Quadratic Voting *

Steven P. Lalley[†] E. Glen Weyl[‡]

December 2015

Abstract

N individuals must choose between two collective alternatives. Under Quadratic Voting (QV), individuals buy vote in favor of their preferred alternative from a clearing house, paying the square of the number of votes purchased, and the sum of all votes purchased determines the outcome. Heuristic arguments and experimental results have suggested that this simple, detail-free mechanism is utilitarian efficient. In an independent private-values environment, we rigorously prove that for any value distribution all symmetric Bayes-Nash equilibria of QV converge toward efficiency in large populations, with waste decaying generically as $^1/_N$.

Keywords: social choice, collective decisions, large markets, costly voting, vote trading

*This paper supersedes the authors' previous joint paper "Nash Equilibria for a Quadratic Voting Game", which contained proofs of the main results of this paper, and a previous working paper by Weyl, "Quadratic Vote Buying", which conjectured the form of the Nash equilibria established here. Weyl is co-founder of a commercial venture, Collective Decision Engines, which is commercializing Quadratic Voting for market research and thus has a financial interest in the success of this mechanism. We are grateful to Nageeb Ali, Bharat Chandar, Jerry Green and Bruno Strulovici, as well as many other colleagues and seminar participations for helpful comments. We acknowledge he financial support of the National Science Foundation Grant DMS - 1106669 received by Lalley and of the Alfred P. Sloan Foundation, the Institut D'Économie Industrielle and the Social Sciences Division at the University of Chicago received by Weyl. Kevin Qian, Tim Rudnicki, Matt Solomon and Daichi Ueda supplied excellent research assistance. We owe a special debt of gratitude to Lars Hansen, who suggested our collaboration, and to Eric Maskin for an excellent formal discussion of the paper. All errors are our own.

[†]Department of Statistics, University of Chicago: 5734 S. University Avenue, Chicago, IL 60637; lalley@galton.uchicago.edu, http://www.stat.uchicago.edu/~lalley/.

[‡]Microsoft Research New England, One Memorial Drive, Cambridge, MA 02142 and Department of Economics, University of Chicago: 1126 E. 59th Street, Chicago, IL 60637; glenweyl@microsoft.com, http://www.glenweyl.com.

1 Introduction

Consider a binary collective-decision problem in which a group of N individuals must choose between two alternatives. Each individual has a privately known value that determines her willingness to pay for one alternative over the other. *Quadratic Voting* (*QV*) is a simple and detail-free mechanism designed to maximize utilitarian efficiency in this setting.[1] In this system, individuals buy votes (either negative or positive, depending on which alternative is favored) from a clearing house, paying the square of the number of votes purchased. The sum of all votes purchased then determines the outcome according to a smoothed version of majority rule described more precisely in Subsection 2.1 below. We prove that, in any symmetric Bayes-Nash equilibrium of a private values environment where valuations are drawn independently and identically according to any smooth value distribution with bounded support, the inefficiency of QV converges as $N \to \infty$ to a fraction 0 of potential welfare, at a rate $1/N$ for generic value distribution parameters. Our analysis is constructive and describes the structure of Bayes-Nash equilibria.

The heuristic rationale for QV is quite simple. The marginal benefit to a voter of an additional vote is her value multiplied by her *marginal pivotality* (roughly, the perceived probability that an additional unit of vote will sway the decision). She maximizes utility by equating this marginal utility to the linear marginal cost of a vote. Therefore, if voters share the same marginal pivotality, they will buy votes in proportion to their values, thus bringing about utilitarian efficiency. Furthermore, the quadratic cost function is the *unique* cost function with this property. The mathematical details behind this argument are given in Subsection 2.3.

Previous research has used variations of this rationale to justify quadratic mechanisms in related collective-decision-making problems discussed below. However, to our knowledge, this heuristic rationale has never been translated into a rigorous argument for efficiency in the sort of non-cooperative, incomplete information game theoretic model in which mechanisms for allocating private goods have been studied at least since the work of Myerson (1981). Although, as we discuss below, one could attempt to make this intuition precise in a variety of potential models, we choose to begin with the simple and canonical setting described above, analogous to the Myerson's auction model. Unfortunately, in this setting, the crucial assumption of this rationale – that in equilibrium all voters will have the same marginal pivotality – is false, as we will show. As a result, formal equilibrium analysis is a far more subtle task than the heuristic argument of the previous paragraph. Nevertheless, we will show that for voters with values in the "bulk" of the distribution, the marginal pivotality is approximately constant,[2] which will

[1] Clearly many other objectives are possible for this problem, and many involve distributional considerations. However, we focus on a utilitarian objective because it is the one most extensively studied in the literature (Bowen, 1943; Groves, 1973).

[2] Theorems of Kahn et al. (1988) and Al-Najjar and Smorodinsky (2000) imply marginal pivotality must converge

allow us to prove convergence to efficiency.

In nearly all previous results on convergence of mechanisms (for private goods) towards efficiency, agents' strategies converge to a simple, determinate continuum equilibrium strategy, typically involving truthful direct demand revelation.[3] Because marginal pivotalities converge to zero in the binary collective decision problem, there is no well-defined continuum analog. In fact, as we will show, when the value distribution has non-zero mean equilibria take an exotic form, with nearly all individuals buying vanishingly few votes but a vanishingly small number of "extremists" buying a relatively large number, enough to single-handedly sway the election. Large$-N$ efficiency emerges because the expected number of extremists in the population decays as $1/N$. By contrast, when the mean value is zero, such extremists will not exist. Here efficiency will depend another subtle property of our equilibrium characterization. Specifically we show that population variability of marginal pivotalities is sufficiently small that deviations in the vote total stemming from nonlinearities of the equilibrium voting strategy vanish compared to variations arising from the sampling variability of the agents' values.

To our knowledge the use of quadratic pricing for collective decision-making was first suggested by Groves and Ledyard (1977a), who proposed it as a Nash implementation of the optimal level of continuous public goods under complete information that avoids the fragility, discussed in the next paragraph, of previously suggested efficient mechanisms.[4] In an unpublished manuscript, Hylland and Zeckhauser (1980) provided the first variant of the heuristic rationale above to uniquely justify quadratic pricing mechanism and proposed an iterative procedure that they conjectured would converge to Groves and Ledyard's complete information optimum in the presence of private information. In a previous version of this paper, Weyl (2012) first proposed the use of QV for binary collective decision problems, and conjectured that it would lead to asymptotically efficient decisions in the environment we consider based on an (independently discovered) extension of Hylland and Zeckhauser's heuristic rationale. Goeree

to zero as $N \to \infty$. Our results will show that, nonetheless with probability approaching 1, the ratio between the marginal pivotalities of two randomly chosen voters will be close to 1.

[3]For example see Roberts and Postlewaite (1976), Rustichini et al. (1994), and Che and Kojima (2010). The only counter-example we are aware of is the work of Ledyard (1984), and Myerson (2000) and Krishna and Morgan (2015) who study the same environment, on costly voting. In that model, however, any individual can make only three choices (vote in either direction or stay home), and thus equilibrium construction is relatively simple. Thus, although the heuristic rationale for QV is similar to that put forward by Ledyard (with the uniform distribution of voting costs independent of values he considers, the cost of a turnout is effectively quadratic in the aggregate), the equilibrium construction with which we are primarily concerned here bears little relationship to that in costly voting models.

[4]Another common criticism of these mechanisms is that they violate individual rationality constraints if individuals are entitled to the status quo outcome. Mailath and Postlewaite (1990) show that, in this environment and if entitlements to one alternative (call it the status quo) is provided to all agents, then as the population grows large as long as any event exists in which the status quo is optimal, the probability of the alternative being adopted tends toward 0. This weakness applies equally to QV as to any of the other mechanisms for this setting. As a result we consider an environment in which the alternatives are symmetric and no individual has an entitlement to any alternative. In this environment individual rationality does not arise as a constraint.

and Zhang (2015) independently suggested using a detail-based, approximately direct variant of QV in the special case where values are sampled from zero-mean normal distributions, and derived an equilibrium in the case $N = 2$.

The subject of this paper is primarily the analytical results quoted above, and we confine speculation on practical applications of QV to a brief motivational Section 4 . However, note that QV is far from the first (asymptotically) efficient mechanism economists have proposed for the binary collective-decision problem. Nonetheless, unlike most other such procedures, such as the expected externality (EE) mechanism of Arrow (1979) and d'Aspremont and Gérard-Varet (1979), the "implement-the-mean preference" mechanism of Krishna and Morgan (2001) and the "costly voting" (CV) mechanism Ledyard (1984), Myerson (2000) and Krishna and Morgan (2015), its description and equilibrium efficiency properties are independent of the distribution of values. It is also precisely budget balanced, unlike the only other detail-free (asymptotically) efficient mechanism of which we are aware, namely, that described by Vickrey (1961), Clarke (1971) and Groves (1973) (VCG). Furthermore, the significant discussion of QV in the past few years in legal (Cariello, 2015) and popular (Ellenberg, 2015) circles, compared to the paucity of such discussions in the decades since Green et al. (1976) and Tideman and Tullock (1976) proposed using VCG for collective decision-making, suggests that QV's rules are simpler to describe, at least to non-economists.

Perhaps a more important question about the attractiveness of QV for practical applications is its robustness to a range of deviations from the baseline model we focus. Some such deviations are known to be crippling for existing (asymptotically) efficient mechanisms. The EE and "implement-the-mean" mechanisms are not defined if aggregate uncertainty exists concerning the value distribution, whereas QV is always defined independent of such distributions. CV is typically inefficient when individuals have non-instrumental or imperfectly-rational motivations to vote, whereas the heuristic rationale for QV relies only on individuals sharing a (on average) common proportional-to-value (but not necessarily rational or instrumental) interest in voting. VCG is extremely sensitive to collusion (Groves and Ledyard, 1977b; Ausubel and Milgrom, 2005) (any two individuals may attain any desired outcome at no cost if they collude), whereas QV appears to require at least some payment by collusive groups. Formal analysis of these robustness concerns is beyond the scope of this paper, given that treating the technical challenges of characterizing equilibria in any of these environments rigorously would require more space than a standard journal article. However, they are active subjects of our on-going research, as we discuss briefly in our conclusion.

2 Model Assumptions and Rationale

2.1 Model

We consider an independent symmetric private values environment with N voters $i = 1, \ldots, N$. Each voter i is characterized by a value, u_i; these values are drawn independently and identically from a continuous probability distribution F supported by a finite interval $[\underline{u}, \overline{u}]$, with associated density f and $\underline{u} < 0 < \overline{u}$. For normalization, we assume the numeraire has been scaled so that $\min(|\underline{u}|, \overline{u}) \geq 1$. We denote by μ, σ^2, and μ_3, respectively, respectively the mean, variance, and raw third moments of u under F, and we assume f is smooth and bounded away from 0 on $[\underline{u}, \overline{u}]$. The two alternatives are 0 and 1. Individuals are risk-neutral, quasi-linear expected utility maximizers who gain $2u_i$ dollars of utility if option 1 is adopted rather than option 0.

We consider a variant on QV[5] where

1. Each individual chooses a number of votes $v_i \in \mathbb{R}$ to buy.
2. Each pays v_i^2 dollars and receives a refund of $\frac{\sum_{j \neq i} v_j^2}{N-1}$ dollars.[6]
3. The outcome is option 1 with probability $\frac{\Psi(V)+1}{2}$ where $V \equiv \sum_i v_i$ and $\Psi : \mathbb{R} \to [-1, 1]$ is an odd, nondecreasing, C^∞ function[7] such that

   (a) for some $0 < \delta < \infty$,
   $$\Psi(x) = \text{sgn}(x) \quad \text{for all } \ |x| \geq \delta;$$

   (b) Ψ has positive derivative[8] ψ, on the interval $(-\delta, \delta)$; and

   (c) $\psi'(x) > 0$ for all $-\delta < x < 0$ and ψ has a single point of inflection in $(-\delta, 0)$.

[5]Although both the discrete binary choice set-up of Weyl (2012) and the continuous public goods model of Hylland and Zeckhauser (1980) helped inspire this model, it differs from both. Consequently, our results have no direct implications for those models. It differs from Weyl's model in that the outcome is smoothed rather than jumping discontinuously at 0. It differs in a variety of respects from Hylland and Zeckhauser's, notably in that utility is linear in the common and bounded outcome, whereas Hylland and Zeckhauser assume strictly concave preferences with heterogeneous ideal points and an outcome that may take values in the full real space. Hylland and Zeckhauser also consider a multidimensional issue space with no access to transfers and an iterative procedure to converge to this outcome, none of which feature in our model.

[6]This particular refund rule is for concreteness and budget balance only, and plays no role in our results. A wide range of refund rules would work equally well.

[7]The payoff function Ψ smooths QV so that close decisions are decided probabilistically, rather than deterministically. While we model it as a feature of the mechanism, it can also be viewed as a product of noise in the environment, possibly arising from a small number of exogenous noise voters or a small, fixed error in the vote-tallying process.

[8]Thus, $\psi/2$ is an even, C^∞ probability density with support $[-\delta, \delta]$ that is strictly positive on $(-\delta, \delta)$. There are infinitely many such probability densities. For example, if $X_1, X_2, \ldots$ is a sequence of independent, identically distributed random variables each with the uniform density on $[-\delta, \delta]$, then the random variable $Y = \sum_{n=1}^{\infty} X_n / 2^n$ has a C^∞ density that meets the requirements (a), (b), (c) in condition 3.

We shall refer to Ψ as the *payoff function*, because it determines the quantity by which values u_i are multiplied to obtain the allocative (as opposed to transfer) component of each individual's utility.

Conditional on the values of $\{v_i\}$, each individual i earns expected utility

$$u_i \Psi(V) - v_i^2 + \frac{1}{N-1} \sum_{j \neq i} v_j^2. \tag{1}$$

Because the last term in this expression is independent of individual i's actions, we will henceforth neglect it. Thus, in a symmetric equilibrium, a voter with value u will maximize[9]

$$E\left[u\Psi\left(V_{-1} + v\right)\right] - v^2, \tag{2}$$

where $V_{-1} \equiv \sum_{i \neq 1} v_i$ is the *one-out vote total*, the sum of all votes cast by all but a single individual.

We define the *expected inefficiency*[10] as $EI \equiv \frac{1}{2} - \frac{E[U\Psi(V)]}{2E[|U|]} \in [0,1]$, where $U \equiv \sum_i u_i$. This measure is the unique negative monotone linear transformation of aggregate utility realized $U\Psi(V)$ that is normalized to have range of the unit interval.[11]

2.2 Existence of Equilibria

Lemma 1. For any $N > 1$ a monotone increasing, type-symmetric Bayes-Nash Equilibrium v exists.

This result follows directly from Reny (2011)'s Theorem 4.5 for symmetric games.[12] We focus on symmetric equilibria because, in large symmetric games, asymmetric equilibria typically differ little from symmetric equilibria and are harder to characterize in detail (Satterthwaite and Williams, 1989; Rustichini et al., 1994; Cripps and Swinkels, 2006).[13]

[9] In the appendix we will prove that Bayes-Nash equilibria are non-randomized strategies.

[10] Given our assumption of quasi-linear preferences, utilitarian welfare is equivalent to efficiency in the sense of Kaldor (1939) and Hicks (1939) and thus we use the term efficiency interchangeably with welfare.

[11] It is also the complement of the ratio of the gap between the expected realized aggregate utility and the worst possible utility that could be achieved to the gap between the expected first-best utility and the worst possible utility that could be achieved.

[12] All of Reny's conditions can easily be checked, so we highlight only the less obvious ones. Continuity of payoffs in actions follows from the smoothed payoffs imposed through Ψ. Type-conditional utility is only bounded from above, not below, but boundedness from below can easily be restored by simply deleting for each value type u votes of magnitude greater $\sqrt{2|u|}$. The existence of a monotone best-response follows from the clear super-modularity of payoffs in value and votes.

[13] However, for this reason, we conjecture that our results extend to all equilibria.

2.3 Rationale

Differentiating expression (2) with respect to v yields the following first-order condition for maximization:

$$u\mathbb{E}\left[\psi\left(V_{-1}+v\right)\right]=2v \implies v(u)=\underbrace{\frac{\mathbb{E}\left[\psi\left(V_{-1}+v(u)\right)\right]}{2}}_{\text{marginal pivotality}}u. \tag{3}$$

The marginal benefit of an additional unit of vote is thus twice the individual's value multiplied by the influence this extra vote has on the chance the alternative is adopted, the vote's *marginal pivotality*. The marginal cost of a vote is twice the number of votes already purchased.

With a large number N of voters, most would reason their votes $v(u)$ have a negligible effect on the vote total $V_{-1}+v(u)$. Taking this logic to an extreme, suppose voters acted as if marginal pivotality were constant across the population, like a price $p>0$ in a market. Then for each possible value u, an individual with that value would buy $v(u)=pu$ votes. This voting strategy would imply $V=\sum_i v_i^\star = p\sum_i u_i$; that is, the vote total would be exactly proportional to the sum of the values. Because the outcome of the election is completely determined by the sign of V except when $V\in[-\delta,\delta]$, this equilibrium would enforce utilitarian efficiency.

Clearly, this argument holds only for a quadratic cost function, because only quadratic functions have linear derivatives. For example, if the cost for v votes were v^4, which has marginal cost $3v^3$, then under the hypothesis of constant marginal pivotality, buying votes in proportion to $\sqrt[3]{u}$ would be optimal for an individual with value u; because the sign of $\sum_i \sqrt[3]{u_i}$ is not generally the same as that of $\sum_i u_i$, utilitarian efficiency would not hold.

3 Main Results

Our main results concern the structure of equilibria in the game described in the previous section when the number N of agents is large, and the implications for the efficiency of QV.

3.1 Characterization of equilibrium in the zero mean case

The structure of a type-symmetric Bayes-Nash equilibrium differs radically depending on whether $\mu=0$ or $\mu\neq 0$. The case of $\mu=0$, although non-generic, is of particular interest because in some elections, for instance, when two candidates are vying for an elected office, the alternatives may be tailored so that an approximate population balance is achieved (Ledyard, 1984).

Theorem 1. For any sampling distribution F with mean $\mu=0$ that satisfies the hypotheses above, constants $\epsilon_N\to 0$ exist such that in any type-symmetric Bayes-Nash equilibrium, $v(u)$ is

C^∞ on $[\underline{u}, \overline{u}]$ and satisfies the following approximate proportionality rule:

$$\left|\frac{v(u)}{p_N u} - 1\right| \leq \epsilon_N \quad \text{where} \quad p_N = \frac{1}{2^{3/4}\sqrt{\sigma}\sqrt[4]{\pi(N-1)}}. \tag{4}$$

Furthermore, constants $\alpha_N, \beta_N \to 0$ exist such that in any equilibrium the vote total $V = V_N$ and expected inefficiency satisfy

$$|\mathbb{E}[V]| \leq \alpha_N \sqrt{\text{var}(V)} \quad \text{and} \quad EI < \beta_N. \tag{5}$$

Thus, in any equilibrium, agents buy votes approximately in proportion to their values u_i, which corresponds to their behavior under price-taking, as described in the previous section. This *approximate proportionality rule* holds because in any equilibrium, each voter perceives approximately the same *marginal pivotality*, that is, roughly, the probability that the vote total V will be in the range $[-\delta, \delta]$, where a small increment to one's vote would affect the utility.

Given approximate proportionality, understanding why the number of votes a typical voter buys should be of order $N^{-1/4}$ is not difficult. If the vote function $v(u)$ in a Bayes-Nash equilibrium follows a proportionality rule $v(u) \approx \beta u$, the constant β must be the consensus marginal pivotality. On the other hand, by the local limit theorem of probability (see Feller (1971), ch. XVI), if $\beta = CN^{-\alpha}$ for some constants $C \neq 0$ and $\alpha \in \mathbb{R}$, the chance that $V \in [-\delta, \delta]$ would be of order $N^{\alpha - \frac{1}{2}}$, and so α must be $1/4$.

Although the relation (4) asserts the ratio $v(u)/u$ is approximately constant, our analysis will show it is not *exactly* constant: different agents will perceive slightly different marginal pivotalities. Consequently, the vote function $v(u)$ is a genuinely *nonlinear* function of u, and so even though $\mathbb{E}[U] = 0$, it need not be the case that $\mathbb{E}[V] = 0$. In particular, if $\mathbb{E}[V] \approx 0$ individuals with very large values will typically buy fewer votes in proportion to their values than those with smaller values, as their votes directly reduce the chance that $V \in [-\delta, \delta]$ by breaking the approximate aggregate tie.

Thus, to establish convergence, we must prove assertion (5), namely, that these non-linearities vanish rapidly enough that the bias created by non-linearity is smaller than the sampling variation in u. We demonstrate these bounds on convergence rates by using the *Edgeworth expansion* of the distribution of V_{-1}. Were $\mathbb{E}[V] = 0$ and the distribution of V_{-1} literally normal, a standard Taylor expansion and the $N^{-1/4}$–decay of $v(u)/u$ could be used directly to show that non-linearities vanish with N^{-1} even relative to the leading term of $v(u)/u$. A detailed application of this argument leads us to conjecture that, under the hypotheses of Theorem 1, the inefficiency of QV decays like $\mu_3^2/(16\sigma^6 N)$. However, given that $\mathbb{E}[V]$ is not exactly 0, nor is V_{-1} literally normal, the arguments we use to establish (5) are subtler and consequently weaker.[14]

[14] Our result implies that inefficiency tends towards zero compared to its maximum magnitude as would occur

3.2 Characterization of equilibrium in the non-zero mean case

When μ is not zero, the nature of equilibrium can be quite different: in particular, if the payoff function is sufficiently sharp (the support of its derivative is sufficiently small) then for sufficiently large N, any type-symmetric Bayes-Nash equilibrium has a large discontinuity in the extreme tail of the sampling distribution. Nevertheless, in all cases the quadratic voting mechanism is asymptotically efficient, as the following theorem shows.

Theorem 2. Assume that the sampling distribution F has mean $\mu > 0$ and that F and Ψ satisfy the hypotheses above. Then there exist constants $\beta_N \to 0$ such that in any type-symmetric Bayes-Nash equilibrium $v(u)$,

$$EI < \beta_N. \tag{6}$$

Furthermore, there is a constant $\alpha \geq \delta$ depending on the sampling distribution F and the payoff function Ψ but not on N such that in any equilibrium $v(u)$, for any $\epsilon > 0$,

$$\sup_{\underline{u}+1/N \leq u \leq \overline{u}} |v(u) - \alpha\mu^{-1}u/N| < \alpha_N/N \quad \text{and hence} \tag{7}$$

$$P\{|V_N - \alpha| > \epsilon\} \leq \epsilon_N, \tag{8}$$

where $\epsilon_N, \alpha_N \to 0$ are constants that depend only on the sample size N, and not on the particular equilibrium.

This theorem allows for two cases. In the first, α, the asymptotic vote total, is equal to δ and thus the vote total is near δ with high probability for large N. This case occurs for large δ and thus relatively smooth payoff functions. In the second, $\alpha > \delta$, so that with high probability the vote total is outside $[-\delta, \delta]$ for large N. This case occurs for small δ and thus payoff functions that are sufficiently sharp.

To see how this dichotomy arises, suppose that for some $\alpha \geq \delta$ there were a value $w \in (-\delta, 0)$ such that

$$(1 - \Psi(w))\,|\underline{u}| > (\alpha - w)^2; \tag{9}$$

then an agent with value u_i near the lower extreme $\underline{u}$ with knowledge that the one-out vote total $V_{-i} = \sum_{j \neq i} v_j$ is near α would find it worthwhile to buy $-\alpha + w$ votes and thus single-handedly move the vote total to w. Consequently, there can be no equilibrium in which V_{-i} concentrates strictly below α if such a w exists, as this would lead a large number of individuals to wish to act a extremists, contradicting the conjectured concentration. Therefore, in any equilibrium the voters

if the wrong decision were always made, and relative to the magnitude that would occur if a random choice were made. It thus dominates any decision that could be made with only the information available to players ex-ante, unlike in the case when $\mu \neq 0$ and making the decision in the direction of μ (if feasible) achieves most potential welfare asymptotically.

with positive values u_j must buy enough votes to guarantee that the vote total concentrates at or above α. The minimal value $\alpha \geq \delta$ at which the advantage of "extremist" behavior in the extreme lower tail disappears thus determines the equilibrium behavior (7). This will be at $\alpha = \delta$ *unless* there is a solution to the following problem.

Optimization Problem. There exists a unique $\alpha > \delta$ and a matching real number $w \in [-\delta, 0]$ such that

$$\begin{aligned}(1 - \Psi(w))\,|\underline{u}| &= (\alpha - w)^2 \quad \text{and} \\ (1 - \Psi(w'))\,|\underline{u}| &\leq (\alpha - w')^2 \quad \text{for all } w' \neq w.\end{aligned} \tag{10}$$

Proposition 1. If $\delta < 1/\sqrt{2}$ then there exists a unique $\alpha > \delta$ and a matching real number $w \in [-\delta, 0]$ satisfying the Optimization Problem (10).

The proof will be given in the Appendix. When the Optimization Problem has a solution, type-symmetric Bayes-Nash equilibria take a rather interesting form in which extremists must appear, with vanishing probability, as the following theorem shows.

Theorem 3. Assume that the sampling distribution F has mean $\mu > 0$ and that F and Ψ satisfy the hypotheses above. Assume further that the Optimization Problem (10) has a unique solution (α, w). Then exists a constant $\zeta > 0$ depending on F such that for any $\epsilon > 0$ and any type-symmetric Bayes-Nash equilibrium $v(u)$, when N is sufficiently large,

(i) $v(u)$ has a single discontinuity at u_*, where $|u_* + |\underline{u}| - \zeta N^{-2}| < \epsilon N^{-2}$;

(ii) $|v(u) + \alpha - w| < \epsilon$ for $u \in [\underline{u}, u_*)$; and

(iii) the approximate proportionality rule (7) holds for all $u \in [u_*, \overline{u}]$.

Theorem 3 implies that an agent with value u will buy approximately $\alpha\mu^{-1}u/N$ votes unless u is in the extreme lower tail of F. Since such exceptional agents occur only with probability $\approx \zeta N^{-1} f(\underline{u})$, it follows by the law of large numbers that with probability $\approx 1 - \zeta N^{-1} f(\underline{u})$, the vote total will be very near α. If, on the other hand, the sample contains an agent with value less than u_* then this agent will buy approximately $\alpha - w \approx -\sqrt{|\underline{u}|}$ votes, enough to move the overall vote total close to w. Agents of the first type will be called *moderates*, and agents of the second kind *extreme contrarians* or *extremists* for short. Because the tail region in which extremists reside has $F-$probability on the order N^{-2}, the sample of agents will contain an extremist with probability only on the order N^{-1}, and will contain two or more extremists with probability on the order N^{-2}. Given that the sample contains no extremists, the conditional probability that $|V - \alpha| > \epsilon$ is $O(e^{-\varrho n})$ for some $\varrho > 0$, by standard large deviations estimates, and so the event that $V < 0$ essentially coincides with the event that the sample contains an extremist.

Why does equilibrium take the somewhat counter-intuitive form described in Theorem 3? An agent i with value u_i in the "bulk" of the sampling distribution F, there is very little information about the vote total V in the agent's value u_i, and so for most such agents the marginal pivotality $\mathbb{E}\left[\psi\left(V_{-i}+v(u_i)\right)\right]$ will be approximately $\mathbb{E}\left[\psi(V)\right]$. This implies that in the bulk of the distribution the function $v(u)$ will be approximately linear in u. Therefore, by the law of large numbers, the vote total will, with high probability, be near $N\mathbb{E}\left[\psi(V)/2\mu\right]$.

Because $\mu > 0$, agents with negative values will, with high probability, be on the losing side of the election. However, if $\mathbb{E}\left[\psi(V)\right]$ were small enough that $N\mathbb{E}\left[\psi(V)/2\mu\right]$ decayed with N, agents with negative values could overcome the votes of all other individuals at sufficiently small cost that all would eventually wish to be extremists. Such a high level of extremist activity in turn would imply $\mathbb{E}\left[\psi(V)\right]$ is large, because any time an extremist exists the conditional expectation $\psi(V)$ must be large enough to satisfy her first-order condition, which involves buying many votes. Consequently, $N\mathbb{E}\left[\psi(V)\mu\right]$ must remain bounded away from 0.

On the other hand, if $N\mathbb{E}\left[\psi(V)\right]$ grows unboundedly with N, no individual could profitably act as an extremist, so no extremist will exist and $\mathbb{E}\left[\psi(V)\right]$ would be exponentially small, which is impossible. Thus, the aggregate number of votes must concentrate near a constant value, and so most voters must buy on the order of $1/N$ votes. For this scenario to occur, $\mathbb{E}\left[\psi(V)\right]$ must decay as $1/N$. But the primary contribution to this expectation must come from the event in which an extremist exists, and so the probability of this event must decay as $1/N$.

3.3 Some Remarks on the Proofs

The heuristic arguments of the preceding two subsections rely on central limit approximations and large deviations estimates for the aggregate vote total. Applying these theorems is not justified *a priori*, however, because despite the fact that the distribution of values is exogeneous, the distribution of votes is endogenous. Moreover, precisely because marginal pivotality is not constant, the distribution of votes may be different from that of values in important respects. This endogeneity severely circumscribes our ex-ante knowledge about how marginal pivotality itself behaves. Thus, we will be forced to squeeze out information about the vote function $v(u)$ by a bootstrapping procedure, at each step using new information about $v(u)$ to make a more precise approximation of the marginal pivotality and thereby obtain even more precise information about the vote function $v(u)$. The key steps, which are established formally in the appendix, are as follows.

(1) *Weak Consensus:* The most crucial step is to show that *most* agents (all but those whose values u are in the tails of the distribution F) will have similar assessments of the conditional distribution (given their knowledge of their own values) of the vote total, and hence the marginal pivotalities. The key to this *weak consensus* principle is a basic result in the theory of random sampling: that the probability of obtaining exactly k individuals in a random sample with values in a given interval J is virtually the same as the probability of obtaining exactly $k+1$ such individ-

uals *unless* J is so small or so far out in the tail of the sampling distribution that the likelihood of obtaining more than a very small number of individuals with values in J is negligible. These notions of "smallness" are quantified in Lemma 6; it implies for any two individuals with values u_i, u_j in the bulk of the distribution F, the ratio $\frac{v(u_i)/u_i}{v(u_j)/u_j}$ is bounded above and below.

(2) *Concentration:* We then employ a *concentration inequality* for sums of i.i.d. random variables that bounds the probability that such a sum will fall in a given interval. This bound, involving the variance, third moment, and N, will lead to bounds on the size of $v(u)$. In a nutshell, the argument is as follows.

For an agent with value u, the marginal pivotality is the expectation $\mathbb{E}\left[\psi(V_{-1} + v(u))\right]$. Because ψ is non-zero only in the interval $(-\delta, \delta)$, the expectation will be large only if the distribution of $V_{-1} + v(u)$ puts a substantial mass on this interval. But the concentration inequality implies the distribution of the sum V_{-1} will be highly concentrated if and only if the individual summands $v(u_j)$ have small variance, which will be the case just when they are (mostly) small. By the necessary condition for an equilibrium, this can occur only if the marginal pivotalities $\mathbb{E}\left[\psi(V_{-j} + v(u_j))\right]$ are (mostly) small. Therefore, by the weak consensus principle, if $v(u)/u$ is even moderately large, the value u must be in one of the extreme tails of the distribution F, because for nearly all other values u the marginal pivotality will be small.

A careful rendition of this argument will show that (i) extremists must have values u within distance $O(N^{-3/2})$ of one of the endpoints $\underline{u}, \overline{u}$ of the support interval, and (ii) the number of votes $|v(u)|$ that any agent in the bulk of the distribution buys cannot be larger than $O(N^{-1/4})$.

(3) *Discontinuities and Smoothness:* At any discontinuity of the vote function $v(u)$, two distinct solutions (the right and left limits $v(u+)$ and $v(u-)$) of the necessary condition (3) must exist. Because ψ is smooth, the derivative ψ' exists and is continuous everywhere, and so the mean value theorem implies that at any such discontinuity, some $\tilde{v} \in [v(u-), v(u+)]$ must exist at which $\mathbb{E}\left[\psi'(\tilde{v} + V_{-1})u\right] = 2$. But because ψ' is non-zero only in the interval $(-\delta, \delta)$, the existence of $\tilde{v}$ implies, once again, that the distribution of V_{-1} is highly concentrated; thus, a variation of the argument in (2) implies that (i) the size of any discontinuity must be bounded below, and (ii) discontinuities can occur only at values u in the extreme tails of F.

Because the vote function $v(u)$ is monotone in any equilibrium, it must be differentiable almost everywhere. At any u where the derivative $v'(u)$ exists, the necessary condition (3) can be differentiated, yielding the identity

$$v'(u) = \frac{\mathbb{E}\left[\psi(v(u) + V_{-1})\right]}{2 - \mathbb{E}\left[\psi'(v(u) + V_{-1})\right] u}.$$

The quantity on the right side varies continuously with u in any interval where $v(u)$ is continuous and $E\psi'(v(u) + V_{-1})u < 2$, and hence v' extends continuously to any such interval. Thus we conclude, by another use of the concentration inequalities, that $v(u)$ is not only continuous but continuously differentiable at all u except in the extreme tails of F.

(4) *Approximate Proportionality:* Weak consensus tells us the ratio $v(u)/u$ does not vary, in relative terms, by more than a bounded amount in the bulk of the distribution F, but to complete the proofs we will need something stronger: that when N is large, the ratio $v(u)/u$ is nearly constant except in the tails of F. To establish this result, we bring the various bits of information gleaned from the analysis in steps 1-3 to bear on the necessary condition (3). Because ψ is smooth and has compact support, it and all of its derivatives are uniformly continuous and uniformly bounded, and so the function $v \mapsto E\psi(v + V_{-1})$ is differentiable, with derivative $E\psi'(v + V_{-1})$. Consequently, by an application of Taylor's theorem to the identity (3), a $\tilde{v}(u)$ intermediate between 0 and $v(u)$ exists such that

$$2v(u) = \mathbb{E}\left[\psi(V_{-1})u\right] + \mathbb{E}\left[\psi'(\tilde{v}(u) + V_{-1})v(u)u\right].$$

Concentration implies that, for u in the bulk of the distribution F, when N is large, the expectation $\mathbb{E}\left[\psi'(\tilde{v}(u) + V_{-1})\right]$ is small; thus, the *approximate proportionality* rule $2v(u) \approx \mathbb{E}\left[\psi(V_{-1})u\right]$ must hold except in the tails of F. A more delicate analysis, using the smoothness of v (Step 3), will show the approximate proportionality rule extends to all but the *extreme* tails, that is to all u not within distance $N^{-3/2}$ of one of the endpoints $\underline{u}, \overline{u}$.

The upshot of approximate proportionality is that the contribution to the one-out vote total V_{-1} of those terms $v(u_i)$ for which u_i is not in the extreme tails of F can be bounded above and below by scalar multiples of the sums of the corresponding values u_i. Since these values are gotten by i.i.d. random sampling from a fixed distribution F with bounded support, standard results of probability theory (central limit theorem, Berry-Esseen bounds, Hoeffding's inequality) can now be brought to bear on the distribution of V_{-1}.

(5) *Non-zero mean case*: In the non-zero mean case, we first rule out discontinuities near the upper endpoint $\overline{u}$ by arguing that, unless the sample contains an extremist with value near $\underline{u}$, an event of vanishingly small probability, the conditional probability that the sum of the moderate votes will exceed δ is exponentially close to 1, so voters with values near $\overline{u}$ need not buy more than $O(N^{-1})$ votes to ensure their side wins the election. From there on the logic essentially follows the heuristic argument given above.

(6) *Zero-mean case*: This case is technically more delicate because the Edgeworth expansion discussed above must be conducted not about 0 as supposed above, but rather about the still unknown $\mathbb{E}[V]$. Thus, another bootstrapping argument is needed to deduce that $\mathbb{E}[V]$ is of a smaller order of magnitude than the standard deviation of V. Given this bound, it then follows

that the lead term in the Edgeworth series (the usual normal approximation) dominates, and the results claimed in Theorem 1 then follow.

4 Potential Applications

The utilitarian efficiency on which we focused is unlikely to be the sole objective in public decision-making.[15] However, Groves and Loeb (1979) argue it is a natural objective within corporations in which distributive concerns are less important and the mechanism will be used repeatedly to make a variety of decisions, so that those adopting the mechanism do so behind at least a partial veil of ignorance. On this basis, Listokin (2015) argues for replacing corporate voting with the VCG mechanism to avoid value-destroying acquisitions of companies in which one large shareholder owns a stake and other forms of "tunneling" that minority shareholder protections are thought to provide insufficient protection against at present (La Porta et al., 2000, 2002).

However, as Listokin concedes, the VCG mechanism is complex, runs a budget surplus, and is highly sensitive to collusion in both theory (Groves and Ledyard, 1977b) and experiments (Attiyeh et al., 2000). By contrast, laboratory (Goeree and Zhang, 2015) and controlled field (Cárdenas et al., 2014) experiments on variants on QV (with a fully sharp Ψ function and different multipliers on cost) have proven more successful.[16] In both cases, these experimental studies find that QV performs quite efficiently, and more efficiently than the one-person-one-vote rule, because individuals buy votes roughly in proportion to their values. In fact, Goeree and Zhang find a large majority of individuals choose QV over majority rule after having multiple rounds of experience with both.[17] Furthermore, as Hansmann and Pargendler (2014) observe, a variety of voting rules, some taking forms not far off from the form prescribed by the quadratic rule, have been used in corporate governance, quite effectively according to the empirical analysis of Bodenhord (2014).

As a result, Posner and Weyl (2014) and Cariello (2015) argue that, unlike previous mechanisms, a version of QV might improve on existing methods of corporate governance and address

[15]Other potential criteria are fairness relative a *status-quo* (Lindahl, 1919) or objectives ex-ante welfare measure that account for differences in the marginal utility of income across individuals (Vickrey, 1945). We are exploring the properties of QV in relation to the first criterion in joint work with Jerry Green and Scott Duke Kominers. Posner and Weyl (2015) discuss the latter informally in the context of applications to democratic politics.

[16]As indicated above, our theoretical results do not directly apply without a smooth Ψ function, though we believe the results could be extended to allow a sharp Ψ, as we discuss in our conclusion.

[17]However, both papers also find that individuals buy a far greater number of votes than called for in the equilibrium we derive here, so our analysis seems inadequate to account for experimentally observed behavior. Goeree and Zhang suggest over-estimation of the chance of being pivotal may explain this behavior, but other explanations, such as a desire to express oneself through voting (Fiorina, 1976), are also plausible, and knowing how QV performs in a broader set of cases than those explored in existing experiments when these motives are present would be useful.

the problems raised by Listokin. They also suggest it could be implemented relatively easily and consistent with existing corporate law, at least in common law countries. Thus, QV seems to be of some practical and not merely academic interest. However, a number of questions would have to be analyzed to determine whether such an application would be advisable. For example, previous analyses of existing voting rules (Grossman and Hart, 1980) and arguments for replacing them with ex-post efficient decision rules like QV (Listokin, 2015) do not rigorously consider impacts on ex-ante investment incentives. In other applications, quite different concerns are more important. For example, in market research, analysis of consumer motivations in filling out surveys is crucial,[18] whereas in potential applications to public democracy, distributional concerns are paramount (Posner and Weyl, 2015). This diversity of practical concerns puts such questions beyond the scope of our present analysis, but we hope future research will treat these challenges seriously and illuminate whether and where QV has practical value.

5 Conclusion

In this paper, we proved convergence of all type-symmetric Bayes-Nash equilibria of QV to efficiency in an independent private-values environment. Our results could be extended in many directions. For example, we assumed value distributions had bounded support and positive density at the boundaries of this support, and we considered a variant of QV with a smooth mapping between aggregate votes and the decision made. We made these assumptions more for tractability than realism; relaxing them would be a useful and, we conjecture, reasonably straightforward, technical contribution.

More broadly, many other models exist in which considering the performance of QV would be informative. In particular, as discussed in the introduction, existing (asymptotically) efficient mechanisms for binary collective decisions are extremely sensitive to at least one of collusion (VCG), aggregate uncertainty (EE and "implement-the mean"), or imperfect rationality or instrumentality of motivations (CV). Furthermore, our finding that in the generic case in which $\mu \neq 0$ extremists exist who buy a large number of votes relative to the total votes in the population arises from the fact that, otherwise, the motivation to vote would be exponentially small. In the presence of aggregate uncertainty, non-instrumental motivations for voting or imperfect rationality there would be much larger motivations to vote even in large populations. Additionally the fear of collusive groups could substitute for the fear of extremist individuals in maintaining the motivation to vote of the broad population. Thus, the precise structure of equilibrium we characterize seems unlikely to persist in such models, though the broader heuristic rationale of Subsection 2.3 could nonetheless remain valid and preserve at least approximate efficiency.

[18] A smooth Ψ is more appropriate in this context than in the examples discussed above, because the impact of survey responses on eventual company decisions is usually noisy and uncertain.

Weyl (2015b) analyzes these issues, without rigorous proof, using analytic approximations similar to those we rigorously show are valid in the independent, non-cooperative private-values setting above. His results suggest QV is fully asymptotically efficient in some, but not all of these cases. For example, he finds it may have limiting expected inefficiency of as high as 4% with aggregate, normal uncertainty. However, he also finds this inefficiency is almost always quantitatively small for reasonable calibrations, compared to that of both the common one-person-one-vote rule and that of other proposed asymptotically efficient mechanisms. Simulations and numerical approximations by Weyl (2015a) also suggest QV achieves results qualitatively close to full efficiency even in small populations and with distributions of values that have unbounded support. If such results can be made precise and shown to hold over a broader range of quantitative calibrations, it would constitute, along with existing experimental and theoretical evidence, a case for larger-scale exploration of QV in practical applications, such as those discussed in the previous section.

References

Al-Najjar, Nabil I. and Rann Smorodinsky, "Pivotal Players and the Characterization of Influence," *Journal of Economic Theory*, 2000, *92* (2), 318–342.

Arrow, Kenneth J., "The Property Rights Doctrine and Demand Revelation under Incomplete Information," in Michael Boskin, ed., *Economics and Human Welfare*, New York: Academic Press, 1979, pp. 23–39.

Attiyeh, Greg, Robert Franciosi, and R. Mark Isaac, "Experiments with the Pivot Process for Providing Public Goods," *Public Choice*, 2000, *102* (1–2), 95–112.

Ausubel, Lawrence M. and Paul Milgrom, "The Lovely but Lonely Vickrey Auction," in Peter Cramton, Richard Steinberg, and Yoav Shoham, eds., *Combinatorial Auctions*, Cambridge, MA: MIT Press, 2005, pp. 17–40.

Bodenhord, Howard, "Voting Rights, Shareholdings, and Leverage at Nineteenth-Century U.S. Banks," *Journal of Law and Economics*, 2014, *57* (2), 431–458.

Bowen, Howard R., "The Interpretation of Voting in the Allocation of Economic Resources," *Quarterly Journal of Economics*, 1943, *58* (1), 27–48.

Cárdenas, Juan Camilo, César Mantilla, and Román David Zárate, "Purchasing Votes without Cash: Implementing Quadratic Voting Outside the Lab," 2014. https://www.aeaweb.org/aea/2015conference/program/retrieve.php?pdfid=719.

Cariello, Vicenzo, "Azioni a Voto Potenziato, "Voti Plurimi Senza Azioni" e Tutela dei Soci Estranei al Controllo," *Rivista delle Societa*, 2015, *1* (1), 164–203.

Che, Yeon-Koo and Fuhito Kojima, "Asymptotic Equivalence of Probabilistic Serial and Random Priority Mechanisms," *Econometrica*, 2010, *78* (5), 1625–1672.

Clarke, Edward H., "Multipart Pricing of Public Goods," *Public Choice*, 1971, *11* (1), 17–33.

Cripps, Martin W. and Jeroen M. Swinkels, "Efficiency of Large Double Auctions," *Econometrica*, 2006, *74* (1), 47–92.

d'Aspremont, Claude and Louis-André Gérard-Varet, "Incentives and Incomplete Information," *Journal of Public Economics*, 1979, *11* (1), 25–45.

Ellenberg, Jordan, "Saving Democracy With Quadratic Equations," *Wall Street Journal*, 2015, *January 16.*

Feller, William, *An introduction to probability theory and its applications. Vol. II.* Second edition, John Wiley & Sons, Inc., New York-London-Sydney, 1971.

Fiorina, Morris P., "The Voting Decision: Instrumental and Expressive Aspects," *Journal of Politics*, 1976, *38* (2), 390–413.

Goeree, Jacob K. and Jingjing Zhang, "One Man, One ~~Vote~~ Bid," 2015. https://goo.gl/G4wnWh.

Green, Jerry, Elon Kohlberg, and Jean-Jacques Laffont, "Partial Equilibrium Approach to the Free-Rider Problem," *Journal of Public Economics*, 1976, *6* (4), 375–394.

Grossman, Sanford J. and Oliver D. Hart, "Takeover Bids, the Free-Rider Problem, and the Theory of the Corporation," *Bell Journal of Economics*, 1980, *11* (1), 42–64.

Groves, Theodore, "Incentives in Teams," *Econometrica*, 1973, *41* (4), 617–631.

_ **and John Ledyard**, "Optimal Allocation of Public Goods: A Solution to the "Free Rider" Problem," *Econometrica*, 1977, *45* (4), 783–809.

_ **and John O. Ledyard**, "Some Limitations of Demand Revealing Processes," *Public Choice*, 1977, *29* (2S), 107–124.

_ **and Martin Loeb**, "Incentives in a Divisionalized Firm," *Management Science*, 1979, *25* (3), 221–230.

Hansmann, Henry and Mariana Pargendler, “Evolution of Shareholder Voting Rights: Separation of Ownership and Consumption,” *Yale Law Journal*, 2014, *123* (4), 100–165.

Hicks, J. R., “The Foundations of Welfare Economics,” *Economic Journal*, 1939, *49* (196), 696–712.

Hoeffding, Wassily, “Probability inequalities for sums of bounded random variables,” *J. Amer. Statist. Assoc.*, 1963, *58*, 13–30.

Hylland, Aanund and Richard Zeckhauser, “A Mechanism for Selecting Public Goods when Preferences must be Elicited,” 1980. Kennedy School of Government Discussion Paper D.

Kahn, Jeff, G. Kalai, and Nathan Linial, “The Influence of Variables on Boolean Functions,” in “29th Annual Symposium on Foundations of Computer Science” 1988, pp. 68–80.

Kaldor, Nicholas, “Welfare Propositions in Economics and Interpersonal Comparisons of Utility,” *Economic Journal*, 1939, *49* (145), 549–552.

Krishna, Vijay and John Morgan, “Asymmetric Information and Legislative Rules: Some Ammendments,” *American Political Science Review*, 2001, *95* (2), 435–452.

_ and _ , “Majority Rule and Utilitarian Welfare,” *American Economic Journal: Microeconomics*, 2015, *7* (4), 339–375.

La Porta, Rafael, Florencio Lopez de Silanes, Andrei Shleifer, and Robert Vishny, “Investor Protection and Corporate Governance,” *Journal of Financial Economics*, 2000, *58* (1–2), 3–27.

_ , _ , _ , and _ , “Investor Protection and Corporate Valuation,” *Journal of Finance*, 2002, *57* (3), 1147–1170.

Ledyard, John, “The Pure Theory of Large Two-Candidate Elections,” *Public Choice*, 1984, *44* (1), 7–41.

Lindahl, Erik, “Positive Lösung,” in “Die gerechtigkeit der Besteuerung” number I, Lund, 1919, chapter 4, pp. 85–98.

Listokin, Yair, “The Vickrey-Clarke-Groves “Pivotal Mechanism” as an Alternative to Voting for Organizational Control,” *Theoretical Inquiries in Law*, 2015, *16* (1), 267–294.

Mailath, George J. and Andrew Postlewaite, “Asymmetric Information Bargaining Problems with Many Agents,” *Review of Economic Studies*, 1990, *57* (3), 351–367.

Myerson, Roger B., “Optimal Auction Design,” *Mathematics of Operations Research*, 1981, *6* (1), 58–73.

— , "Large Poisson Games," *Journal of Economic Theory*, 2000, *94* (1), 7–45.

Posner, Eric A. and E. Glen Weyl, "Quadratic Voting as Efficient Corporate Governance," *University of Chicago Law Review*, 2014, *81* (1), 251–272.

— **and** — , "Voting Squared: Quadratic Voting in Democratic Politics," *Vanderbilt Law Review*, 2015, *68* (2), 441–499.

Reny, Philip J., "On the Existence of Monotone Pure-Strategy Equilibria in Bayesian Games," *Econometrica*, 2011, *79* (2), 499–553.

Roberts, Donald John and Andrew Postlewaite, "The Incentives for Price-Taking Behavior in Large Exchange Economies," *Econometrica*, 1976, *44* (1), 115–127.

Rustichini, Aldo, Mark A. Satterthwaite, and Steven R. Williams, "Convergence to Efficiency in a Simple Market with Incomplete Information," *Econometrica*, 1994, *62* (5), 1041–1063.

Satterthwaite, Mark A. and Steven R. Williams, "The Rate of Convergence to Efficiency in the Buyer's Bid Double Auction as the Market Becomes Large," *Review of Economic Studies*, 1989, *56* (4), 477–498.

Tideman, N. Nicolaus and Gordon Tullock, "A New and Superior Process for Making Social Choices," *Journal of Political Economy*, 1976, *84* (6), 1145–1159.

Vickrey, William, "Measuring Marignal Utility by Reactions to Risk," *Econometrica*, 1945, *13* (4), 319–333.

— , "Counterspeculation, Auctions and Competitive Sealed Tenders," *Journal of Finance*, 1961, *16* (1), 8–37.

Weyl, E. Glen, "Quadratic Vote Buying," 2012. http://goo.gl/8YEO73.

— , "The Efficiency of Quadratic Voting in Finite Populations," 2015. http://ssrn.com/abstract=2571026.

— , "The Robustness of Quadratic Voting," 2015. http://ssrn.com/abstract=2571012.

Appendix

In this appendix we prove Proposition 1 and Theorems 1- 3. We take as given the model assumptions laid out in Subsection 2.1 of the text, though we adopt some slightly altered notation for convenience: in particular, we assume the sampling distribution F is supported by a finite interval $[\underline{u}, \overline{u}]$, that its mean $\mu = \mu_U$ is nonnegative, and that its density f is smooth and strictly positive on this interval; and we assume the payoff function Ψ satisfies the assumptions 3 (a)- (c) of Subsection 2.1. For notational convenience we write $n = N - 1$ for the number of voters minus one and $S_n = V_{-N}$ for the vote total of the first n voters. For brevity, we refer to *type-symmetric Bayes-Nash equilibria* as *Nash equilibria*, and we drop the brackets following the expectation operator E when this does not cause confusion. Except when explicitly noted otherwise, all references to sections and subsections in this appendix are internal.

The proof of Proposition 1, which is independent of the rest of the development, will be given in Section G. The proofs of Theorems 1, 2, and 3 in Sections D and E rely on a number of auxiliary properties of Nash equilibria that will be established in Sections B and C. In Section B we prove that for every mixed-strategy Nash equilibrium an equivalent pure-strategy Nash equilibrium $u \mapsto v(u)$ exists satisfying the necessary condition (3). We then proceed to use this necessary condition to deduce a series of properties that any Nash equilibrium must have, at least when n is sufficiently large. Note that *a priori* we know nothing about the function $v(u)$, and so even though we have made strong assumptions about the distribution of the utility values U_i we cannot appeal to the classical laws of probability governing sums of independent, identically distributed random variables to deduce anything about the sums $S_n = \sum_{i=1}^{n} v(U_i)$ until we know more about the function $v(u)$.

A Terminology and Notation

Under the model assumptions we have stipulated, it would never make sense for an agent to purchase more than $\sqrt{2\overline{u}}$ or fewer than $-\sqrt{2|\underline{u}|}$ votes. Consequently, we shall restrict attention to strategies which result in $V_i \in [-\sqrt{2\,|\underline{u}|}, \sqrt{2\overline{u}}]$ for all agents i. A *pure strategy* is a Borel measurable function $v : [\underline{u}, \overline{u}] \to [-\sqrt{2\,|\underline{u}|}, \sqrt{2\overline{u}}]$; when a pure strategy v is adopted, each agent buys $v(u)$ votes, where u is the agent's utility. A *mixed strategy* is a Borel measurable[19] function $\pi_V : [\underline{u}, \overline{u}] \to \Pi$, where Π is the collection of Borel probability measures on $[-\sqrt{2\,|\underline{u}|}, \sqrt{2\overline{u}}]$; when a mixed strategy π_V is adopted, each agent i will buy a random number V_i of votes, where $V_1, V_2, \ldots$ are conditionally independent given the utilities $U_1, U_2, \ldots$ and V_i has conditional distribution $\pi_V(U_i)$. Clearly, the set of mixed strategies contains the pure strategies.

A *best response* for an agent with utility u to a strategy (either pure or mixed) is a value v such that

$$E\Psi(v + S_n)u - v^2 = \sup_{\tilde{v}} E\Psi(\tilde{v} + S_n)u - \tilde{v}^2, \tag{11}$$

where S_n is the sum of the votes of the other n agents when these agents all play the specified strategy and E denotes expectation. (Thus, under E, the random variables V_i of the n other voters are distributed in accordance with the strategy and the sampling rule for utility values U_i described above.) Because Ψ is continuous and bounded, equation (11) and the dominated convergence theorem imply that for each u the set of best responses is closed and hence has well-defined maximal and minimal elements $v_+(u), v_-(u)$.

[19]The space of Borel probability measures on $[-\sqrt{2\,|\underline{u}|}, \sqrt{2\overline{u}}]$ is given the topology of weak convergence; Borel measurability of a function with range Π is relative to the Borel field induced by this topology. Proposition 2 below implies that in the Quadratic Voting game only pure strategies are relevant, so measurability issues play no role in this paper.

A mixed strategy π_V is a *Nash equilibrium* if for every $u \in [\underline{u}, \overline{u}]$ the measure $\pi_V(u)$ is supported by the set of best responses to π_V for an agent with utility u. Two strategies will be called *equivalent* if they coincide for all u except in a set of Lebesgue measure 0. Because the values U_i of the various agents are sampled from a distribution that is absolutely continuous, if two strategies are equivalent then with probability one they result in exactly the same actions.

Notation. The symbols Ψ, ψ, δ, F, μ, ζ, $\underline{u}, \overline{u}$ will be reserved for the functions and constants specified in Section 2.1, and the letters N, n will be used only for the sample size and sample size minus one. The symbols $\alpha, \beta, \gamma, \epsilon, \varrho$ and C will be used for generic constants whose values might change from one lemma to the next. Because many of the arguments to follow will involve the values of the equilibrium vote function v at points near one of the endpoints $\underline{u}, \overline{u}$, we will use the following shorthand notation, for any $0 < \epsilon < 1$:

$$\overline{u}_\epsilon = \overline{u} - \epsilon \quad \text{and} \quad \underline{u}_\epsilon = \underline{u} + \epsilon.$$

B Necessary Conditions for Nash Equilibrium

Let π_V be a mixed-strategy Nash equilibrium, and let S_n be the sum of the votes of n agents with utilities U_i obtained by random sampling from F, all acting in accordance with the strategy π_V. For an agent with utility u, a best response v must satisfy equation (11), and so in particular for every $\Delta > 0$, if $u > 0$ then

$$\begin{aligned} &E\left\{\Psi(S_n + v + \Delta) - \Psi(S_n + v)\right\} u \leq 2\Delta v + \Delta^2 \quad \text{and} \\ &E\left\{\Psi(S_n + v - \Delta) - \Psi(S_n + v)\right\} u \leq -2\Delta v + \Delta^2 \end{aligned} \tag{12}$$

Similarly, if $u < 0$ and $\Delta > 0$ then

$$\begin{aligned} &E\left\{\Psi(S_n + v - \Delta) - \Psi(S_n + v)\right\} u \leq -2\Delta v + \Delta^2 \quad \text{and} \\ &E\left\{\Psi(S_n + v + \Delta) - \Psi(S_n + v)\right\} u \leq 2\Delta v + \Delta^2 \end{aligned} \tag{13}$$

Because Ψ is C^∞ and its derivative ψ has compact support, differentiation under the expectation is permissible. Thus, we have the following necessary condition.

Lemma 2. If π_V is a mixed-strategy Nash equilibrium then for every u a best response v must satisfy

$$E\psi(S_n + v)u = 2v. \tag{14}$$

Consequently, every pure-strategy Nash equilibrium $v(u)$ must satisfy the functional equation

$$E\psi(S_n + v(u))u = 2v(u). \tag{15}$$

Lemma 3. Let π_V be a mixed-strategy Nash equilibrium, and let $v, \tilde{v}$ be best responses for agents with utilities $u, \tilde{u}$, respectively. If $u = 0$ then $v = 0$, and if $u < \tilde{u}$, then $v \leq \tilde{v}$. Consequently, any pure-strategy Nash equilibrium $v(u)$ is a nondecreasing function of u and therefore has at most countably many discontinuities and is differentiable almost everywhere.

Proof. It is obvious that the only best response for an agent with $u = 0$ is $v = 0$, and the monotonicity of the payoff function Ψ implies a best response v for an agent with utility u must be of the same sign as u. If $v, \tilde{v}$ are best responses for agents with utilities $0 \leq u < \tilde{u}$, then by definition

$$E\Psi(\tilde{v}+S_n)\tilde{u}-\tilde{v}^2 \geq E\Psi(v+S_n)\tilde{u}-v^2 \quad \text{and}$$
$$E\Psi(v+S_n)u-v^2 \geq E\Psi(\tilde{v}+S_n)u-\tilde{v}^2,$$

and so, after re-arrangement of terms,

$$(E\Psi(\tilde{v}+S_n)-E\Psi(v+S_n))\tilde{u} \geq \tilde{v}^2-v^2 \quad \text{and}$$
$$(E\Psi(\tilde{v}+S_n)-E\Psi(v+S_n))u \leq \tilde{v}^2-v^2.$$

Hence,

$$(E\Psi(\tilde{v}+S_n)-E\Psi(v+S_n))(\tilde{u}-u) \geq 0.$$

The monotonicity of Ψ implies that if $0 \leq \tilde{v} < v$ then $E\Psi(\tilde{v}+S_n) \leq E\Psi(v+S_n)$, and so it follows that the two expectations must be equal because $\tilde{u}-u > 0$. But if the two expectations were equal v could not possibly be a best response at u, because an agent with utility u could obtain the same expected payoff $E\Psi(v+S_n)u$ at a lower vote cost by purchasing $\tilde{v}$ votes. This argument proves that if $0 \leq u < \tilde{u}$ best responses $v, \tilde{v}$ for agents with utilities $u, \tilde{u}$ must satisfy $0 \leq v \leq \tilde{v}$. A similar argument shows that if $u < \tilde{u} \leq 0$ best responses $v, \tilde{v}$ for agents with utilities $u, \tilde{u}$ must satisfy $v \leq \tilde{v} \leq 0$.

□

Proposition 2. If a mixed strategy π_V is a Nash equilibrium, the set of utility values $u \in [\underline{u}, \overline{u}]$ for which more than one best response (and hence the set of values u such that $\pi_V(u)$ is not supported by just a single point $v(u)$) is at most countable.

Proof of Proposition 2. For each u denote by $v_-(u)$ and $v_+(u)$ the minimal and maximal best responses at u. Lemma 3 implies that if $u < \tilde{u}$ then $v_+(u) \leq v_-(\tilde{u})$. Consequently, for any $\epsilon > 0$ the set of utilities values u at which $v_+(u) - v_-(u) \geq \epsilon$ must be finite, because otherwise $v_+(u) \to \infty$ as $u \to \overline{u}$, which is impossible because best responses must take values between $-\sqrt{2\,|\underline{u}|}$ and $\sqrt{2\overline{u}}$. □

Because by hypothesis the values U_i are sampled from a distribution F that is absolutely continuous with respect to Lebesgue measure, the probability that one of the votes i will have utility value U_i equal to one of the countably many values where more than one best response exists is zero. Consequently, for every Nash equilibrium an equivalent pure-strategy Nash equilibrium $v(u)$ exists. Henceforth, we consider only pure-strategy Nash equilibria; whenever we refer to a *Nash equilibrium* we mean a pure-strategy Nash equilibrium.

Lemma 4. If $v(u)$ is a Nash equilibrium, $v(u) \neq 0$ for all $u \neq 0$.

Proof. If $v(u) = 0$ for some $u > 0$ then by Lemma 3 $v(u') = 0$ for all $u' \in (0, u)$. Because the density $f(u)$ of the value distribution F is strictly positive on $[\underline{u}, \overline{u}]$, it follows that the probability p that every agent in the sample casts vote $V_i = 0$ is strictly positive. But then an agent with utility u could improve her expectation by buying $\varepsilon > 0$ votes, where $\varepsilon \ll u\psi(0)p$, because the expected utility gain would be at least

$$u\Psi(\varepsilon)p \sim u\psi(0)p\varepsilon$$

at a cost of ε^2. Because by hypothesis $\psi(0) > 0$, the expected utility gain would overwhelm the increased vote cost for small $\varepsilon > 0$. □

Corollary 1. Any Nash equilibrium $v(u)$ is *strictly* increasing on $[\underline{u}, \overline{u}]$.

Proof. Lemmata 2 and 4 imply $E\psi(S_n + v(u)) > 0$ for every $u \neq 0$. Now differentiation of the necessary condition (15) gives

$$E\psi(S_n + v(u)) = (2 - E\psi'(S_n + v(u)))v'(u)$$

at every u where $v(u)$ is differentiable. Because such points are dense in $[\underline{u}, \overline{u}]$, and because ψ and ψ' are C^∞ functions with compact support, it follows that $v'(u) \neq 0$ on a dense set. But $v'(u) \geq 0$ at every point where the derivative exists, so it follows that $v'(u) > 0$ almost everywhere, implies v is strictly monotone. □

C Continuity and Smoothness

C.1 Weak consensus bounds

According to Lemma 2, in any Nash equilibrium the number of votes $v(u)$ an agent with utility u purchases must satisfy the necessary condition (15). It is natural to expect that when the sample size $n + 1$ is large the effect of adding a single vote v to the aggregate total S_n should be small, and so one might expect that the function $v(u)$ should satisfy the approximate proportionality rule

$$2v(u) \approx E\psi(S_n)u.$$

As we will show later, this naive approximation can fail badly for utility values u in the extreme tails of the distribution F. Nevertheless, the idea of approximate population consensus on the expectations $E\psi(v(u) + S_n)$ can be used to obtain weak bounds that we will find useful.

Lemma 5. A $\gamma > 0$ exists such that for all sufficiently large n, in any Nash equilibrium,

$$\max(v(\overline{u}_{1/n}), -v(\underline{u}_{1/n}) \geq \gamma/n. \tag{16}$$

Furthermore, for any $\epsilon > 0$ a $C = C(\epsilon) > 0$ exists such that for all sufficiently large n, in any Nash equilibrium,

$$|v(u)| \geq \frac{C|u|}{n^2} \quad \text{for all } \ |u| \geq \epsilon. \tag{17}$$

Proof. Suppose inequality (16) did not hold; then with probability $\approx (1 - (f(\underline{u}) + f(\overline{u}))/n)^n \approx \exp\{-f(\underline{u}) - f(\overline{u})\} := p$, the values of all agents would lie in the interval $[\underline{u}_{1/n}, \overline{u}_{1/n}]$, and so the vote total would be no more than γ in absolute value. But if γ were sufficiently small, then an agent with value $u = 1$ would find it advantageous to defect from the equilibrium strategy by buying 3γ votes, at cost $9\gamma^2$, thus raising her expected payoff by at least $p(\Psi(2\gamma) - \Psi(\gamma)) \approx p\psi(\gamma)\gamma \gg 9\gamma^2$.

Assume now that $v(\overline{u}_{1/n}) > \gamma/n$; the other case $v(\underline{u}_{1/n}) < -\gamma/n$ can be argued similarly. If this is the case, then by the necessary condition (15) and the monotonicity of v,

$$E\psi(S_n + v(u)) \geq \frac{2\gamma}{n\overline{u}}$$

for all values $u \geq \overline{u}_{1/n}$, and so

$$E(\psi(S_n + v(U)) \,|\, U \geq \overline{u}_{1/n}) \geq \frac{2\gamma}{n\overline{u}}.$$

Fix $\beta > 0$ small, and suppose $|v(u)| < \beta|u|/n^2$ for some value $u \notin [-\epsilon, \epsilon]$. Denote by G the event that $\max_{1 \leq i \leq N} U_i > \overline{u} - n^{-1}$, and let U be independent of $U_1, U_2, \ldots, U_N$, with distribution F; then by Taylor's theorem,

$$\begin{aligned}
E\psi(S_n + v(u)) &\geq E\psi(S_n + v(u))\mathbf{1}_G \\
&= E(\psi(S_{n-1} + v(u) + v(U)) \,|\, U \geq \overline{u}_{1/n})P(G) \\
&= E(\psi(S_{n-1} + v(U)) \,|\, U \geq \overline{u}_{1/n})P(G) \pm 2\|\psi'\|_\infty \gamma n^{-2}.
\end{aligned}$$

(This uses the fact that $|\psi'|$ is uniformly bounded, a consequence of the standing assumption that ψ is C^∞ with compact support). Now another use of Taylor's theorem shows that, for $H = \{U_n \in [-\epsilon, \epsilon]\}$,

$$\begin{aligned}
&E(\psi(S_{n-1} + v(U)) \,|\, U \geq \overline{u}_{1/n}) \\
&\qquad \geq E(\psi(S_n + v(U))\mathbf{1}_H \,|\, U \geq \overline{u}_{1/n}) \pm 2\|\psi'\|_\infty \gamma n^{-2}.
\end{aligned}$$

Because the events G, H both have probabilities bounded away from 0, to complete the proof it will suffice to show that

$$E(\psi(S_n + v(U))\mathbf{1}_H \,|\, U \geq \overline{u}_{1/n}) \geq \frac{1}{4}P(H)E(\psi(S_n + v(U)) \,|\, U \geq \overline{u}_{1/n}). \tag{18}$$

Denote by M the number of voters in the sample $U_1, U_2, \ldots, U_n$ with values between $-\epsilon$ and ϵ; this random variable has the Bernoulli-(n, p_ϵ) distribution, where $p_\epsilon = P(H) = F(\epsilon) - F(-\epsilon)$. Conditioning on the value of M, we obtain (with $q_\epsilon := 1 - p_\epsilon$)

$$\begin{aligned}
E(\psi(S_n + v(U)) \,|\, U \geq \overline{u}_{1/n}) &= \sum_{m=0}^{n} \binom{n}{m} p_\epsilon^m q_\epsilon^{n-m} E_m \quad \text{and} \\
E(\psi(S_n + v(U))\mathbf{1}_H \,|\, U \geq \overline{u}_{1/n}) &= \sum_{m=1}^{n} \binom{n-1}{m-1} p_\epsilon^m q_\epsilon^{n-m} E_m
\end{aligned}$$

where

$$E_m = E(\psi(S_n + v(U))\mathbf{1}_H \,|\, U \geq \overline{u}_{1/n} \text{ and } M = m).$$

The contribution to these sums from those terms with $m \leq p_\epsilon n/2$ is exponentially small in n, by a standard large deviations inequality for the Binomial distribution (Hoeffding, 1963), and hence is asymptotically negligible compared to the first conditional expectation, because this is at least $2\gamma/(n\overline{u})$. On the other hand, for all $m \geq np_\epsilon/2$,

$$\frac{\binom{n-1}{m-1}}{\binom{n}{m}} = \frac{m}{n} \geq \frac{p_\epsilon}{2},$$

and so the inequality (18) follows. $\square$

The following lemma states, roughly, that if it is optimal for some agent in the bulk of the population to buy a moderately large number of votes, then *most* agents will be forced to buy a moderately large number of votes.

Lemma 6. For every $\epsilon > 0$ and all sufficiently small $\alpha > 0$, if n is sufficiently large then in any Nash equilibrium $v(\cdot)$,

$$E\psi(v(u) + S_n) \geq (1 - \alpha)E\psi(v(u') + S_n) \tag{19}$$

for any two values u, u' not within distance ϵ of either $\underline{u}$, or $\overline{u}$, or 0.

Proof. The main idea is that, for an agent with utility $U_i = u$ not in the tails of the distribution F, the empirical distribution of any symmetric monotone aggregation $U_1, U_2, \ldots, U_{n+1}$ conditional on the agent's value $U_i = u$ is not appreciably different from the *unconditional* distribution; that is, the agent gets very little information from knowing her own utility value u.

Fix two non-overlapping intervals J_1, J_2 contained in $(\underline{u}, \overline{u})$, and let $p_i = \int_{J_i} f$ be the probabilities assigned to these intervals by F. By our standing hypotheses on f, these probabilities are *positive*. Denote by M_i the number of points in the sample $U_1, U_2, \ldots, U_n$ that fall in the interval J_i, and let $U = U_{n+1}$ be independent of $U_1, U_2, \ldots, U_n$. Then, with $p_3 = 1 - p_1 - p_2$,

$$E(\psi(v(U) + S_n) \,|\, U \in J_i) = \sum_{m_1, m_2} \binom{n}{m_1, m_2} p_1^{m_1} p_2^{m_2} p_3^{n - m_1 - m_2} E_i(m_1, m_2) \tag{20}$$

where

$$E_i(m_1, m_2) = E(\psi(v(U) + S_n) \,|\, U \in J_i, M_1 = m_1, M_2 = m_2) \quad \text{and}$$
$$\binom{n}{m_1, m_2} = \frac{n!}{m_1! m_2! (n - m_1 - m_2)!}$$

Observe that for each $i = 1, 2$ the random variable M_i is binomial$-(n, p_i)$, and so large deviations estimates (cf. Hoeffding (1963)) for the binomial distribution imply that for any $0 < \varrho < p_i/2$ the probability that $M_i/n \notin [p_i - \varrho, p_i + \varrho]$ decays exponentially in n, at an exponential rate $\beta > 0$ that depends on ϱ but not on p_i. Because the function ψ is bounded, those terms in the sum (20) such that $|m_i/n - p_i| > \varrho$ for either $i = 1, 2$ contribute at most $e^{-\beta n}$ in absolute value to the sum, for some $\beta > 0$ depending only on ϱ.

Now conditional on $M_1 = m_1$ and $M_2 = m_2$, the sample $U_1, U_2, \ldots, U_n$ is obtained by (i) choosing m_1 points at random according to the conditional distribution of U given $U \in J_1$; (ii) independently choosing m_1 points at random according to the conditional distribution of U given $U \in J_1$; and then (iii) choosing the remaining $n - m_1 - m_2$ points according to the conditional distribution of U given $U \notin J_1 \cup J_2$. Consequently,

$$E_1(m_1, m_2) = E_2(m_1 + 1, m_2 - 1).$$

Next, observe that for any pair m_1, m_2 of integers satisfying $|m_i/n - p_i| \leq \varrho$, the ratio

$$\left\{ \binom{n}{m_1, m_2} p_1^{m_1} p_2^{m_2} p_3^{n - m_1 - m_2} \right\} \Big/ \left\{ \binom{n}{m_1 + 1, m_2 - 1} p_1^{m_1 + 1} p_2^{m_2 - 1} p_3^{n - m_1 - m_2} \right\}$$

is bounded above and below by $(1 - \alpha)^{\pm 1}$ for all sufficiently large n, provided $\varrho > 0$ is sufficiently small. It now follows that for all large n,

$$E(\psi(v(U)+S_n)\,|\,U\in J_1)\geq(1-\alpha)E(\psi(v(U)+S_n)\,|\,U\in J_2)-e^{-\beta n}. \tag{21}$$

By Lemma 5, if J_2 does not intersect $[-\epsilon,\epsilon]$ then

$$E\psi(v(U)+S_n)\,|\,U\in J_2)\geq C\epsilon/n^2,$$

and so the exponential error $e^{-\beta n}$ is asymptotically negligible compared to the first term on the right side of inequality (21). Because (21) holds for all choices of intervals J_1, J_2 (subject only to the restriction $p_i/2>\varrho$) it follows that for any $\alpha'>\alpha$,

$$E(\psi(v(U)+S_n)\,|\,U\in J_1)\geq(1-\alpha')E(\psi(v(U)+S_n)\,|\,U\in J_2), \tag{22}$$

provided n is sufficiently large.

Finally, it remains to deduce the bound (19) from (22). For this, we recall that (i) the necessary condition (15) implies $E\psi(v(u)+S_n)=v(u)/u$, provided $u\neq 0$, and (ii) the function $v(u)$ is non-decreasing. Consequently, for any interval J,

$$\begin{aligned}E(\psi(v(U)+S_n)\,|\,U\in J)&=\int_J E(\psi(v(u)+S_n))\,f(u)\,du\Big/\int_J f(u)\,du\\&=\int_J(2v(u)/u)\,f(u)\,du\Big/\int_J f(u)\,du.\end{aligned} \tag{23}$$

Fix $u>0$ such that $(1-\alpha)u>\epsilon$ and $(1-\alpha)^{-1}u<\overline{u}$. Using (23) for each of the intervals $J_1=[(1-\alpha)u,u]$ and $J_2=[u,(1-\alpha)^{-1}u]$ along with the monotonicity of $v(u)$, we obtain

$$(1-\alpha)E(\psi(v(U)+S_n)\,|\,U\in J_1)\leq\psi(v(u)+S_n)\leq(1-\alpha)^{-1}E(\psi(v(U)+S_n)\,|\,U\in J_2).$$

These inequalities also hold, by virtually the same argument, for $u<0$ such that $(1-\alpha)u<\epsilon$ and $(1-\alpha)^{-1}u>\underline{u}$. The inequality (19) now follows from (21) (with a suitable adjustment of α). □

C.2 Concentration and size constraints

Because the vote total S_n is the sum of independent, identically distributed random variables $v(U_i)$ (albeit with unknown distribution), its distribution is subject to concentration restrictions, such as those imposed by the following lemma.

Lemma 7. For any $\epsilon>0$ a constant $\gamma=\gamma(\epsilon)<\infty$ exists such that for all sufficiently large values of n and any Nash equilibrium $v(u)$, if

$$\max\left(v(\overline{u}-\epsilon),-v(\underline{u}+\epsilon)\right)\geq\gamma/\sqrt{n}, \tag{24}$$

then

$$P\{|S_n+v|\leq\delta\}<\epsilon\quad\text{for all } v\in\mathbb{R} \tag{25}$$

and therefore

$$\frac{|2v(u)|}{|u|} \le \epsilon \|\psi\|_\infty \quad \text{for all } u \in [\underline{u}, \overline{u}] \,. \tag{26}$$

We will deduce Lemma 7 from the following general fact about sums of independent, identically distributed random variables.

Lemma 8. Fix $\alpha > 0$. For any $\epsilon > 0$ and any $C < \infty$ there exists $C' = C'(\epsilon, C) > 0$ and $n' = n'(\epsilon, C) < \infty$ such that the following statement is true: if $n \ge n'$ and $Y_1, Y_2, \dots, Y_n$ are independent random variables such that

$$E|Y_1 - EY_1|^3 \le C \mathrm{var}(Y_1)^{3/2} \quad \text{and} \quad \mathrm{var}(Y_1) \ge C'/n \tag{27}$$

then for every interval $J \subset \mathbb{R}$ of length α or greater,

$$P\left\{ \sum_{i=1}^{n} Y_i \in J \right\} \le \epsilon |J|/\alpha. \tag{28}$$

The proof of this lemma, a routine exercise in the use of Fourier methods, is relegated to Section F, at the end of this appendix.

Proof of Lemma 7. Inequality (26) follows from (25), by the necessary condition (15) for Nash equilibria. Hence, it suffices to show that (24) implies (25).

By Lemma 6, constants $\alpha, \beta > 0$ exist such that for any point $u \in [\underline{u}_\epsilon, \overline{u}_\epsilon] \setminus [-2\epsilon, 2\epsilon]$ the ratio $v(u)/u$ is at least $\alpha v(\overline{u}_\epsilon)/\overline{u}_\epsilon$. Because the density f is bounded below, it follows that for suitable constants $0 < C < \infty$ and $\frac{1}{2} > p > 0$, for all sufficiently large n and every Nash equilibrium $v(u)$ an interval $J_+ = [u', \overline{u}_\epsilon]$ of $F-$probability p exists such that

$$v(\overline{u}_\epsilon) \le C v(u'). \tag{29}$$

Similarly, an interval $J_- = [\underline{u}_\epsilon, u^*]$ of probability p exists such that

$$|v(\underline{u}_\epsilon)| \le C |v(u^*)|. \tag{30}$$

Let M be the number of points U_i in the sample $U_1, U_2, \dots, U_n$ that fall in $J_+ \cup J_-$, and let S_n^* be the sum of the votes $v(U_i)$ for those agents whose utility values fall in this range. Observe that M has the binomial-$(n, 2p)$ distribution, and that conditional on the event $M = m$ and $S_n - S_n^* = w$, the random variable S_n^* is the sum of m independent random variables Y_i whose variance is at least $v(u')^2/4$ and whose third moment obeys the restriction (27) (this follows from the inequalities (29)–(30)). Consequently, by Lemma 8, if $v(u_\epsilon)\sqrt{n}$ is sufficiently large then the conditional probability, given $M = n \ge np$ and $S_n - S_n^* = w$, that S_n^* lies in any interval of length δ is bounded above by $\epsilon/2$. Because $P\{M \le np\}$ is, for large n, much less than $\epsilon/2$, the inequality (25) follows.

□

Lemma 7 implies that for any $\epsilon > 0$, if n is sufficiently large then for any Nash equilibrium $v(u)$, the absolute value $|v(u)|$ can assume large values only at utility values u within distance ϵ of one of the endpoints $\underline{u}, \overline{u}$. The following proposition improves this bound to the extreme tails of the distribution.

Lemma 9. For any $0 < C < \infty$ a $C' > 0$ exists such that for all sufficiently large n, every Nash equilibrium $v(u)$ satisfies the inequality

$$|v(u)| \leq C \quad \text{for all} \quad u \in \left[\underline{u} + C'n^{-3/2}, \overline{u} - C'n^{-3/2}\right]. \tag{31}$$

Proof. Fix $C > 0$, and suppose $2v(u_*) \geq C$ for some $u_* > 0$. Because any Nash equilibrium v is monotone, we must have $2v(u) \geq C$ for all $u \geq u_*$, and by the necessary condition (15) it follows that

$$E\psi(v(u) + S_n)u \geq C \quad \Longrightarrow \quad E\psi(v(u) + S_n) \geq C/\overline{u} \quad \forall\, u \geq u_*. \tag{32}$$

Consequently, the probability that $S_n + v(u) \in [-\delta, \delta]$ must be at least $C/\overline{u}\|\psi\|_\infty$. Thus, Lemma 7 implies that for any $\epsilon > 0$ a $\gamma_\epsilon > 0$ exists (depending on both ϵ and C, but not on n) such that

$$\max\left(-v\left(\underline{u}_\epsilon\right), v(\overline{u}_\epsilon)\right) \leq \gamma_\epsilon/\sqrt{n}. \tag{33}$$

In particular, for all sufficiently large n,

$$\begin{aligned} v(\overline{u}/2) \leq \frac{\gamma_{\overline{u}/2}}{\sqrt{n}} \quad &\Longrightarrow \quad E\psi(v(\overline{u}/2) + S_n) \leq \frac{2\gamma_{\overline{u}/2}}{\overline{u}\sqrt{n}} \\ &\Longrightarrow \quad E\psi(S_n) \leq \frac{2\gamma_{\overline{u}/2}}{\overline{u}\sqrt{n}} + \|\psi'\|_\infty v(\overline{u}/2) \\ &\Longrightarrow \quad E\psi(S_n) \leq \frac{C_{\overline{u}/2}}{\sqrt{n}} \end{aligned} \tag{34}$$

for a constant $C_{\overline{u}/2} < \infty$ that may depend on $\overline{u}/2$ and C but not on either n or the particular Nash equilibrium.

Fix C' large, and suppose $2v(u_*) \geq C$ for $u_* = \overline{u} - C'n^{-3/2}$. Let M_* be the number of points U_i in the sample $U_1, U_2, \dots, U_n$ that fall in the interval $[u_*, \overline{u}]$; by our assumptions concerning the sampling procedure, the random variable M_* has the binomial distribution with mean

$$EM_* = n\int_{u_*}^{\overline{u}} f(u)\, du \sim C'f(\overline{u})n^{-1/2} \quad \text{for large } n.$$

This is vanishingly small for large n, so the assumption $v(u_*) \geq C$ implies

$$E\psi(v(u) + S_n)\mathbf{1}\{M_* = 0\} \geq C/2\overline{u} \quad \text{for all } u \geq u_*. \tag{35}$$

This expectation can be decomposed by partitioning the probability space into the event $G = \{U_n \in [\underline{u} + \epsilon, \overline{u} - \epsilon]\}$ and its complement. On the event G, the contribution of $v(U_n)$ to the vote total S_n is at most $\gamma_\epsilon/\sqrt{n}$ in absolute value, by (33). On the complementary event G^c the integrand is bounded above by $\|\psi\|_\infty$. Therefore,

$$\begin{aligned} E\psi(v(u) + S_n)\mathbf{1}\{M_* = 0\} &\leq P(G^c)\|\psi\|_\infty + E\psi(v(u) + S_n)\mathbf{1}\{M_* = 0\}\mathbf{1}_G \\ &\leq P(G^c)\|\psi\|_\infty + E\psi(v(u) + S_{n-1})\mathbf{1}\{M_* = 0\} + \|\psi'\|_\infty(\gamma_\epsilon/\sqrt{n}) \\ &\leq \epsilon' + E\psi(v(u) + S_{n-1})\mathbf{1}\{M_* = 0\} \end{aligned}$$

where $\epsilon' > 0$ can be made arbitrarily small by choosing $\epsilon > 0$ small and n large. This together with inequality (35) implies that for large n,

$$E\psi(v(u)+S_{n-1})\mathbf{1}\{M_*=0\} \geq C/4\overline{u} \quad \text{for all } u \geq u_*. \tag{36}$$

Now consider the conditional distribution of S_n given that $M_* = 1$, which can be simulated by generating S_{n-1} from the conditional distribution of S_{n-1} given that $M_* = 0$ and then adding an independent $v(U)$ where $U = U_n$ is drawn from the conditional distribution of U given that $U \geq u_*$. Consequently, by inequality (36),

$$E(\psi(S_n)\,|\,M_*=1) = E(\psi(S_{n-1}+v(U))\,|\,M_*=0 \,\cap\, U \geq u_*) \geq C/4\overline{u}.$$

But this equation implies

$$E(\psi(S_n)) \geq (C/4\overline{u})P\{M_* \geq 1\} \sim CC'f(\overline{u}))/(4\overline{u}\sqrt{n}).$$

For large C' this arbitrarily precise approximation is incompatible with inequality (34) when n is sufficiently large. □

C.3 Discontinuities

Because any Nash equilibrium $v(u)$ is monotone in the utility u, it can have at most countably many discontinuities. Moreover, because any Nash equilibrium is bounded in absolute value by $\sqrt{2\max(|\underline{u}|,\overline{u})}$ (because no agent will pay more for votes than she could gain in expected utility) the sum of the jumps is bounded by $\sqrt{2\max(|\underline{u}|,\overline{u})}$. We will now establish a *lower* on the size of $|v|$ at a discontinuity.

Lemma 10. Let $v(u)$ be a Nash equilibrium. If v is discontinuous at $u \in (\underline{u}, \overline{u})$ then

$$E\psi'(\tilde{v}+S_n)u = 2 \tag{37}$$

for some $\tilde{v} \in [v_-, v_+]$, where v_- and v_+ are the left and right limits of $v(u')$ as $u' \to u$.

Proof. The necessary condition (15) holds at all u' in a neighborhood of u, so by monotonicity of v and continuity of ψ, Equation (15) must hold when $v(u)$ is replaced by either of $v_\pm$, that is,

$$\begin{aligned} 2v_+ &= E\psi(v_+ + S_n)u \quad \text{and} \\ 2v_- &= E\psi(v_- + S_n)u. \end{aligned}$$

Subtracting one equation from the other and using the differentiability of ψ we obtain

$$2v_+ - 2v_- = uE\int_{v_-}^{v_+} \psi'(t+S_n)\,dt = u\int_{v_-}^{v_+} E\psi'(t+S_n)\,dt.$$

The result then follows from the mean value theorem of calculus. □

Lemma 11. A constant $\Delta > 0$ exists such that for all sufficiently large n, at any point u of discontinuity of a Nash equilibrium,

$$\begin{aligned} v(u_+) &\geq \Delta \quad \text{if } u \geq 0 \quad \text{and} \\ v(u_-) &\leq -\Delta \quad \text{if } u \leq 0. \end{aligned} \tag{38}$$

Consequently, a constant $\beta < \infty$ independent of the sample size exists on the sample size n such that for all sufficiently large n no Nash equilibrium $v(u)$ has a discontinuity at a point u at distance greater than $\beta n^{-3/2}$ from one of the endpoints $\underline{u}, \overline{u}$.

Proof. Because the function ψ has support contained in the interval $[-\delta, \delta]$, equation (37) implies v can have a discontinuity only if the distribution of S_n is highly concentrated: specifically,

$$P\{S_n + \tilde{v} \in [-\delta, \delta]\} \geq \frac{2}{\|\psi'\| \max\left(|\underline{u}|, \overline{u}\right)}. \tag{39}$$

In fact, because ψ' vanishes at the endpoints of $[-\delta, \delta]$, a $0 < \delta' < \delta$ exists such that

$$P\{S_n + \tilde{v} \in [-\delta', \delta']\} \geq \frac{1}{\|\psi'\| \max\left(|\underline{u}|, \overline{u}\right)}. \tag{40}$$

Lemma 7 asserts that such strong concentration of the distribution of S_n can occur only if $|v(u)|$ is vanishingly small in the interior of the interval $[\underline{u}, \overline{u}]$. In particular, if $\epsilon < \left(\|\psi'\| \max\left(|\underline{u}|, \overline{u}\right)\right)^{-1}$ and n is sufficiently large then $|v(u)| < \gamma_\epsilon/\sqrt{n}$ for all $u \in [\underline{u}_\epsilon, \overline{u}_\epsilon]$. But $v(u)$ must satisfy the necessary condition (15) at all such u, so

$$E\psi(v(u) + S_n)|u| \leq 2\gamma_\epsilon/\sqrt{n}$$

for all $u \in [\underline{u}_\epsilon, \overline{u}_\epsilon]$. Because the function ψ is positive and bounded away from 0 in any interval $[-\delta'', \delta'']$ where $0 < \delta'' < \delta$, it follows from (40) that for sufficiently large n,

$$|\tilde{v}| \geq (\delta - \delta')/3 := \Delta.$$

Thus, by the monotonicity of Nash equilibria, at every point u of discontinuity we must have (38). Lemma 9 now implies any such discontinuities can occur only within a distance $\beta n^{-3/2}$ of one of the endpoints $\underline{u}, \overline{u}$. $\square$

C.4 Smoothness

Because Nash equilibria are monotone, by Lemma 3, they are necessarily differentiable almost everywhere. We will show that in fact differentiability must hold at *every* u, except near the endpoints $\underline{u}, \overline{u}$.

Lemma 12. If $v(u)$ is a Nash equilibrium then at every u where v is differentiable,

$$E\psi(S_n + v(u)) + E\psi'(S_n + v(u))uv'(u) = 2v'(u). \tag{41}$$

Proof. Given the smoothness of the function ψ, the result follows from the chain and product rules. $\square$

Equation (41) can be rewritten as a first-order differential equation:

$$v'(u) = \frac{E\psi(S_n + v(u))}{2 - E\psi'(S_n + v(u))u}. \tag{42}$$

This differential equation becomes singular at any point where the denominator approaches 0, but is regular in any interval where $E\psi'(S_n+v(u))u \leq 1$. The following lemma implies regularity on any interval where $|v(u)|$ remains sufficiently small.

Lemma 13. For any $\alpha > 0$ a constant $\beta = \beta_\alpha > 0$ exists such that for any strategy $v(u)$, any $\tilde{v} \in \mathbb{R}$, any $u \in [\underline{u}, \overline{u}]$, and all n,

$$\begin{aligned} E|\psi'(\tilde{v} + S_n)u| \geq \alpha \quad &\Longrightarrow \quad E\psi(\tilde{v} + S_n)|u| \geq \beta \text{ and} \\ E|\psi''(\tilde{v} + S_n)u| \geq \alpha \quad &\Longrightarrow \quad E\psi(\tilde{v} + S_n)|u| \geq \beta. \end{aligned} \tag{43}$$

Proof. Recall that $\psi/2$ is a C^∞ probability density with support $[-\delta, \delta]$ and such that ψ is *strictly* positive in the open interval $(-\delta, \delta)$. Consequently, on any interval $J \subset (-\delta, \delta)$ where $|\psi'|$ (or $|\psi''|$) is bounded below by a positive number, so is ψ.

Fix $\epsilon > 0$ so small that $\epsilon \max(\underline{u}, \overline{u}) < \alpha/2$. In order that $E|\psi'(\tilde{v} + S_n)u| \geq \alpha$, it must be the case that the event $\{|\psi'(\tilde{v} + S_n)| \geq \epsilon\}$ contributes at least $\alpha/2$ to the expectation; hence,

$$P\{|\psi'(\tilde{v} + S_n)| \geq \epsilon\} \geq \frac{\alpha}{2\|\psi'\|_\infty \max(\underline{u}, \overline{u})}.$$

But on this event the random variable $\psi(\tilde{v} + S_n)$ is bounded below by a positive number $\eta = \eta_\epsilon$, so it follows that

$$E\psi(\tilde{v} + S_n)|u| \geq \frac{\eta\alpha}{2\|\psi'\|_\infty \max(\underline{u}, \overline{u})}.$$

A similar argument proves the corresponding result for ψ''. □

Lemma 14. Constants $C, \alpha > 0$ exist such that for all sufficiently large n, any Nash equilibrium $v(u)$ is continuously differentiable on any interval where $|v(u)| \leq C$ (and therefore, by Lemma 9, on $\left(\underline{u} + C'n^{-3/2}, \overline{u} - C'n^{-3/2}\right)$), and the derivative satisfies

$$\alpha \leq \frac{v'(u)}{E\psi(v(u) + S_n)} \leq \alpha^{-1}. \tag{44}$$

Proof. The function $v(u)$ is differentiable almost everywhere, by Lemma 3, and at every point u where $v(u)$ is differentiable the differential equation (42) holds. By Lemma 11, the sizes of discontinuities are bounded below, and so if $C > 0$ is sufficiently small then a Nash equilibrium $v(u)$ can have no discontinuities on any interval where $|v(u)| \leq C$. Furthermore, if $C > 0$ is sufficiently small then by Lemma 13 and the necessary condition (15), we must have $E\psi'(v(u) + S_n) \leq 1$ on any interval where $|v(u)| \leq C$. Because the functions $v \mapsto E\psi(S_n + v)$ and $v \mapsto E\psi'(S_n + v)$ are continuous (by dominated convergence), it now follows from Equation (42) that if $C > 0$ is sufficiently small then on any interval where $|v(u)| \leq C$ the function $v'(u)$ extends to a continuous function. Finally, because the denominator in equation (42) is at least 1 and no larger than $2 + \|\psi'\|_\infty$, the inequalities (44) follow. □

Similar arguments show that Nash equilibria have derivatives of higher orders provided the sample size is sufficiently large. The proof of Theorem 1 in Section E below will require information about the second derivative $v''(u)$, which can be obtained by differentiating under the expectations in (42):

$$v''(u) = \frac{E\psi'(v(u)+S_n)v'(u)}{2-E\psi'(S_n+v(u))u} + \frac{E\psi(v(u)+S_n)(E\psi''(v(u)+S_n)v'(u)u+E\psi'(v(u)+S_n)}{(2-E\psi'(S_n+v(u)u))^2}. \tag{45}$$

A repetition of the proof of Lemma 14 now shows that for suitable constants $C, \beta > 0$ and all sufficiently large n, any Nash equilibrium $v(u)$ is twice continuously differentiable on any interval where $|v(u)| \leq C$ and satisfies the inequalities

$$\beta \leq \frac{v''(u)}{E\psi(v(u)+S_n)} \leq \beta^{-1}. \tag{46}$$

C.5 Approximate proportionality

The information that we now have about the form of Nash equilibria can be used to sharpen the heuristic argument given in Subsection C.1 to support the "approximate proportionality rule". Recall that in a Nash equilibrium the number of votes $v(u)$ purchased by an agent with utility u must satisfy the equation $2v(u) = E\psi(v(u)+S_n)u$. We have shown in Lemma 9 that for any Nash equilibrium, $v(u)$ must be small except in the extreme tails of the distribution (in particular, for all u at distance much more than $n^{-3/2}$ from both endpoints $\underline{u}, \overline{u}$). Because ψ is uniformly continuous, it follows that the expectation $E\psi(v(u)+S_n)$ cannot differ by very much from $E\psi(S_n)$.

Unfortunately, this argument only shows that the approximation $2v(u) \approx E\psi(S_n)u$ is valid up to an error of size $\epsilon_n|u|$ where $\epsilon_n \to 0$ as $n \to \infty$. However, as $n \to \infty$ the expectation $E\psi(S_n) \to 0$, and so the error in the approximation above might be considerably larger than the approximation itself. Proposition 3 makes the stronger assertion that when n is large the relative error in the approximate proportionality rule is small.

Proposition 3. For any $\epsilon > 0$ constants $n_\epsilon < \infty$ and $C < \infty$ exist such that if $n \geq n_\epsilon$ then for any Nash equilibrium $v(u)$ and for all $u \in \left[\underline{u} + Cn^{-3/2}, \overline{u} - Cn^{-3/2}\right]$,

$$(1-\epsilon)E\psi(S_n)|u| \leq |2v(u)| \leq (1+\epsilon)E\psi(S_n)|u|. \tag{47}$$

Furthermore, for all sufficiently large n any Nash equilibrium $v(u)$ with no discontinuities must satisfy (47) for all $u \in [\underline{u}, \overline{u}]$.

Proof of Proposition 3. Because ψ has compact support, it and all of its derivatives are uniformly continuous and uniformly bounded, and so the function $v \mapsto E\psi(v+S_n)$ is differentiable with derivative $E\psi'(v+S_n)$. Consequently, by Taylor's theorem, for every u a $\tilde{v}(u)$ exists intermediate between 0 and $v(u)$ such that

$$2v(u) = E\psi(v(u)+S_n)u = E\psi(S_n)u + E\psi'(\tilde{v}(u)+S_n)v(u)u. \tag{48}$$

We will argue that for all $C > 0$ sufficiently small, if $|2v(u)| \leq C$ then the expectation $E\psi'(\tilde{v}(u)+S_n)$ remains below ϵ in absolute value, provided n is sufficiently large. Lemma 9 will then imply a $C' < \infty$ exists such that (47) holds for all $u \in (\underline{u}, \overline{u})$ at distance greater than $C'n^{-3/2}$ from the endpoints $\underline{u}, \overline{u}$.

If $|2v(u)| \leq C$ then $|E\psi(v(u) + S_n)| \leq C/\max(|\underline{u}|, \overline{u})$, by the necessary condition (15). By Lemma 11, if $C < \Delta$, where Δ is the discontinuity threshold, then $v(u)$ is continuous on any interval $[0, u_C]$ where $|v(u)| \leq C$, and so for each u in this interval a $u' \in [0, u]$ exists such that $\tilde{v}(u) = v(u')$. Consequently, $|E\psi(\tilde{v}(u) + S_n)| \leq C/\max(|\underline{u}|, \overline{u})$. But Lemma 13 implies that for any $\epsilon > 0$, if $C > 0$ is sufficiently small then for all large n and any Nash equilibrium $v(u)$,

$$|E\psi'(\tilde{v}(u) + S_n)| < \epsilon$$

on any interval $[0, u_C]$ where $|v(u)| \leq C$. Thus, the error in the approximation (48) will be small when n is large and $|2v(u)| < C$, for $u > 0$. A similar argument applies for $u \leq 0$. This proves that (47) holds for all $u \in (\underline{u}, \overline{u})$ at distance greater than $C'n^{-3/2}$ from the endpoints $\underline{u}, \overline{u}$.

Finally, suppose $v(u)$ is a Nash equilibrium with no discontinuities. By Lemma 9, for any $C > 0$ a $C' < \infty$ exists such that $|v(u)| \leq C/2$ except at arguments u within distance $C'/n^{3/2}$ of one of the endpoints. Moreover, Lemma 14 implies that if C is sufficiently small then on any interval where $|v(u)| \leq C$ the function v is differentiable, with derivative $v'(u) \leq C''$ for some constant $C'' < \infty$ not depending on n or on the particular Nash equilibrium. Because $v(u)$ is continuous up to $\overline{u}$, if $v(u) \geq C$ for some $u > \overline{u} - C'n^{-3/2}$ then by the intermediate value theorem there would a smallest point $u' \in [\overline{u} - C'n^{-3/2}, \overline{u}]$ at which $v(u') = C$. But then v would be differentiable all the way up to u', with derivative bounded above by C'', and so

$$\begin{aligned}
C = v(u') &= v(\overline{u} - C'n^{-3/2}) + \int_{\overline{u} - C'n^{-3/2}}^{u'} v'(u)\, du \\
&\leq C/2 + \int_{\overline{u} - C'n^{-3/2}}^{u'} v'(u)\, du \\
&\leq C/2 + C''C'n^{-3/2},
\end{aligned}$$

which is impossible for large n. A similar argument shows that for large n, if $v(u)$ has no discontinuities then v cannot attain the value $-C$ near $\underline{u}$. Therefore, for all sufficiently large n, if v has no discontinuities then $|v(u)|C$ for all $u \in [\underline{u}, \overline{u}]$, and so by the preceding argument it follows that $v(u)$ must satisfy the proportionality relations (47) for all $u \in [\underline{u}, \overline{u}]$. □

C.6 Consequences of Proposition 3

Proposition 3 puts strong constraints on the distribution of the vote total S_n in a Nash equilibrium. According to this proposition, the approximate proportionality rule (47) holds for all $u \in [\underline{u}, \overline{u}]$ except those values u within distance $Cn^{-3/2}$ of one of the endpoints $\underline{u}, \overline{u}$. Call this the *extremist range*. Denote by G the event that the sample $U_1, U_2, \ldots, U_n$ contains no values in the extremist range. By Proposition 3, on the event G the approximate proportionality rule (47) will apply for each agent; furthermore, for Nash equilibria with no discontinuities, (47) holds for all $u \in [\underline{u}, \overline{u}]$. Thus, *conditional* on the event G (or, for continuous Nash equilibria, *unconditionally*) the random variables $v(U_i)$ are (at least for sufficiently large n) bounded above and below by $E\psi(S_n)\overline{u}$ and $E\psi(S_n)\underline{u}$, and so Hoeffding's inequality applies.

Corollary 2. Let G be the event that the sample U_i contains no values u in the extremist range.

Then for all sufficiently large n and any Nash equilibrium $v(u)$,

$$P(|S_n - ES_n| \geq tE\psi(S_n) \,|\, G) \leq \exp\{-2t^2/n \max\left(|\underline{u}|^2, \overline{u}^2\right)\}; \tag{49}$$

and for any Nash equilibrium with no discontinuities,

$$P(|S_n - ES_n| \geq tE\psi(S_n)) \leq \exp\{-2t^2/n \max\left(|\underline{u}|^2, \overline{u}^2\right)\}. \tag{50}$$

Proposition 3 also implies uniformity in the normal approximation to the distribution of S_n, because the proportionality rule (47) guarantees that the ratio of the third moment to the $3/2$ power of the variance of $v(U_i)$ is uniformly bounded. Hence, by the Berry-Esseen theorem, we have the following corollary.

Corollary 3. A $\kappa < \infty$ exists such that for all sufficiently large n and any Nash equilibrium $v(u)$, the vote total S_n satisfies

$$\sup_{t \in \mathbb{R}} |P((S_n - ES_n) \leq t\sqrt{\text{var}(S_n)} \,|\, G) - \Phi(t)| \leq \kappa n^{-1/2}; \tag{51}$$

and for any Nash equilibrium with no discontinuities,

$$\sup_{t \in \mathbb{R}} |P((S_n - ES_n) \leq t\sqrt{\text{var}(S_n)}) - \Phi(t)| \leq \kappa n^{-1/2}. \tag{52}$$

Here Φ denotes the standard normal cumulative distribution function.

D Unbalanced Populations: Proofs of Theorems 2–3

D.1 Concentration of the vote total

Lemma 15. If $\mu > 0$ then for all large n no Nash equilibrium $v(u)$ has a discontinuity at a nonnegative value of u. Moreover, for any $\epsilon > 0$, if n is sufficiently large then in any Nash equilibrium the vote total S_n must satisfy

$$\delta - \epsilon \leq ES_n \leq \delta + \epsilon + \sqrt{2\,|\underline{u}|]} \quad \text{and} \tag{53}$$

$$P\{|S_n - ES_n| > \epsilon\} < \epsilon. \tag{54}$$

In addition, a constant $\gamma > 0$ exists such that for any $\epsilon > 0$, if n is sufficiently large and $v(u)$ is a Nash equilibrium with no discontinuities, then

$$P\{|S_n - ES_n| > \epsilon\} < e^{-\gamma n}. \tag{55}$$

Proof. By Lemma 11, a Nash equilibrium $v(u)$ can have no discontinuities at distance greater than $Cn^{-3/2}$ of one of the endpoints $\underline{u}, \overline{u}$. Agents with such utilities are designated *extremists*; if G is the event that the sample $U_1, U_2, \ldots, U_n$ contains no extremists, then

$$P(G^c) \sim C(f(\underline{u}) + f(\overline{u}))/\sqrt{n}.$$

By Proposition 3, Nash equilibria $v(u)$ obey the approximate proportionality rule (47) except in the extremist range. The contribution of extremists to ES_n is vanishingly small for large n, because $P(G^c) = O(n^{-1/2})$, as no voter, even in the extremist range, will buy more than $\max(\sqrt{2\,|\underline{u}|}, \sqrt{2\overline{u}})$ votes. Consequently, (47) and the law of large numbers for the sequence $U_1, U_2, \dots$ imply that for any $\epsilon > 0$, if n is large then

$$E\psi(S_n)\mu(1-\epsilon) \leq ES_n/n \leq E\psi(S_n)\mu(1+\epsilon). \tag{56}$$

Because $\mu > 0$, this implies $ES_n \geq 0$ for all sufficiently large n.

Suppose now that $ES_n < \delta - 2\epsilon'$ for some small $\epsilon' > 0$. If $\epsilon > 0$ is sufficiently small relative to ϵ' then (56) implies $nE\psi(S_n)\mu \leq \delta - \epsilon'$. But then Hoeffding's inequality (49), together with the fact that $P(G^c) = O(n^{-1/2})$, implies

$$P\{S_n \in [-\delta/2, \delta - \epsilon'/2]\} \geq 1 - \epsilon$$

for large n. This conclusion, however, would contradict the hypothesis that $nE\psi(S_n) < \delta - \epsilon'/2$, because we would then have

$$E\psi(S_n) \geq (1-\epsilon) \min_{v \in [-\delta/2, \delta - \epsilon'/2]} \psi(v).$$

(Recall that ψ is bounded away from 0 on any compact sub-interval of $(-\delta, \delta)$.) This reasoning proves that for all large n and all Nash equilibria, $ES_n \geq \delta - 2\epsilon'$.

Next suppose $ES_n > \delta + \sqrt{2\,|\underline{u}|} + 2\epsilon'$, where $\epsilon' > 0$. The proportionality rule (47) (applied with some $\epsilon > 0$ small relative to ϵ') then implies $nE\psi(S_n) > \delta + \sqrt{2\,|\underline{u}|} + \epsilon'$. Hence, by the Hoeffding inequality (49), a $\gamma = \gamma(\epsilon') > 0$ exists such that

$$P(S_n \leq \delta + \sqrt{2\,|\underline{u}|} \,|\, G) \leq e^{-\gamma n},$$

because on the event $S_n \leq \delta + \sqrt{2\,|\underline{u}|}$ the sum S_n must deviate from its expectation by more than $nE\psi(S_n)\epsilon'$. Hence, for all $v \in [-\sqrt{2\,|\underline{u}|}, 0] \leq 0$,

$$E\psi(v + S_n) \leq e^{-\gamma n}\|\psi\|_\infty + P(G^c)\|\psi\|_\infty.$$

Thus, $|v\,(\underline{u})\,|$ must be vanishingly small, and so by Lemma 11 there can be no discontinuities in $[\underline{u}, 0]$. But this implies that the proportionality rule (47) holds for all $u \in \left[\underline{u}, \overline{u} - Cn^{-3/2}\right]$, and so another application of Hoeffding's inequality (coupled with the observation that $v(u)/u \geq (1-\epsilon)E\psi(S_n)$ holds for all $u \in [\underline{u}, \overline{u}]$ if v has no discontinuities at negative values of u) implies

$$P(S_n \leq \delta + \sqrt{2\,|\underline{u}|}) \leq e^{-\gamma n} \quad \Longrightarrow \quad E\psi(S_n) \leq e^{-\gamma n}\|\psi\|_\infty,$$

which is a contradiction. This proves assertion (53).

Because ES_n is now bounded away from 0 and ∞, it follows as before that $nE\psi(S_n)$ is bounded away from 0 and ∞, and so the proportionality rule (47) implies the conditional variance of S_n given the event G is $O(n^{-1})$. The assertion (54) therefore follows from Chebyshev's inequality and the bound $P(G^c) = O(n^{-1/2})$. Given (53) and (54), we can now conclude that there can be no discontinuities at nonnegative values of u, because in view of Lemma 11, the monotonicity of Nash equilibria, and the necessary condition (15), such discontinuities would entail that

$$E\psi(v(\overline{u}) + S_n)\overline{u} \geq 2\Delta,$$

which is incompatible with (53) and (54), because for small $\epsilon > 0$ the function $\psi(w)$ is uniformly small for $w \geq \delta - 2\epsilon$.

Finally, if v is a Nash equilibrium with no discontinuities then Corollary 2 implies the exponential bound (55). □

D.2 Proof of Theorem 2

Proof of Assertion (6). The asymptotic efficiency of quadratic voting in the unbalanced case $\mu > 0$ is a direct and easy consequence of Lemma 15. This implies that for any $\epsilon > 0$, the probability that the vote total $S_N = S_n + v(U_{n+1})$ will fall below $\delta - 2\epsilon$ is less than ϵ for all large n, and so by the continuity of the payoff function Ψ, for any $\epsilon > 0$

$$P\{\Psi(S_N) \leq 1 - \epsilon\} < \epsilon$$

for all sufficiently large N and all Nash equilibria. Moreover, the law of large numbers guarantees that for large N,

$$P\left\{|N^{-1}U - \mu| \geq \epsilon\right\} < \epsilon \quad \text{where} \quad U := \sum_{i=1}^{N} U_i.$$

Because the random variables U and $\Psi(S_N)$ are bounded, it therefore follows that for any $\epsilon > 0$, if N is sufficiently large then in any equilibrium

$$\left|\frac{E[U\Psi(S_N)]}{2E|U|} - 1\right| < \epsilon.$$

□

Proof of Assertions (7)–(8). The second assertion (8) will follow immediately from the first, by the law of large numbers for the sequence $U_1, U_2, \ldots$, and the assertion (7) will follow directly from the approximate proportionality rule (47) and the following estimate for $E\psi(S_n)$.

Claim: Let (α, w) be the solution of the Optimization Problem (10), if one exists, or let $\alpha = \delta$ if not. Then for any $\epsilon > 0$, if n is sufficiently large then in every Nash equilibrium,

$$\left|\frac{1}{2}E\psi(S_n) - \alpha\mu^{-1}\right| < \epsilon. \tag{57}$$

Proof of the Claim. Note first that this claim is equivalent to the assertion that $|ES_n - \alpha| \to 0$, by the proportionality rule (47). We will prove this in two steps, by first showing that for sufficiently large n the expectation ES_n cannot be smaller than $\alpha - 3\epsilon$, and then that it cannot be larger than $\alpha + 3\epsilon$.

Suppose first that $\alpha > \delta$ and that $ES_n < \alpha - 3\epsilon$. (If $\alpha = \delta$ then Lemma 15 implies $ES_n < \alpha - \epsilon$ is impossible for large n.) Then by (10), for any sufficiently small $\epsilon' > 0$,

$$(\Psi(\delta - \epsilon') - \Psi(w))|u| > (\alpha - \epsilon - w)^2 \tag{58}$$

for all u in a neighborhood $[\underline{u}, \underline{u} + \varrho]$, where $\varrho > 0$. By Lemma 15,

$$P\{\delta - \epsilon' \leq S_n \leq \alpha - 2\epsilon\} \longrightarrow 1 \quad \text{as} \;\; n \to \infty.$$

Consequently, a voter with value $u \leq \underline{u} + \varrho$ could, with probability near 1, improve her utility payoff from $\Psi(S_n)u \leq \Psi(\delta - \epsilon')u$ to $\Psi(w)u$, at a cost of $(\alpha - \epsilon - w)^2$, and so in view of (58), all voters with values $u \in [\underline{u}, \underline{u} + \varrho]$ would defect from the equilibrium strategy. This is a contradiction; hence, we conclude that for all large n, in any equilibrium, $ES_n \geq \alpha - 3\epsilon$.

Suppose now that $ES_n > \alpha + 3\epsilon$. Then by (10), for all $w \leq \delta$,

$$(\alpha + 2\epsilon - w)^2 \geq 4\epsilon^2 + (1 - \psi(w))|\underline{u}|.$$

By Lemma 15,

$$P\{\alpha + 2\epsilon \leq S_n\} \longrightarrow 1 \quad \text{as} \;\; n \to \infty;$$

hence, it would be sub-optimal for a voter with value $u = \underline{u}$ to buy more than $\Delta/2$ negative votes, where Δ is the discontinuity threshold. This implies, by Lemma 11, that for large n all Nash equilibria are continuous. But then assertion (55) would imply

$$P\{S_n \leq \alpha - 2\epsilon\} < e^{-\varrho n}$$

which in turn would ensure that

$$E\psi(S_n) \leq \|\psi\|_\infty e^{-\varrho n}.$$

This is impossible, because the proportionality rule (47) would then imply that for some constant $C < \infty$ not depending on n or the particular equilibrium, $\|v\|_\infty \leq Ce^{-\varrho n}$, contradicting Lemma 5. □

D.3 Proof of Theorem 3

Assume now that $\mu > 0$ and that the Optimization Problem (10) has a unique solution (α, w). We will first prove that for large n, every Nash equilibrium has a discontinuity u_* near $\underline{u}$, and then we will argue that this discontinuity must occur very near $\underline{u} + \zeta n^{-2}$ for a constant $\zeta > 0$ depending only on the payoff function Ψ and the sampling distribution F.

We have shown, in Subsection D.2, that for any $\epsilon > 0$, every Nash equilibrium must satisfy $|ES_n - \alpha| < \epsilon$. In addition, we have shown in Lemma 15 that if n is large and $v(u)$ is a Nash equilibrium with no discontinuities, then $P\{|S_n - ES_n| > \epsilon\}$ decays exponentially fast in n. Because $\alpha > \delta$, we may choose $\epsilon > 0$ so small that $\alpha - 3\epsilon > \delta$. It then follows that for a suitable constant $\gamma > 0$, if n is large and $v(u)$ is a Nash equilibrium with no discontinuities, then

$$P\{S_n \leq \delta\} \leq e^{-\gamma n}.$$

But as in the proof of Theorem 2, this would imply $E\psi(S_n)$ decays exponentially with n, which is impossible because, by the proportionality rule, it would contradict Lemma 5. Therefore, for large n every Nash equilibrium $v(u)$ has a discontinuity. Lemma 15 asserts that there are no discontinuities at points $u \in [0, \overline{u}]$, so any discontinuity must be located in $[\underline{u}, 0)$. Lemma 11 implies any such discontinuity must occur at a point within distance $O(n^{-3/2})$ of $\underline{u}$.

Let $v(u)$ be a Nash equilibrium, and let u_* be the rightmost point of discontinuity of v. By Lemma 11, the size of any discontinuity is at least Δ, so $v(u) < -\Delta$ for every $u < u_*$. Obviously, the expected payoff for an agent with utility u must exceed the expected payoff under the alternative strategy of buying no votes. The latter expectation is approximately $\underline{u}$, because S_n is highly concentrated near $ES_n > \alpha - \epsilon$ and so $E\Psi(S_n) \approx 1$. On the other hand, the expected payoff at $u < u_*$ for an agent playing the Nash strategy is approximately

$$\Psi(\alpha - v(u))\,(\underline{u}) - v(u)^2,$$

an improvement over the alternative strategy of buying no votes of about

$$(1 - \Psi(\alpha - v(u))|\underline{u}| - v(u)^2.$$

In order that this difference be nonnegative, it must be the case that $|v(u)| \approx \alpha - w$, because by hypothesis, (α, w) is the unique pair such that relations (10) hold. Because this approximation is valid for all $u \in [\underline{u}, u_*)$, it follows by Lemma 11 that $v(u)$ cannot have another discontinuity in the interval $[\underline{u}, u_*)$. This also proves assertion (ii) of Theorem 3, which states that that

$$v(\underline{u}) = -(\alpha - w).$$

Because u_* must be within distance $Cn^{-3/2}$ of $\underline{u}$, the probability that an extremist exists in the sample $U_1, U_2, \ldots, U_n$ is of order $nf(\underline{u})(u_* - \underline{u}) = O(n^{-1/2})$, and the conditional probability that more than one extremist exists given that at least one does is of order $O(n^{-1/2})$. Consequently, because the distribution of S_n is highly concentrated near α (cf. Lemma 15), where $\psi = 0$, the major contribution to the expectation $E\psi(S_n)$ comes from samples with exactly one extremist; hence,

$$E\psi(S_n + v(\underline{u})) \approx n\psi(w)f\,(\underline{u})\,(u_* + |\underline{u}|) + O(n^{-1/2}(u_* + |\underline{u}|)).$$

On the other hand, because $ES_n \approx \alpha$, the proportionality rule (47) implies $nE\psi(S_n) \approx \alpha$, and so

$$\begin{gathered} n^2\mu\psi(w)f\,(\underline{u})\,(u_* + |\underline{u}|) \approx \alpha \quad \Longrightarrow \\ u_* - \underline{u} \sim \zeta n^{-2}, \end{gathered}$$

where ζ is the unique solution of the equation $\alpha = \zeta\psi(w)f\,(\underline{u})$.

$\square$

E Balanced Populations: Proof of Theorem 1

E.1 Continuity of Nash equilibria

Proposition 4. If $\mu = 0$, then for all sufficiently large values of the sample size n no Nash equilibrium $v(u)$ has a discontinuity in $[\underline{u}, \overline{u}]$. Moreover, for any $\epsilon > 0$, if n is sufficiently large every Nash equilibrium $v(u)$ satisfies

$$\|v\|_\infty \leq \epsilon. \tag{59}$$

Proof. The size of any discontinuity is bounded below by a positive constant Δ, by Lemma 11, so it suffices to prove the assertion (59). By Proposition 7, for any $\epsilon > 0$ a $\gamma = \gamma(\epsilon)$ exists such that if n is sufficiently large any Nash equilibrium $v(u)$ satisfying $\|v\|_\infty > \epsilon$ must also satisfy $|v(u)| \leq \gamma/\sqrt{n}$ for all u not within distance ϵ of one of the endpoints $\underline{u}, \overline{u}$. Hence, the approximate proportionality relation (47) implies

$$E\psi(S_n) \leq \frac{C}{\sqrt{n}} \tag{60}$$

for a suitable $C = C(\gamma)$. Because $v(u)/u$ is within a factor $(1+\epsilon)^{\pm 1}$ of $E\psi(S_n)$ for all u not within distance $C'_\epsilon n^{-3/2}$ of $\underline{u}$ or $\overline{u}$, it follows from Chebyshev's inequality that for any $\alpha > 0$ a $\beta = \beta(\alpha)$ exists such that

$$P\{|S_n - ES_n| \geq \beta\} \leq \alpha.$$

On the other hand, if $\|v\|_\infty \geq \epsilon$, then by the necessary condition (15), some u exists such that

$$P\{S_n + v(u) \in [-\delta, \delta]\} \geq \frac{\epsilon}{\|\psi\|_\infty \max\left(|\underline{u}|, \overline{u}\right)}.$$

Because S_n is concentrated around ES_n, it follows that ES_n must be at bounded distance from $v(u)$, and so the Berry–Esseen bound (51) implies $P\{S_n \in [-\delta/2, \delta/2]\}$ is bounded below. But this in turn implies $E\psi(S_n)$ is bounded below, which for large n is impossible in view of (60). Thus, if n is sufficiently large then no Nash equilibrium $v(u)$ can have $\|v\|_\infty \geq \epsilon$.

□

Because $\|v\|_\infty$ is small for any Nash equilibrium v, the distribution of the vote total S_n cannot be too highly concentrated. This in turn implies the proportionality constant $E\psi(S_n)$ in (47) cannot be too small.

Lemma 16. For any $C < \infty$ a $n_C < \infty$ exists such that for all $n \geq n_C$ and every Nash equilibrium,

$$nE\psi(S_n) \geq C. \tag{61}$$

Proof. By the approximate proportionality rule (47) and the necessary condition (15), for any $\epsilon > 0$ and all sufficiently large n,

$$|ES_n| \leq n\epsilon E\psi(S_n) E|U|.$$

Thus, by Hoeffding's inequality (Corollary 2), if $nE\psi(S_n) < C$ then the distribution of S_n must be highly concentrated in a neighborhood of 0. But if this were so we would have, for all large n,

$$E\psi(S_n) \approx \psi(0) > 0,$$

which is a contradiction. □

E.2 Edgeworth expansions

For the analysis of the case $\mu_U = 0$ refined estimates of the errors in the approximate proportionality rule (47) will be necessary. We derive these from the Edgeworth expansion for the density of a sum of independent, identically distributed random variables (cf. Feller (1971), Ch. XVI, sec.

2, Th. 2). The relevant summands here are the random variables $v(U_i)$, and because the function $v(u)$ depends on the particular Nash equilibrium (and hence also on n), we must employ a version of the Edgeworth expansion in which the error is precisely quantified. The following variant of Feller's Theorem 2 (which can be proved in the same manner as in Feller) will suffice for our purposes.

Proposition 5. Let $Y_1, Y_2, \ldots, Y_n$ be independent, identically distributed random variables with mean $EY_1 = 0$, variance $EY_1^2 = 1$, and finite $2r$th moment $E|Y_1|^{2r} = \mu_{2r} \leq m_{2r}$. Assume the distribution of Y_1 has a density $f_1(y)$ whose Fourier transform $\hat{f}_1$ satisfies $|\hat{f}_1(\theta)| \leq g(\theta)$, where g is a C^{2r} function such that $g \in L^\nu$ for some $\nu \geq 1$ and such that for every $\epsilon > 0$,

$$\sup_{|\theta| \geq \epsilon} g(\theta) < 1. \tag{62}$$

Then a sequence $\epsilon_n \to 0$ depending only on m_{2r} and on the function g exists such that the density $f_n(y)$ of $\sum_{i=1}^n Y_i/\sqrt{n}$ satisfies

$$\left| f_n(x) - \frac{e^{-x^2/2}}{\sqrt{2\pi n}} \left(1 + \sum_{k=3}^{2r} n^{-(k-2)/2} P_k(x) \right) \right| \leq \frac{\epsilon_n}{n^{-r+1}} \tag{63}$$

for all $x \in \mathbb{R}$, where $P_k(x) = C_k H_k(x)$ is a multiple of the kth Hermite polynomial $H_k(x)$, and C_k is a continuous function of the moments $\mu_3, \mu_4, \ldots, \mu_k$ of Y_1.

The following lemma ensures that in any Nash equilibrium the sums $S_n = \sum_{i=1}^n v(U_i)$, after suitable renormalization, meet the requirements of Proposition 5.

Lemma 17. Constants $0 < \sigma_1 < \sigma_2 < m_{2r} < \infty$ and a function $g(\theta)$ that satisfy the hypotheses of Proposition 5 (with $r = 4$) exist such that for all sufficiently large n and any Nash equilibrium $v(u)$ the following statement holds. If $w(u) = 2v(u)/E\psi(S_n)$

(a) $\sigma_1^2 < \mathrm{var}(w(U_i)) < \sigma_2^2$;
(b) $E|w(U_i) - Ew(U_i)|^{2r} \leq m_{2r}$; and
(c) the random variables $w(U_i)$ have density $f_W(w)$ whose Fourier transform is bounded in absolute value by g.

Proof. These statements are consequences of the proportionality relations (47) and the smoothness of Nash equilibria. By Proposition 4, Nash equilibria are continuous on $[\underline{u}, \overline{u}]$ and for large n satisfy $\|v\|_\infty < \epsilon$, where $\epsilon > 0$ is any small constant. Consequently, by Proposition 3, the proportionality relations (47) hold on the entire interval $[\underline{u}, \overline{u}]$. Because $EU_1 = 0$, it follows that for any $\epsilon > 0$, if n is sufficiently large then $|Ew(U_i)| < \epsilon$, and so assertions (a)–(b) follow routinely from (47).

The existence of the density $f_W(w)$ follows from the smoothness of Nash equilibria, which was established in Subsection C.4. In particular, by Lemma 14, inequalities (46), and the proportionality relations (47), if the sample size n is sufficiently large and v is any continuous Nash equilibrium then v is twice continuously differentiable on $[\underline{u}, \overline{u}]$, and constants $\alpha, \beta > 0$ exist not depending on n or on the particular Nash equilibrium such that the derivatives satisfy

$$\alpha \leq \frac{v'(u)}{E\psi(S_n)} \leq \alpha^{-1} \quad \text{and} \quad \beta \leq \frac{v''(u)}{E\psi(S_n)} \leq \beta^{-1} \tag{64}$$

for all $u \in [\underline{u}, \overline{u}]$. Consequently, if U is a random variable with density $f(u)$ the random variable $W := 2v(U)/E\psi(S_n)$ has density

$$f_W(w) = f(u)E\psi(S_n)/(2v'(u)) \quad \text{where } w = 2v(u)/E\psi(S_n). \tag{65}$$

Furthermore, the density $f_W(w)$ is continuously differentiable, and its derivative

$$f'_W(w) = \frac{f'(u)(E\psi(S_n))^2}{4v'(u)^2} - \frac{f(u)(E\psi(S_n))^2 v''(u)}{4v'(u)^3}$$

satisfies

$$|f'_W(w)| \leq \kappa \tag{66}$$

where $\kappa < \infty$ is a constant that does not depend on either n or on the choice of Nash equilibrium.

The last step is to prove the existence of a dominating function $g(\theta)$ for the Fourier transform of f_W. We do this in three pieces: (i) for values $|\theta| \leq \gamma$, where $\gamma > 0$ is a small fixed constant; (ii) for values $|\theta| \geq K$, where K is a large but fixed constant; and (iii) for $\gamma < |\theta| < K$. Region (i) is easily dealt with, in view of the bounds (a)–(b) on the second and third moments and the estimate $|Ew(U)| < \epsilon'$, as these together with Taylor's theorem imply that for all $|\theta| < 1$,

$$|\hat{f}_W(\theta) - (1 + i\theta Ew(U) - \theta^2 \text{var}(w(U))/2| \leq m_3|\theta|^3.$$

Next consider region (ii), where $|\theta|$ is large. Integration by parts shows that

$$\hat{f}_W(\theta) = \int_{w\underline{u}}^{w(\overline{u})} f_W(w)e^{i\theta w}\, dw = -\int_{w\underline{u}}^{w(\overline{u})} \frac{e^{i\theta w}}{i\theta} f'_W(w)\, dw + \frac{e^{i\theta w}}{i\theta} f_W(w)\bigg|_{w\underline{u}}^{w(\overline{u})};$$

because $f_W(w)$ is uniformly bounded at $w\underline{u}$ and $w(\overline{u})$, by (64) and (65), and because $|f'_W(w)| \leq \kappa$, by (66), it follows that a constant $C < \infty$ exists such that for all sufficiently large n and all Nash equilibria,

$$|\hat{f}_W(\theta)| \leq C/|\theta| \quad \forall\, \theta \neq 0.$$

Thus, setting $g(\theta) = C/|\theta|$ for all $|\theta| \geq 2C$, we have a uniform bound for the Fourier transforms $\hat{f}_W(\theta)$ in the region (ii).

Finally, to bound $|\hat{f}_W(\theta)|$ in the region (iii) of intermediate $\theta-$values, we use the proportionality rule once again to deduce that $|w(u) - u| < \epsilon$. Therefore,

$$\begin{aligned}
\hat{f}_W(\theta) &= \int_{\underline{u}}^{\overline{u}} e^{i\theta w(u)} f(u)\, du \\
&= \int_{\underline{u}}^{\overline{u}} e^{i\theta u} f(u)\, du + \int_{\underline{u}}^{\overline{u}} (e^{i\theta w(u)} - e^{i\theta u}) f(u)\, du \\
&= \hat{f}_U(\theta) + R(\theta)
\end{aligned}$$

where $|R(\theta)| < \epsilon'$ uniformly for $|\theta| \leq C$ and $\epsilon' \to 0$ as $\epsilon \to 0$. Because $\hat{f}_U$ is the Fourier transform of an absolutely continuous probability density, its absolute value is bounded away from 1 on the complement of $[-\gamma, \gamma]$, for any $\gamma > 0$. Because $\epsilon > 0$ can be made arbitrarily small (cf. Proposition 3), it follows that a continuous, positive function $g(\theta)$ that is bounded away from 1 on $|\theta| \in [\gamma, C]$ exists such that $|\hat{f}_W)\theta| \leq g(\theta)$ for all $|\theta| \in [\gamma, C]$. The extension of g to the whole real line can now be done by smoothly interpolating at the boundaries of regions (i), (ii), and (iii). $\square$

E.3 Proof of Theorem 1

Because the function ψ is smooth and has compact support, differentiation under the expectation in the necessary condition $2v(u) = E\psi(v(u) + S_n)u$ is permissible, and so for every $u \in [-\underline{u}, \overline{u}]$ a $\tilde{v}(u)$ exists intermediate between 0 and $v(u)$ such that

$$2v(u) = E\psi(S_n)u + E\psi'(S_n + \tilde{v}(u))v(u)u. \tag{67}$$

The proof of Theorem 1 will hinge on the use of the Edgeworth expansion (Proposition 5) to approximate each of the two expectations in (67) precisely.

As in Lemma 17, let $w(u) = 2v(u)/E\psi(S_n)$. We have already observed, in the proof of Lemma 17, that for any $\epsilon > 0$, if n is sufficiently large then for any Nash equilibrium, $|Ew(U)| < \epsilon$. It therefore follows from the proportionality rule that

$$\left| \frac{4\,\mathrm{var}(v(U))}{(E\psi(S_n))^2\sigma_U^2} - 1 \right| \leq \epsilon \quad \text{and} \quad \left| \frac{E|v(u) - Ev(u)|^k}{(E\psi(S_n))^k E|U|^k} \right| < \epsilon \;\; \forall\, k \leq 8. \tag{68}$$

Moreover, Lemma 17 and Proposition 5 imply the distribution of S_n has a density with an Edgeworth expansion, and so for any continuous function $\varphi : [-\delta, \delta] \to \mathbb{R}$,

$$E\varphi(S_n) = \int_{-\delta}^{\delta} \varphi(x) \frac{e^{-y^2/2}}{\sqrt{2\pi n}\sigma_V} \left(1 + \sum_{k=3}^{m} n^{-(k-2)/2} P_k(y) \right) dx + r_n(\varphi) \tag{69}$$

where

$$\begin{aligned} &\sigma_V^2 := \mathrm{var}(v(U)), \\ &y = y(x) = (x - ES_n)/\sqrt{\mathrm{var}(S_n)}, \end{aligned}$$

and $P_k(y) = C_k H_3(y)$ is a multiple of the kth Hermite polynomial. The constants C_k depend only on the first k moments of $w(U)$, and consequently are uniformly bounded by constants C_k' not depending on n or on the choice of Nash equilibrium. The error term $r_n(\varphi)$ satisfies

$$|r_n(\varphi)| \leq \frac{\epsilon_n}{n^{(m-2)/2}} \int_{-\delta}^{\delta} \frac{|\varphi(x)|}{\sqrt{2\pi \mathrm{var}(S_n)}}\, dx. \tag{70}$$

In the special case $\varphi = \psi$, (69) and the remainder estimate (70) (with $m = 4$) imply

$$E\psi(S_n) \leq \frac{1}{\sqrt{2\pi n}\sigma_V} \int_{-\delta}^{\delta} \psi(x)\, dx + o(n^{-1}\sigma_V^{-1}).$$

Because $4\,\sigma_V^2 \approx (E\psi(S_n))^2\sigma_U^2$ for large n, this implies that for a suitable constant $\kappa < \infty$,

$$E\psi(S_n) \leq \frac{\kappa}{\sqrt[4]{n}}. \tag{71}$$

Claim 1. Constants $\alpha_n \to \infty$ exist such that in every Nash equilibrium,

$$|ES_n| \leq \alpha_n^{-1}\sqrt{\mathrm{var}(S_n)} \quad \text{and} \tag{72}$$
$$\mathrm{var}(S_n) \geq \alpha_n^2. \tag{73}$$

Proof of Theorem 1. Before we begin the proof of the claim, we indicate how it will imply Theorem 1. If (72) and (73) hold, then for every $x \in [-\delta, \delta]$,

$$|y(x)| \leq (1 + 2\delta)/\alpha_n \to 0.$$

Consequently, the dominant term in the Edgeworth expansion (69) for $\varphi = \psi$ (with $m = 4$), is the first, and so for any $\epsilon > 0$, if n is sufficiently large,

$$E\psi(S_n) = \frac{1}{\sqrt{2\pi n}\sigma_V}\int_{-\delta}^{\delta}\psi(x)\,dx(1 \pm \epsilon).$$

(Here the notation $(1\pm\epsilon)$ means the ratio of the two sides is bounded above and below by $(1\pm\epsilon)$.) Thus $4\,\sigma_V^2 \approx (E\psi(S_n))^2\sigma_U^2$ implies

$$\sqrt{\pi n/2}\sigma_U(E\psi(S_n))^2 = \int_{-\delta}^{\delta}\psi(x)\,dx(1 \pm \epsilon) = 2 \pm 2\epsilon,$$

proving the assertion (5). $\square$

Proof of Claim 1. First we deal with the remainder term $r_n(\varphi)$ in the Edgeworth expansion (69). By Lemma 16, the expectation $E\psi(S_n)$ is at least C/n for large n, and so by (68) the variance of S_n must be at least C'/n. Consequently, by (70), the remainder term $r_n(\varphi)$ in (69) satisfies

$$|r_n(\varphi)| \leq C''\frac{\epsilon_n\|\varphi\|_1}{n^{(m-2)/2}\sqrt{\mathrm{var}(S_n)}} \leq C'''\frac{\epsilon_n\|\varphi\|_1}{n^{(m-3)/2}}.$$

Suitable choice of m will make this bound small compared to any desired monomial n^{-A}, and so we may ignore the remainder term in the arguments to follow.

Suppose there a constant $C < \infty$ existed such that along some sequence $n \to \infty$ Nash equilibria existed satisfying $\mathrm{var}(S_n) \leq C$. By (68), this would force $C/n \leq E\psi(S_n) \leq C'/\sqrt{n}$. This in turn would force

$$C''\mathrm{var}(S_n)\log n \geq |ES_n|^2 \geq C'''\mathrm{var}(S_n)\log n, \tag{74}$$

because otherwise the dominant term in the Edgeworth series for $E\psi(S_n)$ would be either too large or too small asymptotically (along the sequence $n \to \infty$) to match the requirement that $C/n \leq E\psi(S_n) \leq C'/\sqrt{n}$. (Observe that because the ratio $|ES_n|^2/\mathrm{var}(S_n)$ is bounded above by $C''\log n$, the terms $e^{-y^2/2}P_k(y)$ in the integral (69) are of size at most $(\log n)^A$ for some A depending on m, and so the first term in the Edgeworth series is dominant.) We will show that (74) leads to a contradiction.

Suppose $ES_n > 0$ (the case $ES_n < 0$ is similar). The Taylor expansion (67) for $v(u)$ and the hypothesis $EU = 0$ implies

$$2Ev(U) = E\psi'(S_n + \tilde{v}(U))v(U)U. \tag{75}$$

The Edgeworth expansion (69) for $E\psi'(S_n + \tilde{v}(u))$ together with the independence of S_n and U and the inequalities (74), implies that for any $\epsilon > 0$, if n is sufficiently large then

$$\begin{aligned} &E\psi'(S_n + \tilde{v}(u)) \\ &\qquad = \frac{1}{\sqrt{2\pi \text{var}(S_n)}} \int_{-\delta}^{\delta} \psi'(x) \exp\{-(x + \tilde{v}(u) - ES_n)^2/2\text{var}(S_n)\}\, dx(1 \pm \epsilon). \end{aligned} \tag{76}$$

Now because ψ and ψ' have support $[-\delta, \delta]$, integration by parts yields

$$\begin{aligned} &\int_{-\delta}^{\delta} \psi'(x) \exp\{-(x + \tilde{v}(u) - ES_n)^2/2\text{var}(S_n)\}\, dx \\ &\qquad = \int_{-\delta}^{\delta} \psi(x) \exp\{-(x + \tilde{v}(u) - ES_n)^2/2\text{var}(S_n)\} \frac{x + \tilde{v}(u) - ES_n}{\text{var}(S_n)}\, dx, \end{aligned} \tag{77}$$

and because $x + \tilde{v}(u)$ is of smaller order of magnitude than ES_n, it follows that for large n

$$E\psi'(S_n + \tilde{v}(u)) = -\frac{ES_n}{\text{var}(S_n)} E\psi(S_n)(1 \pm \epsilon). \tag{78}$$

But it now follows from the Taylor series for $2Ev(U_i)$ (by summing over i) that

$$2ES_n = -n\frac{ES_n}{\text{var}(S_n)} E\psi(S_n) Ev(U)U(1 \pm \epsilon), \tag{79}$$

which is a contradiction, because the right side is negative and the left side positive. This proves the assertion (73).

The proof of inequality (72) is similar. Suppose for some $C > 0$ Nash equilibria existed along a sequence $n \to \infty$ for which $ES_n \geq C\sqrt{\text{var}(S_n)}$. In view of (73), this hypothesis implies in particular that $ES_n \to \infty$, and also that $|y(x)| \geq C/2$ for all $x \in [-\delta, \delta]$. Thus, the Edgeworth approximation (76) remains valid, as does the integration by parts identity (77). Because $ES_n \to \infty$, the terms $x + \tilde{v}(u)$ are of smaller order of magnitude that ES_n, and so once again (78) and therefore (79) follow. Again we have a contradiction, because the right side of (79) is negative while the left side diverges to $+\infty$.

□

F Proof of Lemma 8

Lemma 8. *Fix $\delta > 0$. For any $\epsilon > 0$ and any $C < \infty$ a $\beta = \beta(\epsilon, C) > 0$ and a $n' = n'(\epsilon, C) < \infty$ exist such that the following statement is true: if $n \geq n'$ and $Y_1, Y_2, \ldots, Y_n$ are independent random variables such that*

$$E|Y_1 - EY_1|^3 \leq C\text{var}(Y_1)^{3/2} \quad \textit{and} \quad \text{var}(Y_1) \geq \beta/n \tag{80}$$

then for every interval $J \subset \mathbb{R}$ of length δ or greater, the sum $S_n = \sum_{i=1}^n Y_i$ satisfies

$$P\{S_n \in J\} \leq \epsilon |J|/\delta. \tag{81}$$

Proof. It suffices to prove this for intervals of length δ, because any interval of length $n\delta$ can be partitioned into n pairwise disjoint intervals each of length δ. Without loss of generality, $EY_1 = 0$ and $\delta = 1$ (if not, translate and re-scale). Let g be a nonnegative, even, C^∞ function with $\|g\|_\infty = 1$ that takes the value 1 on $[-\frac{1}{2}, \frac{1}{2}]$ and is identically zero outside $[-1, 1]$. It is enough to show that for any $x \in \mathbb{R}$,

$$Eg(S_n + x) \leq \epsilon.$$

Because g is C^∞ and has compact support, its Fourier transform is real-valued and integrable, so the Fourier inversion theorem implies

$$Eg(S_n + x) = \frac{1}{2\pi} \int \hat{g}(\theta) \varphi(-\theta)^n e^{-i\theta x} \, d\theta,$$

where $\varphi(\theta) = Ee^{i\theta Y_1}$ is the characteristic function of Y_1. Because $EY_1 = 0$, the derivative of the characteristic function at $\theta = 0$ is 0, and hence φ has Taylor expansion

$$|1 - \varphi(\theta) - \frac{1}{2} EY_1^2 \theta^2| \leq \frac{1}{6} E|Y_1|^3 |\theta|^3.$$

Consequently, if the hypotheses (27) hold then for any $\gamma > 0$, if n is sufficiently large,

$$|\varphi(\theta)^n| \leq e^{-\beta^2\theta^2/4}$$

for all $|\theta| \leq \gamma$. This bound implies (because $|\hat{g}| \leq 2$) that

$$Eg(S_n + x) \leq \frac{1}{\pi} \int_{|\theta| < \gamma} e^{-\beta^2\theta^2/4} \, d\theta + \frac{1}{2\pi} \int_{|\theta| \geq \gamma} |\hat{g}(\theta)| \, d\theta.$$

Because $\hat{g}$ is integrable, the constant γ can be chosen so that the second integral is less that $\epsilon/2$, and if β is sufficiently large then the first integral is bounded by $\epsilon/2$. $\square$

G Proof of Proposition 1

We shall assume throughout that $\delta < 1/\sqrt{2}$, and that the function $\psi = \Psi'$ satisfies the standing assumptions 3 (a)- (c) of Section 2.1. Thus, $\psi/2$ is an even, C^∞ probability density with support $[-\delta, \delta]$; it has positive derivative ψ' on $(-\delta, 0)$ (and hence negative derivative on $(0, \delta)$); and it has a single point of inflection in the interval $(-\delta, 0)$.

Define

$$H(\alpha, w) = (1 - \Psi(w))|\underline{u}| - (\alpha - w)^2. \tag{82}$$

Proposition 1 asserts that, under the assumption $\delta < 1/\sqrt{2}$, a unique value $\alpha > \delta$ exists such that (i) the maximum value of the function $w \mapsto H(\alpha, w)$ on $w \in \mathbb{R}$ is 0, and (ii) this maximum

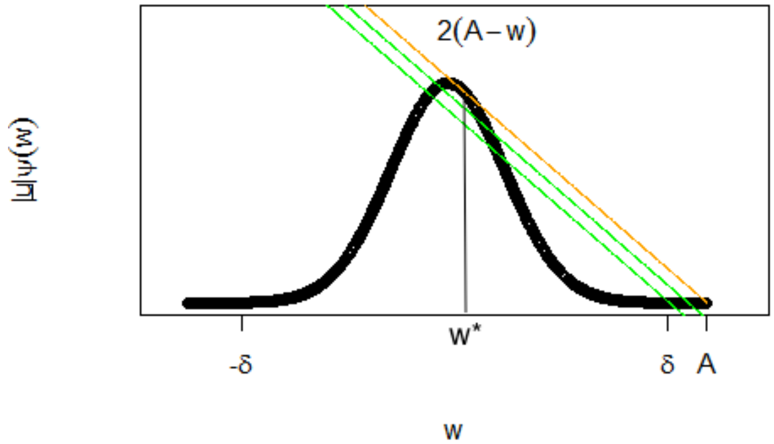

Figure 1: Possible configurations of the optimization problem faced by individuals with extreme negative utilities depending on the size of δ.

is attained at precisely two values of w, one at $w = \alpha$, the other at a point $w \in (-\delta, \delta)$. Local maxima and minima of smooth functions are attained only at *critical points*, that is, roots of the equation

$$\frac{\partial H}{\partial w}(\alpha, w) = -\psi(w)|\underline{u}| + 2(\alpha - w) = 0. \tag{83}$$

Lemma 18. $A \in (\delta, \infty)$ exists such that

(i) $\alpha > A \implies$ a unique critical point exists at $w = \alpha$;
(ii) $\alpha = A \implies$ 2 critical points exist, at $w = \alpha$ and at $w_* \in (-\delta, \delta)$; and
(iii) $\delta \leq \alpha < A \implies$ 3 critical points exist, at $w = \alpha$ and at distinct points $w_-(\alpha), w_+(\alpha) \in (-\delta, \delta)$.

For $\alpha \in [\delta, A)$ the two critical points $w_-(\alpha) < w_+(\alpha)$ vary continuously with α; the function $w_-(\alpha)$ is increasing in α, while $w_+(\alpha)$ is decreasing. Moreover, as $\alpha \to A-$, the critical points $w_\pm(\alpha) \to w_*$.

While the proof of the following lemma is quite involved when presented algebraically, the basic idea can be seen clearly graphically, as shown in Figure 1, which pictures the different arrangements the lemma formally demonstrates are possible.

Proof. It is clear that for every $\alpha \geq \delta$ the equation (83) holds at $w = \alpha$, because $\psi = 0$ outside of the interval $(-\delta, \delta)$. All other critical points are points where the straight line of slope -2 through $(\alpha, 0)$ intersects the graph of $y = |\underline{u}|\psi(w)$. Because $\psi(w) \neq 0$ only for $w \in (-\delta, \delta)$, critical points not equal to α must be located in the interval $(-\delta, \delta)$.

(1) We first show that for $\alpha \geq \delta$ near δ at least two such intersection points exist, one in each of the intervals $(-\delta, 0)$ and $(0, \delta)$. To see this, observe that because $\psi/2$ is a probability density with support $(-\delta, \delta) \subset (-1/\sqrt{2}, 1/\sqrt{2})$, its maximum value, which is assumed at $w = 0$, must exceed $1/\sqrt{2}$, and so $\psi(0) > \sqrt{2}$. But the line $y = 2(\delta - w)$ intersects the y−axis at $y = 2\delta < \sqrt{2}$, so it must cross the graph of $y = |\underline{u}|\psi(w)$ at least once in each of the intervals $(0, \delta)$ and $(-\delta, 0)$. Similarly, if $\alpha - \delta$ is sufficiently small then the line $y = 2(\alpha - w)$ must also cross the graph twice, once on each side of 0.

(2) Next, we show that for any $\alpha \geq \delta$ at most two roots of (83) in $(-\delta, \delta)$ exist, and that two exist if and only if both intersections are *transversal*. Assume at least two distinct roots exist; let $w_+(\alpha) \in (0, \delta)$ be the largest, and let $w_-(\alpha)$ be the second largest. We claim that at $w = w_+(\alpha)$ the intersection between the line $y = 2(\delta - w)$ and the graph of $y = |\underline{u}|\psi(w)$ is *transversal*, that is,

$$|\underline{u}|\psi'(w) \neq -2. \tag{84}$$

It is impossible for $|\underline{u}|\psi'(w_+(\alpha)) > -2$, because $2(\delta - w) > |\underline{u}|\psi(w)$ for all $w_+(\alpha) < w < \alpha$, so we must only show that the line $y = 2(\delta - w)$ cannot be tangent to the graph at $w = w_+(\alpha)$. But if this were the case then $|\underline{u}|\psi''(w_+(\alpha)) < 0$, once again because the line lies above the graph in the interval $w_+(\alpha) < w < \alpha$; because ψ has only a single inflection point in $(0, \delta)$, it would then follow that $w_+(\alpha)$ is the *only* root of (83) in $(-\delta, \delta)$, contrary to our hypothesis. This proves (84) for $w = w_+(\alpha)$.

It now follows from (84) that the intersection between the line $y = 2(\delta - w)$ and the graph of $y = |\underline{u}|\psi(w)$ at $w_-(\alpha)$ is also transversal. To see this, note first that it suffices to consider the case where $w_-(\alpha) > 0$, because at any $w \in (-\delta, 0]$ the slope of the graph of $|\underline{u}|\psi(w)$ is nonnegative. Next, observe that by the Mean Value Theorem a maximal $w_* \in [w_-(\alpha), w_+(\alpha)]$ at which $|\underline{u}|\psi'(w_*) = -2$ exists. This point w_* cannot be $w_+(\alpha)$, because at $w = w_+(\alpha)$ we have (84); furthermore, because $|\underline{u}|\psi'(w_+(\alpha)) < -2$, it must be the case that

$$|\underline{u}|\psi'(w) < -2 \quad \text{for all } \; w \in (w_*, w_+(\alpha)),$$

and so by the Fundamental Theorem of calculus,

$$|\underline{u}|\psi(w_*) > 2(\delta - w_*).$$

This implies $w_* > w_-(\alpha)$. Now because the slope of the graph at w_* is larger than at $w_+(\alpha)$, the second derivative $|\underline{u}|\psi''(w)$ must be negative at all $0 < w \leq w_*$, because of the standing hypothesis that ψ has only one inflection point in $(0, \delta)$. Finally, because $0 < w_-(\alpha) < w_*$, it follows that

$$|\underline{u}|\psi'(w_-(\alpha)) < |\underline{u}|\psi'(w_*) = -2.$$

Thus, both intersections, at $w = w_-(\alpha)$ and $w = w_+(\alpha)$, are transversal provided $w_-(\alpha) < w_+(\alpha)$.

To complete the proof that at most two roots of (83) in $(-\delta, \delta)$ exist, consider the cases $w_-(\alpha) > 0$ and $w_-(\alpha) \leq 0$ separately. In the first case, the intersection at $w_-(\alpha)$ is transversal, and $|\underline{u}|\psi'(w) < -2$ at all $w \in [0, w_-(\alpha)]$. Hence, the line $y = 2(\delta - w)$ is above the graph of $y = |\underline{u}|\psi(w)$ at $w = 0$, and because the line has constant negative slope, it must remain above the graph at all $w < 0$, because $w = 0$ is the unique point where $\psi(w)$ attains its maximum. Thus, there can be no points of intersection to the left of $w_-(\alpha)$. The second case, where $w_-(\alpha) \leq 0$, is even easier: because ψ is increasing on $(-\delta, 0]$, for any $w < w_-(\alpha)$ we must have

$$|\underline{u}|\psi(w) < |\underline{u}|\psi(w_-(\alpha)) = 2(\alpha - w_-(\alpha)) < 2(\alpha - w),$$

and so once again there can be no no points of intersection to the left of $w_-(\alpha)$.

(3) Because equation (83) has the form $G(\alpha, w) = 0$ with G continuously differentiable, the Implicit Function Theorem guarantees that solutions vary continuously in a neighborhood of any solution where $\partial G/\partial w \neq 0$, i.e., where $|\underline{u}|\psi(w) \neq -2$. We have shown in point (2) above

that this is the case for both $w_\pm(\alpha)$, as long as $w_+(\alpha) \neq w_-(\alpha)$; thus, $w_\pm(\alpha)$ have continuous extensions to an open interval with left endpoint δ. We must show that the maximal interval on which the functions $w_\pm$ can be continuously and monotonically continued is such that at the right endpoint A the two roots coalesce, and the intersection becomes non-transversal.

Clearly, the linear function $2(\alpha - w)$ is increasing in α for each w. Consequently, if $A > \delta$ is any point where the line $y = 2(A - w)$ intersects the graph of $|\underline{u}|\psi(w)$ non-transversally at some point w_*, then because this would be the *only* point of intersection in $[-\delta, \delta]$ (by point (2)), the graph would lie entirely below the line and there would be no solutions to equation (83) with $\alpha > A$. Thus, to complete the proof it suffices to establish the following claim.

Claim: The function $w_-(\alpha)$ increases continuously up to the smallest $\alpha = A$ such that $w_-(\alpha) = w_+(\alpha) := w_*$, where w_* is the unique point in $(0, \delta)$ at which the graph of $|\underline{u}|\psi$ has tangent line of slope -2, and lies entirely below its tangent line.

Proof of the Claim. First, observe that because $|\underline{u}|\psi$ is increasing on $[-\delta, 0]$ and decreasing on $[0, \delta]$, the function $w_+(\alpha)$ is decreasing in α, and $w_-(\alpha)$ is increasing as long as $w_-(\alpha) \leq 0$. Let $\alpha_0 = |\underline{u}|\psi(0)/2$; at this value, the line $y = 2(\alpha_0 - w)$ intersects the graph of $|\underline{u}|\,\psi$ transversally at $w = 0$. Thus, $\alpha_0 < A$, and the functions $w_\pm(\alpha)$ are continuous and monotone on $[\delta, \alpha_0]$, and $w_+(\alpha_0) > 0$.

To see that a w_* at which the graph of $|\underline{u}|\psi$ has tangent line of slope -2 exists, observe that the intersections of the line $y = 2(\alpha_0 - w)$ with the graph of $|\underline{u}|\,\psi(w)$ are both transversal: at $w_-(\alpha_0) = 0$ the slope of the graph is > -2, and at $W_+(\alpha_0)$ the slope is < -2. Consequently, by the mean value theorem, there must be a point $w_* \in (0, w_+(\alpha_0))$ at which $|\underline{u}|\psi'(w_*) = -2$. The tangent line at this point has the form $y = 2(A - w)$ for some $A > \delta$. Because the intersection at the point of tangency is non-transversal, it must be the *unique* point of intersection of this line with the graph, and so it follows that the rest of the graph lies *below* the line, as claimed. Hence, the derivative $|\underline{u}|\psi'(w)$ is non-increasing at $w = w_*$; because ψ has only one inflection point in $(0, \delta)$, it follows that

$$|\underline{u}|\psi'(w) > -2 \quad \text{for all } \ 0 \leq w < w_*. \tag{85}$$

For each $w \in [0, w_*)$, the line of slope -2 through the point $(w, |\underline{u}|\psi(w))$ intersects the $w-$axis at a point $\alpha(w) > \delta$. Clearly, the mapping $w \mapsto \alpha(w)$ is continuous, and by (85) it is also increasing in w, with positive derivative. Furthermore, because the intersection of this line with the graph of $|\underline{u}|\psi(w)$ is transversal, there must be a second point of intersection to the right of w. Thus, by point (2),

$$w_-(\alpha(w)) = w.$$

This proves that $w_-(\alpha)$ is continuous and increasing in α on the interval $[\alpha_0, A]$, where A is defined to the unique point where $w_-(A) = w_*$.

It remains to prove that $w_+(A) = w_*$. Recall that w_+ is continuous and decreasing as long as $w_+ > w_*$, because the intersections are transversal at all such points. By point (2), we cannot have $w_+(\alpha) = w_*$ for any $\alpha < A$, because at any such α a distinct second transversal intersection at $w_-(\alpha) < w_*$ exists, while the intersection at $w_+(\alpha) = w_*$ would be non-transversal. Similarly, we cannot have $w_+(A) > w_*$, because this intersection would be non-transversal, whereas the intersection at $w_* = w_-(A)$ would be transversal. Therefore, $w_+(A) = w_-(A) = w_*$.

□

Proof of Proposition 1. First, note that for any $\alpha \geq \delta$ the function $w \mapsto H(\alpha, w)$ satisfies

$$\lim_{w \to -\infty} H(\alpha, w) = \lim_{w \to +\infty} H(\alpha, w) = -\infty.$$

Consequently, for any $\alpha \in [\delta, A]$ neither of the critical points $w = \alpha$ nor $w_-(\alpha)$ can be a local *minimum* of $w \mapsto H(\alpha, w)$. It is easily verified that if $\alpha > \delta$ then $w = \alpha$ is a local maximum, because $\Psi = 1$ in a neighborhood of α and so $\partial^2 H/\partial w^2 = -2$ at all $w > \delta$.

Now consider the behavior of $H(\alpha, w)$ in a neighborhood of $w = w_-(\alpha)$. Because $\delta < 1/\sqrt{2}$, for all α near δ, we have

$$\begin{aligned} H(\alpha, -\delta) &= (1 - \Psi(-\delta))|\underline{u}| - (\alpha + \delta)^2 \\ &= 2|\underline{u}| - (\alpha + \delta)^2 \\ &\geq 2 - (\alpha + \delta)^2 > 0, \end{aligned}$$

because $(\alpha + \delta) \approx 2\delta$, and so the maximum value of $H(\alpha, w)$ for $w \in \mathbb{R}$ must be positive. Because $H(\alpha, \alpha) = 0$, it follows that the global maximum must be attained at one of the other two critical points $w_\pm(\alpha)$, and because $w = \alpha$ is a *local* maximum, it must be that $w_+(\alpha)$ is a local minimum and $w_-(\alpha)$ the global maximum. Thus, $H(\alpha, w_-(\alpha)) = \max_w H(\alpha, w)$ for all $\alpha \in [\delta, A]$ such that $H(\alpha, w_-(\alpha)) > 0$.

Next, observe that for any fixed $w \leq \delta$ the function $H(\alpha, w)$ is *decreasing* in α. Hence, $h(\alpha) := \max_{w \in [-\delta, \delta]} H(\alpha, w)$ is decreasing in α. Because $h(\alpha) = H(\alpha, w_-(\alpha))$ for all α such that $H(\alpha, w_-(\alpha)) \geq 0$, it follows that $h(\alpha)$ decreases *continuously* with α up to the first point α_* where $h(\alpha_*) = 0$, if such a point exists. But it cannot be the case that $h(\alpha) > 0$ for all $\alpha > \delta$, because clearly the definition (82) of H forces $h(\alpha) \to -\infty$ as $\alpha \to \infty$. Finally, α_* cannot be larger than A, because for all $\alpha < \alpha_*$ the global minimum of $w \mapsto H(\alpha, w)$ is attained in $(-\delta, \delta)$, and so at least 2 critical points exist for every such α.

The point $\alpha = \alpha_*$ is the unique point where a solution to the Optimization Problem (10) exists, and $w_-(\alpha_*)$ is the unique matching real number in $[-\delta, \delta]$ where (10) holds.

□