The Equivalence Between Costly and Probabilistic Voting Models

Yuichiro Kamada and Fuhito Kojima*

First Version: June 4, 2009; This Version: March 4, 2013

Abstract

In costly voting models, voters abstain when a stochastic cost of voting exceeds the benefit from voting. In probabilistic voting models, they always vote for a candidate who generates the highest utility, which is subject to random shocks. We prove an equivalence result: In two-candidate elections, given any costly voting model, there exists a probabilistic voting model that generates winning probabilities identical to those in the former model for any policy announcements, and vice versa. Thus many predictions of interest established in one of the models hold in the other as well, providing robustness of the conclusions to model specifications.

*Kamada: Cowles Foundation for Research in Economics, Yale University, New Haven, CT 06511, and Haas School of Business, University of California, Berkeley, Berkeley, CA 94720, e-mail: yuichiro.kamada@yale.edu. Kojima: Department of Economics, Stanford University, Stanford, CA 94305, e-mail: fkojima@stanford.edu. We are grateful to Attila Ambrus, John William Hatfield, Matt Jackson, Martin Osborne, Yuki Takagi, and an anonymous referee. Bobak Pakzad-Hurson and Fanqi Shi provided excellent research assistance.

1 Introduction

Electoral competitions have been analyzed with game theoretic tools, beginning with seminal studies by Hotelling (1929), Downs (1957) and Black (1958). Their models are useful but highly stylized benchmarks, and many subsequent works modified their models to better fit theoretical predictions to the real world. There are at least two assumptions that these subsequent works have relaxed. One is the assumption that no voter abstains from voting, and the other is that it is with certainty that each voter votes for a candidate whose policy is strictly closer to her bliss point.

Costly voting models address the first point. In those models, each voter randomly draws a cost of voting from a certain distribution, and votes for the candidate with a closer policy if and only if the voting cost is smaller than the perceived gain from doing so. Otherwise, a voter abstains from voting. *Probabilistic voting models* tackle the second issue. In those models, voters cannot abstain. Each voter obtains a random utility shock for the candidates in addition to the deterministic utility from announced policies, and votes for the candidate whose overall utility is higher.[1]

These two frameworks have proved useful for predictions, and their properties have been investigated extensively in the literature.[2] However, these approaches have modified different assumptions of the basic Hotelling-Downs framework, and the relationship between them has been unclear in the literature.

In this paper, we show that these two classes of models are equivalent in the context of two-candidate election: For any pair of the two candidates' announced policies, the winning probability of each candidate is identical under costly and probabilistic voting models. More specifically, our main result demonstrates that (1) given any cost distribu-

[1] Standard interpretations of the random utility shocks include ideology or personal characteristics of the party leadership (Persson and Tabellini, 2000).

[2] See Banks and Duggan (2005) and references therein. For example, in probabilistic voting models it is known that with certain regularity conditions the existence of a Nash equilibrium is established. On costly voting models, see an extensive survey by Osborne (1995) as well as more recent studies such as Borgers (2004) for instance.

tion, there exists a distribution of the random utility shocks such that the costly voting model with abstention associated with the former distribution is equivalent to the probabilistic voting model with the latter, and (2) given any random utility shock distribution, there exists a cost distribution such that the probabilistic voting model associated with the former distribution is equivalent to the costly voting model with abstention associated with the latter. Therefore, many predictions of interest established in one of the models carry over to the other, providing robustness of the conclusions to model specifications. The sets of equilibrium policy profiles announced by the two candidates who try to maximize the winning probabilities, for instance, are identical under these two models. As a concrete example, we note that Kamada and Kojima (2011) show that the equilibrium policies can diverge in probabilistic voting if voters have convex utility functions. Our equivalence result implies that their conclusion holds under costly voting models as well.

We also analyze more general hybrid models with both voting cost and random utility shocks. Generalizing our main result we show that, given any such hybrid model, there exist a costly voting model and a probabilistic voting model each of which is equivalent to the given hybrid model. This result implies that there is no loss of generality in assuming away either the cost or random utility shocks from the model. This result can facilitate analysis of voting by giving justification for studying relatively simple models with either cost or utility shocks but not both.

We emphasize that the objective of this paper is not to give a mathematically involved proof (indeed, the proof itself is pretty straightforward), nor to establish the relevance of either or both of the two classes of models. Rather, our objective is to set up the problem in an appropriate way so that the connection becomes clear between the two frameworks that are motivated quite differently in the literature.

The rest of this paper proceeds as follows. Section 2 introduces the environment. Section 3 presents the main theorem. In Section 4, we consider a generalization to hybrid models. Section 5 concludes. Some proofs are relegated to the Appendix.

2 The Environment

A continuum of voters are distributed over a policy space X according to a Borel probability measure μ.[3] There are two political candidates, A and B, who simultaneously choose their policies on X, x_A and x_B, respectively. A voter with position $x \in X$ obtains a deterministic utility $u(x, x')$ from the policy x'. A standard example for this is that the utility depends (only) on the "distance" between x and x', which is defined over the set X.

A **voting model** M is characterized by a corresponding random function $w^M(x_A, x_B) \in \{A, B\}$ that assigns a winner of the election given the policy profile (x_A, x_B). We say that two voting models M and M' are **equivalent** if $\Pr(w^M(x_A, x_B) = A) = \Pr(w^{M'}(x_A, x_B) = A)$ for any given policy profile (x_A, x_B). Note that the definition of a voting model does not specify candidates' objective functions: Our equivalence result is applicable without explicit reference to them.

Two classes of voting models are defined below. In each voting model M, $w^M(x_A, x_B) = i$ if the measure of voters that vote for candidate i is greater than that for the other candidate. If they are exactly equal then $w^M(x_A, x_B)$ takes A and B with equal probabilities.

Costly Voting Models

Each voter experiences a cost $c \geq 0$ from voting.[4] The costs are independent of voters' positions and across voters, and each cost follows a continuous cumulative distribution function $F : \mathbb{R} \to [0, 1]$ such that $F(t) = 0$ for all $t \leq 0$.[5] Let $\mathcal{F}$ be the set of all possible

[3] Few assumptions are imposed on X and μ except that the Borel measure μ can be defined on X. In particular, standard environments such as (multi-dimensional) Euclidean spaces are allowed, but our result applies to other settings. Also note that μ need not be nonatomic.

[4] Our analysis can be extended easily to allow for negative voting cost (Riker and Ordeshook, 1968). An equivalence result holds between such costly voting models and probabilistic voting models in which the probability that a voter receives no utility shock is strictly positive. Similarly, the analysis can be extended to distributions that allow for atoms although we will assume continuous distributions for simplicity.

[5] To be formal, a continuum of voters are distributed over the space $X \times \mathbb{R}_+$ according to a product measure of μ and a measure induced by F, and a voter associated with a pair (x, c) is interpreted as positioning at a point x in the policy space and experiencing a cost c from voting. Note that we assume

cumulative distribution functions with such a restriction.

A voter with position x votes for candidate A with probability 1 if

$$\phi\left(u(x, x_A) - u(x, x_B)\right) > c,$$

and she votes for candidate B with probability 1 if

$$\phi\left(u(x, x_B)) - u(x, x_A)\right) > c,$$

where $\phi : \mathbb{R} \to \mathbb{R}$ is a continuous and strictly increasing function with $\phi(0) = 0$ and $\lim_{t\to\infty} \phi(t) = \infty$ (thus its inverse function ϕ^{-1} exists and is continuous and strictly increasing). Otherwise, the voter abstains from voting.

This class of models subsumes a setup in which the motivation for each voter is to influence the electoral outcome and a voter estimates that, potentially incorrectly, the probability that she is pivotal is a fixed number $p > 0$ (thus $\phi(t) = pt$ for any $t \in \mathbb{R}$). It excludes models in which each voter correctly evaluates the probability of becoming pivotal[6] (see Ledyard (1984) Palfrey and Rosenthal (1985), and McKelvey and Patty (2006)[7]), but many other models in the literature are included. For example, the voter may vote because of "expressive" motives (see Schuessler (2000) and Glaeser et al. (2005)), in which case function ϕ specifies the salience of the political issue in the voter's mind. Denote by $M^C(F)$ the costly voting model with cost distribution $F \in \mathcal{F}$.

---

that the distribution of voters' positions and realized costs is given by μ and F, rather than modeling a continuum of random position-cost pairs. This is a standard modeling approach in the literature, based on the intuitive idea of the "law of large numbers." See however Judd (1985) and Uhlig (1996) for technical issues associated with invoking the law of large numbers for a continuum of random variables. Analogous remarks also apply to the probabilistic voting and hybrid models below.

[6]With a continuum of voters the probability of becoming pivotal is zero, so models with a finite number of voters are more suitable for analyzing this question. As we will show in Remark 2, the equivalence theorem does not extend to the case with a finite number of voters if the candidates' objective is to maximize the winning probabilities. While that remark assumes a fixed pivot probability, it suggests that the equivalence does not hold in the case of a finite number of voters with correct pivot probabilities either.

[7]Feddersen (2004) provides an extensive review of the literature.

Probabilistic Voting Models

Each voter does not have a choice to abstain, while their decision on who to vote for depends on two elements: The first is the deterministic utility of electing the respective candidates defined earlier, and the second is a random shock in the evaluation of the two candidates. Specifically, each voter experiences a random shock ξ in utility that favors candidate A. The random shocks are independent of voters' positions and across voters, and each random shock follows a continuous cumulative distribution function $G : \mathbb{R} \to [0, 1]$ which is symmetric around 0. Let $\mathcal{G}$ be the set of all cumulative distribution functions with such a restriction.

A voter with position x votes for candidate A with probability 1 if

$$u(x, x_A) + \xi > u(x, x_B),$$

and for candidate B with probability 1 if

$$u(x, x_B) > u(x, x_A) + \xi.$$

If $u(x, x_A) + \xi = u(x, x_B)$, the voter votes for the two candidates with probability $\frac{1}{2}$ for each. Denote by $M^P(G)$ the probabilistic voting model with shock distribution $G \in \mathcal{G}$.

3 The Equivalence Theorem

Theorem 1. *Fix ϕ. For any cost distribution $F \in \mathcal{F}$, there exists a utility shock distribution $G \in \mathcal{G}$ such that the costly voting model $M^C(F)$ is equivalent to the probabilistic voting model $M^P(G)$. Conversely, for any $G \in \mathcal{G}$, there exists $F \in \mathcal{F}$ such that $M^C(F)$ is equivalent to $M^P(G)$.*

This is the main theorem of this paper, which establishes the equivalence between a costly voting model and a probabilistic voting model. As we have discussed in the

Introduction, the theorem implies that many predictions of interest established in one of the models carry over to the other, which provides robustness of the conclusions to model specifications. The sets of equilibrium policy profiles announced by the two candidates who try to maximize the winning probabilities, for instance, are identical under these two models.

The proof is simple.[8] We take the difference of expected vote shares of the two candidates in each model, which is a sufficient statistic of the winning probability. Then we construct a probability distribution of random utility shocks (respectively costs of voting) that generates the same expected vote share difference given any policy profile and a probability distribution of cost of voting (respectively random utility shock). It turns out that the restrictions on these probability distributions imposed on each model are sufficient to guarantee that the construction works.

Proof. Fix a policy profile (x_A, x_B). Given a costly voting model or a probabilistic voting model M, let $R(M, x_A, x_B)$ be the "relative vote share," the measure of voters who vote for candidate A minus that for candidate B under M and policy profile (x_A, x_B). To prove the theorem, it suffices to show that (i) for any $F \in \mathcal{F}$ there exists $G \in \mathcal{G}$ such that $R(M^C(F), x_A, x_B) = R(M^P(G), x_A, x_B)$ for every (x_A, x_B), and (ii) for any $G \in \mathcal{G}$ there exists $F \in \mathcal{F}$ such that $R(M^C(F), x_A, x_B) = R(M^P(G), x_A, x_B)$ for every (x_A, x_B).

By definition, for any (x_A, x_B) we have

$$R(M^C(F), x_A, x_B) = \int_X \left[F\left(\phi(u(x, x_A) - u(x, x_B))\right) - F\left(\phi(u(x, x_B) - u(x, x_A))\right)\right] d\mu(x)$$

for any $F \in \mathcal{F}$, and

$$R(M^P(G), x_A, x_B) = \int_X \left[G\left(u(x, x_A) - u(x, x_B)\right) - G\left(u(x, x_B) - u(x, x_A)\right)\right] d\mu(x)$$

[8] As mentioned in the Introduction, mathematical difficulty of the proof is not our focus.

for any $G \in \mathcal{G}$.[9]

Denoting $t = u(x, x_A) - u(x, x_B)$ for notational simplicity, we have:

$$R(M^C(F), x_A, x_B) = \int_X \left[F\left(\phi(t)\right) - F\left(\phi(-t)\right)\right] d\mu(x)$$

and

$$R(M^P(G), x_A, x_B) = \int_X \left[G\left(t\right) - G\left(-t\right)\right] d\mu(x).$$

This means that it suffices to show that (i') for any $F \in \mathcal{F}$ there exists $G \in \mathcal{G}$ such that the following condition (1) holds, and (ii') for any $G \in \mathcal{G}$ there exists $F \in \mathcal{F}$ such that condition (1) holds,[10] where

$$F\left(\phi(t)\right) - F\left(\phi(-t)\right) = G\left(t\right) - G\left(-t\right) \text{ for all } t \in \mathbb{R}. \tag{1}$$

Now we prove (i') and (ii').

Part (i'): Let $F \in \mathcal{F}$ be given. Define function G by

$$G(t) = \begin{cases} \frac{1}{2}F(\phi(t)) + \frac{1}{2}, & t \geq 0, \\ -\frac{1}{2}F(\phi(-t)) + \frac{1}{2}, & t < 0. \end{cases} \tag{2}$$

Since G is nondecreasing and satisfies $\lim_{t\to-\infty} G(t) = -\frac{1}{2}\lim_{t\to-\infty} F(\phi(-t)) + \frac{1}{2} = 0$ and $\lim_{t\to\infty} G(t) = \frac{1}{2}\lim_{t\to\infty} F(\phi(t)) + \frac{1}{2} = 1$, G is a cumulative distribution function over $\mathbb{R}$. Moreover, G is symmetric around zero since $G(0) = \frac{1}{2}F(\phi(0)) + \frac{1}{2} = \frac{1}{2}$ and, for any $t > 0$,

$$\begin{aligned} G(-t) &= -\frac{1}{2}F(\phi(t)) + \frac{1}{2} \\ &= 1 - G(t). \end{aligned}$$

[9] $R(M^P(G), x_A, x_B) = \int_X \left[1 - G\left(u(x, x_B) - u(x, x_A)\right) - G\left(u(x, x_B) - u(x, x_A)\right)\right] d\mu(x)$, and symmetry of G leads to the desired expression.

[10] Note that (i') and (ii') imply (i) and (ii) above, respectively.

Since ϕ and F are continuous and $\lim_{t\to 0-} G(t) = \frac{1}{2} = G(0)$, G is continuous. Hence $G \in \mathcal{G}$. Substituting (2) into the right hand side of (1), we obtain $F(\phi(t))$ if $t \geq 0$ and $-F(\phi(-t))$ if $t < 0$. In either case, the expression is equal to $F(\phi(t)) - F(\phi(-t))$, the left hand side of (1).[11]

Part (ii'): Let $G \in \mathcal{G}$ be given. Define F by

$$F(t) = \begin{cases} 2(G(\phi^{-1}(t)) - \frac{1}{2}), & t \geq 0, \\ 0, & t < 0. \end{cases} \tag{3}$$

F is nondecreasing and satisfies $F(t) = 0$ for all $t \leq 0$, and $\lim_{t\to\infty} F(t) = 2(\lim_{t\to\infty} G(\phi^{-1}(t)) - \frac{1}{2}) = 1$. Thus in particular, F is a cumulative distribution function over $\mathbb{R}$. Since ϕ^{-1} and G are continuous, F is continuous. Hence $F \in \mathcal{F}$. Substituting (3) into the left hand side of (1), we obtain $2(G(t) - \frac{1}{2})$ if $t \geq 0$ and $-2(G(-t) - \frac{1}{2})$ if $t < 0$. In either case, the expression is equal to $G(t) - G(-t)$, the right hand side of (1).[12] This completes the proof. □

Remark 1. The theorem establishes the equivalence in terms of winning probabilities. This in particular implies the equivalence in terms of the sets of Nash equilibria of the games in which each candidate maximizes the winning probability under probabilistic and costly voting models. However, there are models in the literature in which not only the winning probabilities but also the vote shares enter the candidates' objective functions.[13] Our equivalence result does not extend to all of those cases. For example, if each candidate maximizes her own vote share (that is, the total measure of voters voting for her), then the equivalence does not hold due to the possibility of abstention in costly voting models. However, the equivalence does extend to a case in which the vote share affects the candidates' utilities in a particular manner. Specifically, consider the case in which each candidate values having a larger vote share herself while she dislikes the

[11] Recall the value of F is zero in the nonpositive domain.

[12] Since $G(t) = 1 - G(-t)$ by symmetry, $G(t) - G(-t) = 2(G(t) - \frac{1}{2}) = -2(G(-t) - \frac{1}{2})$.

[13] See Zakharov (2012) for such specifications.

opponent having a larger vote share.[14] The proof of the theorem shows that the relative vote shares (the differences of the two candidates' vote shares) are identical under the probabilistic and costly voting models. Therefore, if each candidate's utility is a function of the relative vote share, then the set of Nash equilibria are identical under these two models.[15]

Remark 2. Throughout the paper, we assume that there is a continuum of voters. Our result immediately extends to a model in which there is a finite number of voters if the candidates try to maximize the expected relative vote share, by an essentially identical argument as above. On the other hand, our result does not extend to the case with a finite number of voters if the candidates care about winning probabilities. Appendix A.2 provides an example showing this point.

Remark 3. There is a literature that studies equivalence of the candidates' behavior under different objectives (Hinich (1977), Wittman (1983), Duggan (2000), Patty (2002, 2005), and Zakharov (2012)). This is a different notion of equivalence from what we study in this paper, which studies equivalence between different models of voter behavior while keeping the candidates' objectives fixed.

4 Hybrid Models

In this section we analyze general hybrid models with both voting cost and random utility shocks. We show that, given any such hybrid model, there exist a costly voting model and a probabilistic voting model each of which is equivalent to the given hybrid model.

[14]Zakharov (2012) presents various ways in which each candidate's utility depends on her vote share, and studies models incorporating this dependence. In his models a candidate's utility is increasing in her vote share. Since he assumes that voters do not abstain, the candidate's utility is decreasing in her opponent's vote share.

[15]One could imagine other possibilities for both candidates' vote shares affecting their utilities. For example, the ratio (as opposed to the difference), or some function of the difference and the ratio may affect the utilities. As with our stance throughout the paper we do not intend to take a particular stance on what model to deem as realistic; our primary concern in this paper is to understand the relationships between various election models.

This result implies that there is no loss of generality in assuming away either the cost or random utility shocks from the model. This result can facilitate analysis of voting by giving justification for studying relatively simple models with only either cost or utility shocks but not both.

In a hybrid model, each voter experiences a cost of voting, denoted c, distributed according to $F \in \mathcal{F}$, and a random shock in the evaluation of the two candidates, denoted ξ, distributed according to $G \in \mathcal{G}$. Hence, a voter with position x votes for candidate A with probability 1 if

$$\phi\left((u(x, x_A) + \xi - u(x, x_B)\right) > c$$

and she votes for candidate B with probability 1 if

$$\phi\left(u(x, x_B) - (u(x, x_A) + \xi)\right) > c.$$

Otherwise, the voter abstains from voting. Denote by $M^H(F, G)$ the voting model with cost distribution $F \in \mathcal{F}$ and utility shock distribution $G \in \mathcal{G}$.

Theorem 2. *Fix ϕ. For any cost distribution $F \in \mathcal{F}$ and utility shock distribution $G \in \mathcal{G}$, there exist $\bar{F} \in \mathcal{F}$ and $\bar{G} \in \mathcal{G}$ such that each of the costly voting model $M^C(\bar{F})$ and the probabilistic voting model $M^P(\bar{G})$ is equivalent to the hybrid model $M^H(F, G)$.*

Proof. See the Appendix. □

5 Conclusion

In this paper, we considered costly and probabilistic voting models with two candidates and proved an equivalence result: Given any costly voting model, there exists a probabilistic voting model that generates winning probabilities identical to those in the former model for any policy announcements, and vice versa. This result implies that many predictions of interest established in one of the models carry over to the other, providing robustness of the conclusions to model specifications. The sets of equilibrium policy profiles announced by the two candidates who try to maximize the winning probabilities, for instance, are identical under these two models. We also showed that, given any hybrid model with both voting cost and random utility shocks, there exist a costly voting model and a probabilistic voting model each of which is equivalent to the given hybrid model. This result implies that there is no loss of generality in assuming away either the cost or random utility shocks from the model.

A Appendix

A.1 Proof of Theorem 2

Proof. Extend the definition of $R(M, x_A, x_B)$ in the proof of Theorem 1 to include hybrid models. Again this is well-defined. Let H be the cumulative distribution function of $\phi^{-1}(c) - \xi$.

By definition, we have

$$R(M^C(\bar{F}), x_A, x_B) = \int_X \left[\bar{F}\left(u(x, x_A) - u(x, x_B)\right) - \bar{F}\left(u(x, x_B) - u(x, x_A)\right)\right] d\mu(x)$$

for any $\bar{F} \in \mathcal{F}$ as before, and

$$R(M^H(F, G), x_A, x_B) = \int_X \left[H\left(u(x, x_A) - u(x, x_B)\right) - H\left(u(x, x_B) - u(x, x_A)\right)\right] d\mu(x)$$

for any H as constructed above. Denoting $t = u(x, x_A) - u(x, x_B)$, we have:

$$R(M^C(\bar{F}), x_A, x_B) = \int_X \left[\bar{F}\left(t\right) - \bar{F}\left(-t\right)\right] d\mu(x)$$

and

$$R(M^H(F,G),x_A,x_B) = \int_X \left[H\left(t\right) - H\left(-t\right)\right] d\mu(x).$$

This means it suffices to show that for any $F \in \mathcal{F}$ and $G \in \mathcal{G}$ there exists $\bar{F} \in \mathcal{F}$ such that condition (4) holds, where

$$\bar{F}\left(t\right) - \bar{F}\left(-t\right) = H\left(t\right) - H\left(-t\right). \tag{4}$$

Let $F \in \mathcal{F}$ and $G \in \mathcal{G}$ be given. Define $\bar{F}$ by

$$\bar{F}(t) = \begin{cases} H(t) - H(-t), & t \geq 0, \\ 0, & t < 0. \end{cases}$$

Since H is nondecreasing by construction, $\bar{F}$ is nondecreasing. Moreover, $\bar{F}$ satisfies $\bar{F}(t) = 0$ for all $t \leq 0$ by definition, and $\lim_{t\to\infty} \bar{F}(t) = \lim_{t\to\infty} H(t) - H(-t) = 1$. Therefore $\bar{F}$ is a cumulative distribution function over $\mathbb{R}$ and $\bar{F} \in \mathcal{F}$. It is straightforward by inspection that the above definition of $\bar{F}$ satisfies condition (4).

To complete the proof note that, by Theorem 1, there exists $\bar{G} \in \mathcal{G}$ such that $M^P(\bar{G})$ is equivalent to $M^C(\bar{F})$. Since $M^C(\bar{F})$ is equivalent to $M^H(F,G)$ by the preceding argument, this implies that $M^P(\bar{G})$ is equivalent to $M^H(F,G)$. This completes the proof.

□

A.2 Appendix: Finite Number of Voters

As mentioned in Remark 2, our equivalence result does not extend to the case with a finite number of voters if the candidates care about winning probabilities. The proof of the theorem does not apply because the claim that statements (i') and (ii') are sufficient for the equivalence is not true when the number of voters is finite, although these statements themselves are true. In this section we provide a counterexample that exploits this point, showing that neither of the directions of the theorem holds with a finite number of voters.

Example 1. Consider an environment with policy space $[0, 1)$ and two voters, one at point 0 and another at point 1/2. Each voter's utility function is $u(x, x') = -1/(1 - |x - x'|)$.[16,17] Consider the two candidates' policy profile (x_A, x_B) such that $x_A = 1/2 - y$ and $x_B = 1/2 + y$ where $0 < y < 1/2$.

Consider first a probabilistic voting model. Let $e(z)$ be the probability that the voter at 0 votes for candidate B, where z is the utility difference from the policy positions of A and B (note that z can vary from 0 to infinity as y varies from 0 to 1/2). Then the probability that B wins is $(1/2) \times e(z) + (1/2) \times [(1/2) \times e(z) + (1/2) \times (1 - e(z))] = e(z)/2 + 1/4$.[18]

Second, consider a costly voting model. Let F be the cumulative distribution function of voting cost, and assume that ϕ is an identity function. Since B wins in the costly voting model only when no one votes, and in that case B wins with a conditional probability of 1/2, her winning probability is $(1/2) \times (1 - F(z))$.

Now we show that the equivalence does not hold in this example. To see this, note that z can take any values in $(0, \infty)$ as y varies in $(0, 1/2)$. Suppose first that for any random shock distribution G in a probabilistic voting model, there exists a cost distribution F in a costly voting model such that the costly voting model under F is equivalent to the probabilistic voting model under G. Then we must have $F(z) = 1/2 - e(z)$. Since $e(z) \geq 0$ for all z, this means that F cannot take values exceeding 1/2, which is a contradiction to the property of cumulative distribution functions. Suppose second that for any cost distribution F in a costly voting model, there exists a random utility shock distribution G such that the probabilistic voting model under G is equivalent to the costly voting model

[16] Recall that here, x and x' are the bliss point of the voter and the (implemented) policy, respectively.

[17] We assume that the positions of voters are deterministic. It is straightforward to verify that an analogous argument as below shows impossibility of equivalence when voters are distributed independently and identically according to the specified distribution.

[18] The first term of the left hand side of this expression corresponds to the case in which B gets two votes, and the second term corresponds to the case of receiving one vote. In the second term, the first "1/2" corresponds to the fact that B gets elected with conditional probability 1/2 in the case of tie. The first term inside the large parenthesis accounts for the case in which the voter at 0 votes for B, and the second term accounts for the case in which the voter at 1/2 votes for B.

under F. Then we must have $e(z) = 1/2 - F(z)$. Since F must take values above $1/2$ for sufficiently large z, this means that $e(z)$ takes strictly negative values for sufficiently large z, which contradicts the property of probabilities. $\square$

The reason for the difference between the continuum models and the finite models is that the winning probability is given by the probability that the relative vote share is above zero (plus half the probability that the relative vote share is zero), and the expected relative vote share is a sufficient statistic for the probability of this value only if there are infinitely many voters. In the example, in any probabilistic voting model there is always at least probability $1/4$ that B wins because with probability $1/2$ the voter at $1/2$ votes for B and in that case even when the voter at 0 votes for A, B wins the tie with probability $1/2$. Such a "guaranteed" probability does not exist in costly voting models—in any costly voting model the probability of B's victory is approximated to zero as $y \to 1/2$. This difference comes from the fact that the voter at $1/2$ is "large" in a sense and affects the winning probability too much, so it creates the difference between the expected relative vote share and the probability of winning. The "guaranteed" probability tends to zero when we have n voters at each of points 0 and $1/2$ and $n \to \infty$.

References

Banks, Jeffrey S. and John Duggan, "Probabilistic Voting in the Spatial Models of Elections: The Theory of Office-Motivated Candidates," in David Austen-Smith and John Duggan, eds., *Social Choice and Strategic Decisions: Essays in Honor of Jeffrey S. Banks*, Heidelberg: Springer, 2005.

Black, Duncan, *The Theory of Committees and Elections*, Cambridge: Cambridge University Press, 1958.

Borgers, T., "Costly voting," *The American Economic Review*, 2004, *94* (1), 57–66.

Downs, Anthony, *An Economic Theory of Democracy*, New York: Harper and Row, 1957.

Duggan, John, "Equilibrium Equivalence Under Expected Plurality and Probability of Winning Maximization," 2000. mimeo.

Feddersen, T.J., "Rational choice theory and the paradox of not voting," *The Journal of Economic Perspectives*, 2004, *18* (1), 99–112.

Glaeser, EL, GAM Ponzetto, and JM Shapiro, "Strategic extremism: Why democrats and republicans divide on religious values," *Quarterly Journal of Economics*, 2005, *120* (4), 1283–1330.

Hinich, Melvin J., "Equilibrium in spatial voting: The median voter result is an artifact," *Journal of Economic Theory*, 1977, *16* (2), 208–219.

Hotelling, Harold, "Stability in Competition," *Economic Journal*, 1929, *39*, 41–57.

Judd, K.L., "The law of large numbers with a continuum of IID random variables," *Journal of Economic Theory*, 1985, *35* (1), 19–25.

Kamada, Yuichiro and Fuhito Kojima, "Voter Preferences, Polarization, and Electoral Policies," 2011. mimeo.

Ledyard, J.O., "The pure theory of large two-candidate elections," *Public Choice*, 1984, *44* (1), 7–41.

McKelvey, Richard D. and John W. Patty, "A theory of voting in large elections," *Games and Economic Behavior*, 2006, *57* (1), 155–180.

Osborne, Martin J., "Spatial Models of Political Competition Under Plurality Rule: A Survey of Some Explanations of The Number of Candidates and The Positions They Take," *Canadiann Journal of Economics*, 1995, *2*, 261–301.

Palfrey, T.R. and H. Rosenthal, "Voter participation and strategic uncertainty," *The American Political Science Review*, 1985, pp. 62–78.

Patty, John W., "Equivalence of Objectives in Two Candidate Elections," *Public Choice*, 2002, *112* (1), 151–166.

_ , "Local Equilibrium Equivalence in Probabilistic Voting Models," *Games and Economic Behavior*, 2005, *51* (2), 523–536.

Persson, Torsten and Guido Tabellini, *Political Economics: Explaining Economic Policy*, Cambridge: MIT Press, 2000.

Riker, W.H. and P.C. Ordeshook, "A Theory of the Calculus of Voting," *The American Political Science Review*, 1968, *62* (1), 25–42.

Schuessler, A.A., *A logic of expressive choice*, Princeton Univ Pr, 2000.

Uhlig, H., "A law of large numbers for large economies," *Economic Theory*, 1996, *8* (1), 41–50.

Wittman, Donald A., "Candidate motivation: A synthesis of alternative theories," *American Political Science Review*, 1983, *77* (1), 142–157.

Zakharov, A., "Probabilistic voting equilibria under nonlinear candidate payoffs," *Journal of Theoretical Politics*, 2012, *24* (2), 235–247.