic X", you should question my bias more by disputing whether X as I've defined it is important than by wondering whether the method actually optimizes X.

So what do I think is important in a practical proposal for a multi-winner method? It should:

- Minimize wasted votes — votes that don't help elect a candidate. (Under my rough definition of wasted votes, optimizing this implies proportionality.)
- For those votes which aren't wasted, maximize "similarity" between voter's preferences and candidate's qualities.
- Having looked at many voting methods and many scenarios for each, I find that giving voters breadth of choice does a better job at this than giving them depth of choice. Say I'm voting in a California congressional election, with around 50 seats in play. If I am free to choose my favorite candidate statewide, and then if they lose that vote is transferred based on their preferences, the mismatch between my preferences and theirs introduces less error than if I am able to cast a full ranked ballot in a 5-seat district with 10 times fewer choices.
- Be simple for voters
- Ranked ballots for more than about a dozen candidates are intolerably complex for most voters.
- Retain perceived "advantages" of FPTP, including some guarantees of local representation, as well as a clear concept of "my representative".
- Encourage a moderate number of parties
- Have a relatively weak free-riding incentive
- Be non-disruptive and otherwise "politically viable".
- This is obviously a judgment call, but I think that a method that is any threat to an incumbent of average popularity is a non-starter. Insofar as outcomes are different, the losing incumbents should be among those with below-average popularity.
- Have a precinct-summable counting process
- This is useful for transparency of outcomes and for fraud resistance.

PLACE voting was designed with these characteristics in mind; it does reasonably well on all of them. All other methods I know of fail significantly on several characteristics. (In fact, it took me decades of learning about voting theory, followed by almost a year of concentrated design work for hours a week, to settle on PLACE.)

Down in comments, before I finished this article, there was already a comment criticizing PLACE (from somebody who knows me from elsewhere). I understand that the criticism, that voters may find delegated methods distasteful, is real. I don't think it's as serious as it would be to fail on the other characteristics above.

If you're interested in activism on this, contact me. PLACE is compatible with the US constitution and current law, so it could be done by either state or federal legislation. I'm looking to get this passed somewhere (Somerville, MA?) at a municipal level first (there's a nonpartisan version that's appropriate). My email is firstname dot lastname at google's public email service. I'd also encourage you to support the Center for

Election Science. Even if you're in the UK or Canada (especially BC), I can help hook you up with local movements for reform.