Polarization and Publicity:
Rethinking the Benefits of Deliberative Democracy

David Stasavage
Department of Politics
New York University
david.stasavage@nyu.edu

June 2006¹

¹I would like to thank Gilat Levy, Bernard Manin, Dimitri Landa, Ken Scheve, Thierry Verdier, the editor, four anonymous referees, and seminar participants at the Polarization and Conflict conference in Oslo and the DELTA (Paris) for comments and suggestions. A previous draft of this paper was circulated under the title "Public versus Private Deliberation in a Representative Democracy".
**Abstract**

Though openness in government has obvious benefits, recent scholarship has devoted less attention to the possibility that it might also have costs. I use a formal framework to investigate the effect of public versus private decision-making on opinion polarization. Existing work emphasizes that public debate helps to reduce polarization and promote consensus, but I argue that when debate takes place between representatives the opposite may be true. When representatives make decisions in public, they face incentives to use their actions as a signal of loyalty to their constituents, potentially ignoring private information about the true desirability of different policies. Anticipating this, constituents will not alter their prior policy beliefs following a debate of this type. When representatives instead make policy decisions in private, they are more likely to allow private information to influence their actions. An important consequence is that even if constituents do not observe actions or statements of individual representatives, they can still use the final policy choice to revise their initial beliefs. I suggest that these conclusions have significant implications for both the literature on deliberative democracy and for discussions of polarization in American politics.
There are strong reasons to believe that openness in government is critical for the health of any democracy. When the public can directly observe the actions of its representatives it will find it easier to hold them accountable. Decisions taken in public may also be seen as more legitimate. Given these potential benefits, over the last forty years government agencies in many different countries have taken steps to make themselves more “transparent”, and outside observers have increased pressures for openness in areas where there remains a perceived “democratic deficit”. In this paper I ask how openness in government is likely to influence polarization of opinions between citizens. I develop a model that compares public and private decision making by a body of three representatives. Under public decision making, constituents observe both the final policy choice and the votes of individual representatives. Under private decision making, constituents observe only the final policy choice.

The representatives in the model could be elected members of a legislature, unelected members of a committee, like a central bank governing board or constitutional court, or they could also be members of an international body. I assume that representatives have some degree of private (i.e. expert) information about the effect of different policies. Representatives must vote on a binary action with the knowledge that utility for the public will depend on whether this action corresponds to the realization of an unobserved “state of the world”. To take an example, if the policy in question involves redistribution of income then the “state of the world” could involve uncertainty about the extent to which high taxes will create a disincentive to work (as in Roemer, 1994) or uncertainty about the extent to which individual income levels depend on effort vs. exogenous circumstances (as in Piketty, 1995). There are two types of representative, “unbiased” representatives who share the same utility function as the public and “biased” representatives who always prefer the same action regardless of the realization of the state variable. One prominent reason why a representative might be biased is if their policy preferences are influenced by a special interest group. In the case of income redistribution the special interest group might be composed of individuals who are convinced that a lower or higher level of redistribution is desirable, irrespective of the underlying circumstances involving incentive effects and of effort vs. exogenous circumstances.
Representatives in the model are concerned both about choosing their preferred policy and about developing a reputation for being unbiased.

My conclusions have significant implications for the literature on deliberative democracy. Advocates of public deliberation like Habermas (1996), Bohman (1998), Fishkin (1991), Guttmann and Thompson (1996), and Bentham (1816) emphasize that it can help to achieve greater consensus, as there will be less polarization in terms of opinions after a public discussion. Other authors like Sanders (1997) have been more skeptical. She suggests that the appeal to deliberation has provided an “all-purpose solution” around which both those who think that democratic problems derive from an absence of mass participation and those who think members of the public need better information before making decisions can converge. In contrast to the advocates of deliberative democracy, Sunstein (2002) emphasizes that deliberation may actually lead to an increased polarization of beliefs between groups. This phenomenon of group polarization can involve both a tendency to defer to individuals with more extreme views, as well as a tendency for groups to begin discussions with a narrow range of opinions. Zaller (1992) also considers how greater provision of information about a policy can lead to increased polarization of public opinion. My paper differs from Sunstein’s and Zaller’s contributions in emphasizing how polarization, or more specifically the failure of deliberation to reduce polarization of beliefs, results from the dynamics of political representation. Publicity of debate may prompt representatives to use their actions or statements as signals that they are being faithful to constituent interests. Under these conditions, publicity has the advantage of disciplining representatives, but it also implies that constituents will learn little from a debate in which representatives always take actions or make statements that are expected of them.

My conclusions are also relevant to current discussions about polarization in American politics. In recent years it has been common to hear suggestions that political views in the US are increasingly polarized. A number of scholars have considered this empirically,

---

They are also relevant to several formal contributions on deliberation including Austen-Smith and Feddersen (2006) and Meirowitz (2005a, 2005b) who consider deliberation between actors who exchange private information, as well as Hafer and Landa (2005) and Calvert and Johnson (1998) who consider deliberation without an exchange of information.
providing strong evidence of polarization on the part of political elites, while the assertion of polarization on the part of the broader public seems more debatable.\(^2\) I contribute to this debate by providing a way to address the following puzzle. According to standard accounts, openness in government and publicity of debate should help produce political consensus. By these standards, given the institutional changes to increase openness that have taken place at different levels of government in the United States over the last forty years, we would logically expect polarization to have decreased. If the opposite has taken place, then this may be due in part to the fact that greater publicity of debate between representatives has actually reinforced polarization.

The model I develop draws on several recent formal contributions that consider how an agent may have an incentive to withhold private information about a policy choice if doing otherwise would lead a principal to draw negative inferences about the agent’s type. Maskin and Tirole (2004) compare the relative advantages of having elected vs. unelected representatives choose policies when representatives have expert information about policy choices and there is a risk that representatives are biased. Morris (2001) shows how a “political correctness effect” may exist whereby an advisor gives a false recommendation to a principal in order to avoid having the principal draw the inference that the advisor is “biased”.\(^3\) These authors do not compare private vs. public settings for decision making by representatives (agents).

My paper is also directly related to several recent papers that compare how more transparent and less transparent decision making environments alter incentives for agents. Visser and Swank (2006) have recently presented a model of committee decision making where committee members have varying levels of expertise, and they have an opportunity to engage in communication before a vote. These authors show that if committee members are obliged to make decisions in public then this can minimize incentives for members to distort their private information in order create a reputation for having a high level of expertise. Levy (2005)

\(^2\)See Fiorina (2005), Layman and Carsey (2002), and McCarty, Poole, and Rosenthal (2001).

\(^3\)In an earlier contribution, Austen-Smith (1992) examined a similar problem within the context of voter control over biased and unbiased legislators.
examines decision making by a committee with transparency, where members of the public observe both individual votes and the policy outcome, and without transparency, where the public observes only the outcome. In her model as well, committee members have varying levels of expertise. She demonstrates that secretive procedures may be preferable, but this itself depends on other elements of the decision making process. In particular if the process favors one outcome over another, say because a supermajority is required to overturn a status quo, or if one action is favored by a prior view, then transparency is only beneficial if it is optimal to minimize these other effects, for example so as to facilitate reforms or radical decisions.  

The key difference between the current paper and the contributions by Levy (2005) and Visser and Swank (2006) is that while they consider a scenario where committee members (representatives) have varying levels of expertise, and thus an incentive to develop a reputation for expertise, I consider a scenario where different types of representatives have different preferences, with some biased and some unbiased. As a result, instead of having an incentive to build a reputation for expertise, representatives have an incentive to develop a reputation for having preferences that are congruent with those of constituents. When we consider democratic government it seems clear that members of the public may care about a representative’s level of expertise, but problems of accountability as they have been considered in seminal works like Bentham (1816) also center on the question of whether a representative has similar preferences to those of the public or is biased. As a result, if we want to consider the importance of transparency or publicity in democratic government then we also need to consider a model where representatives have private information but they may have policy preferences that differ from those of their constituents in the same way that has been considered formally by Maskin and Tirole (2004).

In the model developed in this paper the key question for both biased and unbiased repre-

4In a model where there is an individual agent rather than a committee Prat (2005) has demonstrated how non-transparency may actually be optimal in a principal-agent relationship if the consequences of an agent’s actions are difficult to evaluate. Fox (2004) also considers advantages and disadvantages of transparency in a similar formal setup to that used here, but for an individual decision maker. See also Canes-Wrone, Herron, and Shotts (2001), Ferejohn (1999), and Groseclose and McCarty (2001).
sentatives involves their relative incentive to vote in favor of their preferred policy outcome, or alternatively, to use their vote as a signal to convince the public they are unbiased. This relative incentive depends directly on whether decision making occurs in public or in private. Under public decision making, representatives know that their actions will have a certain impact on their reputation but an uncertain impact on the policy outcome, because their vote only changes the policy outcome if it turns out that they are pivotal. In contrast, under private decision making, because the public does not directly observe votes, representatives know that their vote will only have an impact on either their reputation or the policy outcome if they are pivotal. Following this logic I establish the conditions for two types of equilibria: (1) “informative equilibria” which are defined as equilibria where at least one type of representative conditions their vote on their private information about the state of the world and (2) “uninformative” equilibria where no representative conditions their vote on their private information. The word “informative” here refers to the fact that for members of the public, actions taken by representatives provide information about the true state of the world (the optimal policy).

I demonstrate that an informative equilibria is more likely to exist under private, as opposed to public, decision making, because private decision making reduces incentives for representatives to use their vote as a signal that they are unbiased. One important implication of this is that private decision making may, in many instances, actually do more to reduce polarization of opinions in society than will public decision making. This runs contrary to the common suggestion that public discussions will produce greater social consensus. When members of society have divided opinions about the effects of a policy, such as redistribution of income, then in an uninformative equilibrium representatives will articulate the opinions of their constituencies, but the public will not actually learn anything from observing public decision making, because it knows that representatives are not conditioning their actions on their private information about which policy is optimal. In contrast, when an informative equilibrium exists, which is more likely to be the case under private decision making, then even if it does not observe the actions of individual representatives, the public will know
that the policy outcome has depended upon the private information held by representatives. As a consequence, members of the public will be able to draw inferences from the policy outcome, even if these inferences are necessarily imperfect, and they will revise their beliefs about which policy outcome is preferable. If beliefs are initially polarized, then they will tend to converge.

In the main part of the paper, in order to present the model and analyze the effects of transparency in a concise fashion, I assume that decision making takes place in a single stage - representatives vote and the policy outcome is decided by simple majority. However, since the literature on deliberation has emphasized the importance of deliberation as discussion prior to a vote, I then consider an expansion of the model in which representatives exchange messages about desirable policy before voting. Different possibilities are analyzed where either the message stage or the voting stage take place in public, neither takes place in public, or one but not the other take place in public. My conclusions from this exercise are similar to those from the more simple model where deliberation and voting are compressed into a single stage. When the message stage in particular takes place in public this can make it less likely that we will observe an informative equilibrium in which the actions of representatives allow constituents to revise their beliefs about optimal policies. The key difference, however, is that as long as the message stage occurs in private, because this allows representatives to exchange information privately, it may make less difference for constituents whether the voting stage occurs in private or in public.

In what follows I first discuss the relevance of this paper to different literatures that have considered transparency in politics, as well as to broader debates about representative government. I then provide the setup for my model, followed by consideration of public decision making and private decision making. I then extend the model by including an explicit deliberation stage, followed by a general conclusion.
Representation and public decision making

My consideration of the costs and benefits of public decision making touches on a long-standing discussion among democratic theorists about how closely representatives should be bound to their constituents. According to the classic study by Pitkin (1967), modern political representation implies that representatives should act according to the expressed interests of their constituents, but they should also have some latitude to choose alternative actions if they are convinced that the public is mistaken on a particular issue. This view can be distinguished both from theories that emphasize that representatives should act as delegates with mandates, as well as from more “elitist” views of representative democracy, such as those provided by Schumpeter (1942) and Burke (1777 [1963]) who recommended a weaker link between popular desires and choices of representatives. In her work Pitkin (1967) also emphasizes that views on representation will depend on whether one assumes that the questions with which a representative must deal are scientific in form, with objective answers, or alternatively, are based primarily on value judgements. In the former case there is a stronger argument for a weak link between representative and represented.

Transparency is relevant to this fundamental debate in democratic theory, because openness influences decisions by representatives to act according or contrary to popular opinion. As described by Pole (1983), the development of modern ideas about the need for representatives to take account of public opinion went hand in hand with a tendency for bodies like the British House of Commons to shift from closed-door to open-door debates. When the public is able to directly observe what a representative says during deliberations and how he or she votes with regard to policy, then as long as the representative is concerned about maintaining a reputation for acting in the public interest, he or she is more likely to support policies preferred by the public. As one of the most unequivocal early proponents of transparency, Jeremy Bentham (1816) argued that in a parliamentary context, publicity would “constrain members of the assembly to perform their duty”. The potential problem with transparency

5Transparency should also relevant to debates about representation in non-democratic contexts. See Rehfeld (2006) for an interesting effort to develop a more general theory of representation.
is that it can also prompt representatives to ignore any private information they may have which indicates that the public is misinformed. Among nineteenth century observers, John Stuart Mill was a prominent critic of Bentham’s political propositions, suggesting that they went too far in “riveting the yolk of public opinion closer and closer round the necks of all public functionaries” thus excluding the possibility that a representative might use his or her own reason in making a decision. The idea that publicity might have a negative impact on quality of debate can in fact be found in contributions as early as Hobbes (1651, ch.19) who contrasted the sort of confidential consultations that could be undertaken by a monarch with the public consultations of a representative assembly where “long discourses” may “commonly excite men to action but not govern them in it.”

While it has long been suggested that transparency in government can involve costs as well as benefits, recent discussions have focused almost exclusively on the latter phenomena. A prominent example here involves the literature on “deliberative democracy”. Public deliberation, it is suggested, can help transform individual views in order to increase consensus, and it can improve the quality and the legitimacy of decisions taken. What has been less frequently asked by theorists of deliberative democracy is, in the case where deliberation occurs between representatives, whether publicity might prompt participants to refrain from expressing their true opinions, in which case one of the principal goals of deliberation, improving the quality of decisions, is undermined. These issues have often been overlooked despite clear empirical observations of such effects. So, for example, several scholars of committee politics within European Union institutions have suggested that frank exchanges of information and opinions are more likely to occur on EU committees that meet in private, while in EU committees that hold public sessions there is a tendency for members to present pre-prepared speeches and to engage in little real debate (Checkel, 2005). Likewise, a number of observers have argued that the secrecy rule which prevailed during the US Constitutional Convention of 1787 allowed for much more real debate than would have otherwise been the

---

6 For Mill’s opinions on Bentham see Mill (1838, pp.87-88).
case. This was an opinion shared by James Madison.\footnote{James Madison as reported in an interview by Jared Sparks, 1830, cited in Max Farrand’s The Records of the Federal Convention of 1787, vol.III, p.478.} Finally, a recent empirical study by Steiner \textit{et al} (2004) of deliberation in several national legislatures has also produced findings that greater secrecy of legislative sessions is associated with a higher quality of debate.

Two contributions that do consider both the costs and the benefits of public deliberation have been made by Elster (1998, 1991) and Manin (1997).\footnote{See also the study by Chambers (2004), who considers the effect of deliberating in closed-door vs. open-door environments, chapter 3 of Guttman and Thompson (1996), and Barabas (2004).} For Elster public deliberation produces a benefit he refers to as “the civilizing effect of hypocrisy”; when debate takes place in front of an audience, participants will prefer to couch their arguments in terms of appeals to the general interest rather than their own self-interest. While this civilizing effect of hypocrisy may indeed operate, I argue that public debate can also have an informational downside, as constituents will learn little from a debate in which they know in advance what stance participants will take. Elster also emphasizes that public deliberation can have negative consequences, prompting representatives to posture and to refrain from offering opinions out of the fear of being subsequently proven wrong. Manin (1997 p.168) suggests that the public may actually have a better ability to form opinions if deliberations of representatives initially occur in private, followed by a broader public discussion. He contrasts the polarization and posturing that followed debates of the French Constituent Assembly of 1789, which occurred in public from the outset, with the ratification debate over the US Constitution in 1787, which followed the recommendation of a Convention that had met in secret.

In addition to overlooking the potential costs of holding deliberation in public, recent discussions have also tended to overlook the possibility that transparency may be positively related to polarization in politics. In the US context there has been a vigorous debate about the extent to which political positions of political elites and ordinary citizens are more polarized today than in the past. The evidence seems to suggest that this is an accurate description of American political representatives, but the claim that the general public is increasingly polarized is more debatable.\footnote{See Fiorina (2005, 2002), McCarty, Poole and Rosenthal (2001), and Layman and Carsey (2002).} The apparent paradox of increased polarization
in American politics at the same time that political institutions have become more open has not received significant attention, with the prominent exception of Fiorina (2005, 2002). He argues that contrary to what many would suggest, institutional changes over the last forty years to make American political institutions more open and more participatory have actually contributed to polarization in politics, because new possibilities for participation have been seized upon most eagerly by those with more extreme political views. Within this list of institutional changes, Fiorina (2005) specifically includes “open meetings”, “recorded votes”, and “open bureaucracy”. The model I will present in the following sections suggests a related way in which increased openness may have had a polarizing effect by increasing incentives for representatives to use their statements or votes as signals of loyalty to particular constituencies. It is interesting to contrast Fiorina’s discussion of increased openness and elite polarization in American politics with the frequently cited account by Lijphart (1968) of the institutional foundations of “the politics of accommodation” in the post-war Netherlands. According to Lijphart, secrecy was crucial to consensual politics within the Dutch system precisely because during negotiations it shielded political leaders from demands by constituents; this is essentially the opposite trend from that described by Fiorina for the contemporary United States.

**Setup of the model**

In order to explore incentives under public and private decision making, this and the subsequent sections present a model of decision making by a representative body. For the moment, I restrict attention to a case where representatives vote on policy, but there is no formal round of deliberation preceding the vote.

Three representatives must vote by majority rule on a binary action \( y \in \{0, 1\} \) with individual votes denoted by \( x_i \). It is common knowledge that for each representative, with probability \( p \) they are “unbiased”, and with probability \((1 - p)\) they are “biased”. As suggested above, the most prominent reason why a representative might be biased is if their preferences are influenced by a special interest group whose lobbying is not fully observable
to the public. Utility for the public and for unbiased representatives depends upon choosing the action that will match an unobserved state variable $\omega \in \{0, 1\}$. Both the public and representatives have prior beliefs that the two outcomes for the state $\omega$ are equally likely.

I assume that biased representatives always prefer the action $y = 1$, regardless of the state, and each representative receives a private signal about the state, which is accurate with probability $q$ with $1 > q > 0.5$. The signals are assumed to be independent conditional on the state $\omega$. The private signal might be the results of a statistical study or an economic forecast that the public would lack the means to critically evaluate. It could also involve a legal opinion that makes reference to past cases with which the public is not familiar, or a historical analogy with which the public is not familiar. The key point is that the signal cannot be independently verified by members of the public. The existence of this private information also implies that though “unbiased” representatives and their constituents share the same utility function, unbiased representatives may have different beliefs about which action $y$ is most likely to maximize utility for constituents.

In order to make the issue of accountability relevant, I assume that representatives earn utility based on both the policy outcome and on a reputational payoff. Following several papers in the economic literature on agents with “career concerns”, the reputational payoff is modeled simply as the public’s posterior belief that the representative is unbiased.\(^\text{11}\) In this specific model it refers to the public’s posterior belief about the representative’s type after it has observed the representative’s vote and those of other representatives $\Pr(u|x_i, x_j, x_k)$ (in the case of public decision making), or the policy outcome $y$ (in the case of private decision making) $\Pr(u|y)$. For an elected representative this reputational payoff would involve the benefit from being reelected if, as seems plausible, representatives with a greater reputation for being unbiased are more likely to be retained by the public. Unelected officials may not have electoral concerns, but there remain a number of obvious factors that might lead them to be concerned about their reputation. These can involve possibilities for dismissal,

possibilities for reappointment, or career concerns involving subsequent employment. The reputational motivation could also involve a more intrinsic need, psychological or other, to retain the esteem of a constituency by adhering to an ideological line.

Utility for an unbiased representative under public decision making is expressed by the state contingent function in (1) below.

\[
u(y, \omega, x) = \begin{cases} 
\alpha(1 - y) + (1 - \alpha) \Pr(u|x_i, x_j, x_k) & \text{if } \omega = 0 \\
\alpha y + (1 - \alpha) \Pr(u|x_i, x_j, x_k) & \text{if } \omega = 1 
\end{cases}
\]

Utility for an unbiased representative under private decision making will be identical with the exception of the fact that the reputational payoff will be based on the outcome \(\Pr(u|y)\) since individual votes are not observed. The exogenous parameter \(\alpha\) represents the relative weight representatives give to the policy outcome, as opposed to the reputational payoff. Utility for a biased representative would be represented by the second of the two equations in expression (1). One final point worth noting about the model is that it assumes for simplicity that constituents form a homogeneous group with the same utility function, so there is no distinction between the interests of specific groups of constituents and “the national interest”. I make this simplification in order to focus on the potential conflicting incentives for representatives when a specific group of constituents believes one policy to be best for them but their representative believes an alternative policy is best for this same group. This captures a core part of the debate over representational independence referred to in the previous section, though it does assume away a second part of this debate which involved the question whether representatives should pursue the interests of their specific constituents or the national interest.

Given the above assumptions, the game proceeds in the following four stages

1. Nature determines whether each representative is biased, and the state \(\omega \in \{0, 1\}\) is realized but not observed.

2. Each representative receives a private signal \(s_i\) about the state.
3. Representatives vote \( x_i \in \{0, 1\} \), and the policy outcome \( y \) is decided by majority. Under public decision making outsiders can observe both individual votes and the outcome. Under private decision making outsiders observe only the outcome \( y \).

4. Each representative receives a payoff based on the policy outcome, as well as a reputational payoff based on the posterior probability that she is unbiased.

In this game representatives face a potential trade-off between choosing the action they believe will result in the most favorable policy outcome for them, or instead using their vote to signal to the public that they are unbiased. I consider both whether it is possible to have “informative” equilibria, defined as those where at least one type of representative conditions their vote on the private information they hold about the state variable \( \omega \), as well as “uninformative” equilibria, those where representatives do not condition their votes on their private signal. In an informative equilibrium members of the public will be able to update their prior beliefs about the state of the world, either directly by observing representative actions, or indirectly by observing the chosen policy outcome \( y \). In an uninformative equilibrium members of the public will not be able to use representatives actions to update their beliefs about \( \omega \), because they know that these actions do not depend on representatives private information.

In what follows I investigate the conditions for existence of an “informative” equilibrium, in which unbiased representatives vote conditional on their signal \( x_i = s_i \) while biased representatives always vote \( x_i = 1 \).\(^{12}\) The key condition for existence of this informative equilibrium is that an unbiased representative must prefer to vote \( x_i = 1 \) if her signal is \( s_i = 1 \), rather than ignoring her signal by voting \( x_i = 0 \), which would result in a higher reputational payoff but a poorer expected policy outcome. In what follows I show that this informative equilibrium is more likely to exist under private decision making, and that no other informative equilibria exist with pure strategies (all proofs are presented in the technical appendix).

\(^{12}\)I restrict my attention to pure strategies and to symmetric equilibria where all representatives of the same type follow the same strategy.
In addition to considering existence of “informative” equilibria, I also consider the existence of uninformative equilibria. There exist several types of uninformative equilibria (with proofs again presented in the appendix). It is possible to have a pooling equilibrium where both types of representative always vote 0, in order to avoid appearing biased. As this is a pooling equilibrium where all representatives always take the same action, this equilibrium will normally exist for all parameter values, subject to a plausible restriction on the public’s out-of-equilibrium beliefs. It is also possible for certain parameter values to have an uninformative equilibrium in which biased representatives always vote $x_i = 1$ regardless of the signal they receive and in which unbiased representatives always vote $x_i = 0$, regardless of the signal they receive. Existence of this equilibrium is made possible by the fact that by definition, biased representatives are more certain of their preferred policy outcome than are unbiased representatives.

Public decision making

I begin by considering the case of public decision making. When decision making takes place in an open-door context, the public at large is able to observe both the final outcome and the votes of individual representatives. I focus first on the possibilities for informative equilibria. Given the assumption that representatives are limited to pure strategies, there are three possibilities for informative equilibria: (1) both unbiased and biased representatives may always vote conditional on their signal (hereafter “truthfully”) (2) only biased players vote truthfully (3) only unbiased players vote truthfully. As presented in greater detail in the appendix, we can in fact quickly rule out possibilities (1) and (2). An equilibrium in which both types of representative always vote truthfully cannot exist, because in this case a biased representative who received a signal $s_i = 0$ could always costlessly deviate by voting $x_i = 1$. Since either type of representative would be believed to be equally likely to vote for either outcome, there would be no reputational penalty incurred by a biased player who took this action. We can also quickly rule out possibility (2) because a biased representative would always have an incentive to deviate, as shown in the technical appendix.
The final possibility for an informative equilibrium is a strategy profile where an unbiased representative votes conditional on her signal while a biased representative always votes \( x_i = 1 \). There are two constraints for the existence of this equilibrium, that a biased representative not have an incentive to vote \( x_i = 0 \) in order to improve her reputation and that an unbiased representative who receives a signal of \( s_i = 1 \) not have an incentive to vote \( x_i = 0 \) in order to improve her reputation. In fact, the former of these two constraints will always be satisfied when the latter is satisfied. This derives from the fact that an unbiased representative is, by definition, more uncertain than a biased representative of her preferred policy outcome, but she has the same expectations as a biased representative about the reputational payoff. An unbiased representative who receives a signal of 1 will prefer to vote \( x_i = 1 \) rather than 0 as long as the following inequality is satisfied (where "piv" is an indicator for the case where the representative's vote is pivotal).

\[
v \alpha (\Pr(\omega = 1|s_i, \text{piv})) + (1 - \alpha) \Pr(u|x_i = 1) \\
\geq v \alpha (\Pr(\omega = 0|s_i, \text{piv})) + (1 - \alpha) \Pr(u|x_i = 0)
\]

(2)

In this inequality \( v \) represents the probability that a representative's vote is pivotal, and \( \Pr(\omega = 1|s_i, \text{piv}) \) represents her belief that the state is \( \omega = 1 \) given her signal and the fact that her vote is pivotal, \( \Pr(u|x_i = 1) \) represents the public's belief that the representative is unbiased given that she votes \( x_i = 1 \). The expanded form for the reputational payoff \( \Pr(u|x_i = 1) \) is presented in the appendix. It depends on the probability that a representative is unbiased \( p \), the expected accuracy of the signal \( q \), as well as depending on the expected probability for representative \( i \) that representatives \( j \) and \( k \) will each vote 1 or 0. As with any voting game, the unbiased representative knows that her vote will only have an effect on the policy outcome \( y \) if she is pivotal. She also knows that she can draw inferences about the signals that other representatives will have received if it turns out that she is pivotal (this follows Austen-Smith and Banks, 1996). In particular, a representative could be pivotal if

\[\text{The other remaining possibility is that an unbiased representative votes truthfully while a biased representative always votes } x_i = 0, \text{ but this can quickly be ruled out because a biased representative could gain both in terms of reputation and in terms of the expected policy payoff by voting } x_i = 1.\]
both of the other representatives are unbiased, and one received a private signal of 1 while the other received a private signal of 0. In this case an unbiased representative who received a signal of 1 would have a posterior belief about the state of $\Pr(\omega = 1) = q$. But the unbiased representative will also know that she could be pivotal if one of the other two representatives was unbiased and received a signal of 0, while the third representative was biased. In this case her belief about the state would be $\Pr(\omega = 1) = 0.5$. Given these beliefs, as well as the expanded form for the belief $\Pr(\omega = 0|s_i = 1, \text{piv})$, we can expand and rearrange the inequality in (2) above to present the following condition for existence of the informative equilibrium under public decision making.

$$v_1 \frac{\alpha}{1-\alpha} \rho (2q - 1) \geq 1 - \Pr(u|x_i = 1)$$

(3)

The reputational payoff in the case where a representative votes $x_i = 0$ is simply 1 in this equilibrium since biased representatives never take this action. The reputational payoff for a representative who votes $x_i = 1$ is more complicated, since it depends upon the votes registered by the other two representatives, which themselves depend upon whether these representatives are biased and what type of signal they receive. The appendix presents the expanded form for $\Pr(u|x_i = 1)$. The inequality in (3) implies that under public decision making the informative equilibrium will only exist if reputational concerns are weak (if $\alpha$ is high).

**Proposition 1** Under public decision making there is an informative equilibrium if reputational concerns are sufficiently weak that (3) is satisfied.

One immediate implication of Proposition 1 involves the ability of members of the public to learn about the most likely state of the world $\omega$ by observing decision making. In the informative equilibrium the public knows that unbiased representatives will condition their votes on the signals they receive. As a result, the public will update its beliefs about the state variable $\omega$. In strong contrast, if an informative equilibrium does not exist, then by definition in an uninformative equilibrium no representative will condition their vote on their private
information, and as a consequence members of the public will not actually learn anything even if they are able to observe the individual votes of representatives. This leads to the following corollary to Proposition 1.

**Corollary 1**  If representatives make decisions in public, citizens will not alter their prior policy beliefs if reputational concerns for representatives are sufficiently strong that (3) is not satisfied.

For certain ranges of parameters in this model there exist several types of “uninformative equilibria”. In the first of these, a pooling equilibrium, both types of representatives always vote 0 regardless of the private signal they receive. This equilibrium would normally exist for all parameter values as long as the public’s belief regarding the out-of-equilibrium action $x_i = 1$ satisfies the plausible restriction $p \geq \Pr(u|x_i = 1)$. No representative would have an incentive to deviate because voting is unanimous, and each representative will know that her own vote is not pivotal.

In the second type of uninformative equilibrium, which is a separating equilibrium, unbiased representatives always vote $x_i = 0$ and biased representatives always vote $x_i = 1$. For this equilibrium to exist reputational concerns must be neither so strong that the biased representative has an incentive to deviate by voting $x_i = 0$, nor so weak that an unbiased representative who receives a signal of 1 has an incentive to vote $x_i = 1$.

While I have assumed in this section that constituents have a common prior belief that each state of the world is equally likely, one could extend the above model to explicitly consider a case where there are two groups of constituents that have polarized prior beliefs about which state of the world is more likely. The conclusions of this exercise are similar to those presented above under the common prior assumption; if reputational concerns are strong, and no informative equilibrium exists, then the members of the public will not alter their prior policy beliefs even if they observe the actions of individual representatives, and as a consequence public decision making would not help to reduce polarization.\(^\text{14}\)

\(^\text{14}\)It should be noted, however, that in any extension where one assumed that constituents had heterogeneous prior beliefs about the state $\omega$, one would need to specify the basis for these prior heterogenous beliefs.
Private decision making

When decisions are made behind closed doors, incentives for representatives change significantly. Under private decision making the public observes the policy outcome \( y \), but it does not observe the individual votes of representatives. This weakens the incentive for representatives to use their vote to signal that they are unbiased. Under public decision making representatives know that their vote will have a certain impact on their reputation, but it will only have in impact on the policy outcome if it turns out that they are pivotal. In strong contrast, under private decision making, because the public establishes inferences about representative type based exclusively on the policy outcome \( y \), representatives know that their vote will only make a difference for either the outcome or their reputation if they are pivotal. Under private decision making we can again consider the possibilities for informative equilibria, and as presented in the appendix we can again rule out either an informative equilibrium in which both biased and unbiased representatives always vote truthfully, or an informative equilibrium in which biased representatives vote truthfully. This leaves the final possibility of an equilibrium where biased representatives always vote \( x_i = 1 \) and unbiased representatives always vote truthfully. The key condition for existence of this equilibrium again involves an unbiased representative who receives a signal \( s_i = 1 \). She knows that if she is not pivotal then both her policy payoff and her reputational payoff will be unaffected by her vote. If we designate the payoff in the case that the representative is not pivotal by \( U_{\text{nonpiv}} \), then we get the following constraint for a “truthful” vote of \( x_i = 1 \), as shown in (4).

\[
v \alpha \Pr(\omega = 1|s_i \text{ piv}) + (1 - \alpha) \Pr(u|y = 1) + (1 - v)U_{\text{nonpiv}} \\
\geq v \alpha \Pr(\omega = 0|s_i \text{ piv}) + (1 - \alpha) \Pr(u|y = 0) + (1 - v)U_{\text{nonpiv}}
\]

This quickly simplifies as the probability of being pivotal \( v \), as well as the utilities when

Aumann (1976) observed that if players have a common prior belief then it cannot be common knowledge that they disagree, otherwise each player who observed another’s different beliefs would revise their beliefs based on this information and beliefs would converge. The alternative is to assume that beliefs are common knowledge but that players do not share a common prior, as would be the case in this extension of the model. Morris (1995) presents a review of justifications for and against the common prior assumption in different modeling contexts.
the representative is not pivotal, drop out of the expression. In addition, the expected policy gain from deviating by voting \( x_i = 1 \) remains the same as under public decision making (left hand side of the inequality in 5). The appendix presents the expanded form for the reputational payoffs in (5).

\[
\frac{\alpha}{1-\alpha} p(2q - 1) > \Pr(u|y = 0) - \Pr(u|y = 1)
\]  

(5)

**Proposition 2** Under private decision making there will be an informative equilibrium even if reputational concerns are strong, provided (5) is satisfied.

When we compare the inequality in (5), which shows the condition for existence of the informative equilibrium under private decision making, with the inequality in (3), which presented the condition for existence of the informative equilibrium under public decision making, we quickly see that the equilibrium will exist for a greater range of parameters when decision making takes place in private. There are two reasons for this. First, the probability of being pivotal \( v \) will always be less than 1 and this enters into the left hand side of expression (3) but it does not enter into expression (5). Second, under private decision making the representative’s reputation is less sensitive to the observed outcome. The reputational payoff \( \Pr(u|y = 0) \) will be less than 1 because an outcome \( y = 0 \) does not indicate that a specific representative is necessarily unbiased. In addition, the "penalty" for voting \( x_i = 1 \) is lower.

In the informative equilibrium under private decision making, members of the public will update their beliefs about the state \( \omega \) even though they only observe the action taken by the three representatives and not their individual votes. For a given action \( y \), the public can establish inferences about the likely patterns of signals received by representatives that would have led to this outcome. As before, if we extended consideration to a case where members of the public initially had polarized beliefs about the best policy, then we would observe that this initial polarization would diminish in the informative equilibrium but not in an uninformative equilibrium. This leads to the following corollary.
Corollary 2 If reputational concerns for representatives are sufficiently strong that (5) is satisfied but (3) is not satisfied, then citizens may alter their policy beliefs following private decision making between representatives, but they will not alter these beliefs following public decision making.

Figure 1 shows the conditions under which the informative equilibrium will exist under public versus private decision making for different values of $p$ and $\alpha$, based on a realistic parameter assumption of $q = 0.75$ for the accuracy of signals. The region above the dotted line shows the area where an informative equilibrium can exist under both public and private decision making. The region between the dotted and solid lines represents the values of $p$ and $\alpha$ for which an informative equilibrium will exist under private decision making but not under public decision making. Finally, beneath the solid line the informative equilibrium will not exist under either public or private decision making.

In terms of uninformative equilibria, under private decision making there would again continue to exist for all parameters a pooling equilibrium where both types of representative always vote $x_i = 0$, provided the public’s belief about the type of a representative who deviates by voting 1 satisfies $p > \Pr(u|x_i = 1)$. There would also continue to exist a separating equilibrium for certain ranges of parameters where unbiased representatives vote $x_i = 0$ irrespective of their signal and biased representatives vote $x_i = 1$ irrespective of their signal.

Adding a deliberation stage

The two previous sections produced a clear prediction about incentives under private versus public decision making, but they assume away the fact that in the real world representatives commonly have a prior round of discussion before voting. A number of recent contributions have analyzed models of collective decision making where players have an opportunity both to vote and to exchange messages prior to a vote.\(^{15}\) As a result, it seems important to

\(^{15}\)See in particular Austen-Smith and Feddersen (2006), Meirowitz (2005a, b), Gerardi and Yariv (2006), and Visser and Swank (2006).
ask whether the conclusions drawn in the previous section are robust to the addition of a deliberation round to the game, or alternatively whether Propositions 1 and 2 are instead dependent on the useful but unrealistic simplification that deliberation and voting are collapsed into a single stage. In this section I will argue that my conclusions regarding the effect of publicity/transparency can continue to hold when adding an explicit deliberation stage.

Consider an expanded model where each representative sends a message about the state \( m_i \in \{0, 1\} \) before voting. Messages are simultaneous, and for simplicity I assume that not speaking is not an option. The order of play is identical to that presented in the previous sections with the addition of the “deliberation stage” after representatives receive their signals and before they vote. There will now be four possible decision making environments: (i) voting is public and the message stage is public, (ii) voting is public and the message stage is private, (iii) voting is private and the message stage is private, (iv) voting is private and the message stage is public. The question of substantive interest is whether and when it is possible to have equilibria where representatives send “truthful” messages \( m_i = s_i \), and how this varies across the different decision making environments. Since equilibrium possibilities expand significantly when one adds a deliberation stage prior to voting, I concentrate here on examining the robustness of the results from the previous sections by asking whether the following strategies and beliefs can constitute an informative equilibrium:

1. During the deliberation round unbiased representatives send a “truthful” message \( m_i = s_i \) while biased representatives send a message \( m_i = 1 \).

2. During the voting round all representatives vote \( x_i = 0 \) if two or more messages of 0 have been observed.\(^{16}\) Otherwise they vote \( x_i = 1 \).

Since all representatives vote unanimously during the voting stage, in order to fully characterize the equilibrium we also need to specify the public’s out-of-equilibrium belief about the type of any representative who does not vote the same way as the other two represen-

\[^{16}\text{The voting rule prescribed here would be optimal from the point of view of the public provided that two messages of 1 are jointly more informative than two messages of 0. This will be true as long as the following condition is satisfied } \left( \frac{pq + 0.5(1-p)}{1-pq-0.5(1-p)} \right)^2 > \frac{q}{1-q} \].
tatives. The strategy profile can be an equilibrium with the following plausible restriction on out-of-equilibrium beliefs: (1) If one representative votes \( x_i = 0 \) while the other two vote \( x_i = 1 \) this representative receives the same reputational payoff as she would have received if she took the equilibrium action. (2) If one representative votes \( x_i = 1 \) while the other two vote \( x_i = 0 \) this representative receives a lower reputational payoff than she would have received had she taken the equilibrium action.

Under scenario (i), where both the message stage and the voting stage are public, the key constraint for the existence of this equilibrium will be that an unbiased representative who receives a signal \( s_i = 1 \) has an incentive to send a message \( m_i = 1 \) rather than deviating by sending a message \( m_i = 0 \) in order to receive a higher reputational payoff. The expected benefit of pursuing this equilibrium strategy is that if the representative’s message is pivotal, meaning the other two representatives send different messages, then this increases the likelihood of making a policy choice that corresponds to the state \( y = \omega \). An unbiased representative will not deviate from this strategy at the message stage as long as the following condition is satisfied.

\[
v \frac{q}{1 - q} p(2q - 1) \geq 1 - \Pr(u|m_i = 1) \quad (6)
\]

This is identical to the condition for existence of the “informative” equilibrium in the previous sections, apart from the fact that the reputational payoff now depends on a representative’s message rather than her vote. It should be noted that, as in the case where there is a voting stage but no deliberation stage, as long as an unbiased representative has an incentive not to deviate by sending a message \( m_i = 0 \) then this will also hold true for a biased representative, given that the biased representative has the same expectation of the reputational payoff that will result from this action but greater certainty that it will result in a lower policy payoff.

Now consider the conditions for existence of this same equilibrium under scenario (ii) where the voting stage is public but the deliberation stage is private. When the message stage is private this lowers the reputational cost of sending a message \( m_i = 1 \). An unbiased representative knows that if she is not pivotal then neither her reputational payoff nor the
policy payoff will be affected by a decision to send a message \( m_i = 1 \) rather than a message \( m_i = 0 \). This follows the same logic of the public/private trade-off in the previous sections. The condition for existence of the equilibrium will now be the same as that expressed in (5) for the case without a deliberation stage. The same results will apply as in the previous sections regarding existence of an informative equilibrium under public versus private decision making. This conclusion has a significant implication for institutional design; in order to generate conditions for an informative equilibrium where the public can update its prior beliefs about desirable policy, it may be sufficient to shift the deliberation stage to occurring in private even if the voting stage continues to occur in public.

We can also consider equilibrium behavior under scenario (iii). Under scenario (iii) where both messages and voting occur in private, because behavior during the voting stage in this equilibrium is unanimous, the expression in (5) also determines the conditions for existence under scenario (iii) where both the voting and the message round are secret. This result hinges on the fact that though the voting stage now occurs in private, no representative is pivotal, and thus there is no incentive for a biased representative to deviate by voting \( x_i = 1 \). This result corresponds to an observed feature of real-world committees that have shifted from private to public voting but which have retained a private deliberation stage. For example, the US Federal Reserve’s key decision making body, the Federal Open Market Committee, publishes individual votes of its members on interest rate decisions, but authors have noted much higher levels of disagreement at the deliberation stage (which has remained private) than exist during the voting stage where there is a higher degree of unanimity (Meade and Stasavage, 2006). This has significant implications for questions of institutional design. Because the informative equilibrium will exist for the same set of conditions under scenario (ii) as under scenario (iii), if there is a collective decision making body that debates first and then votes, shifting from having both stages occur in private to having only the deliberation stage occur in private may neither help nor hinder chances for the public to update its prior beliefs about which policy.\(^{17}\)

\(^{17}\)Under scenario (iv) voting occurs in private but the message stage occurs in public. While this decisionmaking structure is frequently observed in committees, as well as public discussions prior to elections,
As long as it is believed that representatives should be accountable, then there are advantages to having them deliberate in public, but as long as it is also believed that representatives should exercise a degree of independent judgement, then transparency can also have costs. I have argued that recent discussions of transparency in government have often overlooked this fact. In order to examine this issue, I have compared public versus private decision making by a body of representatives, concluding that public decision making may do less than private decision making to reduce polarization of opinions in society. My conclusions have direct implications for recent discussions of deliberative democracy and for debates about polarization in American politics. With regard to polarization, while one might think that the institutional changes of the past forty years to promote openness in government should logically have reduced opinion polarization, the theoretical model presented here suggests why they may have actually had the opposite effect. With regard to deliberation, the literature on deliberative democracy has often overlooked the fact that much actual decision making occurs between representatives, rather than directly between citizens, a point made by Barabas (2004), and representatives can face conflicting incentives between advocating their preferred policies versus using their actions as a signal that they hold the same interests as those of their constituents. There is empirical evidence from European Union committees, from eighteenth century deliberative bodies, and from other contexts which suggests that representatives are much more likely to engage in a free exchange of opinions and information if they express these opinions in private. My conclusions do not contradict the idea that the free exchange of opinions is a worthwhile goal, but they do point to a potential complication for advocates of “deliberative democracy”, precisely because this literature emphasizes that the benefits of decision making depend upon broad participation in the deliberative process.
References


Hobbes, Thomas (1651 [1996]) Leviathan, Oxford University Press.


Schumpeter, Joseph (1942) Capitalism, Socialism, and Democracy, New York.


Figure 1: Existence of an Informative Equilibrium

- No informative equilibrium
- Exists only under private
- Exists under public and private

\[ \alpha \]

prior belief representative is unbiased