# The Robustness of Quadratic Voting

E. Glen Weyl

February 1, 2015

# The Robustness of Quadratic Voting [*]

E. Glen Weyl[†]

February 2015

## Abstract

I use heuristic analytic approximations to consider the robustness of the Quadratic Voting (QV) mechanism proposed by (Lalley and Weyl, 2015) to collusion and variations in voter behavior. I also consider some examples of aggregate uncertainty in the value distribution and common values. While these variations typically reduce the efficiency of QV, in plausible cases the impacts are small and variations in voter behavior may actually improve QV's rate of convergence towards efficiency. I contrast these results with other (approximately) efficient mechanism proposed by economists which are highly fragile along these dimensions.

*Keywords:* robust mechanism design, Quadratic Voting, collusion, paradox of voting

1

Occam's Razor produces hypotheses that with high probability will be predictive for future observations.

– Anselm Blumer, Andrzej Ehrenfeucht, David Hausler and Manfred K. Warmuth, "Occam's Razor"

# 1   Introduction

Lalley and Weyl (2015, henceforth LW) propose a market-like mechanism for collective decision-making, Quadratic Voting (QV), in which individuals purchase votes, paying the square of the votes they buy. They use a price-taking model and the limit of a simple independent private values game theoretic environment to show that this mechanism leads to efficient outcomes in large population. However, this is only one of many mechanisms economists have proposed that lead to efficient outcomes, especially in large populations. In fact, in the same environment LW consider many of these mechanisms are efficient even in small populations or converge to efficiency more quickly than QV does. LW argue that the principal advantage of QV over these alternatives is that its greater "simplicity" compared to previous mechanisms suggest that it may be more robust than are these other mechanisms, following the machine learning interpretation of Occam's Razor (Blumer et al., 1987). In this paper I explicitly investigate and largely confirm this robustness by considering a variety of other models. I then argue that each other efficient mechanism is extremely sensitive to at least one of these changes in modeling choices. Together these results suggest that QV is uniquely promising as a practical paradigm for efficient collective decision-making among existing mechanisms.

The robustness issues I consider are closely motivated by the leading weaknesses discussed in the literature of previous mechanisms. The most famous alternative mechanism is that proposed by Vickrey (1961), Clarke (1971) and Groves (1973) (VCG) and promoted as a method for collective decision-making by Tideman and Tullock (1976). While VCG has a number of potential weaknesses, it is widely believed that for collective decision-making its greatest limitation is its sensitivity to collusion and fraud (Ausubel and Milgrom, 2005). I therefore begin, in Section 2, by considering the sensitivity of QV to these factors. While collusion and fraud typically benefit their perpetrators, incentives for unilateral deviation and reactions from other participants force collusive groups to be quite large or fraud to be extremely ambitious in order to be efficacious if the broader population is large. Given that such large groups or fraudulent schemes could likely be detected by authorities, it seems unlikely that collusion and fraud would be severe threat to QV's efficiency in large populations.

Another group of mechanisms (Arrow, 1979; d'Aspremont and Gérard-Varet, 1979; Ledyard and Palfrey, 1994, 2002; Krishna and Morgan, 2001; Goeree and Zhang, 2013) is defined only

when the distribution of values is common knowledge and for most of these also requires that this distribution be known by the planner. In Section 3 I therefore consider the performance of QV when there is uncertainty about the value distribution. This typically leads to some inefficiency even in large populations because individuals who favor the ex-ante expected losing alternative will typically expect the chance of a tie to be greater given that their own value draw leads them to update the value of the aggregate parameter towards a tie. This in turn makes those supporting the underdog vote more aggressively, unfairly disadvantaging the ex-ante favorite. However, as I illustrate through several parametric examples, this efficiency loss is unlikely to be very large as the inefficiency inherently relies on the ex-ante favorite in welfare terms remaining the favorite. Thus QV never has more than 5% inefficiency in any examples I study and nearly always outperforms 1p1v, often significantly. I also briefly consider one example of a mixed common values model where individuals, in addition to their heterogeneous preferences, have information about which alternative is in everyone's interests, illustrating that QV may aggregate information even when one-person-one-vote (1p1v) fails to.

Under some extremely special circumstances, 1p1v may approximate efficiency if it is costly to vote (Ledyard, 1984; Myerson, 2000; Krishna and Morgan, Forthcoming). The logic is that so few individuals vote that it becomes possible that the probability of their voting is proportional to the intensity of preference. Clearly this is possible, given rich heterogeneity of preference intensity, only if turnout dwindles to 0 in a large population. This does not appear to be the case in practice. It is therefore natural to ask what models capable of predicting such "anomalously" high levels of voter turnout imply about the efficiency of QV. I study two such models: one in which voters (heterogeneously) over-predict the chance of being pivotal and one in which they are (heterogeneously) motivated to vote to express their preferences rather than merely to influence the outcome. In such models QV may converge more quickly to efficiency, as the "extremist" behavior that accounts for most inefficiency in QV, is more effectively deterred. When the mean of the value distribution is $0$ some limiting inefficiency results, but it is usually quite robustly smaller than under 1p1v.

I summarize and expand on these results in Section 5 by comparing QV to each of the leading alternatives and a few other, less-widely discussed mechanisms. I argue that QV is the only one of the (approximately) efficient mechanism proposed by economists that is plausibly of practical relevance. I then conclude by discussing some of the robustness issues I think would be most interesting to explore in future work.

I do not prove any of the results in this paper fully rigorously. As is evident from LW, any fully rigorous proofs in this environment requires extremely detailed statistical arguments. Instead I use heuristic approximations analogous to those Weyl (2015) uses to obtain conservative upper bounds on the expected inefficiency of QV. These allow me to cover a lot of ground reasonably briefly, but obviously also makes me less confident in the results. Some of the results I

am quite confident I could prove using the methods of LW, but I have not written these proofs out; I refer to such results as "claims" and my confidence in them is at least 90%. Other results I do not presently see how to prove, though I think it is quite likely they are at least roughly right; such results I refer to as "conjectures" and my confidence in them is closer to 75%. Most extended calculations have been grouped into appendices following the main text of the paper.

## 2    Collusion and Fraud

The basic environment I consider is the same as LW, except that the election outcome is determined discontinuously by majority rule rather than smoothly and in each of the following sections I introduce an additional wrinkle. I start by briefly reprising LW's setting to clarify the model I study.

There are $N$ individuals making a binary collective decision about whether an alternative $A$ to the status quo should be adopted. Each individual $i = 1, \ldots N$ is characterized by a quasi-linear *value* $u_i \in \mathbb{R}$ for the alternative being adopted; for normalization I assume individuals gain $u_i$ if the alternative is adopted and lose it the status quo is maintained. Negative values of $u_i$ thus indicate a preference for maintaining the status quo. These values are drawn from a smooth value distribution $F$ with density $f$. This distribution is assumed to have bounded support that includes $0$, exponentially dying tails on both sides or to have Pareto (viz. power law decaying) tails with power decay rate $\alpha_-$ for negative values and $\alpha_+$ for positive values. I assume that $\min\{\alpha_-, \alpha_+\} \geq 1$ so that the first moment of the value distribution exists and denote it by $\mu$.[1]

Each individual chooses a continuous number of *votes* to buy $v_i \in \mathbb{R}$. She pays a net transfer $v_i^2 - \frac{\sum_{j \neq i} v_j^2}{N-1}$. The alternative is implemented if $V \equiv \sum_i v_i \geq 0$. I define the *expected inefficiency* $(EI)$ of QV in the unique manner so that it always lies between $0$ and $1$ (and spans the full range if all decisions move from being made correctly to incorrectly) and is linearly decreasing in social welfare:

$$EI \equiv \frac{1}{2} - \frac{\mathbb{E}\left[U 1_{V \geq 0}\right]}{2\mathbb{E}\left[|U|\right]},$$

where $U \equiv \sum_i u_i$.

LW show that the structure of equilibrium is very different depending on whether $\mu = 0$ or $\mu \neq 0$. When $\mu = 0$, in a large population all individuals buy votes roughly in proportion to their values, yielding efficiency in the limit. When $\mu \neq 0$ (without loss of generality I focus on the case when $\mu > 0$), *almost* all individuals buy votes in proportion to their values; they call such individuals *moderates*. However, a small mass of *extremist* value types near the bottom of

---

[1]For the $\mu = 0$ analyses I assume that $\min\{\alpha_-, \alpha_+\} \geq 3$ and that the third raw moment, which I denote $\mu_3$, is non-zero.

the value distribution buys as many votes as the rest of the population combined. Absent such extremists existing, the chance of individuals being pivotal, and therefore the number of votes any individual chooses to buy, would vanish so quickly that it would be trivial for any individual to buy the whole election. Having some individuals actually do so restores the chance of a tie, raising the number of votes necessary to "steal" the election. This ensures that only the most extreme types act as extremists and thus that they probability of any such extremist existing is vanishingly small in a large population.

On top of this baseline model I now consider the possibility that a collusive group of size $M$ forms. I assume (conservatively) that when such a group forms, the collective learns the values of all participants and the group acts in its common interests, though I consider unilateral incentives to deviate from such optimal collusion in Subsection 2.2 below.

First note that an efficacious collusive group maximizing the joint payoffs of its members will ask all members to purchase the same number of votes.[2] Second note is that a collusive group of size $M$ will buy $M$ times as many votes per unit of aggregate utility as its individual members would buy if they had such a utility. To see this suppose that every individual in the group had the same utility. Each would create positive externalities on the others for each vote she purchased of the same magnitude of the value she obtains from a vote. Collusion internalizes these externalities, magnifying the optimal amount the group would vote. Only this second effect impacts efficiency; the first only reduces revenues.

This sort of collusion is thus most harmful when it involves a large number with large values pointing in the same direction. The worst case scenario is thus that the $M$ individuals involved in collusion are the $M$ most extreme individuals whose values point in the same direction, on either side of the distribution when $\mu = 0$ and in the negative direction when $\mu > 0$.[3] Perhaps this case is also close to the one that would be most plausible in practice. It also, if undetected, clearly can wreak havoc for QV even with a very small number of members when $\mu > 0$ and the distribution has bounded support, because the chance that two individuals together have values in total beyond the extremist threshold is typically quite high even when the chance that any individual is beyond this threshold is small. Thus, at least in this case, QV might seem as sensitive to collusion as other efficient mechanisms. However, as I now examine formally, three forces limit collusion: the size a collusive group must take on to be effective, the unilateral deviation incentives having such a large group colluding creates and the reactions such a group provokes from the rest of the population. The second force is most effective in deterring collusion when $\mu = 0$ and the third when $\mu > 0$.

---

[2]The reason is that the utility of the group depends only on the aggregate number of votes it buys and the aggregate payments it makes. Conditional on the first, the second is minimized when all individuals split evenly the aggregate votes because the quadratic function is convex.

[3]In a previous analysis I considered also an average case with randomly drawn individuals. This case seems fairly unrealistic, however, and my results are strictly better in this case so I did not consider it to be worth including.

For brevity I only discuss rates of growth or decay of various quantities with population size. Constants may also be calculated and I occasionally mention more concrete quantities calibrated using these constants to the gay marriage referendum in California in 2008 that LW use as a motivating example and I analyze in detail in Weyl (2015). As a result, I make heavy use of standard Bachmann-Landau notation: for functions $h$ and $g$ of $N$, $h \in O(g)$ denotes that $h$ is eventually (in $N$) no greater than $\alpha g$ for some constant $\alpha$, $h \in o(g)$ denotes that $h$ is eventually strictly less than $\alpha g$ for any $\alpha$, $h \in \Omega(g)$ if $g \in O(h)$ and $h \in \Theta(g)$ denotes that $h \in O(g)$ and $h \in \Omega(h)$.

Let $u_M$ be the sum of the values of all the individuals in the collusive conspiracy. First I consider the approximate distribution of $u_M$ in the worst case. Without loss of generality, let $\mu \geq 0$ and when $\mu = 0$ that $\alpha_- \leq \alpha_+$ and I suppress the index on $\alpha$. In both cases efficiency is most distorted when all collusive members are drawn as the most extreme members of the lower tail of the distribution, that is the $M$ most negative order statistics. This is true because this maximizes the chance of inefficient extremist behavior when $\mu > 0$ and it creates the largest deviation from proportionality when $\mu = 0$.

For large $N$ and very negative $u^\star$, the probability that the lowest order statistic is less than $u^\star$ is less than $Nk (u^\star)^{-\alpha}$, for some constants $\alpha > 1, k > 0$ using the Pareto tail approximation. By standard results from extreme value theory, the $j$th-to-last order statistic is approximately $\frac{1}{j^{\frac{1}{\alpha}}}$ times the last order statistic. Thus, using the continuous approximation, the sum of the last $M$ order statistics is approximately

$$\int_0^M \frac{1}{x^{\frac{1}{\alpha}}} dx = \frac{\alpha}{\alpha - 1} M^{\frac{\alpha - 1}{\alpha}}$$

times the first-order statistic. Thus the probability that $u_M$ is larger in absolute value than $u$ is approximately

$$Nk \left[ \frac{u(\alpha - 1)}{\alpha M^{\frac{\alpha-1}{\alpha}}} \right]^{-\alpha} = Nk \left[ \frac{u(\alpha - 1)}{\alpha} \right]^{-\alpha} M^{\alpha - 1}.$$

That is, a collusive group raises the probability of extreme values by a factor of order approximately $M^{\alpha - 1}$.

On the other hand the typical value of $u_m$ is $\frac{\alpha}{\alpha - 1} M^{\frac{\alpha - 1}{\alpha}}$ times the last order statistic. This last order statistic is in $\Theta\left(N^{\frac{1}{\alpha}}\right)$. Thus the typical size of $u_M$ is in $\Theta\left(M^{\frac{\alpha - 1}{\alpha}} N^{\frac{1}{\alpha}}\right)$.

## 2.1 Necessary coalition sizes

The case in which collusion poses the greatest threat to the efficiency of QV is when $\mu > 0$ and the distribution of values is bounded. Then if the collusive group is entirely unanticipated by the rest of the population even a very small, extreme collusive group can wreck the efficiency of QV. The

reason is that while the chance of any single individual wishing to act as an extremist is small, the chance of two individuals together wishing to act as an extremist (given the equilibrium voting behavior) is $\Theta(1)$. While it may be rare to have a single individual extremely close to $\underline{u}$, is very common to have two individuals the sum of whose values are larger than $\frac{u}{2}$ in absolute value, which is all that is needed for it to be in their collective interest to buy the approximately requisite $\frac{\sqrt{|u|}}{2}$ votes each. This problem is slightly mitigated when tails are fat, as the single most extreme individual accounts for a greater fraction of the top two most extreme individuals' valuations in this case.

**Claim 1.** If there is a single perfect collusive group of size $M$ and other individuals play as in equilibrium where they believe there is no possibility of collusion taking place, $EI \to 0$ as $N$ grows large as long as

1. $\mu = 0$, colluders are drawn as in the worst case and $M \in O\left(N^{\frac{\alpha-2}{2(2\alpha-1)}}\right)$ so that in the bounded support case $M \in O\left(\sqrt[4]{N}\right)$ or

2. $\mu \neq 0$, colluders are drawn as in the worst case and $M \in O\left(N^{\frac{\alpha-1}{(\alpha+1)(2\alpha+1)}}\right)$ so that in the bounded support case even a fixed number of colluders may stop EI converging to $0$.

*Heuristic justification.* The cost to an extremist coalition of size $M$ of behaving as extremists is only $1/M$ of that of any single individual acting as an extremist. Thus if $u^\star$ is the threshold utility for a single individual to act as an extremist the threshold value for a collusive group to do so is $u^\star/M$. And the extremist threshold is in $\Theta\left(N^{\frac{2}{1+\alpha}}\right)$ according.

Thus the order of the probability of the collusive group finding an extremist strategy attractive is

$$N^{-\frac{2\alpha}{1+\alpha}} N M^\alpha M^{\alpha-1} = N^{-\frac{\alpha-1}{1+\alpha}} M^{2\alpha-1}.$$

Thus, depending on the size of $M$ relative to $N$, the rate of decay of inefficiency changes from $\Theta\left(N^{-\frac{\alpha-1}{\alpha+1}}\right)$ to $\Theta\left(N^{-\frac{\alpha-1}{\alpha+1}} M^{2\alpha-1}\right)$. Inefficiency still dies in the limit as long as $M$ is in $O\left(N^{\frac{\alpha-1}{(\alpha+1)(2\alpha-1)}}\right)$.

On the other hand, when $\mu = 0$, the outcome is somewhat less sensitive to the votes of a few, even quite extreme, individuals buying a large number of votes, as the equilibrium aggregate number of votes is quite large. To see this note that when $\mu = 0$, $u_M$ is of order $M^{\frac{\alpha-1}{\alpha}} N^{\frac{1}{\alpha}}$. By my logic in Weyl (2015), the inefficiency this causes is proportional to the square of the distortion in the sum of the votes normalized by $1/a_N$, which is $M-1$ times the normalized value of the collusive group as they magnify their votes by a factor of $M$, and inversely related to $\sqrt{N}$. $a_N \in \Theta\left(1/\sqrt[4]{N}\right)$, so the squared shift in the mean of votes is $\Theta\left(M^{\frac{2\alpha-1}{\alpha}} N^{\frac{4+\alpha}{4\alpha}}\right)$; dividing by $\sqrt{N}$ gives $\Theta\left(M^{\frac{2\alpha-1}{\alpha}} N^{\frac{4-\alpha}{4\alpha}}\right)$. Squaring this distortion and normalizing it relative to aggregate efficiency implies that the EI created by worst case collusion is $O\left(M^{\frac{2(2\alpha-1)}{\alpha}} N^{-\frac{\alpha-2}{\alpha}}\right)$. This dies with $N$ as long as $M = O\left(N^{\frac{\alpha-2}{2(2\alpha-1)}}\right)$. $\qquad\square$

7

This result offers only a small benefit over VCG and the Expected Externality mechanism in realistic cases. For example, when $\alpha = 3$, the size of the collusive group grows only as $N^{1/10}$ when $\mu = 0$ and only as $N^{1/14}$ when $\mu \neq 0$. In both cases calculating and calibrating constants suggests realistic collusion even in the full California population would only require $3 - 7$ individuals, which is likely quite feasible without detection. While this is a bit better than the $2$ needed for collusion against VCG and EE, it is hardly better.

## 2.2 Unilateral deviation incentives

However, LW show that QV's efficiency occurs at *all* type-symmetric thus its efficiency is coalition-proof, again unlike VCG.[4] This suggests that equilibrium considerations may further limit collusion. The most natural such consideration is incentives for unilateral deviation, which are usually thought to be the primary deterrent to collusion in standard markets when there is not a very large number of competitors (Stigler, 1964). These can be quite powerful in the case when $\mu = 0$.

Note that in the case when $\mu = 0$ the payoff of all individuals is approximately quadratic and thus any deviation from unilaterally optimal actions leads to a) an incentive to deviate, b) a marginal incentive to deviate proportional to the deviation from the unilaterally optimal action and c) an aggregate incentive to deviation that is proportional to the square of the deviation from the unilaterally optimal action. Furthermore, to the extent that the payoff is not perfectly quadratic, large deviations of the votes in any direction diminish the chance of a tie and make it optimal for each individual to purchase fewer votes, further depressing the incentives of the collusive group to buy so many votes.

However, in the case when $\mu \neq 0$ they accomplish little. To see this note that when $\mu \neq 0$, members of an "extremist conspiracy" who band together to behave like extremists have strong strategic complementarity: once one member of the group has purchased, say, half the necessary number of votes, it becomes much cheaper for the other member of the conspiracy to buy the other half and thus swing the decision. Thus while the possibility of unilateral deviation makes collusion a bit harder (requiring both colluding individuals' values, rather than their average value, to exceed a threshold) it does not dramatically change the picture. For the rest of this section all justificatory calculations are reserved for Appendix A.

**Claim 2.** Suppose there is a single perfect collusive group of size $M$ and other individuals play as in equilibrium where they believe there is no possibility of collusion taking place. Then if $\mu = 0$ there are always unilateral deviation incentives for any collusive behavior for large $N$ and the size of marginal deviation incentives is at least proportional to the deviation from unilateral

---

[4]Note, however, that this is only true of "type-symmetric" collusive strategies where it is individuals of particular types, rather than individuals of particular names, that collude.

behavior. If $\mu > 0$ a collusive agreement non-vanishing inefficiency with no unilateral deviation incentives is possible only when $M \in \Omega\left(N^{\frac{1}{2\alpha-1}}\right)$ in the worst case.

Thus when $\mu = 0$, the unilateral incentives for deviation from a collusive group are likely to prevent collusion from significantly denting large-population efficiency. This would be especially true if collusion were made illegal, as vote-buying and collusion in markets are illegal in most developed societies, and was rigorously policed with whistle-blowing incentives that further increase the incentive for unilateral deviations. When $\mu \neq 0$, such incentives do limit the ability to engage in collusion when $\alpha = 3$ and the population is very large; in the calibrated example, a collusive group would have to have nearly 100 individuals. This might well be large enough that the collusive group would find it difficult to conceal their activity from authorities, especially given it would be fairly easy to identify the individuals likely to be the most extreme opponents of the alternative, as they would all be quite wealthy and opinionated. This is a sharp contrast to VCG and EE, where collusion is exactly or approximately individually incentive compatible, requiring only that two individuals coordinate their behavior, as I discuss in the next section. However, QV performs no better in this regard when the value distribution is bounded.

## 2.3 Offsetting reactions

Therefore, the strongest check on extremist collusion may arise not from the explicit enforcement activities of the authorities, but from the implicit enforcement by the public. If the public realizes that an extremist collusive group is operating with significant probability, this will induce them to believe that there is a significant chance of a tie and thus dramatically raise the number of votes they buy, with which the extremist coalition must then compete. This is analogous to the observation that in standard markets collusive arrangements attract entry that undermines them (Stigler, 1964).

**Claim 3.** Suppose that $\mu > 0$ and a single perfectly collusive group of size $M$ drawn from the most extreme individuals with negative utility emerges with common knowledge probability $r$. Then $r$ must vanish in $N$ as long as $M \in O\left(N^{\frac{\alpha-1}{2\alpha-1}}\right)$.

The constants in this result remain essentially the same (except for a small factor equal to approximately $1.4$ when $\alpha = 3$) as in my baseline analysis in Weyl (2015). Thus if we consider the case of Proposition 8 again with $\alpha = 3$ and a collusive group of the 100 strongest gay marriage opponents, inefficiency would be approximately .5%, significantly higher than .058%, but still quite small. Roughly a thousand colluders would be needed to significantly dent efficiency. Thus if participants have rational expectations about collusion and collusive groups are not so

large as to be easily detected, it seems unlikely they will significantly decrease efficiency compared to the non-cooperative baseline, in contrast to VCG and EE, where knowledge of collusion occurring does not change optimal behavior. In the bounded support case ($\alpha \to \infty$), this effect is even stronger.

## 2.4 Fraud

A similar analysis may be applied to a single individual who fraudulently "de-mergers", representing herself as more than a single individual. Such de-merger attacks are known to be highly effective against VCG as we discuss in Subsection 5.2 below.

**Claim 4.** If a single individual can fraudulently misrepresent herself as $L$ individuals, in any equilibrium $EI \to 0$ as $N$ grows large as long as

1. $\mu = 0$, the fraudulent individual is the most extreme individual in the population, all other individuals behave as in the non-cooperative equilibrium and $L \in O\left(N^{\frac{\alpha-2}{2\alpha}}\right)$;

2. $\mu > 0$, the fraudulent individual is the most extreme individual in the population, other individuals behave as in the equilibrium without fraud and $L \in O\left(N^{\frac{\alpha-1}{\alpha(\alpha+1)}}\right)$ or

3. $\mu > 0$, the fraudulent individual is the most extreme individual in the population, all other individuals are aware of the fraudulent behavior, an equilibrium is played given this common knowledge of fraud and $L \in O\left(N^{\frac{\alpha-1}{\alpha}}\right)$.

Again suppose that $\alpha = 3$. Then $L$ would have to be of order $\sqrt[6]{N}$ to make a significant difference, roughly 1-2 dozen in the California population in either the $\mu = 0$ or $\mu > 0$ case. If fraud is at least partly anticipated $L$ would have to be of order $N^{2/3}$ or hundreds of thousands to succeed. This latter number is clearly far too large scale of a fraud to go undetected by authorities. Fraud may cause some damage if not partly anticipated, but the large purchases necessary to do real harm would probably be detected even in a small fraud of an individual representing herself as 1-2 dozen individuals. All of these guarantees (except for undetected fraud when $\mu > 0$) are much stronger for the bounded case. Thus QV appears to be roughly as robust as standard market institutions like the double auction are to collusion, though certainly not more so.

# 3 Aggregate Uncertainty and Common Values

I now assume away collusion and focus on changes to the informational environment.

## 3.1 Aggregate uncertainty

Our baseline model assumes that the distribution of values is commonly known and thus that, except in the knife-edge $\mu = 0$ case, there is no uncertainty about the optimal action. This both seems unrealistic and makes the informational problem somewhat trivial in large populations.[5] It thus seems natural to consider how QV performs when the distribution of valuations is uncertain among voters. Unlike with collusion, this case has a different mathematical structure LW's baseline analysis. However, by analogy to similar models of costly 1p1v I am able to make a detailed set of conjectures in some special cases.

Consider the simplest possible case of aggregate uncertainty, when there is an unknown scalar parameter $\gamma \in (\underline{\gamma}, \overline{\gamma}) \subseteq \mathbb{R}$ that determines the density of valuations, $f(u|\gamma)$, has a prior density distribution $g$ and and is affiliated with $u$, that is it orders $f$ by first-order stochastic dominance (Milgrom, 1981; Milgrom and Weber, 1982). I maintain all of my assumptions on the distribution from above also assume that $g$ is non-atomic and that my assumptions apply to the unconditional distribution of $u$. Assume that $\exists \gamma_-, \gamma_+ \in (\underline{\gamma}, \overline{\gamma}) : \mathbb{E}[u|\gamma_+] > 0 > \mathbb{E}[u|\gamma_-]$.

Analyses of costly 1p1v (Krishna and Morgan, 2012; Myatt, 2012) in this case suggest that equilibrium will have the structure that there exists a unique threshold $\gamma^\star$ such that for $\gamma$ above this threshold the alternative is implemented with probability near 1 in large populations and for $\gamma$ below this threshold the status quo is maintained with high probability in large populations. All pivotal events thus occur when $\gamma$ is very close to $\gamma^\star$ and welfare is by how close $\mathbb{E}[u|\gamma^\star]$ is to 0. Thus I conjecture the following intuitive extension of these results: there exists a threshold $\gamma^\star$ such that if $\gamma > \gamma^\star$ then the alternative is chosen with probability near 1 and if $\gamma < \gamma^\star$ then the alternative is chosen with probability near 0.

**Conjecture 1.** Under the assumptions of this section, there exists a unique $\gamma^\star \in (\underline{\gamma}, \overline{\gamma})$ such that in any equilibrium as $N \to \infty$, $\mathbb{P}(V > 0|\gamma) \to 1$ if $\gamma > \gamma^\star$ and $\mathbb{P}(V > 0|\gamma) \to 0$ if $\gamma < \gamma^\star$.

This conjecture greatly simplifies the analysis of equilibrium for several reasons. First, note that there is also a unique $\gamma_0 : \mathbb{E}[u|\gamma_+] > 0 > \mathbb{E}[u|\gamma_-]$ whenever $\gamma_+ > 0 > \gamma_-$. As a result, perfect limiting efficiency is achieved if and only if $\gamma_0 = \gamma^\star$. Second, by the analysis of Good and Mayer (1975) and Chamberlain and Rothschild (1981), for large $N$ all ties occur when $\gamma$ is very close to $\gamma^\star$. This leads to a very simple description of equilibrium behavior.

**Corollary 1.** At any equilibrium $v_i(u) = \left[ \frac{g(\gamma^\star|u) + o(1/\sqrt{N})}{N} \right] u$.

---

[5]However, as McLean and Postelwaite (Forthcoming) argue, it may be the existence of an efficient mechanism given aggregate certainty that provides correct incentives for individuals to reveal their information to the group and thus create this aggregate certainty. Thus aggregate certainty may be the appropriate framework for analysis of a robust mechanism like QV even if it would admit other, non-robust mechanisms as described in Subsection 5.4 below.

This characterization states that, in large populations individuals buy votes in proportion to the chance they perceive of $\gamma^\star$ realizing. This in turn leads to a simple integral equation for $\gamma^\star$:

$$\mathbb{E}\left[g\left(\gamma^\star|u\right)u|\gamma^\star\right] = 0. \tag{1}$$

I have not been able to derive from this fully general results about efficiency. However, I have studied several examples that admit an analytic solution of Equation 1; others can easily be studied by solving Equation 1 computationally. A common thread running throughout these analyses is the "Bayesian Underdog effect" (BUE) identified by Myatt (2012) in the context of 1p1v when it is costly to turn out.

Suppose, without loss of generality, that $\mathbb{E}[u] > 0$ so that the alternative is the ex-ante "favorite" in welfare terms and that the status quo is the ex-ante "underdog". If efficiency were to result, that is if $\gamma^\star = \gamma_0$, then individuals with $u < 0$ would tend to put a higher probability on $\gamma^\star$ intuitively because their own utility is a poll of one person indicating a lower value of $\gamma$. Because the alternative is the favorite lowering $\gamma$ increases the chance of a tie: Republicans believed that in 2012 a close election was more likely than did Democrats. This BUE thus raises the votes of the ex-ante underdog and thus $\gamma^\star$, leading to inefficiency because there are some values of $\gamma \in (\gamma_0, \gamma^\star)$ when the favorite should win but the alternative does. I have not been able to identify general conditions under which this logic is valid as it is based on a frequentist intuition, while it is the Bayesian probability of $\gamma^\star$ that is relevant. However, it plays an important role in all of the examples I have explored.

By Bayes's rule

$$\mathbb{E}\left[g\left(\gamma^\star|u\right)u|\gamma^\star\right] = \int_u u \frac{f\left(u|\gamma^\star\right)g\left(\gamma^\star\right)}{f(u)} f\left(u|\gamma^\star\right) du = g\left(\gamma^\star\right) \int_u u \frac{f^2\left(u|\gamma^\star\right)}{f(u)} du.$$

Thus any $\gamma^\star$ solving $\int_u u \frac{f^2(u|\gamma^\star)}{f(u)} du = 0$ also solves $\mathbb{E}\left[g\left(\gamma^\star|u\right)u|\gamma^\star\right] = 0$ and thus is an equilibrium value of $\gamma^\star$.

Given that raising the chance of a tie lowers the price, this would lead supporters of the underdog to vote more than those of the favorite. This in turn biases the vote in favor of the underdog, leading to some inefficiency. In every example I have studied where $\mathbb{E}[u] \neq 0$, this effect emerges and leads to some limiting inefficiency in QV. However, this inefficiency is never very large because the very logic that brings it about prevents it from being: if the favorite becomes the underdog because of large inefficiency the bias would reverse.

In Appendix B I consider several examples. I compute efficiency and inefficiency as in the previous section, but with an additional average taken over all possible realizations of $\gamma$ according to the measure over $\gamma$. I find that, over all of my examples, the largest inefficiency of QV is 5%, while 1p1v typically has quite high inefficiency, approaching 1 in some cases and usually above 10%. However, when the conditional distribution of $u$ given $\gamma$ is symmetric, 1p1v is al-

ways fully efficient but QV is not; when the distributions of $\gamma$ and $u$ given $\gamma$ are both normal, for example, QV reaches $2.2\%$ inefficiency in the worst case. In a calibrated gay marriage example, QV has 4% inefficiency and 1p1v 47%. In all cases as the aggregate uncertainty vanishes full efficiency is achieved, an interesting general result to shoot for in the future.

## 3.2  Common values

In our analysis above we assumed that elections served to aggregate preferences. Another role of voting to aggregate information, even when preferences are known to be aligned across individuals. There are two distinct types of such common values settings.

I begin by considering the simplest information aggregation setting, that of pure common interest considered by de Condorcet (1785). McLennan (1998) shows that the optimal information aggregation procedure using the actions available to agents is always an equilibrium. By essentially the same argument this is also true of QV. However, because QV allows expression of cardinal values, it allows the expression of strictly more information and thus, for generic information structures, achieves more efficient information aggregation in some equilibrium than voting does in any equilibrium. The one complicating factor is that under QV individuals must pay for their votes and thus interests are not perfectly aligned.

Without the smoothing through uncertainty ($\Psi$), this could lead to problems of existence, because individual could want to shirk in incorporating their information, given that an arbitrarily small down-scaling of all individuals' reports would be sufficient to implement any given efficient equilibrium. I have to investigate further these issues. With $\Psi$ added for smoothness existence and approximate efficiency should not be a problem, but perfect efficiency may be in a small population because of the aggregate uncertainty created by the smoothing. In either case I believe that QV will significantly outperform 1p1v in almost any case large population where individuals receive signals that are richer than binary.

The second type of common values setting proposed by Feddersen and Pesendorfer (1996, 1997) features individuals who differ both in their information about a common value component and their preferences over a private value component. Feddersen and Pesendorfer (1997) show that 1p1v aggregates information in large populations if there are arbitrarily unbiased, but still well-informed, individuals: these individuals who vote on the basis of information rather than preference. The fact that all information aggregation occurs through the votes of a narrow segment of the population, however, does put important limits on information aggregation. If, for example, all individuals have some minimum intensity of preferences, information does not aggregate and if those who are nearly indifferent also have very poor information, information aggregates very slowly. Under QV, by contrast, information aggregates by small adjustments to *all individuals' vote quantities* rather than large adjustments to a small fraction of the population's

votes. In Appendix B we use this fact to construct an example where information aggregates under QV but not under 1p1v. However, this model has proved challenging to analyze and thus we have not included a more general analysis of it at present.

# 4 Voter Behavior

If voting has any cost, standard instrumental models predict very few individuals should vote (Downs, 1957), contrary to empirical observations. In this subsection we consider how QV would perform if individuals behaved according to models capable of explaining observed turnout. Many such models, including as voting to tell others that one has voted (DellaVigna et al., 2015), would imply that behavior conditional on voting will follow standard rational choice.[6] Thus it may be that LW's baseline analysis is an accurate prediction of voting behavior in QV, given that we assumed universal turnout.

However, we consider the two models discussed in a survey on the paradox of voting by Blais (2000) that we could clearly determine would lead to different behavior in QV even conditional on voting. In the first model individuals overestimate the chance of their being pivotal. In the second, individuals gain some direct, "expressive" utility for each of their votes in addition to a chance of changing the outcome.

First, suppose that individuals misestimate the chance of the density of their being pivotal as a function of the votes they buy, which I denote $g(-v)$ for consistency with Weyl (2015), as $e\left(g(-v)\right)\epsilon$ where $e$ is smooth and $e, e_1, e_2(0) > 0 > e_{11}, e_{12}$ and, but that individuals never take a strictly dominated actions and any individual with a large impact (who subjectively rationally changes the chances of the outcome by 50% or more) has no mispercpetions. $\epsilon$ is a random variable drawn from $(\underline{\epsilon}, \bar{\epsilon}) \subseteq \mathbb{R}$ according to a smooth distribution $h$ that is independent of $u$ and embodies the extent to which individuals over-estimate the chance of being pivotal. $\epsilon$ is assumed to have mean 1 and denote its standard deviation by $\sigma_\epsilon$.

My assumptions on $e$ ensure that over-estimation is greater the small is the chance of being pivotal; individuals may even underestimate the chance when it is very large.[7] This is consistent with experimental evidence reported by Blais. Our assumption of independence of $\epsilon$ from $u$ ensures that no type of individual systematically over-estimates more than others *conditional on the true chance of their being pivotal with a marginal vote*. However, individuals with endogenously

---

[6]In fact, Kawai and Watanabe (2013) and Spenkuch (2014) find empirically that, conditional on voting at all, voters *do* behave quite strategically.

[7]An alternative model that I have also considered by do not report here for brevity is one where individuals overestimate the chance of their being pivotal unless they pay a cost to obtain a better estimate. In this case QV behaves more like 1p1v, thus losing some of its efficiency benefits over 1p1v. However, it may perform better for finite populations as this is the case that most effectively deters extremists and it always continues to outperform 1p1v, at least if the costs of acquiring information about $p$ are excluded. If these are included, 1p1v may perform better.

different chances of being pivotal (because of the number of votes their value induces them to buy) may over-estimate to differing extents. I label this model as the "misperception model".

Second, suppose that, in addition to the instrumental utility each individual earns, each also receives a benefit $\epsilon \left( x + \frac{\xi}{\sum_i |v_i|} \right) uv$ where $x, \xi \geq 0$. $\epsilon$ is again an independent-of-$u$ random variable with the same properties as in the misperception model. $x$ represents a per-unit-of-value expressive utility for each vote she purchases and $\xi$ is a per-unit-of-value expressive utility she earns from the *fraction of total votes cast* that she represents. These two possibilities correspond to two different interpretations of expressive utility in the literature. The first corresponds to traditional expressivist accounts, such as that of Fiorina (1976), where expression creates a personal psychological benefit for the voter. The second corresponds to a more quasi-instrumental motive, suggested by Myerson (2000), where voters vote to influence perceptions of political support, assuming only aggregate vote shares are reported by the media. I label this model as the "expressivist model".

In all cases I evaluate welfare based on the simple instrumentalist welfare I apply in the rest of the paper, as I do not have a coherent approach to incorporating the value individuals gain from expressing themselves or acting on their misperceptions. Furthermore I typically think the latter will be quantitatively small relative to the former. I also assume that the distribution of $\epsilon$ either has thin tails or has a tail index weakly greater than those of the value distribution (has thinner tails). As in my baseline analysis, but even more strongly, the conclusions I reach differ dramatically based on whether I consider the case of $\mu = 0$ or $\mu \neq 0$, though interestingly they do not differ dramatically between the misperception and expressivist models.

## 4.1 $\mu = 0$

First consider the case of $\mu = 0$. First consider either the expressivist model when $x \neq 0$ or the misperception model. In the first case, the expressivist motive must be dominant in large populations as the instrumental motive dies as the population grows large, while the expressivist motive does not. Thus in both cases the variation in votes, orthogonal to $u$, created by $\epsilon$ must be the leading source of inefficiency as it does not die with population size and, given the smoothness of $e$, LW's arguments for other sources of inefficiency vanishing can easily be replicated here.

**Claim 5.** If $F$ has $\mu = 0$ then in any type-symmetric (across the joint type $(u, \epsilon)$) Bayes-Nash equilibrium, then

1.  In the expressivist model if $x = 0$ then there exists an $\underline{N}$ such that $EI \leq \frac{\sigma_\epsilon^2}{4\sqrt{N}}$.

2.  In any other case, EI converges as $N$ grows large to $\frac{1}{2} - \frac{1}{2\sqrt{1+\sigma_\epsilon^2}}$.

*Heuristic justification.* Because $\epsilon$ is independent of $u$, the variance of $\epsilon u$ is $\sigma^2\left(1+\sigma_\epsilon^2\right)$ and the covariance of $u$ and $\epsilon u$ is $\sigma^2$. Thus by the Central Limit Theorem (CLT) logic, letting $\hat{V}$ be an appropriate normalization (by $\mathbb{E}[|\sum_i v_i|]/\mathbb{E}[|U|]$) of $\sum_i v_i$, in equilibrium approximately for large $N$

$$\frac{1}{N}\begin{pmatrix} U \\ \hat{V} \end{pmatrix} \sim \mathcal{N}\left(\begin{array}{c} 0 \\ 0 \end{array}, \frac{1}{N}\begin{bmatrix} \sigma^2 & \sigma^2 \\ \sigma^2 & \sigma^2\left(1+\sigma_\epsilon^2\right) \end{bmatrix}\right).$$

Thus $\hat{V}$ is just $U$ plus mean zero noise of magnitude $\sigma^2\sigma_\epsilon^2/N$. Efficiency is the expected total value of $U$ in events where the vote goes the right way, or in this case approximately

$$2\sqrt{N}\int_0^\infty \frac{xe^{-\frac{x^2}{2\sigma^2}}\Phi\left(\frac{x}{\sigma\sigma_\epsilon}\right)}{\sqrt{2\pi}\sigma}dx = \frac{\left(1+\frac{1}{\sqrt{1+\sigma_\epsilon^2}}\right)\sigma}{\sqrt{2\pi}}.$$

Normalizing this by $\sqrt{2/\pi}\sigma$, the total possible efficiency, yields

$$\frac{1}{2}+\frac{1}{2\sqrt{1+\sigma_\epsilon^2}} \implies EI \approx 1-\left(\frac{1}{2}+\frac{1}{2\sqrt{1+\sigma_\epsilon^2}}\right) = \frac{1}{2}-\frac{1}{2\sqrt{1+\sigma_\epsilon^2}}.$$

In the expressivist model when $x=0$ suppose that everyone behaves as in the limiting equilibrium where the expressivist motive is entirely ignored. Then aggregate votes are on the order of $N^{\frac{3}{4}}$ by LW's arguments in the $\mu=0$ case. Thus the size of the expressivist motive is in $\Theta\left(1/N^{3/4}\right)$. On the other hand my calculations in Weyl (2015), the instrumental voting motive is in $\Theta\left(1/\sqrt[4]{N}\right)$, while the inefficiency arises from deviations from proportionality that are in $\Theta\left(1/N^{5/4}\right)$. Thus the expressivist motive is larger than the sources of inefficiency in this case, dying only at a relative rate $1/\sqrt{N}$, but still dies relative to the vote total. In this case the same logic as in the two cases considered above yields the same constant on the large population inefficiency, but now because this motives dies at a relative rate $1/\sqrt{N}$, EI is bounded above by

$$\frac{1}{2}-\frac{1}{2\sqrt{1+\frac{\sigma_\epsilon^2}{\sqrt{N}}}} \approx \frac{\sigma_\epsilon^2}{4\sqrt{N}},$$

where the approximation is based on the Taylor expansion for large $N$. Furthermore this function is always concave, so this Taylor expansion overstates the limiting inefficiency, implying this is an upper-bound for large $N$. $\qquad\square$

When the expressivist motive dies off with the total number of votes, limiting efficiency still obtains, but more slowly than before; however as I discuss presently a value of $\sigma_\epsilon$ much above unity is implausible so in practice inefficiency in reasonably large populations will be quite small. If its value is unity, for example, then even in a town of 10,000 it would be a quarter

of a percent.

Even in the case when there is limiting inefficiency, QV will often out-perform 1p1v. Goeree and Zhang (2013) show that when $\mu = 0$ the limiting EI of majority rule is

$$\frac{1}{2} - \frac{\mathbb{E}\left[|u|\right]}{2\sigma}.$$

This equals about 10% for the normal distribution and about 24% for my mean-zero gay marriage example with uniform distributions. In order to match these levels of inefficiency the noise from $\epsilon$ have to be quite large: $\sigma_\epsilon$ would have to equal 57% in the case of the normal distribution and 270% in the uniform distribution gay marriage example. In cases with fat-tailed distributions of values the inefficiency of 1p1v is far greater, as it is more important to incorporate the highly heterogeneous intensities of the population into the decision: it is about 43%, requiring an insanely large 5000% value of $\sigma_\epsilon$ to match it to my mean-zero, unbounded support value distribution calibrated example in Weyl (2015). Thus while QV is not perfectly efficient in this case, noise has be extremely large (often much larger than the signal coming from values) in realistic cases for 1p1v to out-perform QV.

Note, however, that this is not always the case; for example, suppose that noise took on a value of $0$ with very high probability and some extremely large value with very low probability to maintain the mean of $1$. Then QV could be worse than 1p1v, especially if intensities are relatively homogeneous, as QV would effectively count only a very small fraction of all valid votes. Intuitively, 1p1v throws away all cardinal signal. Unless this signal is of little value because individuals have homogeneous intensities and/or the noise in individuals' estimations is so large that these signals are overwhelmed, QV will perform better even though it responds to spurious noise that 1p1v neglects.

## 4.2  $\mu > 0$

Now I turn to the $\mu \neq 0$ case and as usual focus on $\mu > 0$ WLOG. I give the heuristic justification of the following claim in Appendix C.

**Claim 6.** If $F$ has $\mu \neq 0$ then in any type-symmetric Bayes-Nash equilibrium, then

1. In the expressivist model, so long as either $\max\{x, \xi\} > 0$, there $EI \in O\left(N^{\alpha-1}\right)$.

2. In the misperception model, if for sufficiently small $p$, $e(g) \in \Omega\left(\sqrt[\beta]{g}\right)$, then there exists a $\underline{N}$ such that for $N \geq \underline{N}$, $EI \in O\left(1/N^\kappa\right)$, where $\kappa = \max\left\{\alpha - 1, \frac{2\alpha\beta-1-\beta}{1+\beta+2\alpha}\right\}$.

In the expressivist model or the misperception model for strong enough misperception ($\beta$ large enough) the rate here is much faster than those reported by LW and myself in Weyl (2015); even in my most fat-tailed case ($\alpha = 3$), EI dies off with the square of $N$. How large does $\beta$ have

17

to be for these faster rates to take hold? Consider my running example of $\alpha = 3$. Then $\beta$ must equal $5$ (the decay of beliefs in pivotality must be no faster than the fifth root of the decay of the true chance of being pivotal) before the new rates dominate. When $\beta$ is greater than $5$, the leading source of inefficiency is insufficient rather than excessive purchases by extreme players. When $\beta$ is smaller convergence is again improved, but not as dramatically. Consider $\beta = 3$, roughly the value necessary according to the data of Gelman et al. (2010) to justify voting in a presidential election worth \$5000 to a citizen who must spend \$10 to vote, perhaps therefore a reasonable ball-park for real-world marginal voters. Then inefficiency decays with $N^{-7/5}$, much faster than the $1/\sqrt{N}$ rate predicted by the fully rational model in this case.

On the other hand in the bounded or thin-tailed distribution case, when $\alpha \to \infty$, excessive extremism is always the leading source of harm, the decay of EI is exponential in the expressivist model and occurs at rate $1/N^\beta$ in the misperception model; if in addition the perceived chance of being pivotal stays significant even when $g$ is truly exponentially small, arguably a finding consistent with Blais (2000)'s results, then the decay of EI is exponential in the misperception model as well.

To summarize:

1. In the Myerson-style expressivist model, the long-term-oriented voter behavior never causes limiting inefficiency, though it does slow its decay when $\mu = 0$, and generically dramatically increases the rate of decay of EI.

2. In the Fiorina-style expressivist model, the non-instrumental voter behavior causes some limiting inefficiency but probably does not make QV less efficient than 1p1v in many realistic cases, even constrained to the non-generic $\mu = 0$ class, and generically dramatically increases the rate of decay of EI.

3. In the Blais-style misperception model, irrational voter behavior causes some limiting inefficiency but probably does not make QV less efficient than 1p1v in many realistic cases, even constrained to the non-generic $\mu = 0$ class, and generically significantly decreases the rate of decay of EI, so long as the overestimation of the chance of being pivotal affects the rate and not just constant of the decay of $g$ near $0$.

Of course if $\epsilon$, rather than being independent of $u$, were correlated with it this would likely cause limiting inefficiency for reasons similar to those in the aggregate uncertainty case of the previous section, though this might be worse than there as there would be little rationality to keep such divergent estimates in check. I do not consider this a particularly realistic case, however, and thus do not discuss it further here. In fact, the limited laboratory (Goeree and Zhang, 2013) and field (Cárdenas et al., 2014) experimentation on QV thus far indicate these forces play an important role in the efficiency of QV in practice and does not appear to interfere with the

efficiency of QV in at least these controlled experimental contexts. Thus I do not think that non-instrumental or irrational voter behavior is a significant threat to, and may even enhance, QV's efficiency.

# 5 Comparison to Other Mechanisms

These conclusions about the robustness of QV appear to be confirmed by the first laboratory (Goeree and Zhang, 2013) and field (Cárdenas et al., 2014) experiments on QV. I now compare these conclusions to the properties of other mechanisms economists have proposed; these comparisons are summarized in Table 1.

## 5.1 One-person-one-vote

Because most of this paper compares QV to majority rule 1p1v, I only cover here a few points not covered above. 1p1v is somewhat more robust to collusion than is QV along some dimensions (a colluding group of extremists may accomplish less unless it is even larger than needed under QV), but along others it is less robust (incentives for unilateral deviation are smaller, reactions by those outside the colluding group are non-existent and the incentive to form a collusive group is stronger given it is the only way to exert more influence). Ledyard and Palfrey (1994, 2002) show that in large populations if the distribution of valuations is known then by choosing the threshold for voting equal to the quantile corresponding to its mean, the limit-efficiency of voting may be restored. This mechanism suffers from the same limitations of the mechanisms, discussed in Subsections 5.3 and 5.4 below, that require the planner to know the distribution of valuations. For any fixed super- or sub-majority rule, voting typically performs worse than it would under a simple majority rule; see Posner and Weyl (Forthcoming) for a detailed discussion.

## 5.2 The Vickrey-Clarke-Groves mechanism

The most canonical mechanism that has been suggested by economists as an alternative to 1p1v is the Vickrey (1961)-Clarke (1971)-Groves (1973) (VCG) mechanism. In its application to binary collective decisions (Tideman and Tullock, 1976; Green et al., 1976), individuals report their cardinal value for the alternative and the decision is chosen based on the sign of the sum of the reports. Any individual who is pivotal in the sense that, had she reported $0$, the decision would have gone the other way pays the amount by which all those other than her preferred the decision she opposed. In addition to sharing with QV its independence from the details of the value distribution, this system is appealing because if every individual plays the weakly-dominant strategy of reporting her valuation truthfully then, in an extremely broad range of

| | Detail-dependence | Standard efficiency | Aggregate uncertainty | Strategic complexity | Collusion and de-mergers | Budget | Other problems |
|---|---|---|---|---|---|---|---|
| QV | Detail-free | Limit-efficient | Limit-efficient with large polls or if $\mathbb{E}[u] = 0$; inefficiency small otherwise | Non-direct | Moderately robust | Balanced | |
| 1p1v | Detail-free | Limit-efficient when mean and median have same definite sign, typically inefficient | Typically limit-inefficient | Strategy-proof | Moderately robust | Transfer-free | |
| VCG | Detail-free | Efficient | Efficient | Strategy-proof | Extremely sensitive | Surplus | Bankruptcy, uncertain payments |
| Expected Externality | Detail-based, requires aggregate certainty | Efficient | Ill-defined | Direct but Bayesian | Highly sensitive | Balanced | Ill-defined |

Table 1: Comparison of binary collective choice mechanisms

circumstances, VCG implements the efficient outcome. VCG is fully efficient (in this equilibrium) in a sense that is very robust to the information structure and the number of participating individuals, unlike QV.

Despite this, somewhat narrow sense of, robustness, the VCG mechanism has almost never been used for collective choices. The reason that VCG is "lovely, but lonely" (Ausubel and Milgrom, 2005) is that a number of other severe failures of robustness make it "not practical" (Rothkopf, 2007). These flaws were recognized by the originators of the mechanism from the start (Vickrey, 1961; Groves and Ledyard, 1977), though their severity and implications for implementing VCG were not well-understood until more recently (Tideman and Tullock, 1977) when the first laboratory experiments based on the simplest Tideman and Tullock environments gave disastrous results (Attiyeh et al., 2000).

Perhaps the most severe defect of the VCG is that under complete information, in addition to its efficient equilibria, VCG has a very large number of other equilibria, including, for any two individuals, equilibria where they attain their desired outcome and make no payments. In particular, any two individuals may announce sufficiently large values in the same direction so that neither is individually pivotal. Anticipating this, other individuals can do no better than to report $0$ (or her true value). Similarly, any individual who can pretend to be two individuals can "break" the mechanism. Even with incomplete information, if value reports are unbounded, for any two individuals this strategy can be implemented in an $\epsilon$ equilibrium for any $\epsilon > 0$ as if each conspirator reports a sufficiently extreme value the chances of the other conspirator being pivotal become negligible.[8]

Among the other practical problems with VCG are that any revenue raised must be destroyed to avoid creating perverse incentives, which may be hard for the government to commit to; absent such a commitment, the scheme falls apart.[9] Even when such commitments are possible, the revenue that must be destroyed can often be greater than the improvement in efficiency over even simple mechanisms like 1p1v (Groves and Ledyard, 1977; Attiyeh et al., 2000; Drexl and Kleiner, 2013). VCG requires much larger liquidity among participants than does QV; individuals must have in cash the full magnitude of their value and place this into escrow when submitting their report. Unless this "bankruptcy" problem is addressed, individuals have an incentive to exaggerate their report and fall back on judgement-proofness if called upon to pay (which occurs with very small probability in a large population). Under QV payments are limited with probability 1 to a very small portion of underlying values and are certain conditional

---

[8]While limiting the report space can help reduce this risk, it either must sacrifice some efficiency to do so (especially in the fat tail case) by reducing its sensitivity to large values or remain open to collusion by slightly larger groups. Furthermore such ad-hoc fixes make the mechanism ever-more complex and eliminate most of its small-sample advantages over QV.

[9]Some have suggested schemes to get around this problem in large populations. See, for example, the work of (Bailey, 1997). Again this adds further complications.

on the report made.[10]

## 5.3    Expected Externality mechanism

The next-most canonical mechanism economists have suggested is the Expected Externality (EE) mechanism of Arrow (1979) and d'Aspremont and Gérard-Varet (1979), which was first applied to binary collective decisions by Goeree and Li (2008). This is similar to VCG, except that individuals pay the planner's ex-ante expectation of their VCG payments rather than their actual payments. Because individuals can thus not affect others' payments, the revenue raised may be refunded much as under QV, though, like QV, the mechanism is Bayesian rather than having a dominant-strategy equilibrium. QV was partly inspired by this mechanism as, in the case when $\mu = 0$, these EE payments are approximately quadratic in large populations.

This intuition was first perceived by Vickrey (1961) and highlighted by Tideman (1983). Vickrey studies how taxes might be used to offset the incentives of an oligopolist to manipulate market prices. Unlike the competitive profits that an oligopolist makes, which are proportional to her quantity, this incentive is typically proportional to the *square* of her quantity, as both the amount by which price moves and the number of units impacted by the price change are proportional to the oligopolist's production. Thus the necessary offsetting "counter-speculation" tax is proportional to the square firm's quantity. In the competitive limit, where each firm is a small part of the market, the quadratic term vanishes, leading to competitive behavior (Satterthwaite and Williams, 1989). On the other hand, in the limit of public goods, where individuals are only concerned with the aggregate outcome rather their own production, the linear term vanishes and the quadratic term becomes the leading term for small individuals.

This quadratic approximation argument is illustrated in Figure 1, where we consider the classic problem of choosing the level of consumption of a good causing a negative externality. Each individual reports her schedule of harms and the optimal level of the externality is determined by equating demand to the vertical sum of the harm schedules. The Vickrey payment is the externality on other individuals of a given individual's report, which is the area between private demand and the cost curve for all other individuals between the quantity that would prevail absent an individual's report and the quantity that prevails with this report. Note that this is a deadweight loss triangle and, as such, grows quadratically in the change in quantity induced by the individual's report as shown in the figure.

However, this is at most a starting point for QV, as when $\mu \neq 0$ EE payments are nothing like quadratic and in the richer information environments we consider they are not even well-defined. However, Goeree and Zhang (2013) take this logic more literally and propose a

---

[10]It seems likely, though I have not tried to show it, that this would lead to strict dominance of QV when agents are risk-averse.
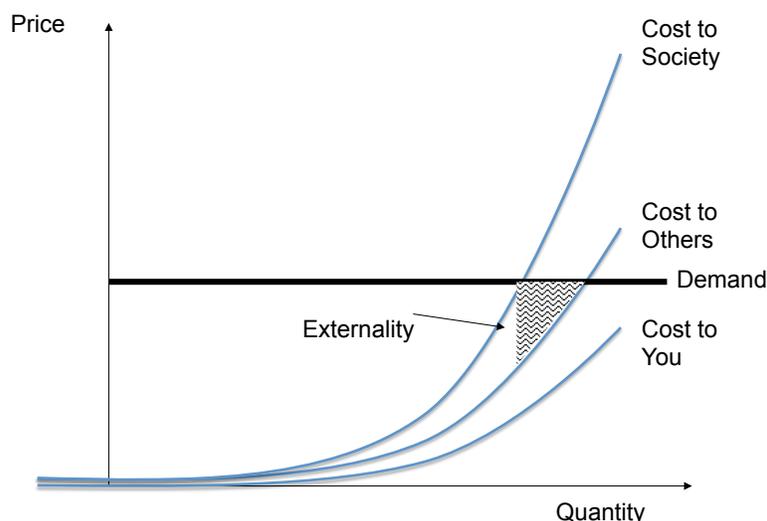
Figure 1: The quadratic growth of the deadweight loss triangle in an individual's influence as an analogy for the optimality of quadratic pricing of votes.

mechanism, limited to the case when $\mu = 0$, in which individuals pay the exact quadratic approximation to their EE payments. This quadratic schedule has a coefficient on it that depends on the number of individuals and the standard deviation of their value and thus, like EE, depends on the planner knowing the distribution of valuations. Thus both the EE and Goeree and Zhang mechanisms are not detail-free and only apply under aggregate certainty (and the later also requires $\mu = 0$). Additionally, when $\mu \neq 0$, the EE mechanism suffers from essentially the same collusion problem as VCG in large populations, though this is less well-known.[11]

## 5.4 Other mechanisms proposed by economists

Many other, even more fragile, mechanisms have been proposed by economists and we briefly discuss a few:

1. *Implement the alternative if and only if $\mu > 0$*: In addition to requiring the planner to know the distribution of valuations, this suggestion places great and easily-abused power in the hands of the authority charged with determining the sign of $\mu$ (Maskin, 1999).

2. *The Krishna and Morgan (2001) mechanism*: In this mechanism, all individuals are asked to report $\mu$ and the alternative is implemented if and only if all report $\mu > 0$. If all do not agree, all agents are "killed" (pay a very large and inefficient penalty). In addition to

---

[11]Suppose two individuals report $-\frac{2\mu N}{3}$. Each will make vanishingly small EE payments as the probability of either of these reports being pivotal is exponentially small in $N$ by standard large deviation theory results. However, together they will ensure the outcome goes in the inefficient direction.

requiring aggregate certainty, this mechanism is likely to be difficult to commit to and has a very large multiplicity of inefficient equilibria (Battaglini, 2002).[12]

3. *Crémer and McLean (1988)-McAfee and Reny (1992)-style mechanisms:* Roughly, individuals are asked (via their report of their type) to guess other individuals' report of their type and are given large rewards for guessing correctly. Like the Krishna and Morgan mechanism, this mechanism has a large multiplicity of equilibria (McLean and Postelwaite, Forthcoming), depends both on the mechanism designer having a very precise knowledge of the distribution of types and on individuals having preferences that are "appropriately" correlated with their beliefs about other individuals' types (Heifetz and Neeman, 2006).

All of these mechanisms are also quite complicated to explain and strain credibility along a variety of other practical dimensions. While there is not the space here to discuss all in detail, a large literature has established their impracticability.[13] An important reason, I believe, for their limitations, and for those of the VCG and EE mechanisms, is the "over-fitting" problem that motivated my investigation of QV's robustness.

## 5.5   Mechanisms used in practice

Existing informal institutions may be more efficient than formal 1p1v is. The fact that voting may be costly is one such institution (Ledyard, 1984; Myerson, 2000; Krishna and Morgan, Forthcoming). Ledyard considers a model where individuals have non-pecuniary costs of voting. If these costs are all strictly positive and follow a non-atomic distribution independent of voters' values, then the "representative voter" with a given value effectively faces a quadratic cost as a function of the fraction of cost types she represents in the limit as the population size grows large and thus both the density of pivotality and thus turn out grow very small.[14] Thus non-pecuniary costs of voting may approximate QV and thus efficiency.

Ledyard argues that this is unlikely to provide a good approximation to reality as it requires turnout to approach zero in large populations, which is rarely observed in practice. If some voters have zero or negative non-pecuniary costs of voting or overestimate their chance of being pivotal to such an extent that turnout remains large in large elections, as much empirical research on voting suggests (Blais, 2000), the result clearly fails, though voluntary voting can somewhat mitigate some of the inefficiency of compulsory voting (Borgers, 2004). On the other hand, I

---

[12]Other mechanisms requiring individuals to have complete information have been proposed (Hurwicz, 1977; Maskin, 1999; Walker, 1981), but are widely viewed as highly fragile (Bailey, 1994). Another mechanism proposed by Thompson (1966) rely very sensitively on the absence of heterogeneity in risk attitudes and beliefs and on detailed knowledge by the planner.

[13]See Tideman (2006) for a detailed and excellent survey from which much of the discussion here is derived.

[14]The argument is same argument underlying the analogy to deadweight loss triangle in Subsection 5.3 above.

showed in Section 4 that QV remains as efficient or even may be more efficient than if voters are "standard". Thus costly voting does not seem a promising approximation to QV in practice.[15]

More comprehensive solutions to the inefficiency of 1p1v are genuinely costly activities undertaken to influence collective decisions, such as log-rolling on committees and legislative bodies and lobbying, campaigning or out-right illegal vote-buying in elections. At least in some contexts, costs may be somewhat convex. In votes on committees it is usually easy for any individual to influence one vote of a colleague, harder to obtain the second, even harder to sway the third and so forth.[16] Exactly how close various contexts come to approximating QV is an interesting empirical question.

# 6   Conclusion

My analysis here suggests that QV is the only plausibly practical, approximately efficient collective decision mechanism economists have proposed. It also suggests it is robust enough to be worth trying out in small-scale but non-trivial practical settings. However I only consider a very limited range of robustness issues in this paper and in some cases consider even these fairly superficially. Much work therefore remains to be done investigating the robustness of QV. I briefly discuss what I consider to be some of the most interesting issues here.

First, my investigations of aggregate uncertainty considered only a few specific examples and my analysis of common values only one extremely special limiting case. A broader treatment of both of these cases, especially the latter, would be highly informative. Second, none of my analyses allowed for endogenous information acquisition or pre-voting communication among voters. Such factors are crucial to how real-world elections operate. Third, both my analysis of collusion and of voter behavior could be substantially enriched by considering richer equilibrium models of collusion and voter behavior, such as factoring in imperfect information among a colluding group (Laffont and Martimort, 1997, 2000) or voting behavior partly motivated by social concerns. Incorporating risk-aversion and budget constraints explicitly would also be interesting in comparing QV to mechanism that I claimed above rely more heavily on quasi-linear utility, like VCG.

---

[15]At a more technical level, note that costly voting is is a wasteful means of vote pricing compared with QV, because the costs are not refunded back.

[16]Similar effects may occur in campaigning, where it is typically much easier for individuals to influence their close friends or at least those they are acquainted with than to influence those further away from them.

# References

ARROW, K. J. (1979): "The Property Rights Doctrine and Demand Revelation under Incomplete Information," in *Economics and Human Welfare*, ed. by M. Boskin, New York: Academic Press, 23–39.

ATTIYEH, G., R. FRANCIOSI, AND R. M. ISAAC (2000): "Experiments with the Pivot Process for Providing Public Goods," *Public Choice*, 102, 95–112.

AUSUBEL, L. M. AND P. MILGROM (2005): "The Lovely but Lonely Vickrey Auction," in *Combinatorial Auctions*, ed. by P. Cramton, R. Steinberg, and Y. Shoham, Cambridge, MA: MIT Press, 17–40.

BAILEY, M. J. (1994): "Lindahl Mechanisms and Free Riders," *Public Choice*, 80, 35–39.

——— (1997): "The Demand Revealing Process: To Distribute the Surplus," *Public Choice*, 91, 107–126.

BATTAGLINI, M. (2002): "Multiple Referrals and Multidimensional Cheap Talk," *Econometrica*, 70, 1379–1401.

BLAIS, A. (2000): *To Vote or Not to Vote: The Merits and Limits of Rational Choice Theory*, Pittsburgh, PA: University of Pittsburgh Press.

BLUMER, A., A. EHRENFEUCHT, D. HAUSSLER, AND M. K. WARMUTH (1987): "Occam's Razor," *Information Processing Letters*, 24, 377–380.

BORGERS, T. (2004): "Costly Voting," *American Economic Review*, 94, 57–66.

CÁRDENAS, J. C., C. MANTILLA, AND R. D. ZÁRATE (2014): "Purchasing Votes without Cash: Implementing Quadratic Voting Outside the Lab," Https://www.aeaweb.org/aea/2015conference/program/retrieve.php?pdfid=719.

CHAMBERLAIN, G. AND M. ROTHSCHILD (1981): "A Note on the Probability of Casting a Decisive Vote," *Journal of Economic Theory*, 25, 152–162.

CLARKE, E. H. (1971): "Multipart Pricing of Public Goods," *Public Choice*, 11, 17–33.

CRÉMER, J. AND R. P. MCLEAN (1988): "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica*, 56, 1247–1257.

D'ASPREMONT, C. AND L.-A. GÉRARD-VARET (1979): "Incentives and Incomplete Information," *Journal of Public Economics*, 11, 25–45.

DE CONDORCET, M. (1785): *Essai sur L'Application de L'Analyse à la Probabilité des Décisions Rendue à la pluralité des Voix*, Paris: L'Imprimerie Royale.

DELLAVIGNA, S., J. A. LIST, U. MALMENDIER, AND G. RAO (2015): "Voting to Tell Others," Http://goo.gl/o8qpS1.

DOWNS, A. (1957): *An Economic Theory of Democracy*, New York: Harper.

DREXL, M. AND A. KLEINER (2013): "Why Voting? A Welfare Analysis," Http://www.coll.mpg.de/sites/www.coll.mpg.de/files/workshop/note.pdf.

FEDDERSEN, T. AND W. PESENDORFER (1997): "Voting Behavior and Information Aggregation in Elections With Private Information," *Econometrica*, 65, 1029–1058.

FEDDERSEN, T. J. AND W. PESENDORFER (1996): "The Swing Voter's Curse," *American Economic Review*, 86, 408–424.

FIORINA, M. P. (1976): "The Voting Decision: Instrumental and Expressive Aspects," *Journal of Politics*, 38, 390–413.

GELMAN, A., N. SILVER, AND A. EDLIN (2010): "What is the Probability Your Vote Will Make a Difference?" *Economic Inquiry*, 50, 321–326.

GOEREE, J. K. AND N. LI (2008): "Inefficient Voting," Http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.192.4763&rep=rep1&type=pdf.

GOEREE, J. K. AND J. ZHANG (2013): "Electoral Engineering: One Man, One ~~Vote~~ Bid," Http://goo.gl/s2SCsP.

GOOD, I. J. AND L. S. MAYER (1975): "Estimating the Efficacy of a Vote," *Behavioral Science*, 20, 25–33.

GREEN, J., E. KOHLBERG, AND J.-J. LAFFONT (1976): "Partial Equilibrium Approach to the Free-Rider Problem," *Journal of Public Economics*, 6, 375–394.

GROVES, T. (1973): "Incentives in Teams," *Econometrica*, 41, 617–631.

GROVES, T. AND J. O. LEDYARD (1977): "Some Limitations of Demand Revealing Processes," *Public Choice*, 29, 107–124.

HEIFETZ, A. AND Z. NEEMAN (2006): "On the Generic (Im)Possibility of Full Surplus Extraction in Mechanism Design," *Econometrica*, 74, 213–233.

27

HURWICZ, L. (1977): "On Informationally Decentralized Systems," in *Studies in Resource Allocation Processes*, ed. by K. J. Arrow and L. Hurwicz, Cambridge, UK: Cabmridge University Press, 425–459.

KAWAI, K. AND Y. WATANABE (2013): "Inferring Strategic Voting," *American Economic Review*, 103, 624–662.

KRISHNA, V. AND J. MORGAN (2001): "Asymmetric Information and Legislative Rules: Some Ammendments," *American Political Science Review*, 95, 435–452.

——— (2012): "Majority Rule and Utilitarian Welfare," Http://papers.ssrn.com/sol3/papers.cfm?abstract

——— (Forthcoming): "Majority Rule and Utilitarian Welfare," *American Economic Journal: Microeconomics*.

LAFFONT, J.-J. AND D. MARTIMORT (1997): "Collusion under Asymmetric Information," *Econometrica*, 65, 875–911.

——— (2000): "Mechanism Design with Collusion and Correlation," *Econometrica*, 68, 309–342.

LALLEY, S. P. AND E. G. WEYL (2015): "Quadratic Voting," Http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2003531.

LEDYARD, J. (1984): "The Pure Theory of Large Two-Candidate Elections," *Public Choice*, 44, 7–41.

LEDYARD, J. O. AND T. R. PALFREY (1994): "Voting and Lottery Drafts as Efficient Public Goods Mechanisms," *Review of Economic Studies*, 61, 327–355.

——— (2002): "The Approximation of Efficient Public Good Mechanisms by Simple Voting Schemes," *Journal of Public Economics*, 83, 153–171.

MASKIN, E. (1999): "Nash Equilibrium and Welfare Optimality," *Review of Economic Studies*, 66, 23–38.

MCAFEE, R. P. AND P. J. RENY (1992): "Correlated Information and Mechanism Design," *Econometrica*, 60, 395–421.

MCLEAN, R. P. AND A. POSTELWAITE (Forthcoming): "Implementation with Interdependent Values," *Theoretical Economics*.

MCLENNAN, A. (1998): "Consequences of the Condorcet Jury Theorem for Beneficial Information Aggregation by Rational Agents," *American Political Science Review*, 92, 413–418.

MILGROM, P. R. (1981): "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12, 380–391.

MILGROM, P. R. AND R. J. WEBER (1982): "A Theory of Auctions and Competitive Bidding," *Econometrica*, 50, 1089–1122.

MYATT, D. P. (2012): "A Rational Choice Theory of Voter Turnout," Http://www2.warwick.ac.uk/fac/soc/economics/news_events/conferences/politicalecon/turnout-2012.pdf.

MYERSON, R. B. (2000): "Large Poisson Games," *Journal of Economic Theory*, 94, 7–45.

POSNER, E. A. AND E. G. WEYL (Forthcoming): "Voting Squared: Quadratic Voting in Democratic Politics," *Vanderbilt Law Review*, 68.

ROTHKOPF, M. H. (2007): "Thirteen Reasons Why the Vickrey-Clarke-Groves Process is Not Practical," *Operations Research*, 55, 191–197.

SATTERTHWAITE, M. A. AND S. R. WILLIAMS (1989): "The Rate of Convergence to Efficiency in the Buyer's Bid Double Auction as the Market Becomes Large," *Review of Economic Studies*, 56, 477–498.

SPENKUCH, J. L. (2014): "(Ir)rational Voters?" Http://www.kellogg.northwestern.edu/faculty/spenkuch/research/voting.pdf.

STIGLER, G. J. (1964): "A Theory of Oligopoly," *Journal of Political Economy*, 72, 44–61.

THOMPSON, E. A. (1966): "A Pareto-Efficient Group Decision Process," *Papers on Non-Market Decision-Making*.

TIDEMAN, N. (2006): *Collective Decisions and Voting: The Potential for Public Choice*, Hampshire, UK: Ashgate.

TIDEMAN, N. N. AND G. TULLOCK (1976): "A New and Superior Process for Making Social Choices," *Journal of Political Economy*, 84, 1145–1159.

TIDEMAN, T. N. (1983): "A Collective Conception of Collective Value," *Perspectives on Local Public Finance and Public Policy*, 1, 3–22.

TIDEMAN, T. N. AND G. TULLOCK (1977): "Some Limitations of Demand Revealing Processes: Comment," *Public Choice*, 29, 125–128.

VICKREY, W. (1961): "Counterspeculation, Auctions and Competitive Sealed Tenders," *Journal of Finance*, 16, 8–37.

WALKER, M. (1981): "A Simple Incentive Compatible Scheme for Attaining Lindahl Allocations," *Econometrica*, 49, 65–71.

WEYL, E. G. (2015): "The Efficiency of Quadratic Voting in Finite Populations," Http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2571026.

# Appendix

## A   Collusion and Fraud

*Heuristic justification of Claim 2.* For an extremist conspiracy to maintain each individual's unilateral incentive to participate, it must be that each individual purchase votes no greater than the square root of the magnitude of her value. This implies that the aggregate vote purchased by an extremist coalition cannot exceed the sum of the squares of values of its members. By my logic in the text, this approximately equals

$$\int_0^M x^{-\frac{1}{2\alpha}} dx = \frac{2\alpha}{2\alpha - 1} M^{\frac{2\alpha-1}{2\alpha}}$$

times the value of the square root of the last-order statistic. The total votes bought by moderates is $\Theta\left(N^{\frac{1}{\alpha}}\right)$ as it is on the order of square root the extremist threshold. Dividing this by $\frac{2\alpha}{2\alpha-1} M^{\frac{2\alpha-1}{2\alpha}}$, we have that, for the extremist conspiracy to succeed, the square root of the first-order statistic must be in $\Theta\left(N^{\frac{1}{\alpha}} M^{-\frac{2\alpha-1}{2\alpha}}\right)$. This implies that that first-order statistic itself must be in $\Theta\left(N^{\frac{2}{\alpha}} M^{-\frac{2\alpha-1}{\alpha}}\right)$. The probability of the first-order statistic being that large is, by my logic in the text

$$\Theta\left(N N^{-2} M^{2\alpha-1}\right) = \Theta\left(\frac{M^{2\alpha-1}}{N}\right),$$

which clearly dies with $N$ if $M \in O\left(N^{1/2\alpha-1}\right)$. □

*Heuristic justification of Claim 3.* It is again instructive to consider this logic in the familiar case of bounded support. Suppose that individuals believe that the chance of an extremist coalition forming and succeeding is $r$. Then the ratio of votes purchased by an average moderate to those purchased by the average member of the extremist coalition will be, by my logic in Weyl (2015), approximately $r\mu/M|\underline{u}|$ because the extremists know they exist when they do and multiply their votes by their coalition size to internalize the externalities to other members of the coalition. For the extremist coalition to win when it does exist, on the other hand, this requires that the total votes purchased by the extremist coalition (more than) balance those purchased by others on average so that

$$r\mu N \le M^2 |\underline{u}| \implies r \le \frac{M^2 |\underline{u}|}{\mu N},$$

so that $M$ must be $\Omega\left(\sqrt{N}\right)$ in order for the rational-expectations chance of the extremist coalition not to vanish as $N$ grows large. Even if the moderates significantly underestimate the chance of such a coalition, as long as they do so only by a constant factor, this will impact only the constant and not the rate at which extremist coalitions must vanish.

Now consider the Pareto tails case. A collusive group of size $M$ that is rationally expected to exist with (non-vanishing) probability $r$ will, by the logic in the text, buy votes relative to those

of population of moderates in

$$\Theta\left(\frac{MM^{\frac{\alpha-1}{\alpha}}N^{\frac{1}{\alpha}}}{r\mu N}\right).$$

But we know that these magnitudes must approximately match or the extremist group to be both effective and not wasteful. Thus, for $r$ not to vanish, we must have

$$M^{\frac{2\alpha-1}{\alpha}}\in\Omega\left(N^{\frac{\alpha-1}{\alpha}}\right)\implies M\in\Omega\left(N^{\frac{\alpha-1}{2\alpha-1}}\right).$$

$\square$

*Heuristic justification of Claim 4.* De-mergers are simpler to analyze as only the value of the most extreme individual is relevant. Repeating my analyses from above, the threshold for an extremist with $L$ identities who is not anticipated is $1/L$ of that for an extremist with a single identity. Thus the threshold is in $\Theta\left(N^{2/1+\alpha}L^{-1}\right)$. Such an individual exists with probability in

$$\Theta\left(N\cdot N^{-\frac{2\alpha}{1+\alpha}}L^{\alpha}\right)=\Theta\left(N^{-\frac{\alpha-1}{\alpha+1}}L^{\alpha}\right),$$

which dwindles with $N$ as long as $L\in O\left(N^{\alpha-1/\alpha(\alpha+1)}\right)$.

For the case of $\mu=0$ the typical size of the votes purchased by the fraudulent extremist is $LN^{\frac{1}{\alpha}}$, which causes inefficiency on the order of $L^2N^{\frac{4-\alpha}{2\alpha}}$ or, normalized by total welfare, $L^2N^{-\frac{\alpha-2}{\alpha}}$. Clearly this dies with $N$ as long as $L\in O\left(N^{-\frac{\alpha-2}{2\alpha}}\right)$.

Finally, a fraudulent extremist rationally expected to exist with non-vanishing probability $r$ will buy votes relative to those of moderate in

$$\Theta\left(\frac{LN^{\frac{1}{\alpha}}}{rN}\right),$$

so that if $L\in O\left(N^{\frac{\alpha-1}{\alpha}}\right)$, $r$ must dwindle to $0$ with $N$. $\square$

# B   Aggregate Uncertainty and Common Values

**Example 1.** Suppose that $u$ is equal to $\gamma$ plus normally distributed noise with standard deviation $\sigma_1^2$ and that $\gamma$ is normally distributed with mean $\mu$ and variance $\sigma_2^2$.

I assume, throughout and without loss of generality given the symmetry of the normal distribution, that $\mu>0$. The marginal distribution of $u$ is $\mathcal{N}\left(\mu,\sigma_1^2+\sigma_2^2\right)$ while the $\gamma$-conditional distribution is $\mathcal{N}\left(\gamma,\sigma_1^2\right)$ by standard properties of the normal distribution. I use this to solve out for $\gamma^\star$:

$$\int_u u\frac{f^2\left(u|\gamma^\star\right)}{f(u)}du=0\iff\int_u ue^{\frac{(u-\mu)^2}{2\left(\sigma_1^2+\sigma_2^2\right)}-\frac{(u-\gamma^\star)^2}{\sigma_1^2}}=0\iff\frac{(u-\mu)^2}{2\left(\sigma_1^2+\sigma_2^2\right)}-\frac{(u-\gamma^\star)^2}{\sigma_1^2}=au^2+b$$

for some constants $a$ and $b$ independent of $u$ as this is the only quadratic form symmetric about $0$, and symmetry about $0$ is clearly necessary to yield a $0$ expectation given the normal

form of the density.

$$\frac{(u-\mu)^2}{2\left(\sigma_1^2+\sigma_2^2\right)}-\frac{(u-\gamma^\star)^2}{\sigma_1^2}=\frac{\sigma_1^2(u-\mu)^2-2\left(\sigma_1^2+\sigma_2^2\right)(u-\gamma^\star)^2}{2\left(\sigma_1^2+\sigma_2^2\right)\sigma_1^2}=au^2+b-2\frac{\sigma_1^2\mu-2\left(\sigma_1^2+\sigma_2^2\right)\gamma^\star}{2\left(\sigma_1^2+\sigma_2^2\right)\sigma_1^2}u.$$

Thus $\gamma^\star$ solves

$$2\frac{\sigma_1^2\mu-2\left(\sigma_1^2+\sigma_2^2\right)\gamma^\star}{2\left(\sigma_1^2+\sigma_2^2\right)\sigma_1^2}=0\iff\gamma^\star=\frac{\sigma_1^2}{2\left(\sigma_1^2+\sigma_2^2\right)}\mu.$$

In a large population, the first-best welfare is proportional to $\mathbb{E}\left[|\gamma|\right]=\sigma_2\sqrt{\frac{2}{\pi}}e^{-\frac{\mu^2}{2\sigma_2^2}}+\mu\left[1-2\Phi\left(-\frac{\mu}{\sigma_2}\right)\right]$.
Welfare loss relative to this occurs in a large population when $\gamma\in(0,\gamma^\star)$ and, in these cases, is
proportional to $|\gamma|$. This loss equals

$$\int_0^{\frac{\sigma_1^2}{2\left(\sigma_1^2+\sigma_2^2\right)}\mu}\frac{\gamma e^{-\frac{(\gamma-\mu)^2}{2n\sigma_2^2}}}{\sigma_2\sqrt{2\pi}}d\gamma$$

which is clearly monotonically increasing in $\frac{\sigma_1^2}{2\left(\sigma_1^2+\sigma_2^2\right)}\mu$, which in turn monotonically increases
in $\sigma_1^2$. I can further compute analytically using Mathematica that

$$\int_0^{\frac{\sigma_1^2}{2\left(\sigma_1^2+\sigma_2^2\right)}\mu}\frac{\gamma e^{-\frac{(\gamma-\mu)^2}{2\sigma_2}}}{\sigma_2\sqrt{2\pi}}d\gamma=\mu\left[\Phi\left(\frac{\mu}{\sigma_2}\right)-\Phi\left(\frac{\mu\left(\sigma_1^2+2\sigma_2^2\right)}{2\sigma_2\left(\sigma_1^2+\sigma_2^2\right)}\right)\right]-\frac{\sigma_2\left(e^{-\frac{\mu^2\left(\sigma_1^2+2\sigma_2^2\right)^2}{8\sigma_2^2\left(\sigma_1^2+\sigma_2^2\right)^2}}-e^{-\frac{\mu^2}{2\sigma_2^2}}\right)}{\sqrt{2\pi}}.$$

Thus EI is

$$\frac{x\left[\Phi\left(x\right)-\Phi\left(\frac{x\left(\sigma_1^2+2\sigma_2^2\right)}{2\left(\sigma_1^2+\sigma_2^2\right)}\right)\right]-\frac{e^{-\frac{x^2\left(\sigma_1^2+2\sigma_2^2\right)^2}{8\left(\sigma_1^2+\sigma_2^2\right)^2}}-e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}}{2\left(\sqrt{\frac{2}{\pi}}e^{-\frac{x^2}{2}}+x\left[1-2\Phi\left(-x\right)\right]\right)}$$

where $x\equiv\frac{\mu}{\sigma_2}$. In the limit as $\sigma_1\to\infty$ this becomes

$$\frac{x\left[\Phi\left(x\right)-\Phi\left(\frac{x}{2}\right)\right]-\frac{e^{-\frac{x^2}{8}}-e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}}{2\left(\sqrt{\frac{2}{\pi}}e^{-\frac{x^2}{2}}+x\left[1-2\Phi\left(-x\right)\right]\right)}$$

and when $\sigma_1=\sigma_2$

$$\frac{x\left[\Phi\left(x\right)-\Phi\left(\frac{3x}{4}\right)\right]-\frac{e^{-\frac{9x^2}{32}}-e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}}{2\left(\sqrt{\frac{2}{\pi}}e^{-\frac{x^2}{2}}+x\left[1-2\Phi\left(-x\right)\right]\right)}.$$
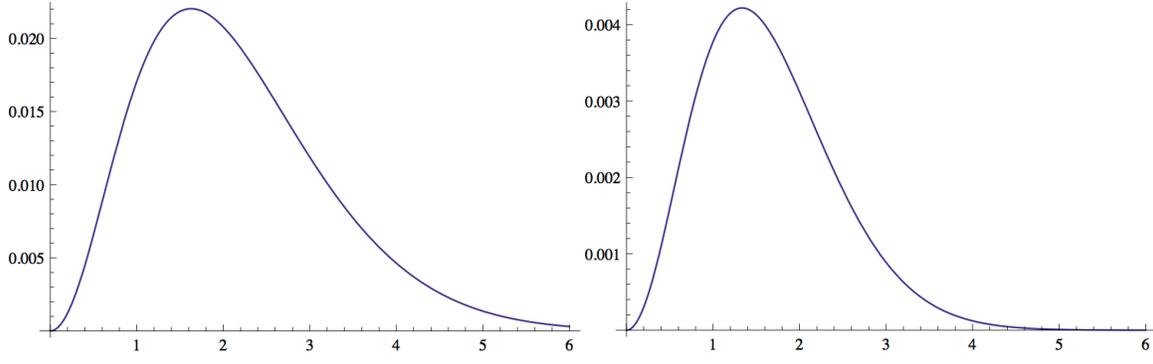
33

Figure 2: EI for the joint normal example when $\sigma_1 \to \infty$ (left) and $\sigma_1 = \sigma_2$ (right) as a function of $\frac{\mu}{\sigma_2}$.

Figure 2 shows the EI expression in both of these cases. Note that 1p1v is limit-efficient by symmetry (1p1v always chooses the sign of the median, which is the same as the sign of the mean) and QV is not. $\frac{\sigma_1^2}{2(\sigma_1^2+\sigma_2^2)}\mu = \gamma^\star > \gamma_0 = 0$. For large $N$ and a fixed $\sigma_2^2$ maximal limit-EI occurs as $\sigma_1^2 \to \infty$; globally maximal limit-EI occurs when $\frac{\mu}{\sigma_2} \approx \pm 1.6$ and equals approximately 2.2%. Typically it is much less; for example if $\sigma_1^2 \to \infty$ but $\frac{\mu}{\sigma_2}$ is less than 75 or greater than 3 inefficiency is below 1% and if $\sigma_1^2 = \sigma_2^2$ then inefficiency is always below .5%.

Because the normal distribution is symmetric, standard voting, which always selects the preference of the median voter, achieves perfect efficiency in this example, while QV is not perfectly efficient. However, even in the worst case, QV still achieves more than 97% efficiency; usually it does much better. A natural intuition is that as the noise of individual values, $\sigma_1^2$, grows large and thus value heterogeneity becomes important relative to the aggregate uncertainty, QV should become perfectly efficient. This turns out to be wrong: the greater $\sigma_1^2$ the lower efficiency, presumably because the combination of extreme values and clearly separated likelihood ratio leads to greater over-weighting of the underdogs.

I now consider an example based on a set-up proposed by Krishna and Morgan (2012) to study costly voting.

**Example 2.** Suppose that $\gamma$ is the fraction of individuals who have positive value, but that the distribution of the magnitude of value conditional on its sign is fixed and commonly known. Let $\mu_+$, $\mu_-$ be respectively the mean magnitude of values for those with positive and negative values respectively.

The average value conditional on $\gamma$ is $\gamma\mu_+ - (1-\gamma)\mu_-$ so that $\gamma_0 = \frac{\mu_-}{\mu_-+\mu_+}$. $f(u|\gamma) = \gamma$ for $u > 0$ and $f(u|\gamma) = 1 - \gamma$ for $u < 0$. As a result, $f(u) = \mathbb{E}[\gamma]$ for $u > 0$ and $1 - \mathbb{E}[u]$ for $u < 0$. Thus $\gamma^\star$ solves

$$\mu_+\frac{\gamma^2}{\mathbb{E}[\gamma]} - \mu_-\frac{(1-\gamma)^2}{1-\mathbb{E}(\gamma)} = 0 \implies \gamma^2 k = (1-\gamma)^2,$$

where $k \equiv \frac{\mu_+(1-\mathbb{E}[\gamma])}{\mu_-\mathbb{E}[\gamma]}$. Solving this quadratic equation yields

$$\gamma = \frac{-1 \pm \sqrt{k}}{k-1}.$$

34

The solution must be in the interval $[0,1]$ , which the negative solution never is and the positive solution always is. Thus

$$\gamma^\star = \frac{\sqrt{k}-1}{k-1} = \frac{1}{\sqrt{k}+1}.$$

Efficiency results if and only if $\gamma_0 = \gamma^\star$, that is if

$$\frac{1}{1+\sqrt{k}} = \frac{\mu_-}{\mu_- + \mu_+} \iff \sqrt{k} = \frac{\mu_+}{\mu_-} \iff \frac{\mu_+^2}{\mu_-^2} = \frac{\mu_+\,(1-E[\gamma])}{\mu_- E[\gamma]} \iff \mu_+ E[\gamma] = \mu_-\,(1-E[\gamma]),$$

hat is the election is an expected welfare tie ex-ante.

$$\frac{\mu_+}{\mu_-\sqrt{k}} = \sqrt{\frac{\mu_+ E[\gamma]}{\mu_-\,(1-E[\gamma])}}$$

so that $\frac{\mu_\pm}{\mu_-} > (<)\sqrt{k} \iff \mu_+ E[\gamma] > (<)\mu_-\,(1-E[\gamma])$. Thus

$$\gamma_0 < (>)\gamma^\star \iff \mu_+ E[\gamma] > (<)\mu_-\,(1-E[\gamma]),$$

the BUE.

Under 1p1v the threshold in $\gamma$ for implementing the alternative with high probability is $1/2$. For each regime I compute EI as

$$1 - \frac{\int_0^{\gamma_t} [\mu_-\,(1-\gamma) - \mu_+\gamma]\,h(\gamma)d\gamma + \int_{\gamma_t}^1 [\mu_+\gamma - \mu_-\,(1-\gamma)]\,h(\gamma)d\gamma}{\int_0^{\gamma_0} [\mu_-\,(1-\gamma) - \mu_+\gamma]\,h(\gamma)d\gamma + \int_{\gamma_0}^1 [\mu_+\gamma - \mu_-\,(1-\gamma)]\,h(\gamma)d\gamma},$$

where $\gamma_t$ is the appropriate threshold value f $\gamma$. Using this method and explicit integration on Mathematica, I computed the relative (to the first best) efficiency of QV and 1p1v assuming $g$ follows a Beta distribution. Note that, if one divides the numerator and denominator by $\mu_-$,

$$\frac{\int_0^{\gamma_t} [\mu_-\,(1-\gamma) - \mu_+\gamma]\,h(\gamma)d\gamma + \int_{\gamma_t}^1 [\mu_+\gamma - \mu_-\,(1-\gamma)]\,h(\gamma)d\gamma}{\int_0^{\gamma_0} [\mu_-\,(1-\gamma) - \mu_+\gamma]\,h(\gamma)d\gamma + \int_{\gamma_0}^1 [\mu_+\gamma - \mu_-\,(1-\gamma)]\,h(\gamma)d\gamma} =$$

$$\frac{\int_0^{\gamma_t} \left[(1-\gamma) - \frac{\mu_\pm}{\mu_-}\gamma\right]\,h(\gamma)d\gamma + \int_{\gamma_t}^1 \left[\frac{\mu_\pm}{\mu_-}\gamma - (1-\gamma)\right]\,h(\gamma)d\gamma}{\int_0^{\gamma_0} \left[(1-\gamma) - \frac{\mu_\pm}{\mu_-}\gamma\right]\,h(\gamma)d\gamma + \int_{\gamma_0}^1 \left[\frac{\mu_\pm}{\mu_-}\gamma - (1-\gamma)\right]\,h(\gamma)d\gamma},$$

So EI depends only on the ratio $r \equiv \frac{\mu_+}{\mu_-}$ and on the parameters of the Beta distribution, not on both $\mu_+$ and $\mu_-$ independently

Figure 3 shows three examples that are representative of the more than 100 cases I experimented with. Whenever $\alpha = \beta$ (the distribution of $\gamma$ is symmetric), QV always dominates 1p1v as it does in the left panel shown, which is $\alpha = \beta = 1$, the uniform distribution. 1p1v obviously performs best when $r$, shown on the horizontal axis, is near to unity. When $\alpha$ is larger than $\beta$, 1p1v may out-perform QV near $r = 1$. This is shown in the center and right panels where $(\alpha, \beta) = (15, 10)$ and $(10, 1)$ respectively. The larger $\alpha$ is relative to $\beta$, the larger the region over which voting outperforms QV. However, it is precisely in these cases where, if $r$ is very small,
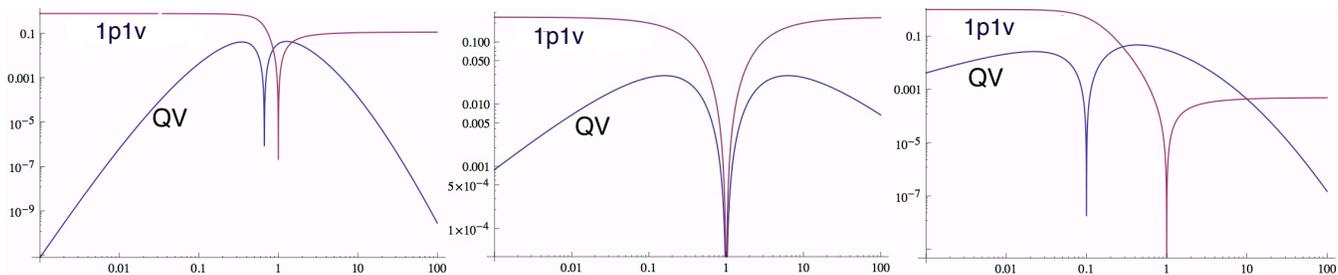
Figure 3: EI of QV compared to 1p1v in the Krishna and Morgan (2012) example with $\gamma$ distributed uniform (left), $\beta(15,10)$ (center) and $\beta(10,1)$ (right). Both axes are on a log-scale, though labeled linearly; the x-axis, measures $r = \frac{\mu_+}{\mu_-}$.

voting is most dramatically inefficient. Intuitively majority may outperform QV by blindly favoring the majority which is almost always in favor of taking the action for $\alpha \gg \beta$, while QV may be a bit too conservative in favoring the action because of the BUE. However this blind favoritism towards the majority view can be highly destructive under 1p1v, but not under QV, when the minority has an intense preference. In fact, while voting becomes highly inefficient when the minority preference becomes very intense, QV actually becomes closer to the first best. In all cases (shown here and that I have sampled) QV's efficiency is above 95% and usually it is well above this.

In summary EI is never greater than 5% for QV and it can be arbitrarily large for majority-rules voting. In the special case when $\gamma$ has a uniform distribution, QV dominates voting, which may have EI as high as 25% while for QV it is never greater than 3%. For "most" parameter ranges QV appears to outperform majority-rules voting, often quite significantly.

Finally I consider an example similar to the previous one but calibrated to the evaluation of Proposition 8 in California I discuss in Weyl (2015).

**Example 3.** Suppose that 4% of the electorate is commonly known to oppose the alternative and is willing to pay on average \$34k to see it defeated. Suppose that the other 96% of the electorate is willing to pay on average \$5k to either support or oppose the alternative with the intensity of their values being independent of $\gamma$. $\gamma$ is the fraction of the 96% that support the alternative and is assumed to have a Beta distribution with parameters set so that on the mean fraction of the population in favor of the alternative is 52%. The average value from implementing the alternative is therefore \$4800(2\gamma - 1) - 1360$ and $\gamma_0 = .64$.

Individuals in the 4% receive no signal about $\gamma$ and thus $f(u|\gamma)$ for this group is simply $f(u)$. For proponents of the alternative among the 96%, $f(u|\gamma)$ is, by the logic of the previous proof, $.96f(u)\gamma$ and for opponents among the 96% $.96f(u)(1-\gamma)$; $f(u)$ is formed by taking expectations over $\gamma$ as in the previous proof. Using the same techniques derivations as there I can solve for $\gamma^\star$.

To calibrate, I assume that a Beta distribution of $\gamma$ and that

$$.96\frac{\alpha}{\alpha + \beta} = .96\mathbb{E}[\gamma] = .52.$$

Solving this out implies that $\alpha = 1.18\beta$. The variance of $\gamma$ is given by the standard formula for the variance of a Beta distributed variable:
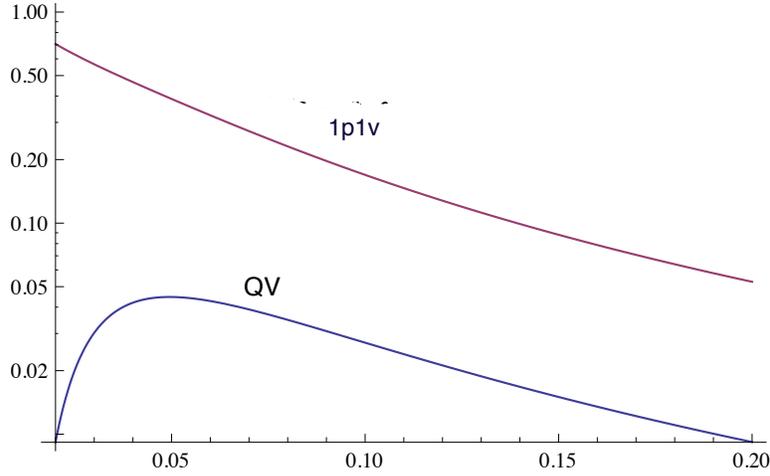
Figure 4: EI as a function of the standard deviation of the total vote share in favor of Proposition 8. The vertical axis is on a logarithmic scale with linear labels.

$$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = \frac{1.18\beta^2}{2.18^2\beta^2(2.18\beta+1)} = \frac{.25}{2.18\beta+1}.$$

Thus the standard deviation of the total fraction of the population supporting the alternative is $\frac{.96\cdot.5}{\sqrt{2.18\beta+1}} = \frac{.48}{\sqrt{2.18\beta+1}}$ and thus if the standard deviation of the vote share for the alternative under standard voting is $\sigma$ then $\beta = \frac{.45(.23-\sigma^2)}{\sigma^2}$. Figure 4 shows the dominance of QV.

Then QV is thus always superior to 1p1v and the gap is larger the smaller is the standard deviation of the vote share. When the standard deviation of the population share supporting the alternative is 5 percentage points (well above the margin of error in most individual polls), QV has 4% EI and 1p1v 47% EI. Even when the standard deviation is 20 percentage points QV achieves 1.2% EI while 1p1v has 7%.

**Example 4.** Suppose that each individual's value $v_i = \mu + \epsilon_i$ where $\mu$ is a common value component and $\epsilon_i$ is the individual's idiosyncratic preference. $\mu$ and $\epsilon_i$ are drawn identically and independently (in the latter case across individuals) from a distribution that equals $-1$ with .5 probability and 1 with .5 probability. Individuals receive signals $s_i$ that are drawn independently and identically conditional on $\mu$, taking on the same value as $\mu$ with probability $p \in (.5, 1)$ and $-\mu$ with probability $1 - p$.

Suppose that (something close to) full information aggregation occurs in 1p1v. Then the decision in a large population must have the same sign as $\mu$: if $\mu$ is negative the alternative is not adopted and if $\mu$ is positive the alternative is adopted. Thus the only event where an individual may be pivotal is when $\mu$ is very close to 0. Whenever this is the case, however, every individual will want to vote purely based on her preference as this makes a bigger difference to the individual than does $\mu$. Thus full information aggregation under voting is impossible. The reason is that there are not any individuals with preferences sufficiently close to 0 so that they will still want to vote based on their information even if aggregation approximately occurs.

Now consider QV. I conjecture an approximate equilibrium where every individual votes $v_i = as_i + k\epsilon_i$. I now verify if an equilibrium of this sort is possible. If this holds, then $V_{-i} =$

37

$a(N-1)\mu + a\sum_{j\neq i}s_i - \mu + k\sum_{j\neq i}\epsilon_i$. For large $N$ this sum is distributed, conditional on $\mu$, as a normal with mean $a(N-1)\mu$ and variance $a^2\sigma_s^2(N-1) + k^2(N-1)$. Thus $\frac{V_{-i}}{a(N-1)}$ is distributed, conditional on $\mu$, normally with mean $\mu$ and variance $\frac{\sigma_s^2 + \frac{1}{\alpha^2}}{N-1}$ where $\alpha \equiv \frac{a}{k}$. Thus $V_{-i}$ is a signal with precision $\frac{N-1}{\sigma_s^2 + \frac{1}{\alpha^2}}$ of $\mu$. Thus $i$ best estimate, by the standard normal updating formula, of $\mu$ is

$$\frac{\frac{N-1}{\sigma_s^2+\frac{1}{\alpha^2}}}{\frac{N-1}{\sigma_s^2+\frac{1}{\alpha^2}} + \frac{1}{\sigma_\mu^2} + \frac{1}{\sigma_s^2}}V_{-i} + \frac{\frac{1}{\sigma_s^2}}{\frac{N-1}{\sigma_s^2+\frac{1}{\alpha^2}} + \frac{1}{\sigma_\mu^2} + \frac{1}{\sigma_s^2}}s_i + \frac{\frac{1}{\sigma_\mu^2}}{\frac{N-1}{\sigma_s^2+\frac{1}{\alpha^2}} + \frac{1}{\sigma_\mu^2} + \frac{1}{\sigma_s^2}}0.$$

Note that this is affine in $s_i$ and $V_{-i}$ so that individual $i$'s best estimate of $u_i$ for the purposes of choosing $v_i$, given that $V_{-i}$ is already incorporated into the sum of votes, is just a constant plus

$$\frac{\frac{1}{\sigma_s^2}}{\frac{N-1}{\sigma_s^2+\frac{1}{\alpha^2}} + \frac{1}{\sigma_\mu^2} + \frac{1}{\sigma_s^2}}s_i + \epsilon_i.$$

Crucially this is true regardless of the value of $V_{-i}$, including conditioning on the individual being pivotal. Thus the ratio of the coefficient on $s_i$ to that on $\epsilon_i$, which must equal $\alpha$ in a large market equilibrium given that individuals buy votes proportional to their utility, is $\frac{\frac{1}{\sigma_s^2}}{\frac{N-1}{\sigma_s^2+\frac{1}{\alpha^2}} + \frac{1}{\sigma_\mu^2} + \frac{1}{\sigma_s^2}}$.
This lets us determine the value of $\alpha$ according to the following equation:

$$\alpha = \frac{\frac{1}{\sigma_s^2}}{\frac{N-1}{\sigma_s^2+\frac{1}{\alpha^2}} + \frac{1}{\sigma_\mu^2} + \frac{1}{\sigma_s^2}}.$$

As $N$ grows large, the denominator must shrink, leading $\alpha$ to shrink, unless $\alpha$ shrinks. Thus I conclude that $\alpha$ must approach zero for large $N$. This allows us to eliminate a bunch for terms and simplify the above equation to

$$\alpha \approx \frac{\frac{1}{\sigma_s^2}}{(N-1)\alpha^2 + \frac{1}{\sigma_\mu^2} + \frac{1}{\sigma_s^2}}.$$

But because $\alpha$ shrinks, the term with $N$ in the denominator must dominate so this simplifies further to

$$\alpha^3 \approx \frac{1}{\sigma_s^2(N-1)} \implies \alpha \approx \frac{1}{\sigma_s^{\frac{2}{3}}\sqrt[3]{N-1}}.$$

Thus $\frac{v_i}{k}$ has, conditional on $\mu$, mean $\frac{1}{\sigma_s^{\frac{2}{3}}\sqrt[3]{N-1}}\mu$ and variance $\frac{\sigma_s^2}{\sigma_s^{\frac{2}{3}}\sqrt[3]{N-1}}+1$ and $\frac{V}{k}$ has mean $\frac{N}{N-1}\left(\frac{N-1}{\sigma_s}\right)^{\frac{2}{3}}\mu$ and variance, for large $N$ approximately, $N$ and standard deviation $\sqrt{N}$. Thus the mean grows relative to the standard deviation in $N$, implying that with probability approaching 1 for large $N$ the decision is made in the direction of $\mu$ so that information is aggregated.

To finish, I need to show that the variance of total votes grows with $N$ so that the approximation that everyone buys nearly linear votes is correct. Conditional on any $s_i$ the variance of $\mu$ is

$\frac{1}{\frac{1}{\sigma_\mu^2}+\frac{1}{\sigma_s^2}}$ and its mean is $\frac{\sigma_\mu^2}{\sigma_\mu^2+\sigma_s^2}s_i$.

$$V_{-i} = k \left[ \alpha(N-1)\mu + \alpha \sum_{j\neq i} s_j - \mu + \sum_{j\neq i} \epsilon_j \right]$$

so the mean of $V_{-i}$ conditional on $s_i$ is $\frac{k\alpha(N-1)\sigma_\mu^2}{\sigma_\mu^2+\sigma_s^2}s_i$ and its conditional variance is $k^2\alpha^2(N-1)^2\sigma_\mu^2 + (N-1)(\sigma_s^2+1)$. Under my assumption that $\sigma_s^2 \gg \sigma_\mu^2$ the mean is approximately $0$ for all but very unlikely and extreme signals. Furthermore for large $N$ the leading term in the expression for the variant is $k^2\alpha^2(N-1)^2\sigma_\mu^2 = k^2\sigma_\mu^2\left[\frac{N-1}{\sigma_s^2}\right]^{\frac{4}{3}}$. Thus the density of a tie is approximately

$$\frac{\sigma_s^{\frac{8}{3}}}{\sqrt{2\pi}k\sigma_\mu^2(N-1)^{\frac{2}{3}}}.$$

This yields a matching-coefficients equation

$$k^2 = \frac{\sigma_s^{\frac{8}{3}}}{2\sqrt{2\pi}\sigma_\mu^2(N-1)^{\frac{2}{3}}} \implies k = \frac{\sigma_s^{\frac{4}{3}}}{2^{\frac{3}{4}}\sqrt[4]{\pi}\sigma_\mu\sqrt[3]{N-1}}.$$

Thus the variance of votes grows with $(N-1)^{\frac{2}{3}}$ and for large $N$ no individual will try to buy the whole election.

Thus as $N \to \infty$ in any QV equilibrium the alternative is implemented if and only if $\mu = 1$, which is efficient. Under 1p1v the alternative is implemented with probability $.5$ regardless of $\mu$.

Why does QV outperform 1p1v here? The reason is that voting only allows individuals to express a binary directional preference. If most (here all) individuals care more about their preferences than their information, information will play little or no role in the social decision. On the other hand, under QV, information can be expressed through a more subtle shading of vote purchases. While on an individual level this expression is small, it cumulates throughout the market allowing information to be aggregated.

The assumption that $\sigma_\mu^2$ is small was useful here as it allowed us to find a solution where vote purchases were linear in signals because the chance of a tie is the same for all types. When $\sigma_\mu^2$ is not small this would fail and those with extreme signals would buy fewer votes than those with signals near $0$. This is closely related to the case of aggregate uncertainty we analyzed in the rest of this appendix. Like there, I are sure this would cause some slowing of information aggregation or even limit inefficiency. But I suspect that it would also leave the basic qualitative conclusions and comparisons to 1p1v in tact.

# C   Voter Behavior

*Heuristic justification of Claim 6 .* In the expressivist model, clearly unless $x = 0$ the expressive motive will become dominant in large populations as the number of votes each individual purchases dies off with $N$, while the expressive motive does not. Even when $x = 0$ note that if

individuals purchase votes primarily from the expressive motive, they will buy votes of magnitude approximately (up to unimportant fluctuations in $\epsilon$ given that the mean is not 0) $\xi/vu$ and thus $V \approx \frac{N\xi\mu}{V}$ by the law of large numbers. Thus $V$ is in $\Theta\left(\sqrt{N}\right)$. Thus only individuals with utility on the order of $N$ will be willing to act as extremists. Thus, by the logic of Subappendix **??** the chance of an extremist existing must die at rate $N^{-(\alpha-1)}$. Furthermore the price of influence $\hat{p}$ perceived by an extremist must satisfy her first-order condition for her value, which will be in $\Theta(N)$ and the votes she buys which are in $\Theta\left(\sqrt{N}\right)$

$$\Theta\left(\frac{N}{\hat{p}}\right) = \Theta\left(\sqrt{N}\right) \implies \Theta\left(\hat{p}\right) = \Theta\left(\sqrt{N}\right).$$

Thus the price perceived by moderates is in $\Theta\left(N^{\alpha-.5}\right)$. This implies that the instrumental motive for buying votes must die be in $O\left(1/N^{.5+\alpha-1}\right)$, that is more quickly than does the expressive motive. So the expressive motive's dominance is confirmed. A reverse calculation, based on the assumption of dominance of the instrumental motive, leads to the same conclusion.

Thus in either case of the expressivist model the expressive motive is dominant for large $N$. Thus, up to the error arising from $\epsilon$, $U$ and $V$ will be proportional to each other in large populations. For $\epsilon$ to make a difference, given that $\mu > 0$, it must create a large deviation or be so close to zero for a significant individual that it cancels out a large deviation driven by that individual's value as in every other event $U$ and $V$ are both positive and efficiency results. However, by my assumption that $\epsilon$ is either thin-tailed or has a tail index weakly greater than that of $u$, we have that the probability of such a large deviation in either $U$ or $V$ is in $O\left(N^{-(\alpha-1)}\right)$, because the tail index of $u\epsilon$ is the minimum of the tail index of $u$ and $\epsilon$ given their independence. Thus a superset of the events causing inefficiency has probability in $O\left(N^{\alpha-1}\right)$ and thus the probability of inefficiency must also be in $O\left(N^{\alpha-1}\right)$.

In the misperception model clearly if overestimation of the chance of being pivotal is severe enough for very small values of $g$ the same logic as in the expressivist model will hold. Suppose, instead, that as $g \to 0$, $e(g) \in \Theta\left(\sqrt[\beta]{g}\right)$ for some $g > 1$ so that the over-estimation is greater than a constant fraction, but not extremely rapid. Under this assumption, I can trace through the logic of my fat-tailed calculations in Weyl (2015) to derive limiting inefficiency. Following the logic there, the price perceived by a rational moderate must now be not in $\Theta\left(N/\sqrt{|u^\star|}\right)$ but instead in $\Theta\left(N^\beta/(|u^\star|)^{\beta/2}\right)$, where $u^\star$ is the threshold for being such an extremist, as this will be seen as $\Theta\left(N/\sqrt{|u^\star|}\right)$ by a misperceiving moderate. Thus the (assumed-rational given that she has significant impact, though this assumption can be dispensed) extremists must perceive a price in $\Theta\left(qN^\beta/(|u^\star|)^{\beta/2}\right)$, where $q$ is the probability an extremist (actually) exists. By extremist rationality, her first-order condition and my arguments about integrating over all ponytail extremists through a power law gives that

$$\Theta\left(\sqrt{|u^\star|}\right) = \Theta\left(\hat{p}\right),$$

where $\hat{p}$ is the price perceived by the average extremist. Combining this with my observation above gives that

$$\Theta\left(\sqrt{|u^\star|}\right) = \Theta\left(\frac{qN^\beta}{(|u^\star|)^{\frac{\beta}{2}}}\right) \implies \Theta\left(q\right) = \Theta\left(\frac{(|u^\star|)^{\frac{1+\beta}{2}}}{N^\beta}\right).$$

I can then set this modified upwards sloping equation equal to the standard downward sloping equation (Weyl, 2015), given this involves no misperceptions.

$$\Theta\left(N\left(|u^\star|\right)^{-\alpha}\right) = \Theta\left(\frac{\left(|u^\star|\right)^{\frac{1+\beta}{2}}}{N^\beta}\right) \implies \Theta\left(N^{1+\beta}\right) = \Theta\left(\left(|u^\star|\right)^{\frac{1+\beta+2\alpha}{2}}\right) \implies \Theta\left(|u^\star|\right) = \Theta\left(N^{\frac{2+2\beta}{1+\beta+2\alpha}}\right).$$

Again applying the demand equation

$$\Theta(q) = \Theta\left(N N^{\frac{-2\alpha(1+\beta)}{1+\beta+2\alpha}}\right) = \Theta\left(N^{-\frac{2\alpha\beta-1-\beta}{1+\beta+2\alpha}}\right).$$

Whether the inefficiency created by extremism or that created by the large deviations relevant in the expressivist model clearly turns on whether the extremist event predicted by this analysis is asymptotically larger or smaller than the large deviations event. □