Condorcet's Jury Theorem under Costly Information Acquisition*

Minoru Kitahara† and Yohei Sekiguchi‡

December 23, 2004

Abstract

Condorcet's Jury Theorem asserts that the probability that the desirable alternative is chosen by the majority rule converges to 1 as the number of voters goes to ∞. Existing various versions of Condorcet's Jury Theorem assume that the probability that each voter predicts the desirable alternative is given exogenously. We endogenize it by considering costly information acquisition. We show the necessary and sufficient condition for Condorcet's Jury Theorem being valid under costly information acquisition. We also show the way to calculate the probability that the desirable alternative is chosen in the limit when Condorcet's Jury Theorem is not valid.

KEYWORDS: Condorcet's Jury Theorem, elections, strategic voting, information aggregation, costly information acquisition.

1 Introduction

Condorcet's Jury Theorem (hereafter CJT) asserts that the probability that the desirable alternative is chosen by the majority rule converges to 1 as the number of voters goes to ∞. Existing various versions of CJT assume that the probability that each voter predicts the desirable alternative is given exogenously. We endogenize it by considering costly information acquisition. We show the necessary and sufficient condition for CJT being valid under costly information acquisition. We also show the way to calculate the probability that the desirable alternative is chosen in the limit when Condorcet's Jury Theorem is not valid. One of the most important function of elections is aggregating private information. Consider the following situation. Suppose that there are two alternatives. Either one of the two is commonly desirable in society. Nobody in the society, however, can predict it perfectly which is desirable. When society faces this binary choice, can an election select the desirable alternative? CJT gave a powerful justification for the majority voting rule: Majorities are more likely to choose the correct alternative than any individual. The most simplest version

*We would like to thank Tomohiko Kawamori, Akihiko Matsui, Ryo Ogawa, Dan Sasaki, and participants of the seminar at University of Tokyo.

†Graduate School of Economics, University of Tokyo. E-mail: ee37019@mail.ecc.u-tokyo.ac.jp

‡Graduate School of Economics, University of Tokyo. E-mail: yohei@grad.e.u-tokyo.ac.jp

of CJT is the following. A group chooses an alternative by the majority voting rule. If each member of the group has same predictability $q > 1/2$ (the probability that she votes for the desirable alternative), then the probability that the desirable alternative is chosen converges to 1 as the number of members goes to infinity. In this version, it is assumed that the distributions of signals received by individuals are independent and identical. More general versions were given by Ladha (1992, 1995) and Berend and Paroush (1998).[1]

CJT requires another important assumption. Voters behave naively, i.e., each voter behaves as if the outcome is determined by only her vote. However, naive voting may not be an equilibrium strategy of the corresponding game. Since each voter's vote affects the outcome only if her vote is pivotal, she must condition her belief about the state on the event that her vote is pivotal. Therefore, there is in general incentive for an individual to manipulate her vote.[2] Then, it is necessary to consider strategic behavior of voters. Feddersen and Pesendorfer (1997) showed full information aggregation under strategic voting.

Although a large number of studies have been made on CJT, all of them assumed that each voter's accuracy of predicting the desirable alternative q is given exogenously. In this paper we consider costly voluntary information acquisitions by voters. That is, initially each voter can guess right with probability 1/2. And in order to raise accuracy she must pay information cost. Each voter chooses her accuracy and votes strategically. We restrict our attention to CSS-equilibria, in which strategies are symmetric with respect to private signals, alternatives and information cost, originally defined by Razin (2003). We show that if the marginal information cost at $q = 1/2$ is sufficiently low, there is a unique CSS-equilibrium, in which each voter votes informatively and pays positive information cost. Why each voter pays positive information cost? We give an intuitive explanations as follows. If voters can predict the desirable alternative perfectly, i.e., $q = 1$, each vote never be pivotal since others vote for the desirable alternative. Then she is not willing to pay information cost. On the contrary, if voters cannot predict the desirable alternative perfectly, i.e., $q < 1$, she becomes the pivotal voter with positive probability. Then she is willing to pay some information cost. These arguments have a same flavor of Grossman-Stiglitz (1980) paradox in rational expectation equilibrium models with discrete choice of costly information acquisition: If information is fully acquired, there is no value of acquiring information. Otherwise, if no information is acquired, there is an incentive of acquiring information. Since in our setting choice of information acquisition is continuous accuracy is determined such that the marginal cost of information and the marginal benefit of information are equivalent in the equilibrium.

Is CJT valid in such a situation? Intuitively, as the number of voters increases, the probability that each vote is pivotal decreases. Then, each voter is willing to reduce her information cost as the number of voters increases. Hence, an increase of the number of voters may lead a poorer decision. Mukhopadhaya (2003) gave an example in which increasing the number of voters decreases the probability that the desirable alternative is chosen. As the number of voter

[1] Ladha (1992, 1995) relaxed the independence assumption and allow for correlated votes. Berend and Paroush (1998) relaxed the identical assumption and provided the necessary and sufficient condition for CJT.

[2] Austen-Smith and Banks (1996) gave a simple example in which naive voting does not consist an equilibrium.

goes to ∞ the probability that a vote is pivotal converges to 0 and hence the accuracy q converges to 1/2. As Berend and Paroush (1998) showed, CJT holds if and only if the speed of convergence of q is sufficiently slow. We show this condition is represented by the second order differential of the cost function at $q = 1/2$ and we provide the necessary and sufficient condition on a cost function for CJT being valid. Namely, CJT holds if and only if the second order differential of the cost function at $q = 1/2$ is zero. We also show the way to calculate the probability that the desirable alternative is chosen in the limit when Condorcet's Jury Theorem is not valid.

The rest of the paper is organized as follows. Section 2 examines the model. Section 3 derives the main results. In section 4, using our results, we investigate the signal with white noise. Section 5 states conclusions.

2 The Model

We study a two alternatives election with the single nontransferable vote. Alternatives are denoted by $j \in \{L, R\}$. Either one of the alternatives is assumed to be commonly desirable. There are $2n + 1$ voters indexed by i. Each voter is uncertain about the desirable alternative. It depends on the state of the world that which alternative is the desirable one. Let $S \in \{L, R\}$ be the state; if the state is L (R), the desirable alternative is L (R). There is a common prior probability about the states; each state will occur with the equal probability. Each voter cannot know the true state but she receives a private signal $\theta \in \Theta$ about the state and guesses the true state. We assume here Θ is a countable set for notational simplicity. The distribution of signals depends on the state and on each voter's accuracy of predicting the state $q_i \in [1/2, 1]$.

Assumption 1 (Consistency of q)
For all $q \in [\frac{1}{2}, 1]$,

$$q = E[\max\{\Pr(L|\theta, q), \Pr(R|\theta, q)\}|q].$$

Since the probability that a voter receiving θ with accuracy q can predict the true state is $\max\{\Pr(L|\theta, q), \Pr(R|\theta, q)\}$, the ex ante probability predicting the true state coincides with accuracy q under Assumption 1.

We restrict our attention to symmetric signal structures.

Assumption 2 (Symmetricity of signals)
For all $\theta \in \Theta$, there exists the unique $\bar{\theta} \in \Theta$ such that for all $q \in (1/2, 1)$,

$$\Pr(\theta|L, q) = \Pr(\bar{\theta}|R, q),$$
$$\Pr(\theta|R, q) = \Pr(\bar{\theta}|L, q).$$

Such $\bar{\theta}$ is called the conjugate signal of θ.

We give an example of the signal structure satisfying our assumptions.

Example 1 (Two signals)
Suppose that $\Theta = \{l, r\}$ and

$$\Pr(l|L, q) = \Pr(r|R, q) = q.$$

This signal structure is the most simplest one which describes the situation supposed by Condorcet. The signal l (r) is the signal "the state is L (R)". $\|$

We define subsets of Θ.

$$\mathcal{L}(q) = \{\theta | \Pr(L|\theta, q) > \Pr(R|\theta, q)\},$$
$$\mathcal{R}(q) = \{\theta | \Pr(L|\theta, q) < \Pr(R|\theta, q)\}.$$

$\mathcal{L}(q)$ ($\mathcal{R}(q)$) is the set of signals under which each one considers that the alternative L (R) will be more reasonable.

Example 2 (Uniform signals)
Suppose that $\Theta = [-1, 1]$ and

$$\theta \sim \begin{cases} U[-1, \frac{1}{q} - 1] & \text{if } S = L \\ U[1 - \frac{1}{q}, 1] & \text{if } S = R. \end{cases}$$

This signal structure violates Assumption 2 since the conjugate signal is not unique. However, we can degenerate it to the signal structure satisfying Assumption 2 as follows. Let $\Theta' = \{l, c, r\}$ and

$$\theta' = \begin{cases} l & \text{if } \theta \in [-1, 1 - \frac{1}{q}) \\ c & \text{if } \theta \in [1 - \frac{1}{q}, \frac{1}{q} - 1] \\ r & \text{if } \theta \in (\frac{1}{q} - 1, 1]. \end{cases}$$

Then, the conjugate signal of l is r and the conjugate signal of c is itself. Thus, Θ' satisfies Assumption 2. Obviously , we have $\mathcal{L} = \{l\}$ and $\mathcal{R} = \{r\}$. $\|$

Proposition 1 *Under Assumption 1,2, for all $q \in [1/2, 1]$,*

$$q = \Pr(\theta \in \mathcal{L}|L, q) + \frac{1}{2}\Pr(\theta \notin \mathcal{L} \cup \mathcal{R}|L, q) = \Pr(\theta \in \mathcal{R}|R, q) + \frac{1}{2}\Pr(\theta \notin \mathcal{L} \cup \mathcal{R}|R, q)$$

Proof:

$$\begin{aligned} q &= E[\max\{\Pr(L|\theta, q), \Pr(R|\theta, q)\}|q] \\ &= \sum_{\theta \in \Theta} \Pr(\theta|q) \max\{\Pr(L|\theta, q), \Pr(R|\theta, q)\} \\ &= \sum_{\theta \in \mathcal{L}} \Pr(\theta|q) \Pr(L|\theta, q) + \sum_{\theta \in \mathcal{R}} \Pr(\theta|q) \Pr(R|\theta, q) + \sum_{\theta \notin \mathcal{L} \cup \mathcal{R}} \Pr(\theta|q) \frac{1}{2} \\ &= \sum_{\theta \in \mathcal{L}} \Pr(L) \Pr(\theta|L, q) + \sum_{\theta \in \mathcal{R}} \Pr(R) \Pr(\theta|R, q) + \frac{1}{2} \sum_{\theta \notin \mathcal{L} \cup \mathcal{R}} \Pr(\theta|L, q) \\ &= \sum_{\theta \in \mathcal{L}} \Pr(\theta|L, q) + \frac{1}{2} \sum_{\theta \notin \mathcal{L} \cup \mathcal{R}} \Pr(\theta|L, q) \\ &= \Pr(\theta \in \mathcal{L}|L, q) + \frac{1}{2} \Pr(\theta \notin \mathcal{L} \cup \mathcal{R}|L, q). \end{aligned}$$

Since, by Assumption 2, $\sum_{\theta \in \mathcal{L}} \Pr(\theta|L, q) = \sum_{\theta \in \mathcal{R}} \Pr(\theta|R, q)$, then the last equality in the proposition is obtained similarly. $\square$

Due to Proposition 1, we can easily calculate q given a signal structure. The accuracy q is equal to the probability that each voter receives the signal under which S will be reasonable when the true state is S. This proposition is valid even if Θ is uncountable by replacing summation with integration.

We consider costly information acquisition. That is, cost is required in order to raise accuracy. Let $C(q_i)$ be i's information cost. We assume $C'(q_i) > 0$, $C''(q_i) > 0$ for all $q_i > 1/2$ and $C(1/2) = 0$.

Example 3 (White noise)
Suppose that $\Theta = \mathbb{R}$ and

$$\theta = \begin{cases} -1 + \epsilon & \text{if } S = L \\ 1 + \epsilon & \text{if } S = R, \end{cases}$$

where $\epsilon \sim N(0, \sigma^2)$. And suppose that the information cost is inverse proportional to the variance of the noise,

$$\tilde{C}(\sigma^2) = \frac{\lambda}{2\sigma^2}.$$

Let $\delta = \frac{1}{\sigma}$. Then, by Proposition 1, $q = \Phi(\delta)$ where $\Phi(\cdot)$ denotes the standard Normal distribution function. Hence, we obtain $C(q)$ implicitly as follows,

$$\begin{cases} \tilde{C}(\delta) = \frac{\lambda\delta^2}{2}, \\ q = \Phi(\delta). \end{cases}$$

Thus, even if the signal structure does not explicitly give the cost of q, we can obtain it easily. $\|$

Example 3 tells us that even if the cost function is not given by the function of q, e.g., the cost is variance of the noise, we can rearrange the cost function to the function of q.

Each voter decides her accuracy q_i and votes strategically. Each voter i's voting strategy is a function $v_i : \Theta \to [0, 1]$ where $v_i(\theta)$ is the probability that a given voter who received the signal θ will vote for L. Since we assume no abstention, the probability of voting for R is $1 - v_i(\theta)$. If the number of voters who choose L is greater than n, then L is the outcome. Otherwise R is the outcome. A generic outcome is denoted by y. The payoff of voter i given an outcome y and her accuracy q_i is given by

$$U_i(y, q_i) = \begin{cases} 1 - C(q_i) & \text{if } y = S, \\ -C(q_i) & \text{if } y \neq S. \end{cases}$$

The expected payoff of voter i given a strategy profile $(\mathbf{q}, \mathbf{v})$ and a signal s is given by

$$\pi_i(\mathbf{q}, \mathbf{v}|\theta) = \Pr(y = S|\mathbf{q}, \mathbf{v}, \theta) - C(q_i).$$

We restrict our attention to a subclass of symmetric perfect Bayesian equilibria, called by CSS-equilibria.[3]

[3]CSS-equilibria was defined by Razin(2003).

Definition 1 *A CSS-equilibrium is a perfect Bayesian equilibrium such that,*

1. *All voters adopt the same voting strategy:*

$$v_i(\theta) = v(\theta),\ for\ all\ i = 1, \cdots, 2n+1\ and\ for\ all\ \theta \in \Theta.$$

2. *The strategy are symmetric with respect to alternatives:*

$$v(\theta) = 1 - v(\bar{\theta}),\ for\ all\ \theta \in \Theta.$$

3. *All voters pay the same information cost:*

$$q_i = q,\ for\ all\ i = 1, \cdots, 2n+1.$$

A strategy profile is CSS if it satisfies these conditions.

Note that if the strategy is CSS, then $v(\theta) = 1/2$, for all $\theta \notin \mathcal{L} \cup \mathcal{R}$ since the conjugate signal of θ is itself.

3 The Main Results

3.1 The Existence and the Uniqueness of the equilibrium

At first we show the existence and the uniqueness of CSS-equilibria. We assume the marginal information cost at $q = 1/2$ is zero.

Assumption 3
$C'(1/2) = 0$.

If Assumption 3 is satisfied, in the unique equilibrium, each one votes informatively and pays strictly positive information cost.

Theorem 1
Under Assumption 3, there exists the unique CSS-equilibrium for all n. It satisfies followings,

$$v^*(\theta) = \begin{cases} 1 & if\ \theta \in \mathcal{L} \\ 0 & if\ \theta \in \mathcal{R} \\ \frac{1}{2} & if\ \theta \notin \mathcal{L} \cup \mathcal{R} \end{cases} \tag{1}$$

$$q^* > \frac{1}{2} \tag{2}$$

$$\binom{2n}{n} q^{*n}(1-q^*)^n = C'(q^*). \tag{3}$$

Proof:
We will prove in several steps.

(i) In CSS-equilibria, $q \neq 1, 1/2$.

Suppose to the contrary that $q = 1$. Then, each vote never become a pivotal since others know the true state and vote for the desirable alternative with probability 1. Hence, each voter has an incentive of reducing her information cost. On the other hand, if $q = 1/2$, each vote become a pivotal with positive probability. Since the marginal information cost at $q = 1/2$ equal zero by Assumption 3, each voter has an incentive of raising q.

(ii) If the strategy profile is CSS, the probability that each one votes for the desirable alternative *ex ante* is independent of the state.

Let $p_j(S, q)$ be the probability that a randomly drawn voter will vote for candidate $j \in \{L, R\}$ given a state S and q. The following Lemma shows that the probability that a randomly drawn voter will vote for the desirable alternative is symmetric between alternatives.

Lemma 1 *If a strategy profile is CSS, then $p_L(L, q) = p_R(R, q)$.*

Proof:
Suppose that $\mathbf{v}$ is the CSS voting strategy profile. Then, we have

$$\begin{aligned} p_L(L, q) &= \sum_{\theta \in \Theta} v(\theta) \Pr(\theta | L, q) \\ &= \sum_{\bar{\theta} \in \Theta} \{1 - v(\bar{\theta})\} \Pr(\bar{\theta} | R, q) \\ &= p_R(R, q). \end{aligned}$$

The second equality is obtained by Assumption 2 and the assumption that $\mathbf{v}$ is CSS. □

(iii) $\Pr(piv|L, q) = \Pr(piv|R, q)$ where piv denotes the event that a vote is pivotal.

The only time a voter can influence the outcome is if a vote is pivotal. A voter will choose L if, conditional on a vote being pivotal, the expected payoff of L is greater than the expected payoff of R.

Due to (ii) the number of votes for L follows the binomial distribution with the parameter $p_L(S, q)$. Hence, given a CSS strategy profile we can compute the probability that a vote is pivotal. This probability is given by

$$\Pr(piv|S, q) = \binom{2n}{n} p_L(S, q)^n p_R(S, q)^n.$$

Consequently, by Lemma 1, we have

$$\Pr(piv|R, q) = \Pr(piv|L, q). \tag{4}$$

(iv) In CSS-equilibria, everyone votes informatively.

The subjective probability distribution over states conditional on being pivotal, the signal θ and the accuracy q is computed by Bayes' rule. This is given by

$$\begin{aligned} \Pr(S|piv, q, \theta) &= \frac{\Pr(piv|S, q) \Pr(\theta|S, q)}{\Pr(piv|L, q) \Pr(\theta|L, q) + \Pr(piv|R, q) \Pr(\theta|R, q)} \\ &= \frac{\Pr(\theta|S, q)}{\Pr(\theta|L, q) + \Pr(\theta|R, q)}. \end{aligned}$$

The last equality is obtained by (4)

Therefore, when a voter receives the signal θ, the difference between the expected payoff from voting for L and for R is given by

$$\Pr(L|piv, q, \theta) - \Pr(R|piv, q, \theta) = \frac{\Pr(\theta|L, q) - \Pr(\theta|R, q)}{\Pr(\theta|L, q) + \Pr(\theta|R, q)}$$

Thus, each one votes for L if he receives signal $\theta \in \mathcal{L}$, i.e., $v^*(\theta) = 1$ for all $\theta \in \mathcal{L}$. Similarly, each one votes for R if he receives signal $\theta \in \mathcal{R}$, i.e., $v^*(\theta) = 0$ for all $\theta \in \mathcal{R}$. Whereas $v(\theta) = 1/2$ for all $\theta \notin \mathcal{L} \cup \mathcal{R}$ since, by Assumption 2, the conjugate signal is itself for all $\theta \notin \mathcal{L} \cup \mathcal{R}$.

(v) If everyone votes informatively, then $p_L(L, q) = p_R(R, q) = q$.

We obtain $p_L(L, q) = q$ as follows.

$$\begin{aligned} p_L(L, q) &= \sum_{\theta \in \Theta} v(\theta) \Pr(\theta|L, q) \\ &= \sum_{\theta \in \mathcal{L}} \Pr(\theta|L, q) + \frac{1}{2} \sum_{\theta \notin \mathcal{L} \cup \mathcal{R}} \Pr(\theta|L, q) \\ &= q. \end{aligned}$$

By (ii), we have $p_L(L, q) = p_R(R, q)$ and we conclude that the above statement is true. Thus, the probability that each one votes for the desirable alternative is equal to the accuracy q in both states.

(vi) There exists unique $q^* > 1/2$.

Finally we consider the optimal information acquisition. Suppose that some voter i deviates to $q \neq q^*$. By (iv), the expected payoff under informative voting is

$$\begin{aligned} \pi_i(\mathbf{v}^*, (q_i, \mathbf{q}^*_{-i})) &= \Pr(y = S|piv, q_i, s) \Pr(piv|q^*, s) + \Pr(y = S, piv^c|q^*, s) - C(q_i) \\ &= q_i \begin{pmatrix} 2n \\ n \end{pmatrix} q^{*n}(1 - q^*)^n + \Pr(y = S, piv^c|q^*, s) - C(q_i). \end{aligned}$$

The first order condition for the payoff maximization is

$$\begin{pmatrix} 2n \\ n \end{pmatrix} q^{*n}(1 - q^*)^n = C'(q_i).$$

Under Assumption 3, for all n there exists $q_i^* > 1/2$ satisfying the F.O.C. and q_i^* is unique since $C'(\cdot)$ is strictly increasing. If $q_i^* = q^*$, q^* is the equilibrium accuracy. Since $q(1 - q)$ is strictly decreasing in q, there exists the unique CSS-equilibrium. □

We give an intuitive proof of Theorem 1 as follows. The only time a voter can influence the outcome is when a vote is pivotal. A voter will choose L if, conditional on her vote being pivotal, the expected payoff of L is greater than the expected payoff of R. Since the signals are symmetric between the two states, if voters adopts the same symmetric voting strategy, the probability that a vote is pivotal is independent of the state. Hence, each one's vote depends only on her private signal. Thus each one votes informatively. That is, v^* is the equilibrium voting strategy if and only if v^* satisfies Eq (1). If all voters vote informatively, the probability that each one votes for the desirable alternative is exactly equal to the accuracy q. Then, the equilibrium accuracy q^* is determined such that the marginal gain of q is equal to the marginal cost of q. The former is the probability that a vote is pivotal and the latter is $C'(q)$. Consequently, under Assumption 3, there exists the unique $q^* > 1/2$ satisfying Eq (3). If, on the contrary, Assumption 3 is violated, for sufficiently large n,

$$\begin{pmatrix} 2n \\ n \end{pmatrix} q^n(1-q)^n < C'(q), \ \forall q \in [\frac{1}{2}, 1].$$

Therefore, if the number of voters is sufficiently large, each voter is not willing to pay positive information cost and $q^* = 1/2$.

We show that the signal structure and the cost function defined in Example 3 satisfies Assumption 3.

Example 3 (Continued)

$$\frac{\partial C(q)}{\partial q} = \frac{\partial \tilde{C}(\delta)}{\partial \delta}\frac{\partial \delta}{\partial q} = \frac{1}{\phi(\delta)}\lambda\delta = \frac{\lambda\delta}{\phi(\delta)},$$

where ϕ is the density function of the standard normal distribution. Since $q = 1/2 \iff \delta = 0$, $C'(1/2) = 0$ holds and Assumption 3 is satisfied. $\|$

Nevertheless, if the information cost is inverse proportional to the standard deviation of the noise, Assumption 3 is violated.

Example 4
Suppose that Θ is the same as in Example 3 and the information cost is inverse proportional to the standard deviation of the noise, that is,

$$\tilde{C}(\sigma^2) = \frac{\lambda}{\sigma}.$$

Similar to Example 3, we obtain $C(q)$ implicitly as follows,

$$\begin{cases} \tilde{C}(\delta) = \lambda\delta, \\ q = \Phi(\delta). \end{cases}$$

Obviously, we have

$$\frac{\partial C(q)}{\partial q} = \frac{\partial \tilde{C}(\delta)}{\partial \delta}\frac{\partial \delta}{\partial q} = \frac{\lambda}{\phi(\delta)}.$$

Hence, $C'(1/2) > 0$ holds and Assumption 3 is violated. $\|$

3.2 Asymptotic Properties

We want to know whether the probability that majorities can predict the desirable alternative converges to 1 as the number of voters becomes large. It depends that how the sequence of the equilibrium accuracies of voters converges. Let q_n^* be the equilibrium accuracy in the election with $2n+1$ voters and consider the sequence $\{q_n^*\}$. As a corollary of Theorem 1, we obtain that $\{q_n^*\}$ is strictly decreasing and converges to 1/2.

Corollary 1 *$\{q_n^*\}$ is a strictly decreasing sequence and* $\lim_{n\to\infty} q_n^* = 1/2$.

Proof:
Note that the equilibrium accuracy q_n^* is a solution of Eq (3). Since the LHS of Eq (3) is strictly decreasing in n for all $q \in [1/2, 1)$, obviously q_n^* is decreasing in n. Moreover, the LHS of Eq (3) converges to 0 as n goes to ∞. Thus, we obtain $q_n^* \to 1/2$ $(n \to \infty)$. □

Due to Theorem 1, everyone votes informatively in the CSS-equilibrium. Then, the probability that each one votes for the desirable alternative is equal to her accuracy q^*. Therefore, in order to investigate the asymptotic behavior of election outcomes, we may consider only monotone decreasing sequences $\{q_n\}$, which converge to 1/2.

Let x_k^n be the independent random variable such that $x_k^n = 1$ with probability q_n and $x_k^n = 0$ with probability $1 - q_n$ and the sum of x_k^n is denoted by $X_n = \sum_{k=1}^{2n+1} x_k^n$. Then X_n denotes the number of voters who vote for the desirable alternative. The desirable alternative is chosen if and only if $X_n > n$. Obviously, we have $E(X_n) = (2n+1)q_n$ and $Var(X_n) = (2n+1)(q_n(1-q_n))$. We define the normalized distance function of q_n,

$$D(q_n) \equiv \frac{(2n+1)(q_n - \frac{1}{2})}{\sqrt{(2n+1)(q_n(1-q_n))}}. \tag{5}$$

$D(q_n)$ is the distance between the expected number of votes for the desirable alternative ($= E(X_n)$) and the half of the total votes ($= \frac{2n+1}{2}$) normalized by the standard deviation of X_n ($= \sqrt{(2n+1)(q_n(1-q_n))}$). Note that Since the order of convergence of $D(q_n)$ is same as that of $\sqrt{2n+1}(q_n - 1/2)$, $D(q_n)$ represents the speed of convergence of $\{q_n\}$.

As Berend and Paroush (1998) showed, CJT holds if and only if the speed of convergence of $\{q_n^*\}$ is sufficiently slow. Let $\Phi(\cdot)$ denote the standard Normal distribution function.

Proposition 2 *Consider the sequence of elections such that there are $2n+1$ voters with accuracy $q_n > 1/2$ and q_n converges to $1/2$. Let y_n and S_n be the outcome and the state in nth election respectively. Then,*

1. $\lim_{n\to\infty} D(q_n) = \infty \Rightarrow \lim_{n\to\infty} \Pr(y_n = S_n | q_n) = 1$.
2. $\lim_{n\to\infty} D(q_n) = c < \infty \Rightarrow \lim_{n\to\infty} \Pr(y_n = S_n | q_n) = \Phi(c) < 1$.

Proof:
1. We obtain the following from Chebychev's Inequality,

$$\begin{aligned}
\Pr(y_n \neq S_n) &= \Pr\left(X_n \leq n\right) = \Pr\left(X_n < n + \frac{1}{2}\right) \\
&= \Pr\left(X_n - (2n+1)q_n < n + \frac{1}{2} - (2n+1)q_n\right) \\
&\leq \Pr\left(|X_n - (2n+1)q_n| > (2n+1)q_n - \left(n + \frac{1}{2}\right)\right) \\
&\leq \left(\frac{\sqrt{q_n(1-q_n)}}{\sqrt{2n+1}(q_n - \frac{1}{2})}\right)^2.
\end{aligned}$$

Therefore, if $\lim_{n\to\infty} D(q_n) = \infty$ holds, we have

$$\lim_{n\to\infty} \Pr(y_n \neq S_n) = 0.$$

2. Similarly, we have

$$\begin{aligned}
\Pr(y_n \neq S_n) &= \Pr\Big(X_n \leq n\Big) = \Pr\Big(X_n < n + \frac{1}{2}\Big) \\
&= \Pr\Big(X_n - (2n+1)q_n < n + \frac{1}{2} - (2n+1)q_n\Big) \\
&= \Pr\left(\frac{X_n - (2n+1)q_n}{\sqrt{(2n+1)q_n(1-q_n)}} < \frac{(2n+1)(\frac{1}{2} - q_n)}{\sqrt{(2n+1)q_n(1-q_n)}}\right)
\end{aligned}$$

By the Berry-Esseen Central Limit Theorem (See, for example, Feller (1971)), we obtain for an arbitrary $0 < \epsilon < \Phi(-c)$ and sufficiently large n,

$$\Phi(-c) + \epsilon \geq \Pr(y_n \neq S_n) \geq \Phi(-c) - \epsilon > 0.$$

Therefore, in the limit the wrong candidate is chosen with the probability $\Phi(-c)$. □

Proposition 2 shows that CJT is valid if and only if $D(q_n)$ converges to ∞, i.e., the speed of convergence of q_n is sufficiently slow. Due to Proposition 2, we can not only check whether CJT is valid but also calculate the limit probability when CJT is not valid. If $D(q_n)$ converges to $c < \infty$, the limit probability is $\Phi(c)$.

Next theorem is our main result. Theorem 2 shows that we can check whether CJT is valid by $C''(1/2)$. The necessary and sufficient condition for CJT is $C''(1/2) = 0$. That is, the probability that majorities can predict the desirable alternative converges to 1 as the number of voters goes to ∞ if and only if $C''(1/2) = 0$. Moreover, if $C''(1/2) > 0$, the limit probability is $\Phi(c)$ such that $\frac{\phi(c)}{c} = C''(1/2)/4$.

Theorem 2
Suppose that $\{q_n^*\}$ is the sequence of CSS-equilibrium accuracies.

1. *If $C''(1/2) = 0$,*

$$\lim_{n\to\infty} \Pr(y_n = S_n | q_n^*) = 1.$$

2. *If $C''(1/2) > 0$,*

$$\lim_{n\to\infty} \Pr(y_n = S_n | q_n^*) = \Phi(c) < 1,$$

such that $\frac{\phi(c)}{c} = \frac{1}{4}C''(\frac{1}{2})$.[4]

Proof:
At first we show the relation between $D(q_n)$ and the probability that a vote is pivotal. Since $\Pr(piv|q_n) \to 0,\ (n \to \infty)$, we must normalize it by multiplying its standard deviation $\sqrt{2n}\sigma(q_n)$ where $\sigma(q_n) = \sqrt{q_n(1-q_n)}$.

[4]Note that $\frac{\phi(c)}{c}$ is strictly decreasing in $[0, \infty)$ with $\lim_{c\to+0} \frac{\phi(c)}{c} = \infty$ and $\lim_{c\to\infty} \frac{\phi(c)}{c} = 0$.

Lemma 2 *Consider the sequence of elections such that there are $2n+1$ voters with accuracy $q_n > 1/2$ and q_n converges to $1/2$.*

1. $\lim_{n\to\infty} D(q_n) = \infty \Rightarrow \lim_{n\to\infty} \sqrt{2n}\sigma(q_n)\Pr(piv|q_n) = 0.$
2. $\lim_{n\to\infty} D(q_n) = c < \infty \Rightarrow \lim_{n\to\infty} \sqrt{2n}\sigma(q_n)\Pr(piv|q_n) = \phi(c).$

Proof:
See Appendix. □

Intuitively, since the distribution of the aggregate votes of n voters converges to a normal one with its standard deviation $\sqrt{2n}\sigma(q_n)$,

$$\frac{\Pr(piv|q_n)}{\Pr\left(c - \frac{1}{2}\frac{1}{\sqrt{2n}\sigma(q_n)} < N(0,1) < c + \frac{1}{2}\frac{1}{\sqrt{2n}\sigma(q_n)}\right)} \to 1 \text{ as } n \to \infty.$$

Then, by the fact that $\frac{\Pr\left(c-\frac{1}{2}x<N(0,1)<c+\frac{1}{2}x\right)}{x} \to \phi(c)$ as $x \to 0$, the result of Lemma 2 follows (See Fig.1).

By applying Lemma 2, we can show that the limit of $D(q_n)$ is characterized by $C''(1/2)$.

Lemma 3 *Suppose that $\{q_n^*\}$ is the sequence of CSS-equilibrium accuracies.*

1. *If $C''(\frac{1}{2}) = 0$,*
$$\lim_{n\to\infty} D(q_n^*) = \infty.$$
2. *If $C''(\frac{1}{2}) > 0$,*
$$\lim_{n\to\infty} D(q_n^*) = c < \infty,$$
such that $\frac{\phi(c)}{c} = \frac{1}{4}C''(\frac{1}{2})$.

Proof:
See Appendix. □

Thus, by Lemma 3 and Proposition 2, we obtain the statement of the theorem. □

We give an intuitive proof as follows. Suppose that $C''(1/2) > 0$. $C'(q)$ is approximated by the linear line with the slope $C''(1/2)$ in the neighborhood of $1/2$. Since q_n^* is determined such that $\Pr(piv|q_n^*) = C'(q_n^*)$, for sufficiently large n, we have $\Pr(piv|q_n^*) \simeq C''(1/2)(q_n^* - \frac{1}{2})$. Hence we obtain,

$$\begin{aligned}\Pr(piv|q_n^*) \simeq C''(\frac{1}{2})(q_n^* - \frac{1}{2}) \quad &\Longrightarrow \quad \sqrt{2n}\Pr(piv|q_n^*) \simeq \sqrt{2n+1}C''(\frac{1}{2})(q_n^* - \frac{1}{2}) \\ &\Longrightarrow \quad \sqrt{2n}\sigma(q_n^*)\Pr(piv|q_n^*) \simeq C''(\frac{1}{2})\sigma(q_n^*)^2 D(q_n^*).\end{aligned}$$

Therefore, by Lemma 2, the limit of $D(q_n^*)$ must be finite since if $D(q_n^*)$ diverges to ∞, $\sqrt{2n}\sigma(q_n^*)\Pr(piv|q_n^*)$ converges to 0. Suppose that $D(q_n^*)$ converges to $c < \infty$. Then, we have $\Pr(piv|q_n^*) \simeq \frac{1}{\sqrt{2n}\sigma(q_n^*)} \times \phi(-c)$ (see Fig.1). Thus,

$$\frac{1}{4}C''(\frac{1}{2}) \simeq \frac{\phi(c)}{c},$$

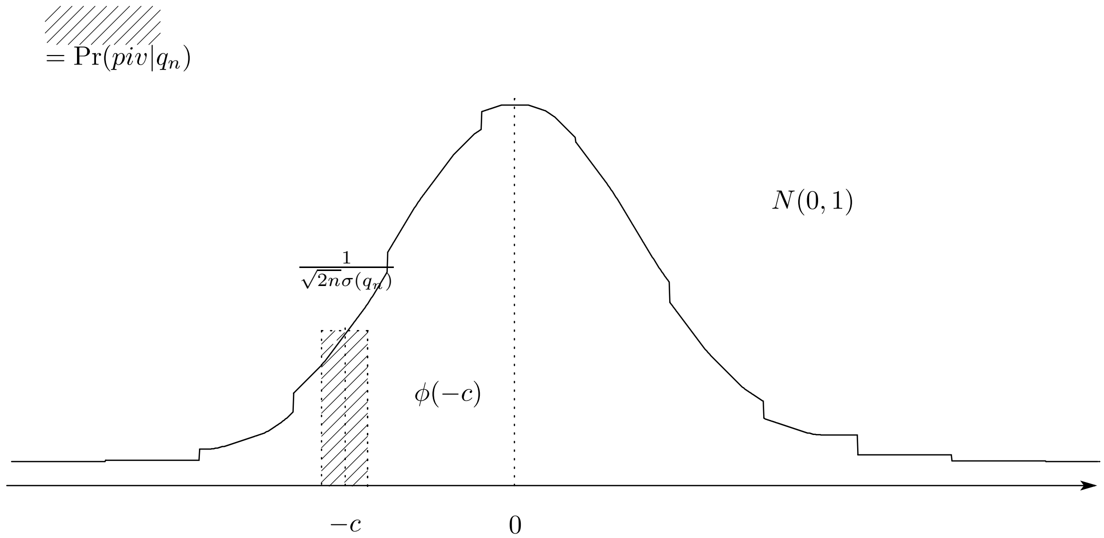

Figure 1:

since $\sigma(q_n^*)$ converges to 1/2. In this way, the limit probability is $\Phi(c)$ such that $\frac{\phi(c)}{c} = \frac{1}{4}C''(1/2)$ by Proposition 2.

On the contrary, if $C''(1/2) = 0$, $C'(q)$ is approximately flat in the neighborhood of 1/2 and hence the speed of convergence of q_n^* is relatively slow compared with that of $\Pr(piv|q_n^*)$. Therefore, CJT is valid.

In the case of the signal with white noise, if the information cost is inverse proportional to the variance of the noise, CJT is not valid.

Example 3 (Continued)
As we showed above, each voter pays positive information cost in the CSS-equilibrium for all n, since $C'(1/2) = 0$. However, CJT is not valid since

$$\left.\frac{\partial^2 C(q)}{(\partial q)^2}\right|_{q=\frac{1}{2}} = \frac{1}{\phi(0)} \frac{\lambda(\phi(0))}{\phi^2(0)} = 2\pi\lambda > 0.$$

Moreover, elections cause the worse outcome than the outcome of the decision made by a single agent though there are infinitely many voters. Suppose that there is only one voter, i.e., $n = 0$. Since the vote is always pivotal, the voter decides q_0^* such that

$$1 = C'(q_0^*) = \frac{\lambda\delta^*}{\phi(\delta^*)}.$$

Therefore, the probability that the desirable alternative is chosen is

$$q_0^* = \Phi(\delta^*) \text{ such that } \frac{\phi(\delta^*)}{\delta^*} = \lambda.$$

On the other hand, consider the limit as n tends to ∞. By Theorem 2, the limit of the probability that the desirable alternative is chosen is $\Phi(c)$ such that

$$\frac{\phi(c)}{c} = \frac{1}{4}C''(1/2) = \frac{\pi\lambda}{2} > \lambda = \frac{\phi(\delta^*)}{\delta^*}.$$

Hence we have $c < \delta^*$ and we conclude $\Phi(\delta^*) > \Phi(c)$. Numerical calculation suggest that the probability that the desirable alternative is chosen decreases monotonically as n increases.

On the other hand, suppose that $\tilde{C}(\delta) = \lambda\delta^3$. In this case, obviously, $C'''(1/2) = 0$ holds and hence CJT is valid. ‖

4 The Signal with White Noise

In this section, we consider the signal with white noise as in Example 3. And we compare the probability that the desirable alternative is chosen in the limit with that in the case of a single voter by applying Theorem 2. We restrict our attention to the class where obtaining the accuracy $q = \Phi(\delta)$ costs $\tilde{C}(\delta)$ with $\tilde{C}'''(0) > 0$, which is equivalent to $C'''(1/2) > 0$. Let all assumptions including Assumption 3 be satisfied, which implies, for example, $\tilde{C}'(\delta) > 0$ for all $\delta > 0$ and $\tilde{C}'(0) = 0$. Hence, by Theorem 2, CJT is not valid.

In addition, we assume $\tilde{C}'''(\delta) > 0$ for all $\delta > 0$. Of course this is a stronger requirement than $C'''(q) > 0$ for all $q > 1/2$. But note that it only requires the increasing marginal cost with respect to the level of separation relative to the standard deviation of the noise.

For convenience, we define $\delta_0(\tilde{C})$ and $\delta_\infty(\tilde{C})$ such that the desirable alternative is chosen with probability $\Phi(\delta_0(\tilde{C}))$ by a single agent and with $\Phi(\delta_\infty(\tilde{C}))$ in the limit. Therefore

$$1 = \frac{\tilde{C}(\delta_0(\tilde{C}))}{\phi(\delta_0(\tilde{C}))}$$

which is the F.O.C. of the single agent and Theorem 2 implies

$$\frac{\phi(\delta_\infty(\tilde{C}))}{\delta_\infty(\tilde{C})} = \frac{\pi}{2}\tilde{C}'''(0).$$

Note that rearranging the first equation yields

$$\frac{\phi(\delta_0(\tilde{C}))}{\delta_0(\tilde{C})} = \frac{\tilde{C}(\delta_0(\tilde{C}))}{\delta_0(\tilde{C})}.$$

Thus the simple condition clarifies the magnitude relationship: let δ_0 be such that $\phi(\delta_0) = \tilde{C}'(\delta_0)$. Then

$$\delta_0(\tilde{C}) \leq \delta_\infty(\tilde{C}) \quad \text{if and only if} \quad \frac{\tilde{C}'(\delta_0)}{\delta_0} \geq \frac{\pi}{2}\tilde{C}'''(0).$$

Using this condition, we find the following properties.

1. Suppose that $\tilde{C}'$ is (weakly) concave. Then $\delta_0(\tilde{C}) > \delta_\infty(\tilde{C})$.
2. Suppose that $\tilde{C}'$ is strictly convex. Then
   (a) if there exists $\bar{\delta} > 0$ such that $\frac{\tilde{C}'(\bar{\delta})}{\bar{\delta}} = \frac{\pi}{2}\tilde{C}'''(0)$, then for $\bar{\alpha} = \frac{\phi(\bar{\delta})}{\tilde{C}'(\bar{\delta})}$,
      i. $\delta_0(\alpha\tilde{C}) < \delta_\infty(\alpha\tilde{C})$ for $0 < \alpha < \bar{\alpha}$,
      ii. $\delta_0(\alpha\tilde{C}) = \delta_\infty(\alpha\tilde{C})$ for $\alpha = \bar{\alpha}$ and

iii. $\delta_0(\alpha\tilde{C}) > \delta_\infty(\alpha\tilde{C})$ for $\alpha > \bar{\alpha}$.

(b) Otherwise, $\delta_0(\alpha\tilde{C}) > \delta_\infty(\alpha\tilde{C})$ for all $\alpha > 0$.

1 contains the case of $\tilde{C}(\delta) = \frac{\lambda}{m}\delta^m$ with $m \leq 2$. Combined with 1 of Theorem 2, we see our policy for the class of $\tilde{C}(\delta) = \frac{\lambda}{m}\delta^m$: rely on limited agents for $m \leq 2$ and allow the full participation for $m > 2$.

Cost structures in which cost becomes prohibitively high at some accuracy level less than 1 may suit 2. For such cases, this result suggests the way we change our policy as the importance of choosing the correct alternative varies while the information structure itself is stable.

5 Conclusions

We showed that the necessary and sufficient condition for CJT is $C''(\frac{1}{2}) = 0$. That is, the probability that majorities can predict the desirable alternative converges to 1 as the number of voters goes to ∞ if and only if $C''(\frac{1}{2}) = 0$. We also showed that the limit probability can be calculated when CJT is not valid.

A Proof of Lemmata

Lemma 2 *Consider the sequence of elections such that there are $2n+1$ voters with accuracy $q_n > 1/2$ and q_n converges to $1/2$.*

1. $\lim_{n\to\infty} D(q_n) = \infty \Rightarrow \lim_{n\to\infty} \sqrt{2n}\sigma(q_n)\Pr(piv|q_n) = 0$.
2. $\lim_{n\to\infty} D(q_n) = c < \infty \Rightarrow \lim_{n\to\infty} \sqrt{2n}\sigma(q_n)\Pr(piv|q_n) = \phi(c)$.

Proof:
Since a vote is pivotal only if the other votes are just split in two,

$$\Pr(piv|q_n) = \frac{(2n)!}{n!n!}\Big\{q_n(1-q_n)\Big\}^n.$$

For convenience define $f(n)$ as $f(n) = 2\sigma^2(q_n)D(q_n)^2$. Then, we have

$$\begin{aligned} f(n) &= 2(q_n(1-q_n))\frac{(2n+1)(q_n-\frac{1}{2})^2}{q_n(1-q_n)} \\ &= 4\left(n+\frac{1}{2}\right)\left(q_n-\frac{1}{2}\right)^2. \end{aligned}$$

And then,

$$1 - \frac{f(n)}{n+\frac{1}{2}} = 1 - \left\{2\left(q_n - \frac{1}{2}\right)\right\}^2 = 2^2(q_n(1-q_n)).$$

Therefore, we obtain

$$
\begin{aligned}
\sqrt{2n}\sigma(q_n)\Pr(\text{piv.}|q_n) &= \sqrt{2n}\frac{(2n)!}{n!n!}\Big\{q_n(1-q_n)\Big\}^{n+\frac{1}{2}} \\
&= \sqrt{2n}\frac{(2n)!}{n!n!}\Big\{q_n(1-q_n)\Big\}^{n+\frac{1}{2}}\frac{2^{2n+\frac{1}{2}}}{2^{2n+\frac{1}{2}}}\frac{\sqrt{\pi}}{\sqrt{\pi}} \\
&= \frac{1}{\sqrt{2\pi}}\frac{(2n)!}{n!n!}\frac{\sqrt{\pi n}}{2^{2n}}\Big\{2^2(q_n(1-q_n))\Big\}^{n+\frac{1}{2}} \\
&= \frac{1}{\sqrt{2\pi}}\frac{\frac{(2n)!}{n!n!}}{\frac{4^n}{\sqrt{\pi n}}}\left\{\left[1-\frac{1}{\frac{n+\frac{1}{2}}{f(n)}}\right]^{\frac{n+\frac{1}{2}}{f(n)}}\right\}^{f(n)}.
\end{aligned}
$$

By the Starling's formula,[5] the 2nd factor of the RHS converges to 1.

Note that $\frac{n+\frac{1}{2}}{f(n)}$ goes to ∞ since q_n converges to $1/2$. Thus, when $D(q_n)$ converges to some finite value c, the RHS converges to $\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{c^2}{2}\right) = \phi(c)$. When $D(q_n)$ goes to ∞, the RHS converges to 0. □

Lemma 3 *Suppose that $\{q_n^*\}$ is the sequence of CSS-equilibrium accuracies.*

1. *If $C''(1/2) = 0$,*

$$\lim_{n\to\infty} D(q_n^*) = \infty.$$

2. *If $C''(1/2) > 0$,*

$$\lim_{n\to\infty} D(q_n^*) = c < \infty,$$

such that $\frac{\phi(c)}{c} = \frac{1}{4}C''(\frac{1}{2})$.

Proof:
Rearranging Eq (3) yields

$$\frac{\sqrt{2n}\sigma(q_n^*)\Pr(piv|q_n^*)}{D(q_n^*)} = \frac{\sqrt{2n}}{\sqrt{2n+1}}\sigma^2(q_n^*)\frac{C'(q_n^*)}{q_n^* - \frac{1}{2}}.$$

Since q_n^* converges to $\frac{1}{2}$, and $C'(\frac{1}{2}) = 0$, we obtain

$$
\begin{aligned}
\lim_{n\to\infty}\frac{\sqrt{2n}\sigma(q_n^*)\Pr(piv|q_n^*)}{D(q_n^*)} &= \lim_{n\to\infty}\frac{\sqrt{2n}}{\sqrt{2n+1}}\sigma^2(q_n^*)\frac{C'(q_n^*) - C'(\frac{1}{2})}{q_n^* - \frac{1}{2}} \\
&= \frac{1}{4}\lim_{n\to\infty}\frac{C'(q_n^*) - C'(\frac{1}{2})}{q_n^* - \frac{1}{2}} \\
&= \frac{1}{4}C''(\frac{1}{2}).
\end{aligned}
$$

Thus, by Lemma 2, if $C''(\frac{1}{2}) = 0$, then $\lim_{n\to\infty} D(q_n^*)$ must be ∞. On the other hand, if $C''(\frac{1}{2}) > 0$, then $\lim_{n\to\infty} D(q_n^*)$ must be $c < \infty$ such that $\frac{\phi(c)}{c} = \frac{1}{4}C''(\frac{1}{2})$. □

[5] Note that $\frac{4^n}{\sqrt{\pi n}} = \frac{\sqrt{2\pi(2n)}(2n)^{2n}e^{-2n}}{(\sqrt{2\pi n}n^n e^{-n})^2}$.

References

[1] Austen-Smith, D. and Banks, J. S. (1996), "Information Aggregation, Rationality, and the Condorcet Jury Theorem," *American Political Science Review*, 90, 1, 34-45.

[2] Black, D. (1958), *The Theory of Communities and Elections*, Cambridge University Press, Cambridge.

[3] Berend, D. and Paroush, J. (1998), "When is Condorcet's Jury Theorem Valid?" *Social Choice and Welfare*, 15, 481-88.

[4] Condorcet, M. de (1785), *Essai sur l'application de l'analyse la probabilit des d cisions redues la pluralit des voix*, Paris.

[5] Feddersen, T. and Pesendorfer, W. (1997), "Voting Behavior and Information Aggregation in Elections with Private Information," *Econometrica*, 65, 5, 1029-58.

[6] Feller, W. (1971), *An Introduction to Probability Theory and Its Applications,* Vol. 2, 2nd ed., John Wiley & Sons, New York, London.

[7] Grossman, S. and Stiglitz, J. (1980), "On the Impossibility of Informationally Efficient Markets," *American Economic Review*, 70, 393-408.

[8] Jackson, M. O. (2003), "Efficiency and Information Aggregation in Auctions with Costly Information," *Review of Economics Design*, 8, 121-41.

[9] Ladha, K. K. (1992), "The Condorcet Jury Theorem, Free Speech and Correlated Votes," *American Journal of Political Science*, 36, 617-34.

[10] Ladha, K. K. (1995), "Information Polling through Majority Rule Voting: Condorcet's Jury Theorem with Correlated Votes," *Journal of Economic Behavior Organization*, 26, 353-72.

[11] Mukhopadhaya, K. (2003), "Jury Size and the Free Rider Problem," *The Journal of Law, Economics, & Organization*, 19, 1, 24-44.

[12] Pesendorfer, W. and Swinkels, J. M. (2000), "Efficiency and Information Aggregation in Auctions," *American Economic Review*, 90, 3, 499-525.

[13] Razin, R. (2003), "Signaling and Election Motivations in a Voting Model with Common Values and Responsive Candidates," *Econometrica*, 71, 4, 1083-1119.

[14] Young, P. (1995), "Optimal Voting Rules," *Journal of Economic Perspective*, 9, 1, 51-64.