

DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

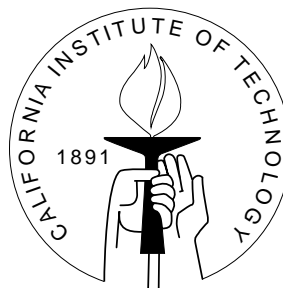
CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

A THEORY OF VOTING IN LARGE ELECTIONS

Richard D. McKelvey

John W. Patty



SOCIAL SCIENCE WORKING PAPER 1056

February 1999

A Theory of Voting in Large Elections*

Richard D. McKelvey[†]
Division of Humanities and Social Sciences
California Institute of Technology
Pasadena, CA 91125

John W. Patty[‡]
Department of Government
Institute for Quantitative Social Science
Harvard University
Cambridge, MA 02138

April 3, 2006

*A previous version of this paper was entitled “Quantal Response Voting”. The financial support of NSF grants #SBR-9631627 and #SES0079301 to the California Institute of Technology are gratefully acknowledged. In addition, Patty acknowledges the financial assistance of the Alfred P. Sloan Foundation. This paper benefited from helpful comments from seminar participants at CERGY-Pontoise (Oct 1998), CORE (Oct 1998), Tilburg (Oct 1998), Caltech Theory Workshop (May, 1999), Public Choice Society (New Orleans, March 1999), the Stan and Cal Show (Pismo Beach, May, 1999), the APSA annual meetings (Atlanta, 1999), and the Wallis Conference on Political Economy (Rochester, NY, October, 1999). We especially thank John Duggan, Tom Palfrey, Norman Schofield, Bob Sherman, two anonymous referees, and an Associate Editor for helpful input.

[†]*Deceased.* Richard passed away on April 22, 2002. He is, and will be, terribly missed as both a friend and a scholar.

[‡]Corresponding Author: jpatty@gov.harvard.edu

Abstract

This paper provides a game-theoretic model of probabilistic voting and then examines the incentives faced by candidates in a spatial model of elections. In our model, voters' strategies form a *Quantal Response Equilibrium* (QRE), which merges strategic voting and probabilistic behavior. We first show that a QRE in the voting game exists for all elections with a finite number of candidates, and then proceed to show that, with enough voters and the addition of a regularity condition on voters' utilities, a Nash equilibrium profile of platforms exists when candidates seek to maximize their expected margin of victory. This equilibrium (1) consists of all candidates converging to the policy that maximizes the expected sum of voters' utilities, (2) exists even when voters can abstain, and (3) is unique when there are only 2 candidates. *Journal of Economic Literature* Classification Numbers: D71, D72, D50, D60.

Keywords: Voting, Probabilistic Voting, Quantal Response Equilibrium.

Proposed Running Head: Voting in Large Elections

1 Introduction

Probabilistic voting is central to both theoretical¹ and empirical² work in political economy. In particular, prediction of an individual's voting behavior is inherently imprecise, even with detailed information about the voter in question. Accordingly, modern empirical studies of voting behavior universally assume some form of probabilistic voting behavior (in the form of, say, a logit or probit regression model). Theoretical models of probabilistic voting incorporate this imprecision into the candidates' strategic calculations. Probabilistic models of voting yield predictions more amenable to empirical testing and have been used to examine many topics, including tax policy, redistribution, and multiparty elections.

Theoretically, models of probabilistic voting are appealing for two reasons. The first is the link between the model and the empirical means of evaluating its predictions. In particular, probabilistic voting models and nearly all statistical analyses of individual vote choice share a common starting point: even the most reliable models of behavior are not guaranteed to predict any given individual's behavior *perfectly*. Probabilistic voting simply incorporates this reality into the strategic calculations of the individuals within the model. This common methodological link has another advantage: as opposed to most models of electoral competition in which voters are presumed to vote in a deterministic fashion,³ pure strategy electoral equilibria typically exist when voters are presumed to vote probabilistically. Indeed, one of the most robust characteristics of models of probabilistic voting is the stability they induce in models of electoral competition, as highlighted by Coughlin [1992], Banks and Duggan [2004], and Schofield [2004].

This paper extends this research in an important way by allowing for strategic behavior within a general formulation of probabilistic voting. Specifically, we investigate the Quantal Response

¹The probabilistic voting literature began with the work of Hinich [1977], and was initially extended by Coughlin and Nitzan [1981a,b]. An excellent overview of the early work is contained in Coughlin [1992]. Recent work in the area includes Lin, Enelow, and Dorussen [1999], Banks and Duggan [2004], Patty [2002, 2006, 2005], and Schofield [2004].

²Recent examples include Alvarez, Nagler, and Bowler [2000]; Lacy and Burden [1999]; Quinn, Martin, and Whitford [1999]; and Schofield, Martin, Quinn, and Whitford [1998], to name only a few.

³For an excellent examination of equilibrium existence in spatial electoral competition, see Banks, Duggan, and Breton [2002].

Equilibrium (see McKelvey and Palfrey [1995, 1998]) within spatial voting games. Quantal Response Equilibrium (QRE) is a theory of behavior in games that assumes that individuals get privately observed random payoff disturbances for each action available to them. A QRE is then just a Bayesian equilibrium of this game of incomplete information. In a QRE, although voters adopt pure strategies, from the point of view of an outside observer who does not know the payoff disturbance, the players choose between strategies probabilistically, choosing actions that yield higher utility with higher probability than actions that yield lower utility. The probability that one action is chosen over another is based on the utility difference between the alternatives. The fact that probabilistic voting is generated through a Bayesian equilibrium implies that voters may vote strategically. This is in contrast to most of the models studied in previous theoretical work on probabilistic voting, which is reviewed in the next section.

As we discuss further below, our results are very similar in spirit to the results obtained by previous scholars. However, it must be noted that the approach taken here is much more game-theoretic than in most previous models of probabilistic voting. In particular, our model allows for voters to take into account the relative likelihoods of different candidates winning the election. As opposed to classical models of probabilistic behavior, the expected utility of casting a particular vote is represented correctly. More precisely, in addition to the candidates' announced platforms and his or her own policy preferences, the notion of QRE presumes that each voter also considers the other voters' strategies when calculating the expected payoff from each possible ballot he or she may cast. Put another way, each voter's (probabilistic) voting behavior generated by a QRE is consistent with that voter being aware that his or her fellow citizens are also voting probabilistically.

Providing an equilibrium derivation of probabilistic voting in large elections is important for three reasons. First, as described above, the assumptions that underpin our theory are taken as given in many empirical analyses of voting behavior in economics and political science. Second, our results indicate that, in equilibrium, public policy outcomes may be governed by voters' preferences even when individual voters' probabilities of being pivotal are infinitesimal. A somewhat

ironic corollary of this fact is that policy outcomes are governed by voters' policy preferences even when any given voter's observed behavior is nearly independent of his or her policy preferences. Finally, and most importantly, the equilibrium platform location that we characterize has several appealing features. As previous authors have shown in other probabilistic voting frameworks, there is an electoral equilibrium in which office-motivated candidates within our model offer identical platforms at the *ex ante* social welfare optimum. This policy is appealing for a number of reasons. First, it is often (but not necessarily) "centrist". Second, and more importantly, this policy is sensitive to the strengths of individual voters' preferences. This stands in contrast to the celebrated median voter theorem (Black [1948]), in which the candidate equilibrium is insensitive to these individual strengths. In addition, while the equilibrium we characterize does require that the electorate be "large enough," it *does not* depend upon restrictions on (1) the dimensionality of the policy space (as opposed to most "median voter" results), (2) the shape of individuals' utility functions (as opposed to median voter results and many classical probabilistic voting models), or (3) the number of candidates (as opposed to both median voter results and many classical probabilistic voting models).

1.1 Related Literature

Many other scholars have studied probabilistic voting (see Coughlin [1992] for a review of this literature). Hinich [1977] showed that the median voter theorem does not always hold in a setting with probabilistic voting, and he constructed examples in a one dimensional space with equilibria at other locations. In particular, with quadratic utility functions, he obtained an equilibrium in two candidate elections at the mean (which is the social welfare optimum with those preferences). Coughlin and Nitzan [1981a,b] (see also Coughlin [1992], p. 96, Theorem 4.2) proved if voters have likelihood of voting functions satisfying the Luce axioms over subsets, there is a local equilibrium at a point maximizing the social log likelihood. While this work was not explicitly rooted in a utility maximization framework, subsequent work (see Coughlin [1992], p. 99-100, Corollaries 4.4 and 4.5, Theorem 4.2) shows how it can be so interpreted. Coughlin [1992] also gives various

conditions on voter likelihood functions or on preferences that result in a global equilibrium. If the likelihood functions are concave, there is a global equilibrium. In a re-distributional model where voters have logarithmic utility functions for income, and candidates use a logistic model to estimate the probability that voters vote for each candidate, there is a global equilibrium at the social utility maximum (p. 57, Theorem 3.7). See Banks and Duggan [2004] for a summary and generalization many of these results. All of the above results are for two candidate competition. Lin et al. [1999] show that one can also obtain equilibrium for multi-candidate elections using probabilistic voting models. They assume that voters' preferences are based on the distance between their own ideal policy and the candidate's announced platform, with a random utility shock, and obtain local equilibria at the social utility maximum. Lin et al. [1999] also find that if the utility shocks have high enough variance, then the expected vote function for each candidate becomes concave, implying the existence of a global equilibrium.

In all of the above cited probabilistic voting literature, game theoretic considerations for the voter are not modeled. Voters are assumed to vote based on their preferences for the candidate policy positions rather than based on the effect their vote will have on the outcome of the election. Ledyard [1984] develops a Bayesian model of two candidate competition that does model the game theoretic considerations for the voter. In his model, voters vote deterministically (there is no random utility shock to preferences), but they can abstain as well as vote for one of the two candidates, and the cost of voting is a random variable. Voter types consist of preferences as well as a cost of voting. He shows that in large elections, if voting costs are non-negative, there is an equilibrium at the social welfare optimum, which under certain restrictive conditions on the distribution of costs, is a global equilibrium. Myerson [2000] extends Ledyard's results in a model where the number of voters is a Poisson random variable, unknown to the voters. He shows that as long as the density function of the costs of voting is positive at zero, there is a global equilibrium in Ledyard's model as the number of voters becomes large. Ledyard's model, as well as Myerson's generalization of it, require that no voters have negative costs of voting.

1.2 Overview of The Model

In this paper, as in Ledyard [1984], we work in a Bayesian framework, and take into account the game theoretic considerations for the voters. Further, we consider elections with an arbitrary number of candidates. Unlike Ledyard [1984], we assume that voters have privately observed payoff disturbances associated with each action. We do not require concavity of preferences, but require that preferences are uniformly bounded. We also impose a regularity condition that the ratio of the variance to expected value of any utility differences is uniformly bounded. Our results basically extend those of the earlier literature. We find that for large enough electorates there is a convergent equilibrium at the alternative that maximizes social welfare. For two candidate contests, the equilibrium is unique. Our equilibrium is global, as in Lin et al. [1999], but in our model, the conditions for a global equilibrium are satisfied by allowing the number of voters to grow large rather than by assuming the utility shock becomes large.

In our model, the payoff disturbances are candidate specific and unrelated to the policy position of the candidate or to whether the candidate wins or not. In other words, a voter gets some randomly drawn (and privately observed) payoff for “pulling the lever” for candidate j . Interpreted in this way, this assumption may appear strange. However, it is standardly and uncontroversially adopted in modern empirical studies of individual vote choice. For example, any “probit” or “logit” analysis of individual voting behavior is based on the assumption that voters respond to random utility shocks that are very similar to those considered here.⁴

In addition, one interpretation of the payoff disturbances is as a representation of *expressive* voting (*e.g.*, Buchanan [1954], Tullock [1971], Brennan and Lomasky [1993], Schuessler [2000]). The individual experiences a subjective payoff from voting for a specific candidate above and beyond the instrumental objective benefits derived from his or her vote choice. While this interpretation may create the appearance that the voters in our model are irrational, it should be noted

⁴The principal difference between our approach and recent empirical work (*e.g.*, Alvarez and Nagler [1995, 1998], Schofield et al. [1998], Lacy and Burden [1999], Quinn et al. [1999], Alvarez et al. [2000], and Schofield [2005]) is that our model assumes that the random utility shocks are independently distributed across the alternatives. Alvarez and Nagler [2001] and Dow and Endersby [2004] discuss considerations surrounding the choice between different multinomial discrete choice models of vote choice.

that the payoff disturbance structure utilized within QRE insures that the likelihood that a voter votes in accordance with his or her *ex ante* preferences increases as the *ex ante* utility difference between his or her choices increases. In other words, a voter is more likely to vote for the candidate whose platform maximizes his or her instrumental expected payoffs as the expected utility gain from doing so increases.

In addition to providing a link between empirical work on voting, expressive theories of vote choice, and the logic of strategic voting, the main contribution of this paper to the existing work in the field is to obtain a global candidate equilibrium in large electorates with very little in the way of assumptions about voter preferences. The main difference between our approach and previous work on probabilistic voting is the way in which we model the probabilistic voting. As in Ledyard [1984], by treating the voter decisions as a game, we explicitly include the pivot probability in the voters' expected utility calculations. In large electorates, because the probability of being pivotal goes to zero, the expected utility difference between any two candidates also goes to zero. Thus, under the QRE assumptions, the voter's choice is determined mainly by the candidate specific payoff disturbance. Hence, in aggregate, voters vote less based on policy, and more based on candidate attributes as the size of the electorate grows. However, even though individuals become less responsive to policy differences, in large electorates, since the total number of voters is also getting large, there is still enough policy voting at the aggregate level to force the candidates to the social optimum. Noting the generic nonexistence of pure strategy equilibria in multidimensional electoral competition when voters' behaviors are perfectly determined by objective payoffs (*e.g.*, McKelvey [1976, 1979], McKelvey and Schofield [1987]), the results of this paper can be seen as demonstrating a stabilizing role of expressive voting in large electorates, *even when voters are strategic and respond to differences between the candidates' proposed platforms.*

2 The Model

We assume the existence of a finite dimensional *policy space*, $X \subseteq \mathbb{R}^m$, where X is bounded, and finite sets N and K of *voters* and *candidates*, respectively. Write $n = |N|$ and $k = |K|$ for the total number of each. We let 0 indicate *abstention*, and write $K_0 = K \cup \{0\}$ for the set of *alternatives*: *i.e.*, the set of all candidates and abstention.

We assume that for each voter, $i \in N$, there is a space T_i of possible characteristics, or *types* of the voter. The space of all type profiles is denoted by $T = \times_{i \in N} T_i$. We assume that T_i can be represented as the Cartesian product of two sets, \mathcal{T} and \mathbb{R}^{K_0} (so that $T_i = \mathcal{T} \times \mathbb{R}^{K_0}$ for all $i \in N$). The first of these sets, \mathcal{T} represents the set of all policy-based determinants of preference, while the second, \mathbb{R}^{K_0} represents the consumption (or expressive) determinants of preference over vote choice. We assume nothing about \mathcal{T} other than that it is a complete separable metric space.⁵ Voters' preferences over the policy space are described by a utility function, $u : X \times \mathcal{T} \rightarrow \mathbb{R}$. Hence, the utility of voter $i \in N$, of type $t_i = (\tau_i, \eta_i) \in T_i$ for the policy $x \in X$ is $u(x, \tau_i)$.⁶ We assume that u is *uniformly bounded* with respect to X and \mathcal{T} , (*i.e.*, there exists a $D \in \mathbb{R}$ such that for all $x \in X$ and $\tau \in \mathcal{T}$, $|u(x, \tau)| < D$).⁷ We also assume that the marginal distribution of voter i 's types is an atomless probability measure, ρ_i , over the Borel sets of T_i , and denote the joint distribution by ρ_N .⁸ Finally, we assume that types are independently distributed: $\rho_N = \times_{i \in N} \rho_i$. While this assumption is stronger than we need, it greatly clarifies the exposition of the model's results.⁹ Note that the assumption of independence does not preclude degenerate distributions of τ_i : in such a case, all voters' policy preferences are common knowledge.

We write η_{ij} to represent the j^{th} component of η_i . For each $i \in N$, each $j \in K_0$, and each

⁵We leave both the topology and metric with which \mathcal{T} is endowed implicit, as they play no substantive role in our analysis.

⁶This assumption implies that no candidate possesses a "valence" advantage (*e.g.*, Ansolabehere and Snyder, Jr. [2000], Groseclose [2001], Aragones and Palfrey [2002], Schofield [2004], and Schofield and Sened [2005]).

⁷Uniform boundedness would follow from the traditional assumptions that u is continuous with respect to both of its arguments and that both X and \mathcal{T} are compact.

⁸It should be noted that the assumption that ρ_i is atomless does not rule out the possibility that the policy-based component of voter i 's type, τ_i , is common knowledge, as we discuss below.

⁹All of the paper's results hold with the weaker assumption that ρ_N is absolutely continuous with respect to the product measure $\times_{i \in N} \rho_i$.

$\tau_i \in \mathcal{T}$, all of the η_{ij} are assumed to be independently distributed random variables with full support, each with a cumulative distribution function that is twice continuously differentiable (and hence atomless). Thus, the assumption that ρ_i is atomless does not preclude the possibility that the marginal distribution of τ_i is degenerate (as in the classical spatial model with commonly known ideal points, for example). Put succinctly, an atom in T_i is a point (τ_i, η_i) . Thus, nonatomicity of ρ_N is assured by the fact that no such point (for any voter i) is assigned positive measure. Finally, we assume that η_{ij} is identically distributed for all electorates N , all $i \in N$, all candidates $j \in K$, and all $\tau_i \in \mathcal{T}$. As we discuss later, it is important to note that we *do not* impose any restrictions on the distribution of τ_i across voters or electorates.

It is important to note and briefly discuss our assumption that the distribution of η_{ij} is identical across all possible *electorates* N .¹⁰ This assumption is made for three reasons. First, it is in keeping with the notion of Quantal Response Equilibrium, as defined by McKelvey and Palfrey [1995, 1998]. Second, many of our results are asymptotic with respect to the size of the electorate, n . Accordingly, this assumption ensures that our results do not depend on an unverifiable assumption about the effect of n on the distribution of individuals' idiosyncratic, policy-independent payoff disturbances. The third reason is more technical, but substantively interpretable: loosely put, our results for candidate competition¹¹ essentially rely on the existence, for each electorate N , each voter $i \in N$, and each candidate $k \in K$, and every profile of policy platforms, of some strictly positive uniform lower bound on the probability that voter i will vote for candidate k in electorate N . From a game-theoretic standpoint, this requirement can be interpreted as requiring that every vote profile can follow from any profile of announced policy platforms.¹²

Any joint distribution ρ_N on T satisfying all of the above conditions is said to be *admissible*.

¹⁰We thank a referee for urging us to clarify this assumption.

¹¹Specifically, Lemma 2 and Proposition 2.

¹²Strictly speaking, we could relax the assumption that the distribution of η_{ij} is invariant across all electorates N : as the discussion indicates, given our assumed uniform bound on u across X and N , we could essentially impose a uniform bound on the “upper tail” of the distribution of η_{ij} (specifically, the probability that $\eta_{ij} \geq D$, where D is the bound on u as stated on p. 9) across all electorates N , for each voter i and candidate j . We omit this added generality for clarity of presentation, since relaxing this assumption would require us to carry further notation through our arguments.

Let μ be the common mean of η_{ij} for $j \in K$, μ_0 be the mean of η_{i0} , and $c = \mu - \mu_0$. Then c is the expected cost of voting.

We now define a game, in which the candidates each simultaneously choose policy positions in X , and then after observing the candidate policy positions, the voters vote for a candidate. Thus, the strategy set Y_i for candidate $i \in K$ is $Y_i = X$, and the set of strategy profiles for the candidates is $Y = \times_{i \in K} Y_i$. The strategy set S_i for voter $i \in N$ is the set of functions $s_i : Y \times T_i \rightarrow K_0$, and the set of strategy profiles for the voters is $S = \times_{i \in N} S_i$. We will use the notation $S_{-i} = \times_{j \neq i} S_j$, and $s_{-i} \in S_{-i}$ to represent strategy profiles for all voters except voter i , with similar notation for candidates.

Given a strategy choice $y = (y_1, \dots, y_k) \in Y$ by the candidates, and $s = (s_1, \dots, s_n) \in S$ of the voters, define for any $j \in K_0$, and $t \in T^n$

$$V_j(y, s; t) = \frac{1}{n} |\{i \in N : s_i(y, t_i) = j\}| \quad (2.1)$$

to be the proportion of the electorate who chose alternative j , and

$$W(y, s; t) = \{j \in K : j \in \operatorname{argmax}_{l \in K} V_l(y, s; t)\} \quad (2.2)$$

to be the set of winners of the election. For any subset of candidates $J \subseteq K$, write

$$P_J(y, s; t_i) = \Pr[\{t_{-i} \in T_{-i} : W(y, s; t) = J\}]. \quad (2.3)$$

to be the probability of a first place tie among the candidates J . We assume that a fair lottery is used to select a winner when there is a tie, so that we can define voter utilities over any nonempty subset of candidates, $J \subseteq K$, by

$$v_J(y, \tau_i) = \frac{1}{|J|} \sum_{j \in J} u(y_j, \tau_i). \quad (2.4)$$

The expected payoff to voter $i \in N$ of type $t_i = (\tau_i, \eta_i)$ given a strategy profile $(y, s) \in Y \times S$ is

$$U(y, s, t_i) = \sum_{J \subseteq K} P_J(y, s; t_i) \cdot v_J(y, \tau_i) + \eta_{is_i(y, t_i)} \quad (2.5)$$

In other words, a voter voting for candidate $j = s_i(y, t_i)$ receives the expected utility of the policy of the winning candidate (the summation term on the right hand side of Equation (2.5)) plus a payoff disturbance $\eta_{is_i(y, t_i)}$ that is associated with the vote, $s_i(y, t_i) \in K_0$, that the voter makes as dictated by the strategy profile (y, s) . We write $U(j; y, s, t_i) = U(y, (j, s_{-i}); t_i)$ for the utility that voter i of type t_i gets from voting for strategy j , given y , and $s_{-i} \in S_{-i}$. Thus, for all $j \in K_0$,

$$U(j; y, s, t_i) = \bar{U}(j; y, s, \tau_i) + \eta_{ij} \quad (2.6)$$

where

$$\bar{U}(j; y, s, \tau_i) = \sum_{J \subseteq K} P_J(y, (j, s_{-i}); t_i) \cdot v_J(y, \tau_i) \quad (2.7)$$

is the expected utility to voter i of type τ_i of voting for candidate j , unconditioned on the payoff disturbance, η_{ij} .

The difference in the expected utility of voting for j over abstaining can be written in the form:¹³

$$\bar{U}(j; y, s, \tau_i) - \bar{U}(0; y, s, \tau_i) = \sum_{k \neq j} \delta_i^{jk}(y, s) \cdot [u(y_j, \tau_i) - u(y_k, \tau_i)] \quad (2.8)$$

where $\delta_i^{jk}(y, s)$ is the *pivot probability* for j over $k \neq j$:

$$\delta_i^{jk}(y, s) = \sum_{J \subseteq K; j, k \in J} \frac{1}{|J|} \left(P_J(y, (0, s_{-i}); t_i) + \frac{P_J(y, (j, s_{-i}); t_i)}{|J| - 1} \right) \quad (2.9)$$

The pivot probability is the probability that by voting for j rather than abstaining, voter i changes the outcome from a win for k to a win for j . To understand Equation (2.9), note that the first

¹³Details about the derivation of pivot probabilities in multicandidate elections can be found in McKelvey and Ordeshook [1972]. In particular, Equation (2.8) follows by reversing the order of summation in the expression for $(E^j - E^0)$ of the Theorem on p. 49.

term, $P_J(y, (0, s_{-i}); t_i)$ is the probability of a tie between a set of candidates J , which includes j and k , conditional on the other voters' strategies, s_{-i} , before accounting for voter i 's vote (which is equivalent to the conditional probability of this tie given abstention by voter i , $s_i = 0$).¹⁴ The probability of victory by candidate k is $1/|N|$ in this case. Recalling that the pivot probability δ_i^{jk} is the probability of changing the election winner from k to j , *conditional on voter i voting for candidate j* and the probability of candidate j winning conditional on receiving voter i 's vote in this case is 1 (since we are studying plurality rule), then the probability of changing the election winner from k to j is $1/|J| * 1 = 1/|J|$.

The second term is the probability of a tie between the same set of candidates, J (the summation occurs over all subsets of candidates including both j and k), conditional on s_{-i} , after voter i 's vote for candidate j has been included. The probability of i 's vote having changed the winner from k to j is the probability of k having won if the set of leading candidates was $J - \{j\}$, which is $1/|J - 1|$, multiplied by the probability of j being selected from J , which is $1/|J|$. Hence, the probability of having changed the election winner from k to j , conditional on $P_J(y, (j, s_{-i}); t_i)$, is $1/(|J||J - 1|)$. Summing over all potential ties yields Equation (2.9).¹⁵

Note that $\delta_i^{jk}(y, s)$ is not necessarily equal to $\delta_i^{kj}(y, s)$ due to the fact that it includes the possibility of creating a tie between the two candidates. In particular, note that j appears on the right hand side of Equation (2.9) but k does not, introducing the potential for asymmetry between any two candidates. Whether this potential for asymmetry is realized for any given voter i depends on s_{-i} : if candidate j (say) is slightly favored over candidate k by the voters other than some voter i , then it is more likely that voter i voting for candidate k will create a tie between j and k than it is that voter i voting for j will create a tie between the two. Of course, the probability of breaking

¹⁴Because we are assuming that the types are independent (p. 9), the t_i argument in $P_J(y, (0, s_{-i}); t_i)$ is superfluous. However, as claimed on p. 2, the assumption of independence is made for expositional purposes. If this assumption is relaxed, then a rational voter should condition on his or her own type, t_i , when calculating his or her probability of casting a pivotal vote. This feature of pivot probabilities is related to recent work on the "swing voter's curse" (e.g., Austen-Smith and Banks [1996] and Feddersen and Pesendorfer [1996], among many others).

¹⁵A potentially confusing aspect of Equation (2.9) is it is pair-specific with respect to candidates, rather than the more traditional "pivot probability" for any single given candidate. This probability for a voter i and a candidate j is given by $\sum_{k \neq j} \delta_i^{jk}$. The pair-specific formulation is necessary because different candidates' platforms may offer different utilities to voter i .

a tie between the two candidates is independent of the “order” of the candidates. The key fact is that this is not the only case in which a vote can be pivotal between the two – creating a tie is also potentially important.

It then follows from Equation (2.8) that the difference in expected utility of voting for j over l is:

$$\begin{aligned} \bar{U}(j; y, s, \tau_i) - \bar{U}(l; y, s, \tau_i) &= \left(\delta_i^{jl}(y, s) + \delta_i^{lj}(y, s) \right) \cdot [u(y_j, \tau_i) - u(y_l, \tau_i)] \\ &\quad + \sum_{k \neq j, l} \left\{ \begin{aligned} &\delta_i^{jk}(y, s) \cdot [u(y_j, \tau_i) - u(y_k, \tau_i)] \\ &+ \delta_i^{lk}(y, s) \cdot [u(y_k, \tau_i) - u(y_l, \tau_i)] \end{aligned} \right\} \end{aligned} \quad (2.10)$$

which, for the case of two candidates, $K = \{j, l\}$, reduces to

$$\bar{U}(j; y, s, \tau_i) - \bar{U}(l; y, s, \tau_i) = \left(\delta_i^{jl}(y, s) + \delta_i^{lj}(y, s) \right) \cdot [u(y_j, \tau_i) - u(y_l, \tau_i)] \quad (2.11)$$

To define the candidate payoff functions, we first define $V_j(y, s)$ to be the *expected* proportion of the electorate choosing alternative j at the profile (y, s) :

$$V_j(y, s) = E_t [V_j(y, s; t)] = \frac{1}{n} E_t [|\{i \in N : s_i(y, t_i) = j\}|] \quad (2.12)$$

Writing E_z to denote expectation with respect to a random variable z , we will find it useful to re-express candidate j 's expected vote share as

$$\begin{aligned} V_j(y, s) &= \frac{1}{n} E_t [|\{i \in N : s_i(y, t_i) = j\}|] \\ &= \frac{1}{n} E_\tau [E_\eta [|\{i \in N : s_i(y, \tau_i, \eta_i) = j\}|]] \\ &= \frac{1}{n} E_\tau \left[\sum_{i \in N} \bar{s}_i(y, \tau_i)(j) \right] = \frac{1}{n} \sum_{i \in N} E_{\tau_i} [\bar{s}_i(y, \tau_i)(j)]. \end{aligned} \quad (2.13)$$

Using candidate j 's expected vote share at (y, s) , we assume that candidate j 's payoff is his or

her *margin of expected victory*, \widehat{V}_j , which is defined as:

$$\widehat{V}_j(y, s) = V_j(y, s) - \max_{l \in N - \{j\}} V_l(y, s). \quad (2.14)$$

Remark 1 Our assumption that u is uniformly bounded rules out the occurrence of the “St. Petersburg paradox.” The paradox occurs if one can construct a sequence of policies, $\{x_k\}_{k \in \mathbb{Z}_{++}}$, such that $u(x_k, \tau_i) > 2^k$ for each $k = 1, 2, \dots$. Such a voter would not trade the lottery that gives outcome x_k with probability $\frac{1}{2^k}$ for *any* policy x – the lottery has an unbounded positive expected payoff. Similarly, if one constructs the sequence $\{x_k\}_{k \in \mathbb{Z}_{++}}$ such that $u(x_k, \tau_i) < -2^k$, the voter prefer *any* policy x over the lottery. Thus, the assumption of bounded utility implies that no voter is subject to the St. Petersburg paradox. The uniformity of the bound on u rules out cases in which, as the electorate grows larger, there exists a sequence of voters the limit of whose preferences is unbounded and hence subject to the paradox.

Remark 2 Our assumptions about the distribution of voter’s types, ρ_N , encompass both the classical spatial voting framework, in which all voter ideal points are known and common knowledge and models (such as Ledyard [1984]) in which all voter types are independent and drawn independently and identically from a common distribution on voter types.

Remark 3 The assumption that the η_{ij} are independently and identically distributed with respect to voters can be viewed as an implicit normalization of utility functions. This is important in interpreting the main theorem, since the weights that individuals are given in the social utility function is determined by this normalization.

3 Voter Equilibrium

In this section, we consider the voter equilibrium to the game defined by Equation (2.6) for any fixed profile of candidate positions, $y \in Y$. Once y is fixed, the strategy space for the voter reduces from S_i (the set of functions $s_i : Y \times T_i \rightarrow K_0$) to the set of functions of the form $s_i(y, \cdot) : T_i \rightarrow K_0$.

We write $S_i(y)$ to designate this conditional strategy space, and $S(y)$ to designate the set of profiles of conditional strategies.

For any fixed $y \in Y$, we define a *voter equilibrium for y* to be a pure strategy Bayesian Nash equilibrium to the voter game defined by Equation (2.6) over the strategy space $S(y)$. This is any profile, $s \in S(y)$, in which voters always choose an action that maximizes expected utility conditional on their type. Thus, s is a voter equilibrium for y if for all $i \in N$, $t_i \in T_i$, and $j \in K_0$,

$$\begin{aligned} s_i(y, t_i) = j &\Leftrightarrow U(j; y, s, t_i) = \max_{l \in K_0} U(l; y, s, t_i) \\ &\Leftrightarrow \bar{U}(j; y, s, \tau_i) + \eta_{ij} = \max_{l \in K_0} [\bar{U}(l; y, s, \tau_i) + \eta_{il}] \end{aligned} \quad (3.1)$$

Note that the structure of the payoffs is essentially the same as used in McKelvey and Palfrey [1998] in defining the agent quantal response equilibrium (AQRE) for extensive form games.¹⁶ So as long as the distribution of the errors, η_{ij} is admissible, a Bayesian Nash equilibrium to the voter game is exactly the same as an AQRE to the game. The following proposition assures us that a voter equilibrium exists for any profile of policy platforms. The proofs for this and all following numbered results are contained in the appendix.

Proposition 1 *For any $y \in Y$, there exists a voter equilibrium for y .*

Of particular interest is the average behavior of a voter i of type t_i , after integrating out η_i . For any $s_i(y, \cdot) \in S_i(y)$, define $\bar{s}_i(y, \cdot) : \mathcal{T} \rightarrow \Delta^{|K_0|}$,¹⁷ as the marginal distribution of s_i with respect to η_i : for any $\tau_i \in \mathcal{T}$ and $j \in K_0$,

$$\bar{s}_i(y, \tau_i)(j) = \Pr[\eta_i : s_i(y, (\tau_i, \eta_i)) = j]. \quad (3.2)$$

We have assumed that the η_{ij} are independently distributed, for all i, j and τ_i , and identically

¹⁶Strictly speaking, our framework is slightly more general than AQRE, since we allow for the distribution of η_{i0} to have a different mean than the distribution of η_{ij} for $j \in K$.

¹⁷Throughout the paper, the notation $\Delta^{|K_0|}$ denotes the $|K_0| - 1$ -dimensional simplex and the notation $\Delta^{|K|}$ denotes the $|K| - 1$ dimensional simplex.

distributed for all $j \in K$. Let $H(\cdot)$ be the cumulative distribution function of η_i , i.e., $H(w) = \Pr[\eta_{ij} \leq w_j \text{ for all } j \in K_0]$ for $w \in \mathfrak{R}^{|K_0|}$. And let $G_j(\cdot)$ be the cumulative distribution function of $z \in \mathfrak{R}^{|K|}$, where $z_l = \eta_{il} - \eta_{ij}$ for $l \in K - \{j\}$, and $z_j = \eta_{i0} - \eta_{ij}$. Thus,

$$G_j(z) = \Pr[\eta_{i0} - \eta_{ij} \leq z_j \text{ and } \eta_{il} - \eta_{ij} \leq z_l \text{ for all } l \neq j] \quad (3.3)$$

for any $z \in \mathfrak{R}^{|K|}$. Under the assumptions we have made on the η_{ij} , for all $j \in K$, both $H(w)$ and $G_j(z)$ are twice continuously differentiable and strictly increasing in all arguments, and everywhere positive. Thus, if s is a Bayesian Nash equilibrium, applying Equation (3.1), for $j \in K$,

$$\begin{aligned} \bar{s}_i(y, \tau_i)(j) &= \Pr[\bar{U}(j; y, s, \tau_i) + \eta_{ij} = \max_{l \in K_0} [\bar{U}(l; y, s, \tau_i) + \eta_{il}]] \\ &= \Pr[\eta_{il} - \eta_{ij} \leq \bar{U}(j; y, s, \tau_i) - \bar{U}(l; y, s, \tau_i) \text{ for all } l \in K_0 - \{j\}] \\ &= G_j \left(\begin{bmatrix} \bar{U}(j; y, s, \tau_i) - \bar{U}(1; y, s, \tau_i) \\ \dots \\ \bar{U}(j; y, s, \tau_i) - \bar{U}(j-1; y, s, \tau_i) \\ \bar{U}(j; y, s, \tau_i) - \bar{U}(0; y, s, \tau_i) \\ \bar{U}(j; y, s, \tau_i) - \bar{U}(j+1; y, s, \tau_i) \\ \dots \\ \bar{U}(j; y, s, \tau_i) - \bar{U}(k; y, s, \tau_i) \end{bmatrix} \right). \end{aligned} \quad (3.4)$$

Example: One example of the above is the logit AQRE, where the density functions of $w_0 = \eta_{i0} + c$ and $w_j = \eta_{ij}$ for $j \in K$ follow a type one extreme value distribution, in which $H_j(w_j) = \exp[-\exp[-\lambda w_j]]$. The independence of η_{ij} across i and j implies $H(w) = \prod_j H_j(w_j)$. This

leads to the logistic formula $G_j(z) = \frac{1}{1 + \exp[\lambda(c + z_j) + \sum_{l \neq j} \exp(\lambda z_l)]}$. In this case, for fixed λ , we get:

$$\begin{aligned} \bar{s}_i(y, \tau_i)(j) &= G_j(\bar{U}^j(y, s, \tau_i)) \\ &= \frac{1}{1 + \exp[\lambda \cdot (c + \bar{U}(0; y, s, \tau_i) - \bar{U}_i(j; y, s, \tau_i))] \\ &\quad + \sum_{l \neq j} (\exp[\lambda \cdot (\bar{U}(l; y, s, \tau_i) - \bar{U}_i(j; y, s, \tau_i)])]} \end{aligned}$$

and in the case of two candidates, where $K = \{j, l\}$,

$$\begin{aligned} \bar{s}_i(y, \tau_i)(j) &= \frac{1}{1 + \exp(\lambda \cdot (c + \delta^{jl}(y, s) \cdot [u(y_j, \tau_i) - u(y_l, \tau_i)])) \\ &\quad + \exp(\lambda \cdot (\delta^{jl}(y, s) + \delta^{lj}(y, s)) \cdot [u(y_j, \tau_i) - u(y_l, \tau_i)])} \end{aligned}$$

We now show that for fixed candidate positions at $y \in Y$, and for *any* voter equilibrium, *all* pivot probabilities go to zero and the probability of voting for any two candidates in K becomes equal as $n \rightarrow \infty$. The reason for this result is simple: one's vote only matters when it is pivotal.¹⁸ Thus, one's vote only matters when the other voters are either evenly split between the two top candidates or when the vote difference between the two top candidates differs by one vote. As n grows large, this becomes a very low probability event. Thus, in general, one's vote rarely affects the outcome of the election. This implies that voters effectively become indifferent with respect to which candidate they vote for as $n \rightarrow \infty$. We formalize this in the following proposition.

Proposition 2 *Fix $y \in Y$, and for each integer n , let ρ_N be any admissible joint distribution over $\times_{i=1}^n T_i$, and let s^n be any AQRE for the voters. Then for any $j, l \in K$ and $i, k > 0$,*

- (a) $\lim_{n \rightarrow \infty} \delta_i^{jl}(y, s^n) = 0$ and
- (b) $\lim_{n \rightarrow \infty} \delta_i^{jl}(y, s^n) / \delta_k^{jl}(y, s^n) = 1$
- (c) $\lim_{n \rightarrow \infty} \delta_i^{jl}(y, s^n) / \delta_i^{lj}(y, s^n) = 1$

¹⁸For more on the logic of pivotal voting, see Myerson and Weber [1993].

$$(d) \lim_{n \rightarrow \infty} [\bar{s}_i^n(y, \tau_i)(j) - \bar{s}_i^n(y, \tau_i)(l)] = 0.$$

Further, in all cases, the convergence is uniform: for any $\varepsilon > 0$, there is an n_ε such that, for all $i, k, j, l, y, \rho_N, s^n$, if $n > n_\varepsilon$,

- $\delta_i^{jl}(y, s^n) < \varepsilon$,
- $\left| \delta_i^{jl}(y, s^n) / \delta_k^{jl}(y, s^n) - 1 \right| < \varepsilon$,
- $\left| \delta_i^{jl}(y, s^n) / \delta_i^{lj}(y, s^n) - 1 \right| < \varepsilon$, and
- $|\bar{s}_i^n(y, \tau_i)(j) - \bar{s}_i^n(y, \tau_i)(l)| < \varepsilon$.

Before continuing, it should be emphasized that Proposition 2 does not impose any requirements on the marginal distribution of τ_i for any voter i as the electorate grows. While we have assumed that the marginal distributions of the payoff disturbances (*i.e.*, $\{\eta_{ij}\}_{j \in K_0}$) are fixed across all electorates N , the marginal distribution of voter 1's type with respect to τ_1 (*e.g.*, voter 1's ideal point) when there are $n + 1$ voters in the electorate is not necessarily equal to the marginal distribution of τ_1 when there are n voters in the electorate. In addition, the choice of a marginal distribution of each “new” voter's policy-based preference type is similarly unconstrained. We feel that this fact illuminates the strength of the proposition. In other words, great regularity is imposed asymptotically upon individual pivot probabilities in *any* AQRE.¹⁹

Remark 4 Note that the requirement that voters adopt a Bayesian equilibrium means that voters vote strategically in multi-candidate elections. Thus, a voter may vote with higher probability for their *a priori* second-ranked candidate than for their *a priori* first-ranked candidate if the pivot probability for the first-ranked candidate is sufficiently low in relation to that for the second-ranked candidate.

¹⁹We thank a referee for helping us clarify this discussion.

4 Candidate Equilibrium

This section examines the incentives of candidates competing for votes in a world populated by voters who play quantal response equilibrium strategies. Making an additional assumption on preferences, we establish that for a large enough electorate, N , all candidates adopting the social optimum constitutes a global equilibrium. Furthermore, the global equilibrium is unique for two candidate elections. The proofs of our results utilize our assumptions that the η_{ij} are *i.i.d.* with full support and that, for each $j \in K_0$, the distribution of the η_{ij} is invariant to the electorate, N .

For a fixed electorate, N , and measure ρ_N on $T = \times_{i \in N} T_i$, let s be any strategy profile for the voters²⁰ such that for any candidate positions, $y \in Y$, $s(y, \tau)$ is a quantal response equilibrium for the voters, as described in the previous section. Then, as discussed earlier, each candidate $j \in K$ is assumed to maximize

$$\widehat{V}_j(y) = \widehat{V}_j(y, s) = V_j(y) - \max_{l \in N - \{j\}} V_l(y). \quad (4.1)$$

For any admissible type distribution ρ_N , let

$$x^*(\rho_N) = \operatorname{argmax}_{x \in X} \sum_{i \in N} E_{\tau_i}[u(x, \tau_i)] \quad (4.2)$$

denote the *expected social optimum* with respect to ρ_N . We assume throughout the remainder of the paper that for each N and ρ_N , that such a point exists, is unique, and lies in the interior of the policy space, X .²¹

Assumption 1 *For all N , the expected social optimum, $x^*(\rho_N)$, exists, is unique, and lies in the interior of X .*

We first show that in general we cannot expect even a local equilibrium to exist at $x^*(\rho_N)$

²⁰To be technically correct, since we are considering N and ρ_N to be variables, we should condition voter and candidate strategies accordingly. To simplify notation, we omit these parameters.

²¹We have not explored the possibility of multiple social optima or the possibility that the social optimum lies on the boundary of X .

without some additional restrictions on preferences or the policy space.

Example 1 (Nonexistence of Equilibrium at $x^*(\rho_N)$.) Consider a simple model of electoral politics: two candidates, with no abstention, and a one dimensional policy space. Specifically, suppose that $K = \{1, 2\}$, $X = [-1, 1]$, and let $\mathcal{T} = \{0, 1\}$, associated with the following (strictly concave) utility functions:

$$u(x, 0) = -x - x^4 \quad (4.3)$$

$$u(x, 1) = x/2 - x^4. \quad (4.4)$$

Now assume that in a small neighborhood of zero, the CDF for $G(z)$ is

$$G(z) = 1/2 + z - z^3.^{22}$$

For some positive integer k , let $n = 3k$. Finally, let ρ_N be a measure that puts all mass on the vector τ^* such that for $1 \leq i \leq k$, $\tau_{3i}^* = 0$ and $\tau_{3i-2}^* = \tau_{3i-1}^* = 1$. Then

$$\frac{1}{n} \sum_i u(x, \tau_i^*) = -x^4,$$

so that the social utility maximizing policy is $x^*(\rho_N) = 0$.

To see that 0 is not an equilibrium, suppose that candidate 2 adopts $x \neq 0$ against $x^*(\rho_N) = 0$. Using Equations (3.4) and (2.13) and letting $\delta_{\tau_i^*} = \delta_i^{12}$, the expected vote for candidate i is

²²Note this function is symmetric around zero, implying that there exists admissible type distributions that are consistent with this.

(omitting a few steps of algebra):

$$\begin{aligned}
V_2(x, 0) &= \frac{1}{n} \sum_i G(\delta_{\tau_i^*} u(x, \tau_i^*)) \\
&= \frac{1}{3} [2G(\delta_1 u(x, 1)) + G(\delta_0 u(x, 0))] \\
&= \frac{1}{3} \left(-\delta_0^3 (-x^4 - x)^3 + \delta_0 (-x^4 - x) + 2 \left(-\delta_1^3 \left(\frac{x}{2} - x^4 \right)^3 + \delta_1 \left(\frac{x}{2} - x^4 \right) \right) \right) + \frac{1}{2}
\end{aligned}$$

Omitting the straightforward calculations, the first derivative of candidate 3's expected vote at $x = 0$ is equal to

$$\frac{dV_2'(0, 0)}{dx} = \frac{1}{3}(\delta_1 - \delta_0). \quad (4.5)$$

Thus, in order for $x = 0$ to be an equilibrium, it must be the case that $\delta_1 = \delta_0$. Suppose that this is the case. Regardless of δ_0 and δ_1 , the second derivative of candidate 2's expected vote at $x = 0$ is equal to 0. Accordingly, for $x = 0$ to be an equilibrium, it must be the case that the *third* derivative of candidate 2's expected vote is negative. However, the third derivative of candidate 2's expected vote at $x = 0$ is equal to

$$2\delta_0^3 - \frac{\delta_1^3}{2}. \quad (4.6)$$

Thus, if $\delta_0 \neq \delta_1$, then the first-order necessary condition for an equilibrium at 0, Equation (4.5), is not satisfied. If $\delta_0 = \delta_1$, then the first-order necessary condition is satisfied, but the second derivative of candidate 2's expected vote is equal to zero, while the third derivative is nonzero, implying that $x = 0$ is a point of inflection. Accordingly, independent of δ_0 and δ_1 , $x = 0$ is not an equilibrium. \triangle

To avoid situations analogous to Example 1, we define the following condition on the sequence of preference distributions as the electorate grows.

Condition (4.7). *There exists an integer \tilde{n} such that, for all N with $|N| > \tilde{n}$, there exists a finite*

number M satisfying the following for each ρ_N on T :

$$M > \sup_{x \in X - \{x^*\}} \frac{-E_\tau [\sum_{i \in N} [u(x, \tau_i) - u(x^*, \tau_i)]^2]}{E_\tau [\sum_{i \in N} [u(x, \tau_i) - u(x^*, \tau_i)]]}. \quad (4.7)$$

Condition (4.7) amounts to a regularity condition on the sequences of preference profiles we consider for the remainder of the paper. The preferences in Example 1 do not satisfy Condition (4.7). The condition requires that, for all electorates exceeding some fixed finite size, there exists a uniform bound (across all policies and large electorates) on the ratio of the variance and mean of the voters' payoffs.

We now provide two sufficient conditions for the satisfaction of Condition (4.7). The lemmas require some additional notation. Let \mathcal{D}^m be the set of unit length direction vectors in \mathbb{R}^m and h be an arbitrary element of \mathcal{D}^m . For any vector $h \in \mathcal{D}^m$, $D_h u(x, \tau)$ and $D_h^2 u(x, \tau)$ denote the first and second directional derivatives of u , respectively, in the direction h . Finally, for any function ϕ , let $|\nabla_x| \phi$ denote the magnitude of the gradient of ϕ with respect to x , and let $H_x[\phi]$ denote the Hessian of ϕ with respect to x at (x, τ) .

The first lemma establishes that compactness of X , smoothness of $u(x, \tau)$ for all τ , and uniform upper bounds on $|\nabla_x| u(x^*(\rho_N), \tau)$ and the eigenvalues of $H_x[E_\tau[\sum_{i \in N} u(x^*(\rho_N), \tau)]]$ are sufficient for the satisfaction of Condition (4.7). The second corollary notes that, when the voters' types are independently and identically distributed (as in Ledyard [1984], for example), the requirement of a uniform bound on the eigenvalues of $H_x[E_\tau[\sum_{i \in N} u(x^*(\rho_N), \tau)]]$ can be replaced with simply requiring that $H_x[E_\tau[\sum_{i \in N} u(x^*(\rho_N), \tau)]]$ be negative definite.

In addition to providing leverage for the application of this paper's results in other settings, Lemma 1 and Corollary 1 are also intended to illustrate the substantive restriction that Condition (4.7) is intended to impose "eventually" on the sequence of preference distributions. Essentially, this requirement is that the sensitivity of individual preferences to policy (*i.e.*, $|\nabla_x| u(x, \tau)$) not be arbitrarily larger than the sensitivity of the sum of individual preferences in a neighborhood of the social welfare optimum, $x^*(\rho_N)$. The distinction here is intimately related to the difference

between (A) strict concavity of a real-valued function and (B) negative definiteness of its Hessian matrix. Of course, for any function, (B) implies (A), but the converse implication does not hold. Example 1, above, demonstrated that strict concavity of preferences is not enough to guarantee equilibrium. Lemma 1 includes Example 1, with the exception that it requires that the eigenvalues of the Hessian of the sum of individuals' utility functions have a strictly negative uniform upper bound at $x^*(\rho_N)$.²³ Quadratic preferences (based on Euclidean distance from an ideal point) over a compact space X would satisfy the necessary assumptions.

Lemma 1 *Assume that X is compact and, for every $\tau \in \mathcal{T}$, let $u(x, \tau)$ be twice continuously differentiable in x . If there exists a finite $B > 0$ such that, for all N , and all $\tau \in \mathcal{T}$,*

1. $|\nabla_x|E_\tau[\sum_{i \in N} u(x^*(\rho_N), \tau)] \in [-B, B]$ and
2. *the maximum eigenvalue of $H_x [E_\tau[\sum_{i \in N} u(x^*(\rho_N), \tau)]]$ is no greater than $-B$,*

then Condition (4.7) is satisfied.

The next corollary follows immediately from Lemma 1.

Corollary 1 *Assume that X is compact and, for every $\tau \in \mathcal{T}$, let $u(x, \tau)$ be twice continuously differentiable in x . If*

1. *each ρ_N is an admissible product measures of the form $\rho_N = \times_{i \in N} \rho^*$, and*
2. $H_x [E_\tau[\sum_{i \in N} u(x^*(\rho_N), \tau)]]$ *is negative definite at $x^*(\rho_N)$,*

then Condition (4.7) is satisfied.

DISCUSSION OF LEMMA 1 AND CONDITION (4.7). Before continuing to the main result, it is worthwhile to comment further on the relationship between Condition (4.7) and Lemma 1. As illustrated by the proof of Lemma 1 (in the appendix), constructing an example in which the sequence of type distributions does not satisfy Lemma 1 but *does* satisfy Condition (4.7) essentially

²³For each electorate N , the Hessian in question is evaluated at the social welfare optimum. Thus, the required uniformity of the upper bound is with respect to the set of all feasible electorates.

requires that one construct a sequence of type distributions for which the limit of the sequence of determinants of $\left[H_x \left[E_\tau \left[\sum_{i \in N} u(x^*(\rho_N, \tau)) \right] \right] \right]$ is zero. Viewed another way, the uniqueness of the social welfare optimum is essentially vanishing in such a sequence of electorates.

Finally, if one assumes that X is compact and $u(x, \tau)$ is continuous in x for each $\tau \in \mathcal{T}$, then by the assumption that $x^*(\rho_N)$ is unique, a necessary condition for (4.7) to be violated is the existence of a subsequence of electorates N satisfying the following:

$$\lim_{x \rightarrow x^*} \frac{-E_\tau \left[\sum_{i \in N} [u(x, \tau_i) - u(x^*, \tau_i)]^2 \right]}{E_\tau \left[\sum_{i \in N} [u(x, \tau_i) - u(x^*, \tau_i)] \right]} = \infty. \quad (4.8)$$

Accordingly, verifying that Equation (4.8) does not hold for any subsequence of electorates is sufficient to ensure that Condition (4.7) is satisfied.

4.1 The Main Result

We now present the main result, Theorem 1, which states that the social optimum is a global equilibrium in large enough electorates so long as preferences are uniformly bounded and the distribution of types across electorates satisfies Condition (4.7).

Theorem 1 *Let u be uniformly bounded, and assume that Condition (4.7) is met. There exists an integer n^* such that for any set of voters N with $|N| = n > n^*$, and any admissible ρ_N on $T = \times_{i \in N} T_i$, $y^* = (x^*(\rho_N), \dots, x^*(\rho_N))$ constitutes a global equilibrium under the margin of expected victory: for any $j \in K$ and $y_j \in X$, $\widehat{V}_j(y) = \widehat{V}_j(y_j, y_{-j}^*) \leq \widehat{V}_j(y^*)$, with the weak inequality becoming strict whenever $y_j \neq x^*(\rho_N)$.*

For the case of two candidates, the equilibrium identified in Theorem 1 is unique.

Corollary 2 *If $k = 2$, then the equilibrium found in Theorem 1 is unique.*

5 Conclusions and Extensions

In this paper we have provided a general framework for probabilistic spatial voting models in large electorates. In particular, we have extended equilibrium results of Coughlin, Ledyard, Banks and Duggan, and other researchers to policy spaces of arbitrary finite dimensionality and elections with both abstention and arbitrary numbers of candidates. In addition, while our model is agnostic as to the cause of probabilistic choice – the probabilistic choice in a QRE model can be assumed to arise either as the result of rational behavior under payoff disturbances (as we have modeled it here), or as the result of boundedly rational behavior – allows for strategic behavior by the voters. In particular, our model incorporates strategic voting within a probabilistic voting setting.

Of course, regardless of the structure of the underlying framework, the primary question that any model must confront is that of empirical validity. Does the equilibrium prediction of convergence comport with observed political platforms? While many readers have (and undoubtedly will) doubt whether policy convergence is observed in real-world elections, adequately testing this prediction is difficult for (at least) three reasons. First, the issue of whether the platforms offered by the major political parties differ from one another is debatable on several levels. This is because the platforms in this paper represent the credible commitments of the parties regarding which policies they will implement if elected. Thus, differences in announced platforms that will not translate into differences in policy outcomes are not inconsistent with the convergent equilibrium constructed here.²⁴ This is further complicated by the mechanisms of policymaking in real world democracies. The President of the United States does not set policy unilaterally. Similarly, party leaders in parliamentary systems generally serve at the pleasure of their party's MPs. Further theorizing about electoral competition within richer models of policymaking is necessary before a definitive conclusion can be reached about the effective amount of divergence between observed

²⁴In addition, our results are based on the assumption that the social welfare maximizing policy is unique. While we do not feel that this assumption is restrictive when the space of feasible platforms possesses nonempty interior, it becomes much more restrictive if one allows for finite policy spaces and indifference about certain components of policy by substantial proportions of the electorate. We have not explored the impact of multiple social welfare optima on electoral competition within a quantal response voting framework.

party platforms.

Secondly, the results presented here are asymptotic. Furthermore, the electorate size that is required to guarantee that the social welfare maximizing policy is a convergent equilibrium in the electoral competition game depends upon the distribution of payoff disturbances. Speaking somewhat loosely, the required size is decreasing in the variance of this distribution. In other words, as “expressive” motivations become a larger determinant of individuals’ vote choices, the number of voters required for the social welfare optimum to be an equilibrium decreases.²⁵ Thus, properly gauging whether one should expect convergence in electoral competition requires an estimation of the relative strength of expressive versus instrumental benefits in determining individual vote choice within a particular polity.

Finally, even if convergence is not observed, our results do not rule out the existence of other equilibria in races with more than two candidates. On a related note, the motivations of real-world political parties may include more than plurality maximization. Candidates and/or party leaders may seek to maximize their own policy-based payoff functions (*e.g.*, Wittman [1983], Duggan and Fey [2005]). Similarly, minor parties may play out-of-equilibrium strategies for a variety of reasons. Finally, it is entirely plausible that equilibrium platforms are divergent because of nonpolicy (*i.e.*, “valence”) advantages accruing to one or more of the parties (as examined in Ansolabehere and Snyder, Jr. [2000], Groseclose [2001], Aragonés and Palfrey [2002], Schofield [2004], and Schofield and Sened [2005]). This type of setting is ruled out in our model by the assumption that voters’ preferences over the election outcome depend only on the policy chosen by the winning candidate and not on his or her identity. It should be noted, however, that the predictions of a “valence” model of electoral politics diverge from those presented here in an interesting way only if the candidates are assumed to have policy preferences as well, as discussed by Wittman [1983], Calvert [1985], Duggan and Fey [2005], and others. Accordingly, inclusion of different candidate, party, and/or voter motivations represent promising avenues for future research.

²⁵This is in accordance with the findings of Lin et al. [1999].

References

- R. Michael Alvarez and Jonathan Nagler. Correlated Disturbances in Discrete Choice Models: A Comparison of Multinomial Probit Models and Logit Models. Mimeo, California Institute of Technology, 2001.
- R. Michael Alvarez and Jonathan Nagler. Economics, Issues, and the Perot Candidacy: Voter Choice in the 1992 Presidential Election. *American Journal of Political Science*, 39:714–744, 1995.
- R. Michael Alvarez and Jonathan Nagler. When Politics and Models Collide: Estimating Models of Multiparty Competition. *American Journal of Political Science*, 42:55–96, 1998.
- R. Michael Alvarez, Jonathan Nagler, and Shaun Bowler. Issues, Economics, and the Dynamics of Multiparty Elections: The British 1987 General Election. *American Political Science Review*, 94:131–149, 2000.
- Stephen Ansolabehere and James M. Snyder, Jr. Valence Politics and Equilibrium in Spatial Election Models. *Public Choice*, 103(3):327–336, 2000.
- Enriqueta Aragonés and Thomas Palfrey. Mixed Equilibrium in a Downsian Model with a Favored Candidate. *Journal of Economic Theory*, 103(1):131–161, 2002.
- David Austen-Smith and Jeffrey S. Banks. Information Aggregation, Rationality and the Condorcet Jury Theorem. *American Political Science Review*, 90:3445, 1996.
- Jeff Banks, John Duggan, and Michel Le Breton. Bounds for Mixed Strategy Equilibria and the Spatial Model of Elections. *Journal of Economic Theory*, 103:88–105, 2002.
- Jeffrey Banks and John Duggan. Probabilistic Voting in the Spatial Model of Elections: The Theory of Office-Motivated Candidates. In David Austen-Smith and John Duggan, editors, *Social Choice and Strategic Decisions*. Springer, New York, NY, 2004.

- Rabindra Nath Bhattacharya and Ramaswamy Ranga Rao. *Normal Approximation and Asymptotic Expansions*. R. E. Krieger Publishing Co, Malabar, FL, 1986.
- Duncan Black. On the Rationale of Group Decision-making. *Journal of Political Economy*, 56: 23–34, 1948.
- H. Geoffrey Brennan and Loren E. Lomasky. *Democracy and Decision*. Cambridge University Press, Cambridge, 1993.
- James M. Buchanan. Individual Choice in Voting and the Market. *Journal of Political Economy*, 62:334–343, 1954.
- Randall Calvert. Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence. *American Journal of Political Science*, 28(1):127–146, 1985.
- Peter J. Coughlin. *Probabilistic Voting Theory*. Cambridge University Press, Cambridge, 1992.
- Peter J. Coughlin and Shmuel Nitzan. Electoral Outcomes with Probabilistic Voting and Nash Social Welfare Maxima. *Journal of Public Economics*, 15:113–122, 1981a.
- Peter J. Coughlin and Shmuel Nitzan. Directional and Local Electoral Equilibria with Probabilistic Voting. *Journal of Economic Theory*, 24:226–240, 1981b.
- Jay K. Dow and James W. Endersby. Multinomial Probit and Multinomial Logit: A Comparison of Choice Models for Voting Research. *Electoral Studies*, 23:107–122, 2004.
- John Duggan and Mark Fey. Electoral competition with policy-motivated candidates. *Games and Economic Behavior*, 51(2):490–522, 2005.
- Timothy Feddersen and Wolfgang Pesendorfer. The Swing Voters Curse. *American Economic Review*, 86(3):408–424, 1996.
- Tim Groseclose. A Model of Candidate Location When One Candidate Has a Valence Advantage. *American Journal of Political Science*, 45(4):862–886, 2001.

- Melvin J. Hinich. Equilibrium in Spatial Voting: The Median Voter Result is an Artifact. *Journal of Economic Theory*, 16:208–219, 1977.
- Dean Lacy and Barry C. Burden. The Vote-Stealing and Turnout Effects of Ross Perot in the 1992 U. S. Presidential Election. *American Journal of Political Science*, 43:233–255, 1999.
- John Ledyard. The Pure Theory of Large Two-Candidate Elections. *Public Choice*, 44:7–41, 1984.
- Tse-Min Lin, James Enelow, and Han Dorussen. Equilibrium in Multicandidate Probabilistic Voting. *Public Choice*, 98:59–82, 1999.
- Richard McKelvey. Intransitivities in Multidimensional Voting Models and Some Implications for Agenda Control. *Journal of Economic Theory*, 12:472–484, 1976.
- Richard McKelvey. General Conditions for Global Intransitivities in Formal Voting Games. *Econometrica*, 47:1085–1111, 1979.
- Richard McKelvey and Norman Schofield. Generalized Symmetry Conditions at a Core Point. *Econometrica*, 55:923–933, 1987.
- Richard D. McKelvey and Peter C. Ordeshook. A General Theory of the Calculus of Voting. In J. Herndon and J. Bernd, editors, *Mathematical Applications in Political Science VI*, pages 32–78, Charlottesville, 1972. University Press of Virginia.
- Richard D. McKelvey and Tom R. Palfrey. Quantal Response Equilibria in Normal Form Games. *Games and Economic Behavior*, 10:6–38, 1995.
- Richard D. McKelvey and Tom R. Palfrey. Quantal Response Equilibria in Extensive Form Games. *Experimental Economics*, 1:9–41, 1998.
- Mark M. Meerschaert and Hans-Peter Scheffler. *Limit Distributions for Sums of Independent Random Vectors*. John Wiley and Sons, New York, 2001.

- Paul Milgrom and Robert Weber. Distributional Strategies for Games with Incomplete Information. *Mathematics of Operations Research*, 10:619–632, 1985.
- Roger Myerson. Large Poisson Games. *Journal of Economic Theory*, 94:7–45, 2000.
- Roger Myerson and Robert Weber. A Theory of Voting Equilibria. *American Political Science Review*, 87:102–114, 1993.
- John W. Patty. Equivalence of Objectives in Two Candidate Elections. *Public Choice*, 112(1): 151–166, 2002.
- John W. Patty. Generic Difference of Expected Vote Share and Probability of Victory Maximization in Simple Plurality Elections with Probabilistic Voters. *Social Choice and Welfare*, Forthcoming, 2006.
- John W. Patty. Local Equilibrium Equivalence in Probabilistic Voting Models. *Games and Economic Behavior*, 51(2):523–536, 2005.
- Kevin M. Quinn, Andrew D. Martin, and Andrew B. Whitford. Voter Choice in Multiparty Democracies: A Test of Competing Theories and Models. *American Journal of Political Science*, 43: 1231–1247, 1999.
- Norman Schofield. Local Political Equilibria. In David Austen-Smith and John Duggan, editors, *Social Choice and Strategic Decisions*. Springer, Heidelberg, 2004.
- Norman Schofield. A Valence Model of Political Competition in Britain, 1992-1997. *Electoral Studies*, 24(3):347–370, 2005.
- Norman Schofield and Itai Sened. Modeling the Interaction of Parties, Activists and Voters: Why Is the Political Center So Empty? *European Journal of Political Research*, 44(3):355–390, 2005.

Norman Schofield, Andrew D. Martin, Kevin M. Quinn, and Andrew B. Whitford. Multiparty Electoral Competition in Netherlands and Germany: A Model based on Multinomial Probit. *Public Choice*, 97:257–293, 1998.

Alexander A. Schuessler. *A Logic of Expressive Choice*. Princeton University Press, Princeton, 2000.

Gordon Tullock. The Charity of the Uncharitable. *Western Economic Journal*, 9:379–392, 1971.

Donald A. Wittman. Candidate Motivation: A Synthesis of Alternative Theories. *American Political Science Review*, 77:142–157, 1983.

A Proofs of Numbered Results

Proof of Proposition 1

Proof: This is a game of incomplete information, with action spaces $A_i = K_0$ and type space T_i for each $i \in N$. The action spaces are finite, and the distribution of types is equal to (and hence absolutely continuous with respect to) the product distribution of the marginal distributions of types across individuals. Thus, we can apply Theorem 1 of Milgrom and Weber [1985] to conclude that there exists an equilibrium in distributional strategies. Further, since the distribution of player i 's types, ρ_i , is assumed atomless, it follows from Theorem 4 in the same paper that the equilibrium can be purified to be in pure strategies. ■

Proof of Proposition 2

To prove Proposition 2, we first need a Lemma.

Lemma 2 Fix $\varepsilon^* > 0$, and let \mathcal{Z}^n be the set of sequences $Z = (Z_1, \dots, Z_n)$ of independent random vectors $Z_i \in \mathbb{R}^{|K_0|}$ of the form

$$Z_i = \left\{ \alpha_j \text{ with probability } p_{ij} \text{ for } j \in K_0 \right.$$

where α_j is the j^{th} unit basis vector in $\mathbb{R}^{|K_0|}$, and $p \in (\Delta^{|K_0|})^n$ satisfies $p_{ij} \geq \varepsilon^*$ for all i, j . For any $J \subseteq K$, define

$$B_J = \{z \in \Delta^{|K_0|} : z_j = z_k > z_l \text{ for all } j, k \in J, l \in K \setminus J\}.$$

Write $\bar{Z} = \frac{1}{n} \sum_i Z_i$, and define

$$\delta_J^{n*} = \max_{Z \in \mathcal{Z}^n} \Pr[\bar{Z} \in B_J] \quad (\text{A.1})$$

Then for any $J \subseteq K$ with $|J| \geq 2$,

$$(a) \lim_{n \rightarrow \infty} \delta_J^{n*} = 0$$

$$(b) \lim_{n \rightarrow \infty} \delta_{J'}^{n*} / \delta_J^{n*} = 0 \text{ for any } J \subsetneq J'$$

Proof: An element $Z = (Z_1, \dots, Z_n) \in \mathcal{Z}^n$ consists of independent, but not identically distributed random vectors, and is characterized by a vector $p = (p_1, \dots, p_n)$, where $p_i = (p_{i0}, p_{i1}, \dots, p_{iK}) \in \Delta^{|K_0|}$. The mean of Z_i is $\mu_i = (p_{i1}, \dots, p_{iK})$ which consists of all but the first component of p . Pick $Z^n = (Z_1^n, \dots, Z_n^n) \in \mathcal{Z}^n$ to attain the maximum in Equation (A.1). Since $\Pr[\bar{Z} \in B_J]$ is continuous as a function of p , which ranges over a compact set, it follows that such a δ_J^{n*} and Z^n exist. Define $X_i^n = Z_i^n - \mu_i$. Then the X_i^n form a triangular array (see Meerschaert and Scheffler [2001], Definition 3.2.1, p. 52), where each random variable X_i^n has zero mean, and for each n , the X_i^n are independent.

Define V_{ni} to be the variance covariance matrix of Z_i^n and let $V_n = \frac{1}{n} \sum_i V_{ni}$ denote the variance-covariance matrix of the random variable $\sum_i X_i^n$. By the assumption that $p_{ij} > \varepsilon^*$ for all $j \in K_0$, it follows that V_n is strictly positive definite and hence invertible. Let T_n denote the

symmetric, positive definite matrix satisfying $T_n^2 = V_n^{-1}$. Then

$$\begin{aligned}
\delta^{n*} &= \Pr[\bar{Z}^n \in B_J] \\
&= \Pr \left[\begin{array}{l} \sum_i Z_{ij}^n - \sum_i Z_{ik}^n = 0 \text{ for } j, k \in J, \text{ and} \\ \sum_i Z_{ij}^n - \sum_i Z_{il}^n > 0 \text{ for } j \in J, l \in K \setminus J \end{array} \right] \\
&= \Pr \left[\begin{array}{l} \sum_i (X_{ij}^n - X_{ik}^n) = \sum_i (p_{ik} - p_{ij}) \text{ for } j, k \in J, \text{ and} \\ \sum_i (X_{ij}^n - X_{il}^n) > \sum_i (p_{il} - p_{ij}) \text{ for } j \in J, l \in K \setminus J \end{array} \right] \\
&= \Pr \left[\begin{array}{l} \frac{1}{\sqrt{n}} T_n \sum_i (X_{ij}^n - X_{ik}^n) = \frac{1}{\sqrt{n}} T_n \sum_i (p_{ik} - p_{ij}) \text{ for } j, k \in J, \text{ and} \\ \frac{1}{\sqrt{n}} T_n \sum_i (X_{ij}^n - X_{il}^n) > \frac{1}{\sqrt{n}} T_n \sum_i (p_{il} - p_{ij}) \text{ for } j \in J, l \in K \setminus J \end{array} \right] \quad (\text{A.2})
\end{aligned}$$

Writing Q_i^n for the cumulative distribution function of X_i^n , the random vectors satisfy the following multivariate Lindeberg condition: For every $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_i \int_{\|T_n X_i\| > \epsilon \sqrt{n}} \|T_n X_i\|^2 dQ_i^n(X) = 0 \quad (\text{A.3})$$

To see this, note that Z_i^n is in the simplex $\Delta^{|K|}$. Hence, $\|X_i^n\| \leq 2$. The probability that $Z_{ij}^n = 1$ is $p_{ij} \geq \epsilon^*$. Further, the variances and covariance of V_{ni} are all uniformly bounded away from zero and one, since $p_{ij} \geq \epsilon^*$ for all i, j . Thus, the same will be true of V_n . So V_n will be invertible, and for any ϵ , we can pick large enough n so that $\|T_n X_i\| < \epsilon \sqrt{n}$. So each term in the summation of Equation (A.3) goes to zero with n , which establishes (A.3). It follows by Lindebrgh's multivariate version of the central limit theorem for triangular arrays (see Bhattacharya and Rao [1986], Corollary 18.2, p 183) that the distribution of $\frac{1}{\sqrt{n}} T_n \sum_i X_i^n$ converges weakly to a multivariate unit normal distribution. Hence the probability it falls in a subset of any lower dimensional subspace goes to zero. Thus, when $|J| \geq 2$, the right hand side of Equation (A.2) converges to 0 with n . That is, $\lim_{n \rightarrow \infty} \delta^{n*} = 0$, proving (a). To prove (b), we note that B_{J^*} describes a lower dimensional subspace than B_J . Hence, an argument similar to above shows that for all sequences, the $\Pr[\bar{Z} \in B_{J^*}]$ goes to zero faster than $\Pr[\bar{Z} \in B_J]$, establishing the result. \blacksquare

Proof of Proposition 2.

Proof: To prove (a), define $D = 2 \cdot (|K| - 1) \cdot \sup_{x,y,\tau} [u(x, \tau) - u(y, \tau)]$, and $\varepsilon^* = \min_{j \in K} G_j(-\mathbf{1} \cdot D)$, where $\mathbf{1} = (1, \dots, 1)$ is the unit vector of length $|K|$. The assumption (p. 10) that, for each $j \in K_0$, the marginal distribution of η_{ij} possesses full support on \mathfrak{R} and is identical for all voters $i \in N$ and all electorates N implies that $\varepsilon^* > 0$. Then from Equation (2.8), using the fact that $\delta_i^{jl} \leq 1$ for all i, j, k , we have $-D \leq \bar{U}(j; y, s, \tau_i) - \bar{U}(l; y, s, \tau_i) \leq D$ for all $j, l \in K$, which implies that $\bar{s}_i(y, \tau_i)(j) = G_j(\bar{U}^j(y, s, \tau_i)) \geq G_j(-\mathbf{1} \cdot D) \geq \varepsilon^*$.

Now, given any sequence $\tau = (\tau_1, \dots, \tau_n)$ with $\tau_i \in \mathcal{T}$ for all $i > 0$, define the random variable

$$Z_{ni}(\tau_i) = \alpha_j \text{ if } s_i^n(y, (\tau_i, \eta_i)) = j$$

So $Z_{ni}(\tau_i) \in \mathcal{Z}^n$, with $p_{ij} = \bar{s}_\alpha^n(y, \tau_\alpha)(j)$.

Then, letting $(0, s_{-i}^n)$ be the profile where the voter i abstains, and (j, s_{-i}^n) be the profile where voter i votes for candidate j , we have, from Equation (2.9):

$$\delta_i^{jl}(y, s^n) = \sum_{j,k \in J \subseteq K} \frac{1}{|J|} \left(P_J(y, (0, s_{-i}); t_i) + \frac{P_J(y, (j, s_{-i}); t_i)}{|J| - 1} \right) \quad (\text{A.4})$$

But, from Equation (2.3), for any $J \subseteq K$,

$$P_J(y, (0, s_{-i}); t_i) = \Pr[\{t_{-i} \in T_{-i} : W(y, s; t) = J\}].$$

Reexpressing $\Pr[\{t_{-i} \in T_{-i} : W(y, s; t) = J\}]$ as $E_{t_{-i}} [\mathbf{1}[W(y, (0, s_{-i}^n); t_{-i}) = J]]$, where $\mathbf{1}$ denotes the indicator function that is equal to 1 if the condition is true and zero otherwise, we

obtain

$$\begin{aligned}
\Pr[\{t_{-i} \in T_{-i} : W(y, s; t) = J\}] &= E_{t_{-i}} [\mathbf{1}[W(y, (0, s_{-i}^n); t_{-i}) = J]] \\
&= E_{t_{-i}} \left[\mathbf{1} \left[\sum_{l \neq i} Z_{nl}(\tau_l) \in B_J \right] \right] \\
&= E_{\tau_{-i}} E_{\eta_{-i}} \left[\mathbf{1} \left[\sum_{l \neq i} Z_{nl}(\tau_l) \in B_J \right] \right] \\
&\leq E_{\tau_{-i}} [\delta_J^{n*}] = \delta_J^{n*}, \tag{A.5}
\end{aligned}$$

where the inequality follows from the definition of δ_J^{n*} in Lemma 2. A similar argument shows the second term in Equation (A.4) is less than or equal to δ_J^{n*} . Thus,

$$\delta_i^{jl}(y, s^n) \leq \sum_{j, k \in J \subseteq K} \left(\frac{1}{|J| - 1} \right) \delta_J^{n*} \leq \left(\sum_{j, k \in J \subseteq K} \frac{1}{|J| - 1} \right) \delta^{n*},$$

where $\delta^{n*} = \max_{J \subseteq K} \delta_J^{n*}$. By Lemma 2, $\lim_{n \rightarrow \infty} \delta^{n*} = 0$, which proves (a). Since δ^{n*} is independent of i, j, l, y , the convergence is uniform in all arguments.

To show (b), recall the proof of Lemma 2 and note that for each $J \subseteq K$, we can write

$$P_J(y, (0, s_{-i}); t_i) = E_{t_{-i}} \left[\mathbf{1} \left[\sum_{l \neq i} Z_{nl}(\tau_l) \in B_J \right] \right],$$

and the corresponding expression for voter j :

$$P_J(y, (0, s_{-j}); t_i) = E_{t_{-j}} \left[\mathbf{1} \left[\sum_{l \neq j} Z_{nl}(\tau_l) \in B_J \right] \right].$$

The right hand sides of these two expressions differ only by the i^{th} and j^{th} terms. Thus, by once again applying Lindebrgh's multivariate version of the central limit theorem for triangular arrays

(Bhattacharya and Rao [1986], Corollary 18.2, p 183), it follows that, for each $J \subseteq K$,

$$\begin{aligned} \lim_{n \rightarrow \infty} P_J(y, (0, s_{-i}); t_i) &= \lim_{n \rightarrow \infty} E_{t_{-i}} \left[\mathbf{1} \left[\sum_{l \neq i} Z_{nl}(\tau_l) \in B_J \right] \right] \\ &= \lim_{n \rightarrow \infty} E_{t_{-j}} \left[\mathbf{1} \left[\sum_{l \neq j} Z_{nl}(\tau_l) \in B_J \right] \right] \\ &= \lim_{n \rightarrow \infty} P_J(y, (0, s_{-j}); t_i) \end{aligned}$$

since both $\sum_{l \neq i} Z_{nl}$ and $\sum_{l \neq j} Z_{nl}$ converge weakly to the multivariate standard normal distribution. Thus, all terms in the sum in (A.4) converge. Thus, conclusion (b) follows. An analogous argument suffices to establish (c).

To show (d), we have from Equation (3.2) that

$$\bar{s}_i^n(y, \tau_i)(j) = \Pr[\max_{l \neq j} \bar{U}(l; y, s^n, \tau_i) + \eta_{il} \leq \bar{U}(j; y, s^n, \tau_i) + \eta_{ij}].$$

Now, in the first part of the proposition we showed all pivot probabilities go to zero uniformly as n gets large. Hence, using Equation (2.10) we get that as $n \rightarrow \infty$, for $j, l \in K$, $\bar{U}(l; y, s^n, \tau_i) - \bar{U}(j; y, s^n, \tau_i) \rightarrow 0$ uniformly in i, j, l, y, τ . But then we get

$$\begin{aligned} \lim_{n \rightarrow \infty} [\bar{s}_i^n(y, \tau_i)(j) - \bar{s}_i^n(y, \tau_i)(l)] &= \Pr[\max_{\alpha \neq j} \eta_{ia} - \eta_{ij} \leq 0] - \Pr[\max_{\alpha \neq l} \eta_{ia} - \eta_{il} \leq 0] \\ &= G_j(0) - G_l(0) = 0. \end{aligned} \tag{A.6}$$

Since the convergence of $\bar{U}(l; y, s^n, \tau_i) - \bar{U}(j; y, s^n, \tau_i)$ is uniform in all arguments, it follows that the convergence in Equation (A.6) is also. ■

Proof of Lemma 1

Proof: Define the following function:

$$m_N(x) = \frac{-E_\tau \left[\sum_{i \in N} [u(x, \tau_i) - u(x^*, \tau_i)]^2 \right]}{E_\tau \left[\sum_{i \in N} [u(x, \tau_i) - u(x^*, \tau_i)] \right]}, \tag{A.7}$$

To prove the lemma, it is sufficient to show that $\limsup_{n \rightarrow \infty} [\sup_{x \in X} [m_N(x)]] < \infty$.

After normalizing so that $u(x^*(\rho_N), \tau) = 0$ for all $\tau \in \mathcal{T}$ we can write Equation (A.7) as

$$m_N(x) = \frac{-E_\tau [u(x, \tau)^2]}{E_\tau [u(x, \tau)]}$$

for each N . Thus, we want to show that

$$\limsup_{n \rightarrow \infty} \left[\sup_{x \in X} \frac{-E_\tau [\sum_{i \in N} u(x, \tau_i)^2]}{E_\tau [\sum_{i \in N} u(x, \tau_i)]} \right] < \infty. \quad (\text{A.8})$$

By the fact that $u(x, \tau)$ is twice continuously differentiable in x for any τ , we can take any unit length direction vector $h \in \mathcal{D}^m$ and apply L'Hopital's rule twice, obtaining

$$m_h(x^*(\rho_N)) \equiv \lim_{\varepsilon \rightarrow 0} m_N(x^*(\rho_N) + \varepsilon h) = -\frac{E_\tau [\sum_{i \in N} (D_h u(x^*(\rho_N), \tau_i))^2]}{E_\tau [\sum_{i \in N} D_h^2 u(x^*(\rho_N), \tau_i)]}, \quad (\text{A.9})$$

for each N . If Condition (4.7) is violated, then Equation (A.9) will be infinite for some N . Accordingly, satisfaction of Equation (A.8) is equivalent to the following:

$$\limsup_{n \rightarrow \infty} \left[\sup_{h \in \mathcal{D}^m} \frac{-E_\tau [\sum_{i \in N} (D_h u(x^*(\rho_N), \tau_i))^2]}{E_\tau [\sum_{i \in N} D_h^2 u(x^*(\rho_N), \tau_i)]} \right] < \infty.$$

Since

$$\limsup_{n \rightarrow \infty} \left[\sup_{h \in \mathcal{D}^m} \frac{-E_\tau [\sum_{i \in N} (D_h u(x^*(\rho_N), \tau_i))^2]}{E_\tau [\sum_{i \in N} D_h^2 u(x^*(\rho_N), \tau_i)]} \right] < \frac{\limsup_{n \rightarrow \infty} -E_\tau [\sum_{i \in N} (D_h u(x^*(\rho_N), \tau_i))^2]}{\liminf_{n \rightarrow \infty} E_\tau [\sum_{i \in N} D_h^2 u(x^*(\rho_N), \tau_i)]},$$

and

$$\frac{\limsup_{n \rightarrow \infty} -E_\tau [\sum_{i \in N} (D_h u(x^*(\rho_N), \tau_i))^2]}{\liminf_{n \rightarrow \infty} E_\tau [\sum_{i \in N} D_h^2 u(x^*(\rho_N), \tau_i)]} \leq \frac{B^2}{B} = B < \infty,$$

it follows that Condition (4.7) is satisfied for any finite $M \geq B$, as was to be shown. ■

Proof of Theorem 1

Proof: For any set of voters N , and admissible ρ , let $y = (y_j, y_{-j}^*)$, where $y_l^* = x^*(\rho_N)$ for all

$l \neq j$ and $y_j \neq x^*(\rho)$. We first show that for large enough n , $V_j(y) = V_j(y_j, y_{-j}^*) \leq V_j(y^*)$.

Given an individual $i \in N$, and using Equations (2.8) and (2.10), the probability of a vote for candidate j is given by

$$\begin{aligned} s_i(y, \tau_i)(j) &= \Pr \left[\max_{l \in K_0 - \{j\}} [U(l; y, s, t_i) - U(j; y, s, t_i)] \leq 0 \right] \\ &= \Pr \left[\begin{array}{l} \eta_{ik} - \eta_{ij} \leq \Delta_i^k(y, s) \cdot [u(y_j, \tau_i) - u(x^*(\rho_N), \tau_i)] \text{ for } k \in K - \{j\} \\ \text{and } \eta_{i0} - \eta_{ij} \leq \Delta_i^j(y, s) \cdot [u(y_j, \tau_i) - u(x^*(\rho), \tau_i)] \end{array} \right] \\ &= G_j(\Delta_i(y, s) \cdot [u(y_j, \tau_i) - u(x^*(\rho), \tau_i)]) \end{aligned} \quad (\text{A.10})$$

where $\Delta_i(y, s) = (\Delta_i^1(y, s), \dots, \Delta_i^k(y, s))$, $\Delta_i^l(y, s) = 2\delta_i^{lj}(y, s) + \sum_{\alpha \neq j, l} \delta_i^{\alpha j}(y, s)$, for all $l \in K - \{j\}$, and $\Delta_i^j(y, s) = \sum_{\alpha \neq j} \delta_i^{j\alpha}(y, s)$.

Using Equation (2.13) we can express the vote for candidate j as

$$V_j(y) = \frac{1}{n} \sum_{i \in N} E_{\tau_i} [\bar{s}_i(y, \tau_i)(j)] \quad (\text{A.11})$$

Then, from Equation (A.10), we have that

$$V_j(y) = \frac{1}{n} \sum_{i \in N} E_{\tau_i} [G_j(\Delta_i(y, s) \cdot [u(y_j, \tau_i) - u(x^*(\rho_N), \tau_i)])]$$

Without loss of generality, we can assume utility functions are normalized with $u(x^*(\rho_N), \tau_i) = 0$ for all $i \in N$ and $\tau_i \in T$. Write $u_i = u(y_j, \tau_i) \in \mathbb{R}$, and $\Delta_i = \Delta_i(y, s)$. Then, the above can be written as:

$$V_j(y) = \frac{1}{n} \sum_{i \in N} E_{\tau_i} [G_j(\Delta_i(y, s) \cdot u(y_j, \tau_i))] = \frac{1}{n} \sum_{i \in N} E_{\tau_i} [G_j(\Delta_i \cdot u_i)] \quad (\text{A.12})$$

Normalize the Δ_i by Δ_1 in the following manner. For $i \in N$, let

$$\chi_i = \left(\frac{\Delta_i^1}{\Delta_1^1}, \dots, \frac{\Delta_i^k}{\Delta_1^k} \right)^T,$$

and

$$D = \begin{bmatrix} \Delta_1^1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \Delta_1^k \end{bmatrix}.$$

It is easily shown that $\Delta_i^j > 0$ for all $i \in N$ and $j \in K$, so that χ_i is well defined. Then, applying Taylor's theorem, we can write

$$\begin{aligned} V_j(y) - V_j(y^*) &= \frac{1}{n} \sum_{i \in N} [E_{\tau_i} [G_j (D \cdot \chi_i \cdot u_i)] - E_{\tau_i} [G_j (0)]] \\ &= \frac{1}{n} \sum_{i \in N} E_{\tau_i} \left[(D \cdot \chi_i \cdot u_i)^T G'_j (\mathbf{0}) + \frac{1}{2} (\Delta_i \cdot u_i)^T G''_j (z_i(y)) (\Delta_i \cdot u_i) \right] \end{aligned} \quad (\text{A.13})$$

where $z_i(y) = \alpha \mathbf{0} + (1 - \alpha)(0, \dots, u_i, 0, \dots, 0)$ for some $\alpha \in (0, 1)$ for each $i \in N$, and $\mathbf{0}$ denotes a k dimensional vector of zeros.

Now, by Proposition 2, it follows that for any $\epsilon > 0$, we can find a value n^* such that $\max D < \epsilon$, $\max_{i \in N} [\max [\chi_i]] < (1 + \epsilon)$ (where the inner max is with respect to the components of χ_i), and $\max_{i \in N} [\max [\Delta_i]] < \epsilon$ for all $n > n^*$ (where the inner max is with respect to the components of Δ_i). Using these facts and continuing the derivation of $V_j(y) - V_j(y^*)$,

$$\begin{aligned} V_j(y) - V_j(y^*) &= \frac{1}{n} \sum_{i \in N} E_{\tau_i} \left[G'_j (\mathbf{0}) D \cdot \chi_i \cdot u_i + \frac{1}{2} (\Delta_i \cdot u_i)^T G''_j (z_i(y)) (\Delta_i \cdot u_i) \right] \\ &= \frac{1}{n} G'_j (\mathbf{0}) \cdot \sum_{i \in N} E_{\tau_i} [D \cdot \chi_i \cdot u_i] + \frac{1}{2n} \sum_{i \in N} E_{\tau_i} [\Delta_i^T \cdot G''_j (z_i(y)) \cdot \Delta_i \cdot u_i^2] \\ &\leq \frac{k}{n} G_j^* \epsilon (1 + \epsilon) \sum_{i \in N} E_{\tau_i} [u_i] + \frac{k^2}{2n} \epsilon^2 G_j^{**} \sum_{i \in N} E_{\tau_i} [u_i^2] \end{aligned} \quad (\text{A.14})$$

where $G'_j(\mathbf{0})$ is a k dimensional vector consisting of the gradient of G_j evaluated at $\mathbf{0}$, $G''_j(\mathbf{0})$ is a $k \times k$ symmetric matrix of second partial derivatives of G_j evaluated at $\mathbf{0}$, G_j^* is the smallest

element of $G'_j(\mathbf{0})$, and G_j^{**} is defined as

$$G_j^{**} = \sup_{z \in \mathbb{R}^k} [G''_j(z)].$$

We construct G_j^* as the smallest element of $G'(\mathbf{0})$ since the first term of Equation (A.14) is negative, which follows because y^* is the unique social welfare maximizer. The construction of G_j^{**} is similarly motivated: it is chosen so as to maximize the impact of $E_{\tau_i} [u_i^2]$, which is positive. The final step in the derivation of Equation (A.14) follows from substitution of kG_j^* for $G'_j(0)$, ϵ for D , $(1 + \epsilon)$ for χ_i , ϵ for Δ_i , and $k^2 G_j^{**}$ for $G''_j(z_i(y))$. It should be noted that (1) G_j^* is positive by the assumption that the distribution of η_{ij} possesses full support for all i and j , (2) G_j^{**} must be finite since G_j is a cumulative distribution function (and hence is bounded) and is twice continuously differentiable, and (3) G_j^{**} is positive because (i) the marginal distribution of η_{ij} possesses full support for all i , (ii) G_j is bounded between 0 and 1, and (iii) G_j is twice continuously differentiable, so that it must be strictly convex on some open interval of \mathfrak{R} . Thus, it follows that $G_j^*/G_j^{**} > 0$.

We now show that there exists n^* such that for all N with $n > n^*$, the right-hand side of Equation (A.14) becomes less than zero for all $y \neq x^*(\rho_N) \in X$. For a given $y \neq x^*(\rho_N)$,

$$\begin{aligned} \frac{k}{n}(1 + \epsilon)\epsilon G_j^* \sum_{i \in N} E_{\tau_i} [u_i] + \frac{k^2}{2n}\epsilon^2 G_j^{**} \sum_{i \in N} E_{\tau_i} [u_i^2] &< 0 \\ \frac{k^2}{2n}\epsilon^2 G_j^{**} \sum_{i \in N} E_{\tau_i} [u_i^2] &< -\frac{k}{n}(1 + \epsilon)\epsilon G_j^* \sum_{i \in N} E_{\tau_i} [u_i] \\ \frac{-\sum_{i \in N} E_{\tau_i} [u_i^2]}{\sum_{i \in N} E_{\tau_i} [u_i]} &< \frac{2(1 + \epsilon)}{k\epsilon} \frac{G_j^*}{G_j^{**}}, \end{aligned} \quad (\text{A.15})$$

The inequality in Equation (A.15) is satisfied for sufficiently small $\epsilon > 0$. Of course, this is for a given $y \neq x^*(\rho_N)$. In order to satisfy Equation (A.15) for all $y \neq x^*(\rho_N)$, we must take the supremum of the left-hand side over all $y \neq x^*(\rho_N)$. This supremum is defined to be finite and denoted by M in Condition (4.7),²⁶ resulting in the following requirement for Equation (A.15) to

²⁶If necessary, we can choose n^* to exceed \tilde{n} as defined in the statement of Condition (4.7).

be satisfied for all $y \neq x^*(\rho_N)$:

$$M < \frac{2(1+\epsilon)}{k\epsilon} \frac{G_j^*}{G_j^{**}}. \quad (\text{A.16})$$

As with Equation (A.15), for ϵ sufficiently small, Equation (A.16) is satisfied. Thus, for any $y_j \in Y_j$, $V_j(y) = V_j(y_j, y_{-j}^*) \leq V_j(y^*)$ with strict inequality whenever $y_j \neq x^*(\rho)$.

Next, we show that for some $l \neq j$, $V_l(y_j, y_{-j}^*) \geq V_l(y^*)$. We pick $l \in K - \{j\}$ for which $\delta^{jl}(y, s)$ is maximized. For $z \in \mathbb{R}^{|K|}$, write $G_j(z) = G_l(z)$, where G_l is as defined in Equation (3.3). In this case, the probability of a vote for candidate l is given by the following:

$$\begin{aligned} s_i(y, \tau_i)(l) &= \Pr \left[\begin{array}{l} U(0; y, s, \tau_i) - U(l; y, s, \tau_i) \leq 0, \text{ and} \\ U(j; y, s, \tau_i) - U(l; y, s, \tau_i) \leq 0, \text{ and} \\ \max_{k \in K - \{l, j\}} [U(k; y, s, \tau_i) - U(l; y, s, \tau_i)] \leq 0, \end{array} \right] \\ &= \Pr \left[\begin{array}{l} \eta_{i0} - \eta_{il} \leq \Delta_i^l(y, s) \cdot [u(x^*(\rho_N), \tau_i) - u(y_\alpha, \tau_i)], \text{ and} \\ \eta_{ij} - \eta_{il} \leq \Delta_i^j(y, s) \cdot [u(x^*(\rho_N), \tau_i) - u(y_j, \tau_i)], \text{ and} \\ \eta_{ik} - \eta_{il} \leq \max_{k \in K - \{l, j\}} (\Delta_i^k(y, s) \cdot [u(x^*(\rho_N), \tau_i) - u(y_j, \tau_i)]) \end{array} \right] \\ &= G_j(\Delta_i(y, s) \cdot [u(x^*(\rho_N), \tau_i) - u(y_j, \tau_i)]) \end{aligned}$$

where $\Delta_i(y, s) = (\Delta_i^1(y, s), \dots, \Delta_i^k(y, s))$, with $\Delta_i^l(y, s) = \sum_{\alpha \neq l} \delta_i^{l\alpha}(y, s)$, $\Delta_i^j(y, s) = 2\delta_i^{lj}(y, s) + \sum_{\alpha \neq j, l} \delta_i^{j\alpha}(y, s)$, and $\Delta_i^k(y, s) = \delta_i^{jl}(y, s) - \delta_i^{kj}(y, s)$ for all $k \in K - \{l, j\}$.

Using Equation (2.13) we can express the vote for candidate l as

$$V_l(y) = \frac{1}{n} \sum_{i \in N} E_{\tau_i} [s_i(y, \tau_i)(l)] = \frac{1}{n} \sum_{i \in N} E_{\tau_i} [G_j(\Delta_i(y, s) \cdot [u(x^*(\rho_N), \tau_i) - u(y_j, \tau_i)])] \quad (\text{A.17})$$

As above, we can assume utility functions are normalized with $u(x^*(\rho_N), \tau_i) = 0$ for all $i \in N$

and $\tau_i \in T$. As before, write $u_i = u(y_j, \tau_i)$, and $\Delta_i = \Delta_i(y, s)$. Then, the above can be written as:

$$V_l(y) = \frac{1}{n} \sum_{i \in N} E_{\tau_i} [G_j(-\Delta_i(y, s) \cdot u(y_j, \tau_i))] = \frac{1}{n} \sum_{i \in N} E_{\tau_i} [G_j(-\Delta_i u_i)] \quad (\text{A.18})$$

Note that the above takes exactly the same form as Equation (A.12) above, with the exception of the negative sign. Consequently, an analogous argument to that in (A.14) establishes that we can find large enough n so that $V_l(y) - V_l(y^*)$ is positive. Thus, for any $y_j \in Y_j$, $V_l(y) = V_l(y_j, y_{-j}^*) \geq V_l(y^*)$ with strict inequality whenever $y_j \neq x^*(\rho_N)$. We have shown that $V_j(y_j, y_{-j}^*) \leq V_j(y^*)$ and $V_l(y_j, y_{-j}^*) \geq V_l(y^*)$. So $\widehat{V}_j(y_j, y_{-j}^*) \leq \widehat{V}_j(y^*)$. So y^* is a global equilibrium for the objective function \widehat{V} . ■

Proof of Corollary 1

Proof: Suppose there is another equilibrium, y . Then for at least one candidate j , $y_j \neq x^*(\rho_N)$. Assume W.L.O.G. that $j = 2$. By Theorem 1, $\widehat{V}_1(y_1, y_2) \geq \widehat{V}_1(x^*(\rho_N), y_2) > 0$. It follows that $\widehat{V}_2(y_1, y_2) < 0$. But this cannot be an equilibrium for candidate 2, Since $\widehat{V}_2(y_1, x^*(\rho_N)) \geq 0 > \widehat{V}_2(y_1, y_2)$. This yields a contradiction. Hence the equilibrium is unique. ■