Deliberative democracy or market democracy: Designing institutions to aggregate preferences and information

Adam Meirowitz[1]

Dept. of Politics

Princeton University

ameirowi@princeton.edu

March 23, 2005

I consider the design of policy making institutions to aggregate preferences and information. The mechanism design approach makes it possible to consider a large set of institutions or game forms in which participants take observable actions prior to voting. A pervasive incentive problem is found; participants that expect to have the minority preference type will have an incentive to misrepresent their information. Consequentially, if some policy relevant information is observed by fewer than three individual participants and ideological types are not highly correlated no institution can fully aggregate the information and preferences without distributing transfers. Contrary to conventional wisdom, diversity may hurt deliberation as the incentives for information transmission are worse in groups with heterogenous sources of information or low levels of ideological correlation. Institutions that distribute transfers conditional on either the validity of agent reports of facts (like information markets), or the frequency of each type of report (like clubs) can truthfully implement the full information majority rule core policy. Overall, expectations of full information and preference aggregation with strategic participants require either strong correlation of preferences, the presence of external interests to structure incentives or information structures in which each piece of information is observed by several participants.

[1]I appreciate comments from John Duggan, Mark Fey and Roger Myerson on earlier work that motivated this paper. In addition I benefited from discussions with and comments from Scott Ashworth, David Baron, Jon Bendor, Josh Clinton, Keith Krehbiel, Nolan McCarty, Tom Palfrey, Kris Ramsay and Tom Romer. I thank Raymond Hicks for his assistance.

1 Introduction

> There are again two methods of removing the causes of faction: the one, by destroying the liberty..., the other, by giving every citizen the same opinions, the same passions, and the same interests. .... The second expedient is as impractical as the first would be unwise. As long as the reason of man continues fallible, and he is at liberty to exercise it, different opinions will be formed. As long as the connection subsists between his reason and his self-love, his opinions and his passions will have a reciprocal influence on each other; and the former will be objects to which the latter will attach themselves. The diversity in the faculties of men, from which the rights of property originate, is not less an insuperable obstacle to a uniformity of interests.... The latent causes of faction are thus sown in the nature of man; and I see them everywhere brought into different degrees of activity, according to the different circumstances of civil society. A zeal for different opinions concerning religion, concerning government, and many other points, as well as speculation as of practice; an attachment to different leaders ambitiously contending for pre-eminence and power; or to persons of other descriptions whose fortunes have been interesting to the human passions, have, in turn, divided mankind into parties, inflamed them with mutual animosity, and rendered them much more disposed to vex and oppress each other than to cooperate for their common good. (James Madison 1788 , Federalist 10, p. 131-32).

Consider a collective choice problem in which the members of a townhouse community must choose between keeping their snow removal service with the incumbent firm or contracting with a rival firm. The cost for the service is passed on to members of the community. Given the uncertainty about snowfall in the upcoming year contracts with either snow removal service specify a price per visit, and a maximum amount of snow that may remain on the individual driveways before the removal service comes out to shovel. While all firms contract for the same maximum snow depth and price per visit there is slippage, some firms have high capacity and come out as much as possible, while others are less aggressive, allowing more snow to build up before coming out to shovel. The incumbent removal service's frequency of visits has been observed over the last few years, but the rival firm's level of aggressiveness is not known by the community members. Some residents in the community don't mind snow on their driveways and are price sensitive – preferring less frequent visits. Other residents in the community dislike snow on their driveways and are price insensitive – preferring a service that makes more visits. Moreover, the members of the community are not particularly close so members are uncertain about the price sensitivities of their neighbors. Following the lead of contemporary scholars of deliberative democracy (e.g., Fishkin, 1991, Gutmann and Thompson, 1996) a deliberative meeting is held and then the following week the decision is made via referendum. It would be satisfying to conclude that with a free and open exchange any private information that the community members have about the competing firm will be shared and the resulting decision will make a majority of the community members happy.[2] Is this conclusion justified in the example? Suppose Mr. Smith stands up during the board meeting and announces that he has a friend, Mr. Jones, in a neighboring townhouse community that is serviced by the rival firm in question and that this friend reports, "The service is excellent. The snow is frequently removed. In big storms they come by three times a day." What should the other members of the community take away from Smith's speech? Should they trust Smith? If Smith believes that he will be trusted might he have an incentive to lie?

In this paper, we focus on the simplest collective choice settings in which both private information and preferences are to be aggregated into a policy decision. These are problems in which participants possess private information about the consequences of a policy decision as well as uncertainty about how their private information would affect other members' preferences over policy. The analysis offers an explanation for why behavior in deliberative institutions might not meet the expectations of contemporary scholars of deliberative democracy. Returning to the case of Mr. Smith, suppose he is price sensitive, and believes that it is likely that the majority of community members are not price sensitive. After being persuaded by his friend that the rival firm would provide fewer visits than the incumbent firm, Mr. Smith would gain from convincing his neighbors that the rival firm would provide more visits (say by giving the speech quoted above). This suggests that the value of the information that Smith provides might depend on the extent to which he believes his preferences are aligned with a majority. But if this is the case then another community member,

[2]Much of the literature on deliberative democracy seems to make an even stronger claim: through discourse the community members will reach a consensus on the decision that serves the common good.

unaware of Smith's price sensitivity might be hard-pressed to take Smith at his word. Now, if several other members of Smith's community are on the same bowling team as Smith's friend Jones then things may be different. In this case it is possible that Smith's dishonesty would be pointed out. Alternatively, if Smith could be made to care about the accuracy of his predictions then the incentive to be right may swamp the incentive to affect policy. This paper addresses the incentive problem illustrated in the townhouse example and isolates the types of institutions that can overcome this problem.

If it is possible for participants to believe that they are likely to have a minority interest, participants care only about the policy outcome from the collective choice body, and a condition known as strong nonexclusivity of information is violated, then there are no institutions which use all of the available information to select the policy that a majority of the participants would favor if the information available to the individuals in the group were public.[3] However, if all types of participants believe that they possess the majority state-contingent preference type (as in the case of strongly correlated ideologies) or strong nonexclusivity is satisfied, then there exist simple institutions (including deliberation) with equilibria that fully aggregate preferences and information. As such, expectations that deliberation will reach desirable policies with strategic participants would seem to require either the absence of exclusively observed information, or a common belief among all participants that they are likely to possess the majority preference type. This conclusion stands in contrast to the conventional wisdom of contemporary deliberative democrats that heterogeneity of experience and diversity of opinion help deliberation.[4]

When the incentive problem is present, it is still possible to design institutions to aggregate preferences and information, but external incentives or transfers are needed. A leading

[3]Informally, strong nonexclusivity requires that all pieces of relevant information are observed by at least three distinct agents. Baron and Meirowitz (2004) introduce the concept of strong nonexclusivity in considering signalling and screening games. For early uses of the slightly weaker concept of nonexclusivity in the mechanism design literature see Postlewaite and Schmeidler (1986) and Palfrey and Srivastava (1987).

[4]One positive note for deliberative democrats is that a positive explanation for information sharing does not require common conceptions of the good. The relevant condition is weaker, requiring only that everyone believe their conception of the good is in agreement with the majority conception of the good.

example of this type of institution involves voting after a market for suitably defined Arrow (state-contingent) securities has cleared. While the idea of legislators, policy experts or members of a townhouse community trading securities prior to policy selection may seem far-fetched, institutions of this form are beginning to surface. Recently, Hewlett-Packard has tested inhouse information markets and administered scoring rules, which are mechanisms that condition payments on a respondent's predictive accuracy, to generate better sales forecasts. Chen and Plott (2002) present evidence on the markets and Chen, et. al. (2003) present evidence on the scoring rules. Scholars at the University of Iowa have been administering prediction markets for the influenza risk in Iowa, which can be used to make decisions about the vaccination rate for at-risk populations and staffing decisions for hospitals. Neumann (2005) discusses the results for the 2004-2005 influenza season market.[5] A novel finding of the institutional analysis with transfers is the conclusion that the transfers do not need to be conditioned on the actual state (or accuracy of participant messages). Instead, it is sufficient to create incentives for participants to forecast the private information held by other participants. This insight can be quite useful in policy areas in which the state is not observed until the distant future, if ever. An example of an institution of this form involves offering participants the opportunity to choose between exclusive clubs and then distributing rents to club members based on the size of the club.

While deliberative democracy is a central piece of the normative theory literature, two strands of literature develop positive theories of how participants behave in deliberative institutions. The first draws on the traditions of political psychology and political behavior and the second draws on the traditions of noncooperative game theory.[6] Work in the behavioral literature assumes that communication is honest. For example in recent work, Barabas (2004) treats deliberation as an opportunity for political communication and focuses

[5]The well publicized market on terrorism futures proposed by the Defence Advanced Research Projects Agency (DARPA) in the summer of 2003 suggests how information markets might be incorporated in a national policy making setting. This example also demonstrates some difficulties associated with initiating institutions of this form. Meirowitz and Tucker (2004) present an account of the proposal and the politics behind its demise. Wolfers and Zitzewits (2004) present a broad review of recent information markets.

[6]See Mendelberg (2002) for a recent review of the normative and behavioral literatures.

on understanding how differences in signal quality and prior precision affect posterior beliefs. In work of this kind an unstated but critical assumption is that participants freely and honestly reveal their private information to the group. Drawing on cheap talk signalling games, the game theoretic tradition also considers deliberation as a forum for information transmission/sharing. Here too, belief updating is an important feature of the analysis, but unlike the behavioral work it does not treat the precision, accuracy or value of communication as exogenous. Whereas Barabas (2004) treats the precision of information as a key explanatory variable in his simulation model, in the game theoretic work the value of communication is endogenous and shown to depend on the rules used to aggregate votes and the size of the group. Specifically, Austen-Smith and Feddersen (2003a,b) show that unanimity rule induces less information sharing than majority rule, and Gerardi and Yariv (2003) establish an equivalence between the amount of information aggregation in the sequential equilibrium sets for communication games under all simple rules other than unanimity. Meirowitz (2003) shows that in models in which private preference types are correlated, equilibria that reach the full information majority rule core policy are more likely to exist for small groups than for large groups.[7] In addition, Lipman and Seppi (1995), Glazer and Rubinstein (2001) and Hafer and Landa (2003) consider models of communication that do not involve cheap talk.

The remainder of the paper is organized as follows. In section 2, I present the primitives of the class of collective choice problems and highlight the relevant mechanism design results that will be used in the subsequent analysis. In section 3 the possibility of aggregating preferences and information in institutions that do not involve transfers is investigated. In section 4, I consider institutions that involve transfers. Section 5 concludes with a discussion.

[7] In addition the literature on the strategic Condorcet Jury Theorem (e.g., Austen-Smith and Banks, 1996; Feddersen and Pessendorfer, 1998; Witt, 1998; Duggan and Martinelli 2001; Meirowitz, 2002) has connections with the formal work on deliberation. These papers consider common values problems and no communication. Coughlan (2000) demonstrated that cheap talk communication prior to voting can improve the aggregation properties of equilibria in a common values setting, and Kim (2004) considers jury voting when preferences can be diametrically opposed, but without communicaiton.

2 A model of collective choice with private beliefs and values

2.1 The primitives

I consider a class of simple collective choice problems with asymmetric information about preferences and policy relevant facts. In more complicated problems with larger policy and state spaces even stronger assumptions are needed to rule out the incentive problem isolated in these simplest of problems.[8] A collective choice problem involves a collective of agents $N = \{1, 2, ..., n\}$ (n odd) that must make a binary group decision, choosing a policy $p \in P = \{a, b\}$. Each agent has a binary preference type $\theta_i \in \Theta = \{-1, 1\}$ and there is an unknown state of the world $x \in X = \{a, b\}$. So in the snow removal example, policy a corresponds to retaining the incumbent and policy b to replacing the incumbent with the rival firm. State a corresponds to the world in which the rival firm provides fewer removal trips (and costs less) and state b corresponds to the world in which the rival firm provides more trips. A participant with type $\theta_i = 1$ has a low price sensitivity (and thus demands more trips) and a type -1 participant has a high price sensitivity. Agent i has a Von Neumann-Morgenstern utility function over lotteries. Her Bernoulli (state-contingent) utility function depends on the policy $p \in P$, the state $x \in X$, and her type $\theta_i \in \Theta$ and is of the form

$$u_i(p; x, \theta_i) = \theta_i \cdot 1_{\{p=x\}} \tag{1}$$

where $1_{\{p=x\}}$ is the indicator function taking value 1 if $p = x$ and 0 otherwise. So type $\theta_1 = 1$ agents want to match x and p while type $\theta_i = -1$ agents want x and p unmatched.[9] The potential for agents of both types captures a stark case of potentially opposed preferences over state contingent policies (as in the case of a community consisting of both high and low price sensitive participants). In addition to the motivating snow removal example, 3 other

[8] See Meirowitz (2005) for work on deliberation and bargaining in the spatial model.

[9] We use the same names for the possible states, x, and policies, p, to make it easy to keep track of the preference types. Type 1 participants want to match the policy and state, while type -1 participants don't want to match.

examples demonstrate some of the types of choice problems that have this structure.

Example 1: The spatial model with a noisy alternatives to the status quo. Suppose that all participants have symmetric single-peaked preferences on the real line and that type $\theta_i = -1$ participants have ideal point -1 and type $\theta_i = 1$ participants have ideal point 1. Let policy a correspond to the status quo which results in the outcome $\frac{1}{8}$ with certainty and let policy b result in an alternative policy which involves some uncertainty. If $x = a$ then the outcome that results from policy b is $-\frac{1}{4}$ and if $x = b$ then the outcome that results from policy b is $\frac{3}{8}$. In this case if $x = a$ then type $\theta_i = -1$ participants prefer policy b and if $x = b$ then type $\theta_i = -1$ participants prefer policy a. While, if $x = a$ then type $\theta_i = 1$ participants prefer policy a and if $x = b$ then type $\theta_i = 1$ participants prefer policy b.[10]

Example 2: Group decision making with heterogenous risk attitudes. Consider two policies a and b and a corporate board that must decide which policy to adopt. Suppose that both policies involve risk but participants do not know which policy actually has the higher risk and expected return and which policy has the lower risk and lower expected return. Let x denote the identity of the lower risk and lower return project. A board member with type $\theta_i = 1$ is very risk averse while a $\theta_i = -1$ type board member is risk neutral.

Example 3: Fishkin's deliberative poll with saboteurs. Consider a party primary with 2 viable candidates. Let p denote the identity of the party nominee for the general election, while x is the identity of the party nominee that general election voters are more likely to prefer to the incumbent. In this setting $\theta_i = 1$ is a preference type that wants to unseat the incumbent while $\theta_i = -1$ is a preference type that wants to see the incumbent retained. The possibility of both preference types participating is a risk in a caucus or deliberative poll without a strong screening technology.

[10]Similar examples can be constructed in an arbitrary dimensional spatial model, by assuming that a is a policy with little uncertainty (say the status quo) and b is a policy that involves more uncertainty, with the state x determining whether the outcome associated with policy b is preferred to a for type 1 or type -1 participants.

In order to complete the description of the choice problem we must characterize the informational environment. The common prior probability distribution over the unknown state is given by $pr(x = a) = \pi \geq \frac{1}{2}$. To capture situations in which agents are uncertain about the preferences of the other members of the deliberative body, I assume that only agent i knows her type and that there is common knowledge about the random process that generates the profile $\theta = (\theta_1, ..., \theta_n)$. A convenient probability model to keep in mind involves θ_i's being drawn from the binomial distribution with $pr(\theta_i = 1) = z \in (\frac{1}{2}, 1]$. However, there is no need to rule out the possibility of some correlation in types. An example of a probability model of this form involves a mixture model in which with probability $1 - c$ the types are generated as in the binomial example and with probability c nature randomizes between giving all participants $\theta_i = 1$ with probability z and all types $\theta_i = -1$ with probability $1 - z$. Returning to the snow removal problem, under the first probability model, Smith's knowledge that he is a high price sensitivity type would not affect his beliefs about the types of the other participants, whereas in the second probability model this would not be true. In the remainder of the paper, I am agnostic about the process generating θ except for the symmetry assumption that the θ_i's are identically (but not necessarily independently) distributed. For $t \in \{-1, 1\}$, let η_t^+ denote the probability that at least $\frac{n+1}{2}$ of the $N \backslash i$ participants have type $\theta_j = t$ conditional on $\theta_i = t$. Let η_t^- denote the probability that at least $\frac{n+1}{2}$ of the $N \backslash i$ participants have type $\theta_j = -t$ conditional on $\theta_i = t$. We assume that $\eta_1^+ > \eta_{-1}^+$ and $\eta_{-1}^- > \eta_1^-$, to capture the case where $\theta_i = 1$ is more likely than $\theta_i = -1$. For the remainder of the analysis, the probabilities $(\eta_1^+, \eta_{-1}^+, \eta_1^-, \eta_{-1}^-)$ will be sufficient summaries of the joint distribution of θ.[11]

To capture the case of private information about the state, I assume that agents do not observe x, but instead each agent receives an informative private signal $s_i \in \{a, b\}$ about x. For now assume that these signals are independent conditional on x, with $pr(s_i = x) = g > \frac{1}{2}$. This assumption means that information violates nonexclusivity - there are

[11]In the binomial example with $z > \frac{1}{2}$, $\eta_t^+ > \eta_t^-$ for $t \in \{-1, 1\}$. Meirowitz (2004) shows that in the mixture example $\eta_t^+ > \eta_t^-$ for $t \in \{-1, 1\}$ if $c \geq \frac{1}{2}$ or $z = \frac{1}{2}$ but if $c < \frac{1}{2}$ and n is sufficiently big then $\eta_{-1}^+ < \eta_{-1}^-$.

pieces of information observed by only one agent. In a subsequent section the relevance of this assumption is demonstrated when the possibility of perfect conditional correlations is considered. I assume that the generation of agent types is independent of both x and $s = (s_1, ..., s_n)$. The above describes a lottery over the space $\Omega := \{a, b\} \times \{-1, 1\}^n \times \{a, b\}^n$. The first $n + 1$ dimensions represent payoff relevant information and the last n dimensions represent the imperfect signals that agents learn. In this setting each agent's type is a double $\phi_i = (\theta_i, s_i) \in \Phi = \{-1, 1\} \times \{a, b\}$. I sometimes use the notation $\phi = (\phi_1, ..., \phi_n)$ and $\Phi^n = \times_{i=1}^n \Phi$. It is convenient to let s_{-i}, θ_{-i} and ϕ_{-i} denote the appropriate vectors of values for $N \backslash i$. Finally I assume that utility is transferable, in the sense that participants have separable preferences over policy and a transferable resource,

$$U_i(p, x, t_i, \theta_i) = u_i(p; x, \theta_i) + t_i. \tag{2}$$

2.2 Revelation principles

2.2.1 Direct Mechanisms

In this setting a mechanism consists of action spaces H_i for each participant, a policy mapping $p : \times_{i=1}^n H_i \rightarrow [0, 1]$ and n transfer mappings $t_i : \times_{i=1}^n H_i \rightarrow \mathbb{R}$. The value $p(h)$ is the probability that a is chosen when actions h are chosen. A choice function $c : \Phi^n \rightarrow [0, 1]$ is a rule that selects a lottery over policy for each realization of ϕ. Mechanism design is the study of whether careful selection of the mechanism can result in equilibrium play that results in the relationship between ϕ and the policy choice specified by desirable choice functions .

Definition 1 *A choice function $c(\cdot)$ is Bayesian Nash implementable with transfers if there exists a mechanism $\langle H_1, ..., H_n, p(\cdot), t_1(\cdot), ..., t_n(\cdot)\rangle$ with a Bayesian Nash equilibrium $h_1^*(\cdot), ..., h_n^*(\cdot)$ to the mechanism in which for every $\phi \in \Phi^n$, $p(h_1(\phi_1), ..., h_n^*(\phi_n)) = c(\phi_1, ..., \phi_n)$. A choice function $c(\cdot)$ is Bayesian Nash implementable without transfers if there exists a mechanism $\langle H_1, ..., H_n, p(\cdot), t_1(\cdot), ..., t_n(\cdot)\rangle$ with transfer schedules that are identically 0, $(t_i(\phi_i) = 0$,*

$\forall \phi_i \forall i$) *in which there is a Bayesian Nash equilibrium* $h_1^*(\cdot), ..., h_n^*(\cdot)$ *to the mechanism in which for every* $\phi \in \Phi^n$, $p(h_1^*(\phi_1), ...ha_n^*(\phi_n)) = c(\phi_1, ..., \phi_n)$.

The phrase *equilibrium to the mechanism* means nothing more than *equilibrium to the game in which participants choose strategies in* H_i *and receive the payoffs that correspond with the choice* $p(h)$. So a choice function is implementable when there exists a mechanism that has an equilibrium which results in the same relationship between ϕ and policy as the choice function. This notion is weaker than the notion of fully implementable (Palfrey and Srivastava, 1989) which involves a mechanism in which all equilibria match the choice function. While some authors use implementation in exactly the sense that I do, others have used the terms achievability or truthful implementation to capture what I call implementable. I do not address the question of full implementation, other than noting that the conditions of theorem 3 in Palfrey and Srivastava (1989) are not generally satisfied by the problems considered here, suggesting that even when positive implementation results attain it is not likely that positive strong implementation results can be attained.[12]

The revelation principle (Gibbard 1973; Green and Laffont 1977; Myerson 1979; Dasgupta, Hammond and Maskin 1979) justifies focus on a smaller set of mechanisms. A direct mechanism with transfers (or without transfers) is a mechanism in which $H_i = \Phi$ for all i. In considering direct mechanisms, m_i^s denotes the s_i coordinate of i's report and m_i^θ denotes the θ_i coordinate of i's report. In analyzing direct mechanisms the focus is on determining when truthful strategies, $m_i(\phi_i) = \phi_i$, form a Bayesian Nash equilibrium.

Definition 2 *A direct mechanism* $(p(\cdot), \{t_i(\cdot)\}_{i \in N})$ *is Bayesian Nash incentive compatible if* $m_i(\phi_i) = \phi_i$ *for each* $i \in N$ *is a Bayesian Nash equilibrium to the mechanism. This is true if and only if the following incentive compatibility condition is satisfied: for each* $i \in N$, ϕ_{-i}, ϕ_i *and* ϕ_i'

[12] See example 3 of Palfrey and Srivastava (1989) for a discussion of why full implementation generally fails in problems without purely private values.

$$\sum_{\phi_{-i}} \sum_{x} \left[\begin{array}{c} p(\phi_i, \phi_{-i}) \left(u_i(a; x, \theta_i) + t_i(\phi_i, \phi_{-i})\right) + \\ \left(1 - p(\phi_i, \phi_{-i})\right) \left(u_i(b; x, \theta_i) + t_i(\phi_i, \phi_{-i})\right) \end{array} \right] pr(x, \phi_{-i} \mid \phi_i) \geq \tag{3}$$

$$\sum_{\phi_{-i}} \sum_{x} \left[\begin{array}{c} p(\phi_i', \phi_{-i}) \left(u_i(a; x, \theta_i) + t_i(\phi_i', \phi_{-i})\right) + \\ \left(1 - p(\phi_i', \phi_{-i})\right) \left(u_i(b; x, \theta_i) + t_i(\phi_i', \phi_{-i})\right) \end{array} \right] pr(x, \phi_{-i} \mid \phi_i).$$

The Revelation principle allows us to focus only on choice functions that can be implemented by direct mechanisms. I state without proof this well-known result.

Theorem 1 *(Revelation Principle) The choice function $c(\cdot)$ is Bayesian Nash implementable with transfers (without transfers) if and only if there exists a Bayesian Nash incentive compatible direct mechanism $\langle \Phi^n, p(\cdot), t_1(\cdot), ..., t_n(\cdot) \rangle$ with transfers (without transfers) in which $p(\phi) = c(\phi)$ for every $\phi \in \Phi^n$.*

To see that a direct mechanism implements a choice function if any mechanism implements the choice function, suppose there is a mechanism that implements the choice function $c(\cdot)$, and let $a_i^*(\cdot)$ denote the equilibrium strategies which result in choice according to the choice function $c(\cdot)$. Suppose instead of playing this mechanism, participants were asked to submit their private types ϕ_i to a disinterested mediator, that would then reliably play the strategy $a_i^*(\cdot)$ in the mechanism. Since $a_i^*(\phi_i)$ is a best response given the type ϕ_i when all participants play the equilibrium in the original mechanism, truthfully reporting ϕ_i to the mediator must be a best response when all other participants truthfully report their type. Accordingly, the direct mechanism in which $p^d(\phi_1, ..., \phi_n) = p(a_1^*(\phi_1), ..., a_n^*(\phi_n))$ for each $\phi \in \Phi$ is Bayesian Nash incentive compatible.

Given the collective choice problem described, the natural benchmark is the first best or efficient policy that an aggregate welfare maximizing planner would select,

$$p^{+}(\phi)=\begin{cases} 1 \text{ if } \left\{\mu(\phi)>\frac{1}{2} \text{ and } |i:\theta_i=1|\geq\frac{n+1}{2}\right\} \text{ or } \left\{\mu(\phi)<\frac{1}{2} \text{ and } |i:\theta_i=-1|\geq\frac{n+1}{2}\right\} \\ 0 \text{ if } \left\{\mu(\phi)<\frac{1}{2} \text{ and } |i:\theta_i=1|\geq\frac{n+1}{2}\right\} \text{ or } \left\{\mu(\phi)>\frac{1}{2} \text{ and } |i:\theta_i=-1|\geq\frac{n+1}{2}\right\} \\ \frac{1}{2} \text{ if } \mu(\phi)=\frac{1}{2} \end{cases} \tag{4}$$

where $\mu(\phi)$ is the posterior probability of $x=a$ given the profile ϕ. Given the probability model of conditionally i.i.d private signals, this is

$$\mu(\phi)=\frac{\pi g^{a(s)}(1-g)^{b(s)}}{\pi g^{a(s)}(1-g)^{b(s)}+(1-\pi)(1-g)^{a(s)}g^{b(s)}} \tag{5}$$

with $a(s)$ denoting the number of private signals with value a and $b(s)$ denoting the number of private signals with value b. The benchmark also corresponds to the full information majority rule (FIMR) outcome and satisfies what Feddersen and Pesendorfer (1997) term full information equivalence.

2.2.2 Direct public communication mechanisms

Certain types of institutions may be considered unacceptable. For example, in studying mechanisms that are representations of deliberative democracy, a natural restriction is that the process be democratic – the policy choice is ultimately the result of non-coerced voting and aggregation by a *nice* rule (say majority rule). In the current setting it turns out that a strengthening of the revelation principle is possible. Namely, there exists a game with public communication and then uncoerced voting under majority rule that has a perfect Bayesian equilibrium with weakly undominated voting that selects policy according to the function $p^{+}(\cdot)$ if and only if there exists a direct mechanism that implements $p^{+}(\cdot)$. The remainder of this section develops this result.

Let Γ^{v} denote a simple Bayesian game in which the participants learn their private types ϕ_i and then simultaneously cast ballots $v_i\in\{a,b\}$ and the outcome is chosen by simple majority rule (thus the game is characterized by type space $\{-1,1\}\times\{a,b\}$ action space

$\{a, b\}$ and payoffs $u_i(\cdot, \cdot, \cdot, \cdot)$). This game form is a canonical model of democracy but it is inappropriate as a model of deliberative democracy because it has no communication. It is possible, however, to consider extensions to the game that involve communication between participants prior to the voting. The possibilities for pre-vote communication are endless – participants may be allowed to only speak to certain other participants, many rounds of argument are possible, agents may play correlated strategies, etc. Instead of considering all of the possible ways to augment a game with pre-play communication, Myerson (1982) shows (Proposition 2) that it is sufficient to focus on direct coordination mechanisms. A direct coordination mechanism (formally defined in the appendix) differs from a direct mechanism in that it asks players to privately report their types (as in a direct mechanism) and then makes a private suggestion to each player about how they should play the game (instead of just selecting the payoffs directly). Players are free to disregard the suggestion. In the spirit of the revelation principle, Myerson's result establishes an equivalence between choice functions that are implementable in communication games and choice functions that are attained as equilibria to direct coordination mechanisms. In the current paper, then the interest is in drawing connections between choice functions that can be implemented in direct coordination mechanism and choice functions that can be implemented in direct mechanism. It is not surprising that if $p^+(\cdot)$ can be implemented by a direct coordination mechanism then it can be implemented by a direct mechanism. The less obvious implication, that if $p^+(\cdot)$ can be implemented by any direct mechanism (with transfers) then it can be implemented in a direct coordination mechanism (with transfers) is true if the voting rule in the voting game that generates the direct coordination mechanism is majority rule (recall that $p^+(\cdot)$ is the full information *majority rule* core). This result is formally stated as lemma 1 with proof in the appendix and justifies focus on direct mechanisms even though institutions that involve communication and democratic voting are desired. One shortcoming of the concept of Bayesian incentive compatibility in multiple period mechanisms is that behavior may violate sequential rationality. To this end a strengthening of the equilibrium concept is required. First, it is useful to define a canonical public communication and voting game.

Definition 3 *Given a voting game* Γ^v *a direct public communication mechanism (with transfers) is a two stage game of the following form: In period 1 participants simultaneously submit messages* $m_i \in \{-1, 1\} \times \{a, b\}$ *and then in period 2 participants observe the profile of messages about* s_i, $m^s \in \{a, b\}^n$ *and then cast ballots* $v_i \in \{a, b\}$. *If the game involves transfers then* i *receives a transfer* $t_i(m)$.

Direct public communication mechanism do not involve a mediator making recommendations, instead the messages are publicly observed. The following result is established in the appendix.

Proposition 1 *If there exists a Bayesian incentive compatible direct mechanism (with transfers) that implements the choice function* $p^+(\cdot)$ *then there exists a direct public communication game (with transfers) that possesses a perfect Bayesian Nash equilibrium in weakly undominated voting which reaches the same policy decision as* $p^+(\phi)$ *for every* ϕ.

Following proposition 1, I will investigate whether $p^+(\cdot)$ can be implemented by examining incentive compatibility conditions in direct mechanisms. This approach involves verifying whether truthful response is a best response if everyone else is truthful given the direct mechanism in question. The incentive compatibility condition is then nothing more than the requirement that truthfulness be a mutual best-response. When I attain positive implementation results with transfers I will consider the institutional design question of what types of institutions can decentralize desirable direct mechanisms. These institutions will typically be direct public communication mechanisms. Proposition 1 tells us that in these decentralized institutions appeal to the standard refinement of perfect Bayesian equilibrium in weakly undominated voting is innocuous.

3 Mechanism design without transfers

I first investigate whether there is a mechanism that implements $p^+(\cdot)$ without transfers. By proposition 1 it is sufficient to investigate whether a direct mechanism implements $p^+(\cdot)$. Equivalently, it is sufficient to investigate whether the direct mechanism that selects $p^+(\cdot)$ and distributes no transfers induces a game in which truthful strategies constitute a Bayesian Nash equilibrium. To avoid trivial problems, I assume that the private information is potentially useful. The requirement that information can be useful is formalized by requiring that there are two distinct profiles of private information s and s' in which under s, $x = a$ is more likely and under s', $x = b$ is more likely. Specifically, since the number of agents with $s_i = a$ is a sufficient statistic for s, this condition requires that for some α and $\beta = n - 1 - \alpha$

$$\frac{\pi g^{\alpha+1}(1-g)^{\beta}}{\pi g^{\alpha+1}(1-g)^{\beta} + (1-\pi)(1-g)^{\alpha+1}g^{\beta}} < \frac{1}{2} < \frac{\pi g^{\alpha}(1-g)^{\beta+1}}{\pi g^{\alpha}(1-g)^{\beta+1} + (1-\pi)(1-g)^{\alpha}g^{\beta+1}}. \qquad (6)$$

A useful interpretation of (6) is that if α of $n-1$ participants observe $s_i = a$ and the remaining β of $n-1$ participants observe $s_i = b$ then the remaining observation will affect which state is more likely given s. The ordering in (6) is true for some α and β with $\alpha + \beta = n - 1$ iff

$$\left(\frac{1-g}{g}\right)^{n} < \frac{\pi}{1-\pi} < \left(\frac{g}{1-g}\right)^{n}. \qquad (7)$$

Throughout I assume that (7) is satisfied.

Proposition 2 *If (7) is satisfied then the mapping $p^{+}(\cdot)$ can be implemented in Bayesian Nash strategies without transfers if and only if for $t \in \{-1, 1\}$*

$$\eta_t^{+} \geq \eta_t^{-}. \qquad (8)$$

Proof: The proof proceeds in two parts.

($\Longrightarrow$). The proof is by contraposition. Since individual signals s_i are conditionally independent and (7) is satisfied, selection according to the choice rule $p^{+}(\cdot)$ requires that the mechanism select policy by the function $p^{+}(\cdot)$.[13] Assume that $p^{+}(\cdot)$ is used and that $\eta_t^{+} < \eta_t^{-}$ for $j \in \{-1, 1\}$. If all agents other than i are truthful then the s_i coordinate of ϕ_i will only affect the outcome if the remaining private signals are summarized by α and β satisfying (6). In this case a report of a will result in policy a if and only if more participants report type 1 than -1

[13] Subsequently when information is assumed to satisfy nonexclusivity this claim will not hold as the policy need not depend on each participant's message.

and a report of b will result in b if and only if more participants report type 1 than -1. It is sufficient to focus on profiles of s for which (6) is satisfied. If at least $\frac{n+1}{2}$ of the other participants have a particular type then i's report about θ_i is irrelevant. If exactly $\frac{n-1}{2}$ of the participants $N\backslash i$ have each type then θ_i is relevant. So lying about both both s_i and θ_i results in the same outcome as a truthful report if $\frac{n-1}{2}$ of the participants $N\backslash i$ have each type. In the remaining cases, lying about both s_i and θ_i results in a better policy if at least $\frac{n+1}{2}$ of the other participants have type not equal to θ_i, and lying about both s_i and θ_i results in a worse policy if at least $\frac{n+1}{2}$ of the other participants have type not equal to θ_i. Accordingly, for agent i with fixed s_i, θ_i truthful revelation is not optimal if $\eta_t^+ < \eta_t^-$ for $t = \theta_i$.

($\Longleftarrow$). Assume that (8) is satisfied and that the direct mechanism $p^+(\cdot)$ is used. There are three possible deviations from a truthful strategy (lying only about s_i, lying only about θ_i or lying about both). If all agents other than i are truthful then the s_i coordinate of ϕ_i will only affect the outcome if the remaining private signals are summarized by α and β satisfying (6). From the argument in the first part of the proof, $\eta_t^+ \geq \eta_t^-$ for $t = \theta_i$ is sufficient for truthfulness to be better than lying about both s_i and θ_i. Since i's report of θ_i is only relevant if exactly $\frac{n-1}{2}$ of the participants $N\backslash i$ have each type given that (6) is satisfied and i's report of s_i is dishonest, truthfully reporting θ_i is less desirable than dishonestly reporting θ_i and thus the deviation of only dishonestly reporting s_i is less desirable than the deviation of dishonestly reporting both quantities. Finally it is clear that given $p^+(\cdot)$ if i truthfully reports s_i then dishonestly reporting θ_i is less desirable than honestly reporting both quantities. Thus no agent has a unilateral incentive to deviate from a truthful message strategy.■

Informally, the result has a straightforward interpretation. First-best aggregation can only occur if both preference types believe that conditional on their own type they are likely to be the majority type in the collective. In contrast if one type believes that they are likely to be in the minority then they will have an incentive to manipulate the outcome by incorrectly revealing their private information.

Corollary 1 *When the profile of types θ is known and a minimal diversity condition ($\exists i, j \in N$ s.t. $\theta_i \neq \theta_j$) is satisfied, $p^+(\cdot)$ cannot be implemented without transfers.*

Remark 1 *Proposition 2 demonstrates that in contrast to the claims of some scholars, designing institutions to aggregate information and preferences is harder in the presence of ideological diversity.*

For the remainder of the paper I focus on the challenging case where $\eta_t^+ \geq \eta_t^-$ is not satisfied for some $t \in \{-1, 1\}$. More precisely it is assumed that

$$\eta_{-1}^+ < \eta_{-1}^-. \tag{9}$$

It is natural to ask whether there are any reasonably good mechanisms without transfers or whether it is impossible to aggregate the information any more efficiently than through communication-free voting if transfers are not feasible. An example of a direct mechanism which partially aggregates the available information is

$$p^w(\phi) = \begin{cases} 1 \text{ if } \mu(\phi) < \frac{1}{2} \text{ and } |i : \theta_i = -1| \geq \frac{n+1}{2} \\ 0 \text{ if } \mu(\phi) > \frac{1}{2} \text{ and } |i : \theta_i = -1| \geq \frac{n+1}{2} \\ w \text{ if } \mu(\phi) > \frac{1}{2} \text{ and } |i : \theta_i = 1| \geq \frac{n+1}{2} \\ 1 - w \text{ if } \mu(\phi) < \frac{1}{2} \text{ and } |i : \theta_i = 1| \geq \frac{n+1}{2} \\ \frac{1}{2} \text{ if } \mu(\phi) = \frac{1}{2} \end{cases} \tag{10}$$

where w solves the condition

$$w = 1 - \frac{(\eta_{-1}^- - \eta_{-1}^+)}{2\eta_{-1}^-}.^{14} \tag{11}$$

Remark 2 *The mechanism $p^w(\cdot)$ can be interpreted as an institution in which participants reveal their preferences to a moderator and then if $\theta_i = -1$ is the minority type the moderator randomly decides whether to select policy to make majority happy (with probability w) or to make the minority happy (with probability $1 - w$), and if $\theta_i = -1$ is the majority type the moderator selects policy to make the majority happy.*

[14] To see that this mechanism is Bayesian Nash incentive compatible, consider a participant with $\theta_i = -1$. Note that if all participants other than i are truthful and i is truthful, conditional on m_i^s being consequential the probability that the policy that i prefers is chosen is $\eta_{-1}^- w + 1 - \eta_{-1}^-$. Alternatively, conditional on m_i^s being consequential and the best deviation, the probability that the policy that i prefers is chosen is $\eta_{-1}^-(1 - w) + (1 - \eta_{-1}^- - \eta_{-1}^+)$. The value of w that equates these expressions is in the unit interval when (9) is satisfied, and thus is feasible.

3.1 A digression on the relevance of nonexclusivity of information

Thus far, I have maintained the assumption that private signals s are independent conditional on the state x. In this section I consider the relevance of Postlewaite and Schmeidler's (1986) condition of nonexclusivity and Baron and Meirowitz's (2004) strong nonexclusivity condition. With binary signals that are positively correlated with the state x, the conditions can be defined as follows.

Definition 4 *The information environment satisfies strong nonexclusivity if for all* $i \in N$ *there are two distinct* $j, k \in N \backslash i$ *s.t.* $pr(s_i = s_j = s_k) = 1$.

Definition 5 *The information environment satisfies nonexclusivity if for all* $i \in N$ *there is a* $j \in N \backslash i$ *s.t.* $pr(s_i = s_j) = 1$.

As before I assume that s_i and θ_i are independent, but for this section I do not specify further the joint distribution of (x, s), and directly assume that the following extension of (7) is satisfied. (i) For any $i \in N$ there exists an s and s' which differ only in s_i s.t.

$$pr(x = a \mid s) > \frac{1}{2} > pr(x = a \mid s') \tag{12}$$

and (ii) for every $i \in N$

$$pr(s_i = a \mid x = a) \geq pr(s_i = a \mid x = b). \tag{13}$$

Since I have relaxed the assumption that signals are conditionally independent the posterior on x based on s is no longer of the form in (5). For this section let $\mu(s)$ denote the posterior probability of x conditional on s derived via Bayes' rule. Accordingly, $p^+(\cdot)$ denotes the rule that uses $\mu(s)$ and θ in the natural manner. The next result uses the fact that when strong nonexclusivity is violated and (i) is satisfied, any mechanism that selects $p^+(\phi)$ for every ϕ is potentially responsive to each signal, whereas when strong nonexclusivity is satisfied it is always possible to ignore a single participant (in the face of truthful messages from everyone else).

Proposition 3 *Assume $\eta_{-1}^{+} < \eta_{-1}^{-}$ and that conditions (i) and (ii) hold. The choice function $p^{+}(\cdot)$ can be implemented in Bayesian Nash strategies without transfers if and only if the information environment satisfies strong nonexclusivity.*

Proof:

($\Longrightarrow$) Assume that (i) and (ii) hold, that strong nonexclusivity is not satisfied and that $p^{+}(\phi)$ can be implemented in Bayesian Nash strategies without transfers. Let $p(\cdot)$ denote the mechanism that implements $p^{+}(\cdot)$. Let $p(m_i^s, m_j^s, s_{-ij}, m^{\theta})$ denote the lottery over policy that results from m^{θ}, m_i^s, m_j^s, and m_{-ij}^s—the profile of messages from $N\backslash\{i,j\}$. Since strong nonexclusivity is not satisfied, for some $i, j \in N$ there is no third participant k with $pr(s_i = s_k) = 1$. This and (i) imply that for every θ and some s' either for $m_i^s \neq s_i'$, $p(m_i^s, s_j', s_{-ij}', m^{\theta}) \neq p(s_i', s_j', s_{-ij}', m^{\theta})$ or for $m_j^s \neq s_j'$, $p(s_i', m_j^s, s_{-ij}, m^{\theta}) \neq p(s_i', s_j', s_{-ij}', m^{\theta})$, or both. Assume that $N\backslash\{i,j\}$ are truthful. Since $\eta_{-1}^{+} < \eta_{-1}^{-}$, the assumption that truthful messages are mutual best responses to $p(\cdot)$ requires that $p(b, a, s_{-ij}', m^{\theta}) \geq p(a, a, s_{-ij}', m^{\theta})$ if $m_i^{\theta} = -1$ and $p(b, a, s_{-ij}', m^{\theta}) \leq p(a, a, s_{-ij}', m^{\theta})$ if $m_i^{\theta} = 1$. Similarly, it must be the case that $p(b, a, s_{-ij}', m^{\theta}) \geq p(b, b, s_{-ij}', m^{\theta})$ if $m_j^{\theta} = 1$ and $p(b, a, s_{-ij}', m^{\theta}) \leq p(b, b, s_{-ij}', m^{\theta})$ if $m_j^{\theta} = -1$. But the first and last of these four conclusions imply that $p(b, b, s_{-ij}', m^{\theta}) \geq p(a, a, s_{-ij}', m^{\theta})$ if $m_i^{\theta} = -1$ and $m_{ij}^{\theta} = -1$. However, if more than $\frac{n+1}{2}$ of the participants have $\theta_k = 1$ then we must have $p(b, b, s_{-ij}', m^{\theta}) < p(a, a, s_{-ij}', m^{\theta})$ since $p(\cdot)$ and $p^{+}(\cdot)$ coincide when $m_i^s = m_j^s$. Thus, it is not possible to construct the required mechanism $p(\cdot)$.

($\Longleftarrow$) The proof is by construction. By strong nonexclusivity for each $i \in N$ there exists a set $N_i \subset N$ s.t. $pr(s_i = s_j \forall j \in N_i)$ and N_i contains 3 participants (including i). For each $i \in N$ and thus each N_i define the following function $\varrho^i : \{a,b\}^3 \to \{a,b\}$

$$\varrho^i(m^s) = \left\{ \begin{array}{c} m_i^s \text{ if for } i,j,k \in N_i \ m_j^s = m_i^s = m_k^s \\ b \text{ if for } t,j,k \in N_i \ m_t^s = a \text{ and } m_j^s = m_k^s = b \\ a \text{ if for } t,j,k \in N_i \ m_t^s = b \text{ and } m_j^s = m_k^s = a. \end{array} \right. \tag{14}$$

Now define the mapping, $\varrho(m^s) : \{a,b\}^n \to \{a,b\}^n$ that translates m^s into $(\varrho^1(m^s), ..., \varrho^i(m^s), ..., \varrho^n(m^s))$. Given the mechanism

$$p^{\wedge}(\phi) = p^{+}(\varrho(m^s), \theta) \tag{15}$$

truthfulness results in the same outcome as $p^{+}(\phi)$ and by strong nonexclusivity for any m^s no unilateral deviation in m_i^s will affect $\varrho(m^s)$ and thus the mechanism is Bayesian Nash incentive compatible.■

Existence of a perfect Bayesian Nash equilibrium in a direct communication game is not guaranteed by propositions 1 and 3 because proposition 1 holds for the case where private signals s are conditionally independent. Specifically, in the proof of proposition 1 the fact that every m^s is consistent with truthful strategies for some ϕ is used in the construction of beliefs. The following corollary extends the result to the case of strongly nonexclusive environments.

Corollary 2 *Assume $\eta_{-1}^{+} < \eta_{-1}^{-}$, that conditions (i) and (ii) hold and that strong nonexclusivity is satisfied. In the direct public communication game there is a perfect Bayesian Nash equilibrium in weakly undominated voting strategies in which the policy mapping is $p^{+}(\cdot)$.*

Proof: That there is a Bayesian Nash equilibrium of this form follows from lemma 1 and proposition 3, but the fact that the equilibrium also satisfies sequential rationality given consistent beliefs requires a little extra work. In the described game with truthful messages every history h of messages satisfying $m_i = m_j$ if $m_j \in N_i$ is feasible for some s and thus equilibrium beliefs are given by Bayes' rule. Following any history in which the above condition is violated assume the beliefs assign probability 0 to an s in which two messages from the same N_i are incorrect. Given these beliefs, weak dominance and sequential rationality are satisfied by any voting strategy with $v_i(h, \theta_i = 1) = 1$ if $\mu(x \mid h) > \frac{1}{2}$, $v_i(h, \theta_i = -1) = 0$ if $\mu(x \mid h) > \frac{1}{2}$, $v_i(h, \theta_i = 1) = 0$ if $\mu(x \mid h) < \frac{1}{2}$ and $v_i(h, \theta_i = -1) = 1$ if $\mu(x \mid h) < \frac{1}{2}$. Given that the initial decisions fully reveal s, all such voting strategies result in choice according to $p^{+}(\cdot)$ under majority rule. Accordingly, it is sufficient to show that announcement of truthful messages is a best response given the belief that policy will be selected by $p^{\wedge}(\cdot)$. But since the equilibrium outcome does not depend on a unilateral deviation in m_i^s the result follows.■

Thus, if $\eta_{-1}^{+} < \eta_{-1}^{-}$, without transfers strong nonexclusivity is necessary and sufficient for implementability of $p^{+}(\cdot)$. In signalling games with fixed and known preferences, nonexclusivity can be sufficient for fully-revealing equilibria. Krishna and Morgan (2001) demonstrate this point in a two sender game in which the sender's private signals are identical. In the current setting, the weaker condition of nonexclusivity is sufficient if the policy space is perturbed slightly to include a policy which is undesirable to all participants. It is straightforward to see that if there is one additional feasible policy which each type of participants finds less desirable than a or b in either realization of x, then nonexclusivity is sufficient. The intuition is that if the mechanism can select this punishment policy whenever it is clear that at least one participant lied then no one will have an incentive to lie if a unilateral lie can be detected. Nonexclusivity is the minimal condition necessary to always be able to detect a unilateral lie. One conclusion of the above result is that in considering the value of adding an additional participant to a deliberative process, it may be better to add one that observes the same information as some other participant than one that has new information. While the new information has potential value to the group it may not be of use if incentives for information transmission are not present. The redundant information, however, may create better incentives for information transmission of the information already possessed by group members.

4 Mechanism design with transfers

For the remainder of the paper, the focus is on informational environments in which aggregation is not possible without transfers. To that end, assume that as before s_i is conditionally independent with $pr(s_i = x) = g > \frac{1}{2}$ (thus nonexclusivity is violated) and that $\eta_{-1}^{+} < \eta_{-1}^{-}$.

4.1 State dependent transfers

Now suppose that while the policy must depend on only ϕ, the transfers can depend on both ϕ and x. In this case I construct a direct mechanism that implements the first-best outcome.

Proposition 4 *There exist values $v_a^+, v_a^-, v_b^+, v_b^-$ (characterized by (IC-i and IC-ii) below) for which the mechanism that selects $p^+(\cdot)$ and distributes transfers*

$$t_i(m_i^s, x) = \begin{cases} v_a^+ \text{ if } m_i^s = a = x \\ v_a^- \text{ if } m_i^s = a \neq x \\ v_b^+ \text{ if } m_i^s = b = x \\ v_b^- \text{ if } m_i^s = b \neq x \end{cases} \tag{16}$$

is Bayesian Nash incentive compatible.

Proof: Suppose that the transfers to i depend only on i's report about s_i and x and are given by the transfer schedule in (16). Furthermore suppose that the policy is chosen by $p^+(\cdot)$. Since this is the rule that selects the optimal policy for the majority θ_i type given knowledge of s, if in fact all agents are truthfully revealing m_i^s, truthful revelation of θ_i is clearly a best response. Since the deviation to lying about both s_i and θ_i is better than the deviation to lying only about s_i it remains only to establish that truthful revelation is better for i than the deviation in both dimensions when all other agents are truthful. I first show that the binding incentive compatibility conditions apply for $\theta_i = -1$ types and then derive the values $v_a^+, v_a^-, v_b^+, v_b^-$ to induce truthfulness by $\theta_i = -1$ types. Let λ_a denote the probability that the profile s_{-i} is such that s_i is pivotal in affecting the policy decision given that $x = a$. This is the probability that exactly α of the $n-1$ other signals are supportive of a conditional on $x = a$ with α solving

$$\frac{\pi g^{\alpha+1}(1-g)^{n-\alpha-1}}{\pi g^{\alpha+1}(1-g)^{n-\alpha-1} + (1-\pi)(1-g)^{\alpha+1}g^{n-\alpha-1}} > \frac{1}{2} > \frac{\pi g^{\alpha}(1-g)^{n-\alpha}}{\pi g^{\alpha}(1-g)^{n-\alpha} + (1-\pi)(1-g)^{\alpha}g^{n-\alpha}} \tag{17}$$

By the assumption that (7) is satisfied such a value α exists, and since $g, \pi \geq \frac{1}{2}$, $\alpha \leq \frac{n-1}{2}$. Similarly, let λ_b denote the probability that the profile s_{-i} is such that s_i is pivotal given that $x = b$. Formally,

$$\begin{aligned} \lambda_a &= \begin{pmatrix} n-1 \\ \alpha \end{pmatrix} g^{\alpha}(1-g)^{n-1-\alpha} \\ \lambda_b &= \begin{pmatrix} n-1 \\ \alpha \end{pmatrix} (1-g)^{\alpha} g^{n-1-\alpha}. \end{aligned} \tag{18}$$

Since $\alpha \leq \frac{n-1}{2}$ and $g \geq \frac{1}{2}$, $\lambda_a < \lambda_b$. Given truthful reports by $N \backslash i$ the incentive compatibility condition for truthfully reporting as opposed to lying about both s_i and θ_i when $s_i = a$ is (dividing through by a constant)

$$\begin{aligned} &\pi g(\lambda_a \eta_{-1}^{+} + v_a^{+}) + (1-\pi)(1-g)(\lambda_b \eta_{-1}^{-} + v_a^{-}) \geq \\ &\pi g(\lambda_a \eta_{-1}^{-} + v_b^{-}) + (1-\pi)(1-g)(\lambda_b^{+} \eta_{-1}^{+} + v_b^{+}). \end{aligned} \tag{IC1}$$

Similarly the incentive compatibility condition for truthfully reporting when $s_i = b$ is

$$\begin{aligned} &\pi(1-g)(\lambda_a \eta_{-1}^{-} + v_b^{-}) + (1-\pi)g(\lambda_b \eta_{-1}^{+} + v_b^{+}) \geq \\ &\pi(1-g)(\lambda_a p \eta_{-1}^{+} + v_a^{+}) + (1-\pi)g(\lambda_b \eta_{-1}^{-} + v_a^{-}). \end{aligned} \tag{IC2}$$

The incentive compatibility condition for a $\theta_i = 1$ type that observes $s_i = a$ is

$$\pi g(\lambda_a \eta_1^+ + v_a^+) + (1-\pi)(1-g)(\lambda_b \eta_1^- + v_a^-) \geq \tag{IC3}$$
$$\pi g(\lambda_a \eta_1^- + v_b^-) + (1-\pi)(1-g)(\lambda_b^+ \eta_1^+ + v_b^+).$$

The incentive compatibility condition for a $\theta_i = 1$ type that observes $s_i = b$ is

$$\pi(1-g)(\lambda_a \eta_1^- + v_b^-) + (1-\pi)g(\lambda_b \eta_1^+ + v_b^+) \geq \tag{IC4}$$
$$\pi(1-g)(\lambda_a p \eta_1^+ + v_a^+) + (1-\pi)g(\lambda_b \eta_1^- + v_a^-).$$

It remains only to show that there is a pair of values $v_a^+, v_a^-, v_b^+, v_b^-$ satisfying this system. Since $\pi g > (1-\pi)(1-g)$,and $\eta_{-1}^+ < \eta_1^+, \eta_{-1}^- > \eta_1^-$ if IC1 is satisfied then IC3 is (with strict inequality). Note that (IC2) and (IC4) can be rewritten as

$$(\eta_{-1}^+ - \eta_{-1}^-)((1-\pi)g\lambda_b - \pi(1-g)\lambda_a) \geq \tag{19}$$
$$\pi(1-g)(v_a^+ - v_b^-) + (1-\pi)g(v_a^- - v_b^+)$$

$$(\eta_1^+ - \eta_1^-)((1-\pi)g\lambda_b - \pi(1-g)\lambda_a) \geq \tag{20}$$
$$\pi(1-g)(v_a^+ - v_b^-) + (1-\pi)g(v_a^- - v_b^+)$$

Since $(\eta_1^+ - \eta_1^-) > (\eta_{-1}^+ - \eta_{-1}^-)$ the first constraint (IC2) is more restrictive if $(1-\pi)g\lambda_b > \pi(1-g)\lambda_a$ and the second constraint (IC4) is more restrictive if $(1-\pi)g\lambda_b < \pi(1-g)\lambda_a$. Assuming that the second inequality holds and substituting for λ_a, λ_b and multiplying by a constant yields

$$\frac{(1-\pi)(1-g)^\alpha g^{n-\alpha}}{\pi g^\alpha (1-g)^{n-\alpha} + (1-\pi)(1-g)^\alpha g^{n-\alpha}} < \frac{\pi g^\alpha (1-g)^{n-\alpha}}{\pi g^\alpha (1-g)^{n-\alpha} + (1-\pi)(1-g)^\alpha g^{n-\alpha}}. \tag{21}$$

But the left hand side is just the difference between 1 and the right hand side and by (17) α is defined to make the right hand side less than $\frac{1}{2}$ implying that in fact the second inequality cannot hold. Thus the first inequality holds and thus it is (IC2) and not (IC4) that binds. Thus the relevant constraints are (IC1) and (IC2). I can then express the relevant IC constraints

$$(\eta_{-1}^+ - \eta_{-1}^-)(\pi g\lambda_a - (1-\pi)(1-g)\lambda_b) \geq \tag{IC-i}$$
$$\pi g(v_b^- - v_a^+) + (1-\pi)(1-g)(v_b^+ - v_a^-)$$

$$(\eta_{-1}^{+} - \eta_{-1}^{-})((1-\pi)g\lambda_b - \pi(1-g)\lambda_a) \geq \\ \pi(1-g)(v_a^{+} - v_b^{-}) + (1-\pi)g(v_a^{-} - v_b^{+}). \tag{IC-ii}$$

Finally, to show that it is possible to simultaneously satisfy these conditions note that the system of equations is satisfied with equality if

$$v_a^{+} - v_b^{-} = v_a := \lambda_a \left[\eta_{-1}^{-} - \eta_{-1}^{+}\right] \\ v_b^{+} - v_a^{-} = v_b := \lambda_b \left[\eta_{-1}^{-} - \eta_{-1}^{+}\right]. \tag{22}$$

This completes the proof.■

For the remainder of this section I will focus on transfer schedules that satisfy (22) with equality. A natural question to ask is how large is the expected transfer volume when $v_a^{-} = v_b^{-}$. If $x = a$ and the group size is n then the expected transfer to any participant is

$$Et_a^n = \left[\eta_{-1}^{-} - \eta_{-1}^{+}\right] g\lambda_a. \tag{23}$$

It is clear that this term goes to 0 as $n \to \infty$. Moreover, by conditional independence the expected sum of transfers conditional on $x = a$ is

$$E(\sum t_a^n) = nEt_a^n. \tag{24}$$

Since both $\left[\eta_{-1}^{-} - \eta_{-1}^{+}\right]$ and λ_a tend to 0 faster than $\frac{1}{n}$ it must be the case that $nEt_a^n \to 0$. A similar argument establishes the conclusion for the $x = b$ conditional sum.

Corollary 3 *For the mechanism with $v_a^{-} = v_b^{-} = 0$, $v_a^{+} = v_a$ and $v_b^{+} = v_b$, as $n \to \infty$ the expected sum of transfers, $E(\sum t^n) = \pi E(\sum t_a^n) + (1-\pi)E(\sum t_b^n)$, converges to 0.*

A second natural question to ask is whether this mechanism can be modified to satisfy ex-ante budget balance for finite populations. In other words can I construct a transfer schedule $t_i'(\cdot,\cdot)$ s.t. $E(\sum t_a') = 0$. Unlike the schedule in proposition 4 this will require negative transfers to some participants. Consider

$$t_i'(m_i^s, x) = \begin{cases} v_a - \varphi \text{ if } m_i^s = x = a \\ v_b - \varphi \text{ if } m_i^s = x = b \\ -\varphi \text{ otherwise.} \end{cases} \tag{25}$$

Expected budget balance requires only that

$$\varphi = n^{-1}\left[\pi E(\sum t_a^n) + (1-\pi)E(\sum t_b^n)\right] = \tag{26}$$
$$\left[\eta_{-1}^{-} - \eta_{-1}^{+}\right] g\left(\pi\lambda_a + (1-\pi)\lambda_a\right)$$

Since $t_i'(\cdot)$ satisfies equations (IC-i) and (IC-ii) (set $v_a^- = v_b^- = \varphi$) it is incentive compatible.

Corollary 4 *The mechanism that selects $p^+(\cdot)$ and distributes transfers $t_i'(m_i^s, x)$ defined in (25) is Bayesian Nash incentive compatible and satisfies expected budget balance.*

Corollary 5 *The mechanism described in corollary 4 can be decentralized as a two stage game. In period 1 each agent must select one of two Arrow securities (one that pays v_a if the state is $x = a$ and 0 otherwise and one that pays v_b if the state is $x = b$ and 0 otherwise). These securities have a price of φ and each agent must select exactly 1 security. In the second period the participants get to observe the number of securities of each type that are selected and then they vote and the policy is chosen by majority rule. In this game there is a perfect Bayesian Nash equilibrium in weakly undominated voting strategies in which the policy mapping is $p^+(\cdot)$.*

Proof: The result follows from propositions 1 and 4.■

Experimental work (Chen et. al. 2003) on the use of scoring rules to aggregate information with and across departments in corporations demonstrates the feasibility and apparent efficiency of using mechanism similar to the one described in corollary 5.

4.1.1 A digression on market trading

A natural interpretation of the mechanism described in corollary 5 is of a market maker that sells securities at a fixed price. An alternative type of market is one where participants are endowed with risk free and risky securities and they trade at endogenously determined prices. I follow the literature on strategic market games (see Jackson and Peck 1999 for recent developments) and specify an explicit market mechanism in which the participants select strategies. I first describe a Bayesian game in which participants trade in a discrete version of the market game of Shapely and Shubik (1977) and then participants vote after observing the equilibrium price. I then use proposition 4 to show that there is an equilibrium that selects policy according to the choice function $p^+(\cdot)$. In this game each participant is endowed with 1 unit of a risk free asset and 1 unit of a risky Arrow security that pays 1 if $x = a$ and 0 if $x = b$. First, participants learn their private information and then decide which side of the market they want to be on (sellers of the risk free asset or sellers of the risky asset), a market maker then sets the price to clear the market and makes the trades, participants observe the market price and finally vote. Specifically, the action space in the market is $\sigma_i \in \{0, 1\}$ with $\sigma_i = 1$ (0) interpreted as a decision to trade risk free assets for the risky asset (risky assets for the risk free asset) at the market clearing price $P(\sigma)$. Thus if $\sigma_i = 1$ then the participant has $1-B$ units of the the risk free asset and $1+\frac{B}{P(\sigma)}$ units of the risky asset. If $\sigma_i = 0$ then the participant has $1+QP(\sigma)$ units of the risk free asset and $1-Q$ units of the risky asset. I will characterize values of Q and B to satisfy incentive compatibility conditions and thus support a separating equilibrium. The price $P(\sigma)$ is chosen to clear the market. Specifically, letting n^0 and n^1 denote the number of participants selecting $\sigma_i = 0$ and $\sigma_i = 1$ respectively, I define

$$P(\sigma) = \frac{Bn^1}{Qn^0} \tag{27}$$

as long as $n^0, n^1 > 0$. If $n^0 = 0$ or $n^1 = 0$ then no trades are made and the price is not well defined.

Proposition 5 *Consider the game in which (1) participants are endowed a risk free security that pays 1 and an Arrow security that pays off iff $x = a$.; (2) each agent decides whether to sell*

$$B = \frac{v_b(1 - g + gn)}{n} \tag{28}$$

units of the risk free security or sell

$$Q = \frac{v_b(1 - 2g + 2gn + g^2 - 2g^2n + g^2n^2) + v_a(g + n - 2gn - g^2 + gn^2 + 2g^2n - g^2n^2)}{n(g + n - gn)} \tag{29}$$

units of the risky security; (3) the transactions are then made at the market clearing price, $P(\sigma)$; and (4) participants observe the price and then vote. This game has a perfect Bayesian Nash equilibrium with weakly undominated voting strategies which selects policy according to the choice function $p^+(\cdot)$.

Proof: In an equilibrium in which market actions are separating, any deviation results in a history that is possible under play of the equilibrium strategies so beliefs about $(\phi_{-i} \mid \phi_i)$ are given by Bayes' rule for every feasible observation of P as well as the possible histories where no trade occurs. Additionally, note that $P(\sigma) = \frac{Bn^1}{Qn^0}$ is monotone in n^1 and n^0 so that if market decisions are separating in s_i then posteriors based on P (or a history in which no trade occurs and σ_i is known to i) will be as informative as posteriors based directly on s. Moreover, if market actions fully reveal s, then sequentially rational and weakly undominated voting will result in choice according to $p^+(\cdot)$. This implies that there is a separating perfect Bayesian Nash equilibrium if Q and P to solve an incentive compatibility condition for selecting σ_i to reveal s_i. I consider the strategy

$$\sigma_i(s_i) = \begin{cases} 1 \text{ if } s_i = a \\ 0 \text{ otherwise} \end{cases}. \tag{30}$$

In the candidate game, the market based payment to i associated with a profile of trades, σ,(which thus defines the values n^0, n^1 recording the number of participants reporting b and a respectively) is

$$t_i^m(\sigma, x) = \begin{cases} 1 - B + 1 + B\frac{Qn^0}{Bn^1} \text{ if } \sigma_i = 1 \text{ and } x = a \\ 1 - B \text{ if } \sigma_i = 1 \text{ and } x = b \\ 1 + Q\frac{Bn^1}{Qn^0} \text{ if } \sigma_i = 0 \text{ and } x = b \\ 1 + Q\frac{Bn^1}{Qn^0} + 1 - Q \text{ if } \sigma_i = 0 \text{ and } x = a. \end{cases} \tag{31}$$

If $N \backslash i$ use the strategy in (30) then the expected market based income as a function of σ_i is given by

$$Et_i^m(\sigma_i, x) = \begin{cases} 1 - B + 1 + B\frac{(n-1)Q(1-g)}{(n-1)Bg+B} \text{ if } \sigma_i = 1 \text{ and } x = a \\ 1 - B \text{ if } \sigma_i = 1 \text{ and } x = b \\ 1 + Q\frac{(n-1)B(1-g)}{(n-1)Qg+Q} \text{ if } \sigma_i = 0 \text{ and } x = b \\ 1 + Q\frac{(n-1)Bg}{(n-1)Q(1-g)+Q} + 1 - Q \text{ if } \sigma_i = 0 \text{ and } x = a. \end{cases} \tag{32}$$

From proposition 4, use of (30) satisfies the incentive compatibility conditions with equality if. The values of Q and B characterized in (28) and (29) satisfy this system. This completes the proof.■

4.2 Transfers that only depend on actions

One shortcoming of the transfer mechanisms described above is that the compensation needs to depend on the state x. In some applications, this may be problematic if the time difference between policy making and observation of x is large bringing into question the assumption that the participants will survive until x is realized. Alternatively, it is not always reasonable to assume that x is ever observed. Using the logic in the proof of proposition 4, I can construct a mechanism in which the transfers do not depend on x. Instead of betting on x participants bet on each others bets.

Proposition 6 *For a fixed k with $n \geq k \geq \frac{n+1}{2}$ there exist a pair of values ξ, ψ s.t. the mechanism that selects $p^+(\cdot)$ and distributes transfers*

$$t_i'(m^s) = \begin{cases} \frac{v_a}{\xi} - \psi \textit{ if } m_i^s = a \textit{ and } \#\{j : m_j^s = a\} \geq k \\ \frac{v_b}{\xi} - \psi \textit{ if } m_i^s = b \textit{ and } \#\{j : m_j^s = b\} \geq k \\ -\psi \textit{ otherwise} \end{cases} \tag{33}$$

is Bayesian Nash incentive compatible and satisfies expected budget balance.

Proof: Suppose again that policy is given by $p^+(\cdot)$, but this time the transfers are given by

$$t_i(m_i^s, x) = \begin{cases} u_a \text{ if } m_i^s = a \text{ and } \#\{j : m_j^s = a\} \geq k \\ u_b \text{ if } m_i^s = b \text{ and } \#\{j : m_j^s = b\} \geq k \\ 0 \text{ otherwise} \end{cases} \tag{34}$$

with $n \geq k \geq \frac{n+1}{2}$. (A natural example to keep in mind is $k = \frac{n+1}{2}$.) Let

$$\xi_c = \sum_{j=k}^{n-1} \binom{n-1}{j} g^j (1-g)^{n-1-j} \tag{35}$$
$$\xi_f = \sum_{j=k}^{n-1} \binom{n-1}{j} (1-g)^j g^{n-1-j}.$$

denote the probability that at least k of the remaining $n-1$ participants have received correct and incorrect private signals respectively. Following the approach used above, I consider the incentive of a type $\theta_i = -1$ agent that has observed $s_i = a$ to reveal (s_i, θ_i) instead of lying about both coordinates if all other agents are truthful. Truthfulness requires that

$$\begin{aligned} \pi g \left(\lambda_a pr_{-1}^+ + u_a \xi_c\right) + (1-\pi)(1-g)(\lambda_b(1-pr_{-1}^+) + u\xi_f) \geq \\ \pi g(\lambda_a(1-pr_{-1}^+) + u\xi_f) + (1-\pi)(1-g)(\lambda_b pr_{-1}^+ + +u_b\xi_c). \end{aligned} \tag{IC-i'}$$

Similarly the incentive compatibility condition for truthfully reporting when $s_i = b$ is

$$\begin{aligned} \pi(1-g)(\lambda_a(1-pr_{-1}^+) + u_b\xi_f) + (1-\pi)g(\lambda_b pr_{-1}^+ + u_b\xi_c) \geq \\ \pi(1-g)(\lambda_a pr_{-1}^+ + u_a\xi_c) + (1-\pi)g(\lambda_b(1-pr_{-1}^+) + u_a\xi_f). \end{aligned} \tag{IC-ii'}$$

By the arguments above these are the binding incentive compatibility constraints. A sufficient condition for incentive compatibility is

$$u_a\xi_c - u_a\xi_f = v_a \tag{36}$$
$$u_b\xi_c - u_b\xi_f = v_b.$$

This means that for fixed k and thus (η_c^k, η_f^k) setting

$$u_a = \frac{v_a}{\xi_c - \xi_f} \tag{37}$$
$$u_b = \frac{v_b}{\xi_c - \xi_f}$$

will result in a transfer scheme that implements $p^+(\cdot)$.

As before, I can modify the transfers to satisfy expected budget balance. Namely I can set

$$\psi = \frac{\left[\eta_{-1}^{-} - \eta_{-1}^{+}\right](\lambda_a(\pi\xi_c + (1-\pi)\xi_f + \lambda_b((1-\pi)\xi_c + \pi\xi_f)}{\xi_c - \xi_f} \tag{38}$$

to be the expected average transfer and set $\xi = \xi_c - \xi_f$ and use the transfer schedule in (34) to satisfy expected budget balance.■

Corollary 6 *This mechanism can be decentralized by first giving each participant the choice of two securities (each at a price ψ), one that pays off $\frac{v_a}{\xi}$ iff at least k others choose this security and one that pays $\frac{v_a}{\xi}$ iff at least k others choose the security, and then letting participants vote after observing the payoff from their security. Note that the term ξ depends on k. In this game there is a perfect Bayesian Nash equilibrium in weakly undominated voting strategies in which the policy mapping is $p^+(\cdot)$.*

Proof: The claim follows from propositions 1 and 6.■

Remark 3 *One example would be endogenous affiliation with parties or clubs that distribute a subsidy to members if the club membership is sufficiently large. With $k = \frac{n+1}{2}$ the interpretation is of a contest to join one of two clubs and a prize is distributed to the bigger club. The multiplicity of equilibria problem is most striking in the mechanism of proposition 6 which also has pooling equilibria in which all participants send the same message about s_i.*

5 Discussion

Recent studies of strategic behavior in institutions with communication and voting are less sanguine about the effectiveness of deliberation in settings with private beliefs and values than the traditional literature on deliberative democracy. The presence of exclusive private information and the possibility that a participant will believe that she has minority interests renders institutions that do not involve transfers incapable of always selecting the first-best policy that a majority would enact if all of the private information were public. However, in groups that are likely to be more homogenous with respect to preference types or information sources transfers are not needed.[15]

If the institutions are allowed to make transfers to individuals then an institution designer can create incentives for participants to share their private information. These transfers need not be very large at all. Moreover, it is not necessary to condition the transfers on the realized state. Asking participants to forecast each other's forecasts can be as informative as asking them to forecast the state. The structure of institutions that work well is amenable to some simple decentralizations. Examples include the sale of state contingent securities at fixed prices, trading of state contingent and risk free securities at market determined prices, and the creation of clubs (or parties) that distribute member subsidies according to membership.

These findings offer guidance into what is needed for a justification of efficient information sharing in deliberative settings by strategic participants. Expectations of full preference and information aggregation require strong assumptions about commonality of interests, information structures in which no information is privately possessed, or the presence of externally motivated incentives. While the reply of a dedicated deliberative democrat may be that "participants in ideal deliberation will value truthfulness and thus be opposed to lying", the analysis sheds light on when aggregation may not work well and how institutions can be amended to improve aggregation in the presence of participants that are less noble than the "ideal" of deliberative democrats and more akin to Madison's expectation "–more disposed to vex and oppress each other than to cooperate for their common good."

[15] While, the analysis demonstrates that preference divergence can destroy incentives for information sharing, the analysis also sheds light on a debate among normative theorists about the idea of a "common good". The weaker ideal that all participants think their conception of the good is shared by a majority turns out to be sufficient for information sharing. This condition seems less problematic and more likely to be satisfied in interesting policy making settings.

6 Appendix

First, the concept of a direct coordination mechanism is defined.

Definition 6 *Given the voting game* Γ^v *a direct coordination mechanism is a multi stage game: first each participant simultaneously submits a message* $m_i \in \{-1,1\} \times \{a,b\}$ *to a mediator, next the mediator makes a private recommendation* $r_i(\cdot) : \{-1,1\}^n \times \{a,b\}^n \to [0,1]$ *(with* r_i *interpreted as the proposed probability that* i *should vote for* a*) to each participant, and finally each participant selects a ballot* $v_i \in \{a,b\}$ *(or randomizes), the policy is chosen by majority rule, and utility is assigned based on the outcome,* x*, the policy,* p*, chosen (and if transfers are involved, the transfers* $t_i(m_i)$*).*

For any ϕ, let $p^c(\phi)$ denote the probability that a is chosen when participants truthfully reveal their types and adhere to the recommendations and let $p_i^c(m_i, \phi_{-i}, v_i(r_i(m_i, \phi_{-i}))$ denote the probability that a is chosen when $N \backslash i$ are truthful and adhere to the recommendations but i reveals m_i and responds to recommendation $r_i(m_i, \phi_{-i})$ by selecting $v_i(r_i(m_i, \phi_{-i}) \in \{a,b\}$.

Definition 7 *A direct coordination mechanism* $(r_1(\cdot), ..., r_i(\cdot), ..., r_n(\cdot), t_1(\cdot), ..., t_i(\cdot), ..., t_n(\cdot))$ *is perfect Bayesian incentive compatible if for each* $i \in N$ *and each* $\phi_i \in \{-1,1\} \times \{a,b\}$ *and every* $m_i \in \{-1,1\} \times \{a,b\}$ *and every function* $v' : [0,1] \to \{a,b\}$

$$\sum_{\phi_{-i}} \sum_{x} \left[\begin{array}{c} p^c(\phi_i, \phi_{-i}) \, (u_i(a; x, \theta_i) + t_i(\phi_i, \phi_{-i})) + \\ (1 - p^c(\phi_i, \phi_{-i})) \, (u_i(b; x, \theta_i) + t_i(\phi_i, \phi_{-i})) \end{array} \right] pr(x, \phi_{-i} \mid \phi_i) \geq \quad (39)$$

$$\sum_{\phi_{-i}} \sum_{x} \left[\begin{array}{c} p_i^c(m_i, \phi_{-i}, v_i'(r_i(m_i, \phi_{-i}))) \, (u_i(a; x, \theta_i) + t_i(m_i, \phi_{-i})) + \\ (1 - p_i^c(m_i, \phi_{-i}, v_i'(r_i(m_i, \phi_{-i})))) \, (u_i(b; x, \theta_i) + t_i(m_i, \phi_{-i})) \end{array} \right] pr(x, \phi_{-i} \mid \phi_i)$$

Now a useful lemma is established.

Lemma 1 *There exists a Bayesian incentive compatible direct coordination mechanism (with transfers) that implements the choice function* $p^+(\cdot)$ *if and only there exists a Bayesian incentive compatible direct mechanism (with transfers) that implement the choice function* $p^+(\cdot)$*.*

Proof: $(\Longrightarrow)$Assume that there is a direct coordination mechanism (with transfers) $\langle r(\cdot), t(\cdot) \rangle$ that implements $p^+(\cdot)$. Consider the direct mechanism that mimics the direct coordination mechanism except that instead of making recommendations to participants about how to vote it selects policy according to the mapping

$$p^*(\phi) = \left\{ \begin{array}{c} 1 \text{ if } \ |i : r_i(\phi) = 1| \geq \frac{n+1}{2} \\ 0 \text{ otherwise} \end{array} \right. \quad (40)$$

Furthermore, note that for each $i \in N$, if $v_i'(r_i(m_i, \phi_{-i})) = r_i(m_i, \phi_{-i})$ then $p_i^c(m_i, \phi_{-i}, v_i'(r_i(m_i, \phi_{-i}))) = p^*(m_i, \phi_{-i})$. Fix $v_i'(\cdot)$ to be the identity mapping. In this case the fact that $\langle r(\cdot), t(\cdot)\rangle$ is Bayesian incentive compatible (and thus satisfies (3)) implies that for each $i \in N$ and each $\phi_i \in \{-1, 1\} \times \{a, b\}$ and every $m_i \in \{-1, 1\} \times \{a, b\}$

$$\sum_{\phi_{-i}} \sum_{x} \left[\begin{array}{c} p^c(\phi_i, \phi_{-i}) \left(u_i(a; x, \theta_i) + t_i(\phi_i, \phi_{-i})\right) + \\ (1 - p^c(\phi_i, \phi_{-i})) \left(u_i(b; x, \theta_i) + t_i(\phi_i, \phi_{-i})\right) \end{array} \right] pr(x, \phi_{-i} \mid \phi_i) \geq \tag{41}$$

$$\sum_{\phi_{-i}} \sum_{x} \left[\begin{array}{c} p^c(m_i, \phi_{-i}) \left(u_i(a; x, \theta_i) + t_i(m_i, \phi_{-i})\right) + \\ (1 - p^c(m_i, \phi_{-i})) \left(u_i(b; x, \theta_i) + t_i(m_i, \phi_{-i})\right) \end{array} \right] pr(x, \phi_{-i} \mid \phi_i)$$

which is just the incentive compatibility condition for the direct mechanism.

($\Longleftarrow$) Assume that there is a direct mechanism (with transfers) $\langle p(\cdot), t(\cdot)\rangle$ that implements the choice function $p^+(\cdot)$. Consider the direct coordination mechanism that distributes $t(\cdot)$ and makes the following recommendation

$$r_i(\phi) = \begin{cases} p(\phi) \text{ if } \sum_x p(\phi) u_i(a, x, \theta_i) pr(x \mid \phi) \geq \sum_x (1 - p(\phi))\, u_i(a, x, \theta_i) pr(x \mid \phi) \\ 1 - p(\phi) \text{ if } \sum_x p(\phi) u_i(a, x, \theta_i) pr(x \mid \phi) < \sum_x (1 - p(\phi))\, u_i(a, x, \theta_i) pr(x \mid \phi) \end{cases} \tag{A1}$$

Since $\langle p(\cdot), t(\cdot)\rangle$ implements the choice function $p^+(\cdot)$, and $r_i(\phi)$ suggests that i vote for the policy that $p^+(\phi)$ chooses iff this policy maximizes i's expected utility given ϕ, the decisiveness of a majority under majority rule implies that: (i) if everyone adheres to the recommendation then the lottery over $\{a, b\}$ corresponds to $p^+(\phi)$ for every ϕ and (ii) if everyone is truthful then following $r_i(\phi)$ is a best response in weakly undominated strategies for every ϕ. By (i) and the fact that $\langle p(\cdot), t(\cdot)\rangle$ is Bayesian incentive compatible no one has a unilateral incentive to send $m_i \neq \phi_i$. By (ii) if everyone is truthful then no one has a unilateral incentive to deviate from $v_i = r_i(\phi)$. To check that no agent-type has a unilateral incentive to deviate with $v_i \neq r_i(\phi)$ and $m_i \neq \phi_i$ simultaneously, note that if this deviation results in a different lottery over utility then a deviation in only m_i or v_i then it must be the case that $|j : \theta_j = \theta_i| > \frac{n+1}{2}$ and either $\mu(m_i, \phi_{-i}) < \frac{1}{2} < \mu(\phi)$ or $\mu(m_i, \phi_{-i}) > \frac{1}{2} > \mu(\phi)$. But the first condition implies that $p(\phi)$ is optimal for i and thus no deviation can be desirable.■

A few observations are in order. First, for arbitrary choice functions and arbitrary voting games it is not generally the case that implementability in direct mechanisms implies implementability in direct coordination mechanisms. This result hinges on the fact that $p^+(\cdot)$ is the full information majority rule core and the direct coordination games considered build on Γ^v which uses majority rule. The argument does not involve appeal to strategies that are weakly dominated in the voting stage. Gerardi and Yariv (2004) establish an equivalence across choice functions that are implementable in direct communication mechanisms that use non degenerate voting rules (all quota rules other than unanimity). Their construction hinges on giving all agents the same recommendation and then noting that no agent is pivotal. In contrast the construction used to establish lemma 1 does not hinge on strategies that are optimal simply because pivot probabilities are 0. Here, in the direct coordination game that mimics a particular direct mechanism each agent's recommendation is a weakly undominated strategy in the voting stage. The proof of proposition 1 builds on lemma 1.

Proof of Proposition 1: Assume that a Bayesian incentive compatible direct mechanism implements $p^+(\cdot)$. By lemma 1, a direct coordination mechanism exists that has a Bayesian Nash equilibrium in which choice corresponds to $p^+(\cdot)$. Let $r_i(\cdot)$ denote the recommendation function of this direct coordination mechanism. Note (from the proof above) that $r_i(\phi)$ depends only on s and θ_i. Consider the direct public communication mechanism in which the transfer function is identical to that in the incentive compatible direct coordination mechanism. I characterize a perfect Bayesian Nash equilibrium in weakly undominated voting to the direct public communication mechanism. Consider the strategy $m_i(\phi_i) = \phi_i$ and $v_i(m^s, \theta_i) = r_i(m)$ where the latter is feasible since (as noted) $r_i(\phi)$ depends only on s and θ_i. Finally consider the belief $\mu(x \mid m^s)$ that uses Bayes' rule given the conjecture that m^s is generated by truthful messages. Given this message function the beliefs satisfies Bayes' rule for every possible m^s. Given the belief, the voting rule selects i's expected utility maximizing alternative because in (A1) $r_i(\cdot)$ is constructed to select the best alternative for i given ϕ. Thus, the voting strategy is sequentially rational. Given $v_i(\cdot)$ and truthful messages by $N\backslash i$, the mapping from m_i to expected utility is the same in this equilibrium as in the direct mechanism and thus, since by assumption the direct mechanism is Bayesian incentive compatible $m_i(\phi_i) = \phi_i$ is a best response and thus sequentially rational.■

References

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

[2] Austen-Smith, David, and Timothy Feddersen. 2003a. "Deliberation and Voting Rules." Northwestern University Typescript.

[3] Austen-Smith, David, and Timothy Feddersen. 2003b. "The Inferiority of Deliberation under Unanimity Rule." Northwestern University Typescript.

[4] Banks, Jeffrey. 1990. "Equilibrium Behavior in Crisis Bargaining Games." *American Journal of Political Science* 34(3):599-614

[5] Condorcet, Marquis de. [1785] 1976. "Essay on the Application of Mathematics to the Theory of Decision-Making." Reprinted in *Condorcet: Selected Writings*, Keith Michael Baker, ed., 33-70. Indianapolis: Bobbs-Marrill Co.

[6] Coughlan, Peter, J. 2000. "In Defense of Unanimous Jury Verdicts: Mistrials, Communication, and Strategic Voting." *American Political Science Review* 94:375-393

[7] Duggan, John, and César Martinelli. 2001. "A Bayesian Model of Voting in Juries." *Games and Economic Behavior* 37:259-94.

[8] Barabas, Jason. 2004. "How Deliberation Affects Policy Opinions." *American Political Science Review* 98(4):687-702.

[9] Baron, David P. 2000. "Legislative Organization with Informational Committees." *American Journal of Political Science* 44: 506-522.

[10] Baron, David, P., and Adam Meirowitz. 2004. "Fully revealing equilibria of multiple-sender signaling and screening models" *Social Choice and Welfare*, forthcoming.

[11] Chen, Kay-Yut, Leslie R. Fine, and Bernardo A. Huberman. 2003. "Eliminating Public Knowledge Biases in Small Group Predictions." *Management Science,* 50(7):983-994.

[12] Chen, Kay-Yut, and Charlie Plott. 2002. Information Aggregation Mechanisms: Concept, design and implementation for a sales forecasting problem. *Social Science Working Paper 1131. California Institute of Technology.*

[13] Cohen, Joshua. 1997. "Procedure and Substance in Deliberative Democracy" in James Bohman, and William Rehg, eds., *Deliberative Democracy: Essays on Reason and Politics*, Cambridge: MIT Press

[14] Dahl, Robert.A. 1989. *Democracy and its Critics.* New Haven: Yale University Press.

[15] Dasgupta, P., P. Hammond, and E. Maskin. 1979. "The Implementation of Social Choice Rules: Some general results on incentive compatibility." *Review of Economic Studies.* 46:185-219.

[16] Dubey, Pradeep, John Geanakopolis, and Martin Shubik. 1987. "The revelation of information in strategic market games." *Journal of Mathematical Economics* 16:105-137.

[17] Feddersen, Timothy, and Wolfgang Pesendorfer. 1997. "Voting Behavior and Information Aggregatin in Elections with Private Information" *Econometrica* 65:1029-58.

[18] Feddersen, Timothy, and Wolfgang Pesendorfer. 1998. "Convicting the Innocent: the Inferiority of Unanimous Jury Verdicts." *American Political Science Review* 92:23-35.

[19] Fishkin, James S. 1991. *Democracy and Deliberation: New Directions for Democratic Reform*. New Haven: Yale University Press.

[20] Gerardi, Dino, and Leeat Yariv. 2003. "Putting Your Ballot Where Your Mouth Is: An Analysis of Collective Choice with Communication." Yale University Typescript.

[21] Gibbard, A. 1973. "Manipulation of Voting Schemes" *Econometrica* 41:587-601.

[22] Glazer, J, and A. Rubinstein. 2001. "Debates and Decisions: on a rationale for argumentation rules." *Games and Economic Behavior* 36:158-73.

[23] Green J.R., and J.-J. Laffont. 1977. "Characterization of Satisfactory Mechanisms for the Revelation of Preferences for Public Goods." *Econometrica* 45:427-38.

[24] Gutmann, Amy, and Dennis Thompson. 1996. *Democracy and Disagreement*. Cambridge: Belknap, Harvard Press.

[25] Hafer, Catherine, and Dimitri Landa. 2003. "Deliberation as Self-discovery." New York University Typescript.

[26] Hamilton, Alexander, James Madison, and John Jay [1788] 1961. *The Federalist: The famous papers on the principles of American government,* ed. Benjamin F. Wright. New York: Metrobooks.

[27] Jackson, Mathew O., and James Peck. 1999. "Asymmetric information in a competitive market game: Reexamining the implications of rational expectations." *Economic Theory* 13:603-628.

[28] Kim, Jaehoon. 2004. "A Model of Adversarial Committees" University of Rochester Typescript.

[29] Krishna, Vijay, and John Morgan. 2001. "Asymmetric Information and Legislative Rules: Some Amendments." *American Political Science Review* 95(2):435-452.

[30] Lipman, Bart, and D. Seppi. 1995. "Robust Inference in Communication Games with Partial Proveability." *Journal of Economic Theory* 66:370-405.

[31] Meirowitz, Adam. 2002. "Informative voting and Condorcet jury theorems with a continuum of types." *Social Choice and Welfare* 19:219-236.

[32] Meirowitz, Adam. 2003. "In defence of exclusionary deliberation: Communication and voting with private beliefs and values" Princeton University Typescript.

[33] Meirowitz, Adam. 2005. "Deliberation and bargaining in the spatial model" Princeton University Typescript.

[34] Meirowitz, Adam, and Joshua Tucker. 2004. "Learning from terrorism markets." *Perspectives on Politics* 2(2):331-336.

[35] Mendelberg, Tali. 2002. "The Deliberative Citizen: Theory and Evidence." *Political Decision Making, Deliberation and Participation* 6:151-93.

[36] Myerson, R. 1979. "Incentive Compatibility and the Bargaining Problem." *Econometrica* 47:61-73.

[37] Myerson, Roger. 1982. "Optimal coordination mechanisms in generalized principal-agent problems." *Journal of Mathematical Economics* 10:67-81.

[38] Neumann, George. 2005. "Operating With Doctors: Results From the 2004 and 2005 Influenza Markets". University of Iowa Typescript.

[39] Palfrey, Thomas R. and Sanjay Srivastava. 1987. "On Bayesian Implementable Allocations." *Review of Economic Studies* 54: 193-208.

[40] Palfrey, Thomas R. and Sanjay Srivastava. 1989. "Mechanism design with incomplete information: A solution to the implementation problem." *Journal of Political Economy* 97(3): 668-691.

[41] Postlewaite, Andrew, and David Schmeidler. 1986. "Implementation in Differential Information Economies." *Journal of Economic Theory* 39: 14-33.

[42] Shapely, Loyd, and Martin Shubik. 1977. "Trade using one commodity as a means of payment." *The Journal of Political Economy* 85(5):937-968.

[43] Sunstein, Cass R. 1988. "The Republican Civic Tradition: Beyond the Republican Revival." *Yale Law Journal* 97(July):1539-90.

[44] Williams, David. 1997. *Probability with Martingales.* Cambridge: Cambridge University Press.

[45] Witt, Jörgan. 1998. "Rational Choice and the Condorcet Jury Theorem." *Games and Economic Behavior* 22:364-376.

[46] Wolfers, J. and E. Zitzewits. 2004. "Prediction Markets." *Journal of Economic Perspectives* 18(2):107-126.