On Optimal Voting Rules with Homogeneous Preferences*

Arnaud Costinot† Navin Kartik‡

First version: November 2005; This version: January 2009.

Abstract

This paper analyzes the choice of optimal voting rules under various theories of voting behavior: strategic voting, informative or sincere voting, level-k or cognitive hierarchy voting, and mixtures among all of these. We show that if preferences are homogeneous, then the choice of the optimal voting rule is independent of which of these assumptions is made on voters' behavior.

Keywords: Pivotal voting, Strategic Voting, Informative Voting, Naive Voting

J.E.L. Classification: C7, D7

*We thank Nageeb Ali, Vince Crawford, Wiola Dziuda, Tom Palfrey, and Joel Sobel for helpful comments.

†Department of Economics, MIT and NBER. E-mail: costinot@mit.edu.

‡Department of Economics, Columbia University and UC San Diego. E-mail: nkartik@gmail.com.

1 Introduction

A fundamental role of voting is to aggregate private information. The "Jury Theorem" first proposed by Condorcet (1785) illustrates this idea. Consider a society which needs to select one of two mutually exclusive alternatives. All members of the society share the same preferences, but have different information regarding which of the two alternatives is better. The Jury Theorem states that under majority rule (and some other conditions), groups of voters are more likely to take the correct decision than any single individual; moreover, as the number of voters grows, the probability of the collective decision being correct approaches one.

In its original statistical formulation, this result and subsequent extensions (e.g. Berg 1993; Ladha, 1992) implicitly rely on one important behavioral assumption: voters will mechanically reveal their private information through their vote. This has come to be known as *informative voting*. More recently, starting with the seminal work of Austen-Smith and Banks (1996), Feddersen and Pesendorfer, (1996), and Myerson (1998), scholars have argued that such behavior may be inconsistent with Nash Equilibrium behavior, or *strategic voting*. In particular, a rational voter should condition on being pivotal, i.e. vote under the presumption that her vote actually matters in deciding the outcome of the election, taking into full account any information that this hypothesis reveals. Under general conditions, mechanically revealing one's information—informative voting—is often inconsistent with pivotal or strategic voting.

This suggests that different assumptions on voting behavior can lead to very different conclusions. Indeed, Feddersen and Pesendorfer (1998) show how strategic voting may have surprising implications. In the context of jury procedures, the authors find that, contrary to the standard intuition, unanimity rule for conviction may not be the optimal way of protecting the innocent. They prove the following result: when voters behave strategically and use symmetric strategies, if a jury is large enough, then a unanimous verdict requirement will be worse than some other voting rules with regards to both the probability of acquitting a guilty defendant and convicting an innocent one. Therefore, the authors suggest that in order to avoid "the terrible consequences of convicting an innocent [...], it would be better to combine a supermajority rule with a larger jury"(Feddersen and Pesendorfer, 1998, p. 31).

The objective of our paper is to offer a new perspective on the implications of informative or strategic voting. Following much of the existing literature, we focus on situations where all voters have *homogeneous preferences* under complete information; hence, the only role of voting is to aggregate information. Our point of departure is to assume that prior to the voting game, a possibly fictitious social planner or mechanism designer has to select an aggregation rule for a given group of voters. The designer's goal is to maximize the expected welfare of the voters, or more generally, society. The question we are interested in is: would different voting rules be chosen depending on the (ex-ante) beliefs about how voters would behave under each rule? In other words, do different assumptions about voters' behavior imply different optimal voting rules?

We address this question in a framework that generalizes the canonical model of the literature. Specifically, while we maintain a simple structure with a binary state of the world (Guilt or Innocence in the jury context) and binary private signals, we permit signals to be conditionally correlated across voters, and for voters to be informationally asymmetric in the sense that some voters could have more precise information than others.[1] In this setting, we show that for a variety of assumptions regarding how voters behave—including the aforementioned polar cases where either all voters vote strategically or all voters vote informatively—*a benevolent mechanism designer would (uniquely) choose the same voting rule*. In particular, this demonstrates that even though strategic voting can dramatically affect voters' behavior under an arbitrarily chosen voting rule, it does not generate different implications for institutional design relative to the traditional presumption of informative voting.

The logic behind our result is simple. It is well-known that there are no direct incentive conflicts in a setting with homogeneous preferences (McLennan, 1998). Thus, if voting informatively under some voting rule achieves full information aggregation—i.e. for every profile of private information, the outcome is the same as what would be optimal under complete information—it is an equilibrium for all voters to vote informatively in a voting game where all voters are fully strategic and the designer has chosen the above voting rule (Austen-Smith and Banks, 1996).

We build on this insight by making three observations. First, full information aggregation is feasible in the environments usually considered.[2] Second, while previous authors have focussed on the cases where either all voters are naive or there is common knowledge of strategic voting, these polar cases are unnecessarily restrictive. In particular, we allow for a range of uncertainty, both on the part of the mechanism designer and the voters themselves, about whether any voter is non-strategic or strategic, and also consider the possibility of bounded-rationality in the form of *level-k* thinkers of hetero-

[1] Other extensions, such as relaxing the binary structure of private information (even to a continuum, as in Duggan and Martinelli, 2001), also are possible; we discuss these in the conclusion.

[2] In general, what this requires is that the vote space for each voter be at least as large as her signal space; as already noted, the focus of literature is on a binary signal and vote space.

geneous sophistication (Stahl and Wilson, 1995; Nagel, 1995). Finally, we show that in our informational environment, the *only* way full information aggregation is achieved is if the designer chooses a unique voting rule and all voters vote informatively—without one or the other, full information aggregation fails. Our results directly follow from these observations.

To sum up, the main contribution of this paper is to offer a new perspective on the importance of strategic considerations in voting. By taking a mechanism design approach, we show that, under the prevalent assumption of homogeneous preferences, differences between various assumptions on voting behavior are irrelevant. From a technical perspective, our result is elementary: the proof merely consists of recognizing that, under mild conditions, there is a unique way to obtain full information aggregation. Nevertheless, it has the strong implication that for game-theoretic insights to have interesting implications for the design of optimal voting rules, heterogeneous and/or uncertain preferences are key (e.g., Chwe, 1999; Gerardi, 2001).

Before proceeding to the formal analysis, we wish to point out that in independent and simultaneous work, Ben-Yashar and Milchtaich (2007) have derived some related results.[3] Their focus, however, is on *anonymous* voting rules, where the identity of voters cannot be taken into account. The full generality of our main result makes essential use of non-anonymous voting rules, and we identify broader conditions under which there is a unique optimal rule, while also considering richer specifications of strategic sophistication.[4]

2 Theoretical Framework

For expositional purposes, we maintain the terminology of jury decision making, although it should be clear that the model has applications to various committees, boards, and other group decision problems where voting is the means of aggregating private information.

Voting Environment. There is a set $\mathcal{N} = \{1, \ldots, N\}$ of $N \geq 1$ jurors who must decide to either acquit or convict a defendant. The defendant is either guilty (G) or innocent (I). Jurors are unsure whether the defendant is guilty or not: the true state of the world $\omega \in \{G, I\}$ is unobservable. There is a common prior over the states.

[3] We were made aware of their paper after our research was completed.

[4] In particular, we permit conditional correlation of voters' signals and uncertainty about how strategically sophisticated any particular voter is. On the other hand, we do not consider the restriction to anonymous voting rules nor abstention.

Each juror gets a private signal that may reveal some information about the true state. Specifically, a juror j gets a binary signal, $s_j \in \{i, g\}$, where i indicates a signal for innocence and g indicates a signal for guilty. Denote the profile of signals for all jurors by $s = (s_j)_{j=1}^n$; as usual, we use s_{-j} to denote the profile of signals for all voters except j. We assume that the distribution of signals has full support, i.e. there is positive probability that each signal profile is realized.

After privately observing her signal, each juror casts a vote to either acquit or convict the defendant. The profile of N votes determines a collective outcome of either conviction (C) or acquittal (A), via a voting or aggregation rule. We describe a voting rule, V, as a collection of distinct subsets of $\mathcal{N}$, such that the collective outcome is C if and only if the set of voters who vote to convict is some $X \in V$.[5] We restrict attention to *monotonic voting rules*, where if $\emptyset \neq X \in V$ and $X \subseteq Y$, then $Y \in V$. Note that a voting rule, V, need not treat all jurors symmetrically or anonymously. For example, juror 1 is a dictator if $V = \{X : 1 \in X\}$. The simple anonymous majority rule is the case of $V = \{X : |X| > \frac{N}{2}\}$; unanimity rule for conviction is the case of $V = \{\mathcal{N}\}$; $V = \emptyset$ means that regardless of the vote profile, the outcome is acquittal; and so forth.

Preferences. Our focus is on situations where jurors have *homogeneous preferences* over the collective outcome and the true state of the world, in the sense that were all information about the state public, there would be no disagreement among voters about the optimal decision. Specifically, we assume there exists a parameter $q \in (0, 1)$ such that every voter's preferences can be represented by the following utility function: $u(A, I) = u(C, G) = 0$; $u(C, I) = -q$; and $u(A, G) = -(1 - q)$.[6] Thus, if a guilty defendant is convicted or an innocent defendant is acquitted, the payoff to any voter is 0. The payoffs to the two possible erroneous decisions are $-q$ and $-(1 - q)$ respectively. The parameter q can be interpreted as a threshold of reasonable doubt: society wants to convict (resp. acquit) the defendant if the probability of guilt is greater (resp. lower) than q.[7]

[5] This precludes aggregation rules that involve randomness given a profile of votes. However, as will be clear from the subsequent analysis, our conclusions would not change even if we permitted such rules, because an optimal voting rule will never involve randomness in our setting.

[6] While we assume it for simplicity, it is not important for our results that that the parameter q be identical across all votes. We could permit an idiosyncratic q^j for each voter j, so long as all q^j's are "sufficiently close".

[7] If the probability of guilt is p, the expected utility from conviction is $-q(1-p)$, whereas the expected utility from acquittal is $-(1 - q)p$. Conviction is strictly preferred if $-q(1 - p) > -(1 - q)p$, which is equivalent to $p > q$.

Information. To describe the information structure, some notation is helpful. Let $\beta(X)$ be the posterior on $\omega = G$ when the set $X \subseteq \mathcal{N}$ of voters receive guilty signals, and all voters in $\mathcal{N} \setminus X$ receive innocent signals. Given a profile of signals, s, let $\Gamma(s)$ denote the set of voters who have guilty signals in s. Let $D(s)$ denote a socially optimal decision if the signal profile is s, i.e. $D(s) = C$ if $\beta(\Gamma(s)) > q$ and $D(s) = A$ if $\beta(\Gamma(s)) < q$. We make the following minimal assumptions.

A1. [*Monotonicity*] if $X \subseteq Y$ then $\beta(X) \leq \beta(Y)$.

A2. [*Non-Degeneracy*] $\beta(\mathcal{N}) > q > \beta(\emptyset)$.

A3. [*Genericity*] For every $X \subseteq \mathcal{N}$, $\beta(X) \neq q$.

A4. [*Relevance*] For every $j \in \mathcal{N}$, there exists some s_{-j} such that $D(g, s_{-j}) = C$ whereas $D(i, s_{-j}) = A$.

(A1) embodies a natural monotonicity condition: holding all other signals constant, if a juror receives a guilty signal rather than an innocent signal, the (common) posterior on the defendant's guilt should not decrease. (A2) simply rules out situations where jurors prefer to always acquit or always convict, regardless of the profile of signals; in such cases there is no information aggregation problem. (A3) is a genericity assumption: it says that regardless of the profile of signals, the posterior is never exactly q. Since there are only a finite number of signal profiles, it is clear that this must hold for almost all prior/signal distribution combinations. Lastly, (A4) requires that each juror be informationally relevant in the sense that for *some* profile of signals of the other jurors, the collective preference for conviction or acquittal (under complete information) should be responsive to her signal. We find this reasonable because a juror who is not relevant in this sense can simply be dropped from the jury altogether.

All of these assumptions are satisfied by Austen-Smith and Banks (1996), Feddersen and Pesendorfer (1998), and other papers in the jury literature: typically, jurors are ex-ante symmetric and receive signals that are independent conditional on the state of the world. In contrast, we permit conditional correlation of signals. This is appealing because it is often reasonable that jurors form their individual opinions having observed at least some of the same information. Moreover, our assumptions allow voters to be asymmetric with regards to the precision of their signals: it could be that whenever voter 1 receives a guilty signal, society's preferred decision no longer depends on the signals of all the remaining $N - 1$ voters. Such asymmetries apply to many committee contexts where, while all members have useful information, some members have far greater expertise than others, reflected by the precision of their private information.

Behavior. The traditional, non-game-theoretic approach is to assume that a voter simply votes informatively, i.e. votes to convict if she observes $s_j = g$ and votes to acquit if she observes $s_j = i$.[8] The standard game-theoretic alternative is to assume common knowledge of fully strategic behavior. We define a game that spans these polar cases.

After the state of the world ω and private signals s_j have been determined, Nature draws a type for each voter, $t_j \in T \equiv \{naive, strategic\}$, resulting in a profile $t = (t_j)_{j=1}^n$. While *strategic* voters aim to maximize their utility taking as given the strategy of the other players, *naive* voters simply vote informatively, as assumed in the early Condorcet literature. We assume a common prior, τ, of an arbitrary probability distribution over the set of profiles, T^N.[9] Each voter j only observes his type t_j.

In this Bayesian game, a (possibly mixed) strategy for a voter j can be represented by a pair of mappings from signals to probability of convicting, one for the case her type is $t_j = naive$ and one for the case her type is $t_j = strategic$. Because of our interpretation of $t_j = naive$, we can simplify matters and represent a strategy for voter j as a single mapping $\sigma_j : \{i, g\} \to [0, 1]$, which maps her signal into a probability of voting to convict when $t_j = strategic$. When $t_j = naive$, we require the voter to vote informatively, i.e. use the mapping σ^{inf} where $\sigma^{inf}(i) = 0$ and $\sigma^{inf}(g) = 1$. Accordingly, when referring to a voter's strategy hereafter, we refer to her play when strategic. A profile of strategies is denoted by $\sigma = (\sigma_j)_{j=1}^N$. An equilibrium refers to a Bayesian Nash equilibrium (BNE) of this game. As is common in voting games, there will often by multiple equilibria; we will assume below that the best possible BNE (maximizing the utility of all players) is played.[10] Pareto dominance seems a reasonable selection criterion, particularly since preferences are homogeneous.

---

[8] Throughout, our focus is on comparing strategic voting with informative voting, which we refer to as naive behavior. A different kind of naivety would be to always vote *sincerely*, where a juror votes for the best alternative conditioning only on her own signal and the prior. It is common to assume that sincere and informative voting coincide, i.e. if there were only one juror, then she should convict if she receives a guilty signal, and acquit if she receives an acquit signal. With this assumption, all references to informative voting in what follows can be replaced with sincere voting.

[9] As will be clear from the analysis, this can also be weakened quite a bit, because beliefs and higher order beliefs about t will play no role in our conclusions.

[10] We use the terminology "the best possible" somewhat loosely: if there is more than one BNE that maximizes utility, any of them can be considered.

3 Optimal Voting Rules

Suppose that a fictitious mechanism designer is considering the choice of a voting rule to maximize ex-ante expected utility (since agents have common preferences, the problem is unambiguous). The question she must consider is how the jurors will behave given a chosen voting rule. We can capture this in our framework through the probability distribution τ on the set T^N. For example, if the mechanism designer believes that all voters are strategic, then he believes that voters will play a game where τ puts all mass on the profile $t = \{strategic, \ldots, strategic\}$. Similarly, if he believes that all voters are naive, this is the case of τ being degenerate on $t = \{naive, \ldots, naive\}$. More generally, the framework allows for each voter to be naive with some independent probability, or even for there to be some correlation in the strategic sophistication of (an arbitrary subset of) voters, etc.[11]

The question we are interested in is: does it matter what beliefs the designer has about voters' behavior? Formally, we formulate the question as follows. For any type distribution τ and any aggregation rule, V, let $\mathcal{U}(V, \tau)$ denoted the ex-ante expected utility in the game defined by V and τ when voters play an optimal BNE of this game. That is,

$$\mathcal{U}(V, \tau) = \max_{\sigma} EU(V, \sigma; \tau) \text{ subject to } \sigma \text{ being a BNE given } V \text{ and } \tau,$$

where $EU(V, \sigma; \tau)$ is the ex-ante expected utility when the voting rule is V, the type distribution is τ, and voters play the profile σ (recall, this means that voter j plays σ_j if strategic, and votes informatively if not). Given his belief, τ, the mechanism designer's problem is to choose an aggregation rule, V, to solve

$$\max_{V} \mathcal{U}(V, \tau). \tag{1}$$

Among others, Austen-Smith and Banks (1996) and Feddersen and Pesendorfer (1998) have shown that for arbitrary voting rules, strategic jurors' behavior at an (optimal) BNE may be quite different from the behavior anticipated by a designer who believes that they are all naive. In our language, their point is that if $\tau \neq \tau'$, then for arbitrarily chosen V, $\mathcal{U}(V, \tau) \neq \mathcal{U}(V, \tau')$. Accordingly, the maximization program in (1) certainly changes as τ is varied. However, this does not imply that designers with different views about τ would select distinct voting rules, as we now demonstrate.

[11] For example, perhaps there is some identifiable characteristic that determines strategic sophistication, such as education.

Theorem 1. *There exists V^* such that, for any τ, the unique solution to maximization problem* (1) *is V^*.*

The intuition is simple. First, by choosing the voting rule appropriately, full information aggregation can be achieved when all voters vote informatively: this defines V^*. Second, if the designer chooses V^*, then because naive voters are voting informatively, the unique optimal BNE consists of all strategic voters voting informatively as well. Third, for any τ, no other voting rule can achieve full information aggregation, regardless of how voters behave.

Formally, Theorem 1 is proved by the following three Lemmas.

Lemma 1. *There exists a unique voting rule, V^*, such that if all signals were public, the socially optimal decision is to convict if and only if $\Gamma(s) \in V^*$. Moreover, $\emptyset \notin V^*$ and $\mathcal{N} \in V^*$.*

Proof. Immediate consequence of Assumptions (A1)–(A3). □

Lemma 2. *For any τ, if the voting rule is V^*, it is a BNE for all jurors to vote informatively, achieving the first-best outcome for every signal profile realization.*

Proof. By the definition of V^*, if all jurors vote informatively under V^*, first-best outcomes are achieved for every signal profile realization. Since jurors have homogeneous preferences, no juror has a profitable deviation given this strategy profile. □

Lemma 3. *Suppose $V \neq V^*$. Then, regardless of the strategy profile played, there exists a profile of signals that occurs with positive probability for which the socially optimal outcome is not attained with positive probability.*

Proof. Fix a $V \neq V^*$, and consider any strategy profile, σ. It suffices to show that there is a signal profile for which the socially optimal outcome is not attained with positive probability, since all signals profiles have positive probability of being realized. There are two cases:

1. Every strategic voter votes informatively. Then all voters are voting informatively. There are two possibilities: either there exists an $X \in V$ such that $X \notin V^*$, or there exists an $X \in V^*$ such that $X \notin V$. In either case, the signal profile s where $\Gamma(s) = X$ results in an outcome—conviction in the first case and acquittal in the second—that is not socially optimal given s.

2. Not every strategic voter votes informatively. Let n be a voter who is not voting informatively (when strategic). Let s_{-n} be a profile such that $D(g, s_{-n}) = C$ and $D(i, s_{-n}) = A$; this exists by Assumption (A4). If V and σ_{-n} are such that given s_{-n}, there is positive probability that the outcome is independent of n's vote, then since the socially optimal decision given s_{-n} depends non-trivially on s_n, there is positive probability that the socially optimal decision is not achieved for the profile $s = (s_n, s_{-n})$, for any $s_n \in \{i, g\}$. So suppose that V and σ_{-n} are such that given s_{-n}, the outcome is determined by n's vote with probability 1. Since the voting mechanism, V, is monotonic, a vote by n to convict (resp. acquit) must lead to the outcome of conviction (resp. acquittal), holding fixed s_{-n} and σ_{-n}. But because n is not voting informatively, either $\sigma_n(i) > 0$ or $\sigma_n(g) < 1$. In the first (resp. second) case, there is positive probability of not achieving the socially optimal outcome when the signal profile is $s = (i, s_{-n})$ (resp. $s = (g, s_{-n})$). □

Let us relate Theorem 1 to the previous literature. Both Austen-Smith and Banks (1996) and Mclennan (1998) have noted that a population of fully strategic voters have no incentive to deviate from a strategy profile that attains the first best. (In our notation, this is the case of the distribution τ putting point mass on the type profile $t = \{strategic, \ldots, strategic\}$.) Lemma 2 extends this logic to arbitrary τ's. In addition, Lemma 3 shows that full information aggregation cannot be achieved unless voters reveal their information and the voting rule efficiently aggregates their reports, a combination that only be attained under a unique voting rule, regardless of τ.

Communication. At this point, it is appropriate to discuss the role of communication—or lack thereof—in our model. In many contexts, such as large elections, communication between all voters may simply be infeasible; similarly, communication can be quite costly even in small group settings. Nevertheless, some institutions, such as juries, do combine communication and voting. Coughlan (2000) points out that under homogeneous preferences, a round of cheap talk prior to voting can insure efficient information aggregation under any voting rule, so long as voters are strategic. This can be interpreted as a result about the "irrelevance of voting rules" under strategic voting and communication.[12] Theorem 1 says quite the opposite: without communication, it establishes uniqueness of optimal voting rule, for a variety of possible voting behavior. One may therefore wonder whether the result unravels once communication is introduced.

While there is a conceptual question of what it means for voter to behave naively in

[12] Gerardi and Yariv (2006) prove a rather general result along these lines.

the presence of communication, a natural benchmark is to suppose that she continues to vote informatively at the voting stage. Then, it is not hard to see that the irrelevance of voting rules result is knife-edged because it relies on common knowledge of fully strategic behavior. Formally, regardless of the assumptions made on behavior in the communication stage, Theorem 1 extends for any τ that does not put probability one on $t = \{strategic, \ldots, strategic\}$. The reason is that the mere possibility of even one voter voting naively suffices for the logic behind Theorem 1 (in particular, Lemma 3) to have force, regardless of how much information sharing voters engage in during communication.

Boundedly Rational Voters. A burgeoning recent literature on strategic but non-equilibrium behavior, step-level reasoning or *level-k* models (e.g., Stahl and Wilson, 1995; Nagel, 1995; Crawford, 2003), explores the idea that players in games belong to one of a hierarchy of decision-rules or types. These types are distinguished by the level of sophistication they embody: type Lk for $k > 0$ anchors its beliefs in a non-strategic $L0$ type and adjusts them via thought-experiments with iterated best responses: $L1$ best responds to $L0$, $L2$ to $L1$ and so forth. $L1$ and all higher types are rational in the sense that they best respond to their beliefs: the departure from equilibrium is that rather than having accurate beliefs about the decision-rules of all other players, a Lk type (for $k > 0$) simply behaves as would be optimal when all other players are $L(k-1)$ types.

In our context, it is natural to specify the anchoring type, $L0$, as informative voting.[13] Let the vector $\{p_0, p_1, \ldots, \}$ denote (a designer's beliefs about) the distribution of Lk types in the population, i.e. p_k represents the probability of type Lk.

Theorem 2. *For any Lk type distribution, $\{p_0, p_1, \ldots\}$, the unique voting rule that maximizes ex-ante expected utility is V^*.*

Proof. Fix any $\{p_0, p_1, \ldots\}$. Observe that under V^*, type $L1$ plays informatively, since if all other voters are voting informatively, it is a unique best response to vote informatively under V^*. By induction, all types vote informatively. Thus, V^* achieves full information aggregation, and is an optimal voting rule. That V^* is uniquely optimal follows from the same argument as that of uniqueness in Lemma 3. □

[13] An alternative specification used in experimental literature is to have the $L0$ types play a completely random action. Our result does not extend to such a specification. However, the choice of how to best specify $L0$ is a situation-specific matter (see the discussion in Crawford and Iriberri, 2007), and at this point it is not clear which one of the two specifications is more appropriate in voting games.

The proof of the Theorem makes it clear that this result would hold under numerous variations. For example, the same argument works if there are Lk types who do not simply best respond to a belief that all others are of type $L(k-1)$, but rather, best respond to an arbitrary conjecture distribution over all $L0, ..., L(k-1)$ types in the population. This implies that Theorem 2 also holds when agents belong to a *cognitive hierarchy* (Camerer, Ho, and Chung, 2004). Similarly, a "sophisticated" type who best responds to the true population distribution of Lk types can also be accommodated.

4 Concluding Remarks

Strategic voting may differ from informative voting in complex and surprising ways under arbitrary voting mechanisms. However, our results demonstrate that the two notions may not generate different implications for the design of optimal voting rules. If a policy maker believes that the voting game satisfies the assumptions of Section 2—which subsumes the focus of the previous literature—then her recommendation for design will be completely unaffected by her knowledge (or ignorance) of game-theoretic insights. In particular, if unanimity rule is the optimal voting rule from the point of view of "naive" policy maker, who ignores strategic voting, then it is also the optimal voting rule for a "sophisticated" designer, who believes that voters will behave strategically.

Although it may appear that this conclusion contradicts Feddersen and Pesendorfer's (1998, p. 31) finding that "in large juries the unanimity rule is inferior to a variety of other rules", this is not the case. It is crucial to recognize that because they assume that the threshold for reasonable doubt, q, is strictly between 0 and 1, no institution designer will find it optimal to choose unanimity rule when the jury size is large enough, regardless of the designer's beliefs about the fraction of naive jurors. Our result says that for a *given jury size*, the choice of the optimal voting rule is invariant to beliefs about the fraction of naive and strategic jurors. Since a naive designer will not choose unanimity rule for large jury sizes, neither will a sophisticated designer. A useful implication of our results is that even when solving the joint problem of choosing both an optimal jury size and an optimal voting rule (say, when there is a fixed cost of adding an additional juror to the jury), the designer can simply assume that jurors behave naively. Given the mixed evidence on voters' actual behavior (e.g. Guarnaschelli, McKelvey, and Palfrey, 2000), this is a striking and perhaps reassuring result.

Does the main result of this paper extend to more general environments? As mentioned in the introduction, only one feature of the model is crucial: voters have homogeneous preferences.[14] As long as this condition holds, our basic insight is robust to various voting environments. In particular, the naive designer still chooses a voting rule that is also optimal for the sophisticated designer if mistrials are allowed (Coughlan, 2000), or when voters receive more than two signals (Duggan and Martinelli, 2001).[15] The reason is that in all these cases, homogeneous preferences imply that informative voting is a Nash equilibrium under the naive designer's optimal voting rule. Conversely, it is also the case that in any of these generalized settings, any optimal voting rule for the sophisticated designer induces the same mapping from signal profiles to final decisions as the optimal rule for the naive designer does. Therefore, the main message regarding the equivalence between the two designer's choices continues to apply, albeit in weaker form.

If voters have heterogeneous preferences, then the basic logic underlying the main insight of this paper no longer holds. In such a setting, the revelation principle for mechanism design still guarantees that there exists an optimal mechanism for the sophisticated designer in which all voters are asked to vote informatively—i.e. disclose their signal truthfully—and it is a Nash equilibrium for them to do so. However, this mechanism would in general *not* map the revealed signals into the ex-post optimal decision for the designer. Accordingly, it will typically be the case that a naive designer (incorrectly) expects a higher ex-ante utility under a different mechanism, viz. the one that maps revealed signals into the ex-post optimal decisions.

We believe that our results demonstrate in a very transparent manner that imposing homogeneous preferences, though a common practice in the previous literature, is an extremely strong restriction. Once it is imposed, fundamental differences in voters' behavior become irrelevant for the choice of optimal voting rules.

[14]Recall, however, from fn. 6 that a little heterogeneity in the thresholds of reasonable doubt among voters does nothing to upset the arguments presented here. When discussing non-homogeneous preferences, we mean "sufficiently heterogeneous" preferences.

[15]In this setting with more than two signals, the appropriate notion is that voters choose their vote from a set at least as large as the signal space, and that the mechanism designer chooses amongst all possible mappings from the profile of votes to final decisions.

References

[1] Austen-Smith, D. and J. Banks [1996], "Information Aggregation, Rationality, and the Condorcet Jury Theorem." *American Political Science Review,* **90-1**, pp34-45.

[2] Ben-Yashar, R. and I. Milchtaich [2007], "First and second best voting rules in committees. " *Social Choice and Welfare*, **29**, pp453-486.

[3] Berg, S. [1993], "Condorcet's Jury Theorem, Dependency among Jurors." *Social Choice and Welfare*, **10-1**, pp87-95.

[4] Camerer, C. F., T. Ho, and J. Chong [2004], "A Cognitive Hierarchy Model of Games ." *Quarterly Journal of Economics*, **119-3**, pp861-898.

[5] Chwe, M. [1999], "Minority Voting Rights Can Maximize Majority Welfare." *American Political Science Review*, **93-1**, pp85-97.

[6] Crawford, V. P. [2003], "Lying for Strategic Advantage: Rational and Boundedly Rational Misrepresentation of Intentions." *American Economic Review*, **93**, pp133-149.

[7] Crawford, V. P. and N. Iribberi [2007], "Fatal Attraction: Salience, Naivete, and Sophistication in Experimental Hide-and-Seek Games." *American Economic Review* **97-5**, pp1731-1750.

[8] Condorcet, A. [1785], *Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix*. Imprimerie Royale, Paris.

[9] Coughlan, P. [2000], "In Defense of Unanimous Jury Verdicts: Mistrials, Communication, and Strategic Voting." *American Political Science Review,* **94-2**, pp375-393.

[10] Duggan, J. and C. Martinelli [2001], "A Bayesian Model of Voting in Juries." *Games and Economic Behavior*, **37**, pp259-294.

[11] Feddersen, T. and W. Pesendorfer [1996], "The Swing Voter's Curse." *American Economic Review*, **86-3**, pp408-424.

[12] Feddersen, T. and W. Pesendorfer [1998], "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting." *American Political Science Review*, **92-1**, pp23-35.

[13] Gerardi D. [2000], "Jury Verdicts and Preference Diversity." *American Political Science Review,* **94-2**, pp395-406.

[14] Gerardi D. and L. Yariv [2006], "Deliberative Voting." *Journal of Economic Theory,* forthcoming.

[15] Guarnaschelli, S., McKelvey, R. and T. Palfrey [2000], "An Experimental Study of Jury Decision Rules." *American Political Science Review*, **94-2**, pp407-23.

[16] Ladha, K. [1992], "The Condorcet Jury Theorem, Free Speech, and Correlated Votes." *American Journal of Political Science*, **36**, pp617-634.

[17] McLennan, A. [1998], "Consequences of the Condorcet Jury Theorem for Beneficial Information Aggregation by Rational Agents." *American Political Science Review*, **92-2**, pp413-418.

[18] Myerson, R. [1998], "Extended Poisson Games and the Condorcet Jury Theorem." *Games and Economic Behavior*, **25-1**, pp111-131.

[19] Nagel, R. [1995], "Unravelling in Guessing Games: An Experimental Study." *American Economic Review*, **85**, pp1313-1326.

[20] Stahl, D. O. and P. W. Wilson [1995], "On Players' Models of Other Players: Theory and Experimental Evidence." *Games and Economic Behavior*, **10-1**, pp218-254.