Deliberation as Coordination through Cheap Talk*

Randall Calvert
Washington University, St. Louis

11 Oct. 2005; this version 12 April 2006

Abstract

In contrast with other formal theory treatments of democratic deliberation in terms of (1) strategic information pooling or (2) relaxing the assumption of logical omni-science, this paper considers the role of deliberation in coordinating citizens on a higher-welfare equilibrium. The equilibrium in question belongs to a post-deliberation game of fundamental action vis-à-vis other individuals: giving, taking, harming, or producing private or collective goods. This action game has multiple equilibria, some of which achieve preferred outcomes for most or all individuals through coordinated actions. The present paper portrays this problem as a multi-player version of a battle-of-the-sexes (BoS) game, and deliberation is represented as a sequence of pre-play proposals for various coordinated action patterns. Failure to reach sufficient agreement on these leads to play of a less favorable action profile, stylized as a mixed-strategy equilibrium. Thus the model extends Farrell’s (1987) model to the setting of more than two players, examining issues of how actions respond to proposals and how this connection bears on the aggregated effect of proposals—issues that do not arise in the two-player case but are fundamental to deliberation (as well as to endogenous-institution voting models).

The BoS game for multiple players is first specified, and relevant equilibria derived. As a first cut at pre-play deliberation, we then add one pre-play stage of simultaneous proposals for action. Depending on whether the participants are swayed by agreement among a plurality of players, or instead require that a consensus is expressed, deliberation yields different results in terms of strategic proposals, likelihood of coordination, and expected payoff. The same questions are then addressed in a deliberation process consisting of multiple periods of such simultaneous proposals. The results match Farrell’s two-player conclusions in terms of expected payoff, but differ from his conclusions concerning deliberative strategies. Moreover, the length of deliberation necessary to approach the achievable payoffs depends strongly on the deliberation criterion (one-time consensus, or one-time plurality).

*Prepared for presentation at Midwest Political Science Association annual meeting, April 20-23 2006, Chicago. A preliminary version was presented at the “Conference on Deliberation and Collective Choice,” Kellogg School, Northwestern University, May 20-21, 2005; thanks to David Austen-Smith and Timothy Feddersen for providing this opportunity. Thanks as well to seminar participants at the Hoover Institution, Stanford University.
1 Introduction

Democratic theorists have come to see deliberation as a fundamental feature of democracy. In this view, democratic politics does not simply aggregate preferences; more importantly, they promote the reaching of a sufficient degree of general agreement on government policies through the airing of public-regarding reasons for choosing one policy rather than another. Through deliberation, the polity pools knowledge about the results of policies and reaches agreements about trade-offs among the benefits and costs imposed on different individuals or groups. The making of such trade-offs is influenced, in part, by principles or values held in common, so deliberation consists also of airing the connection of such principles with policies. Theorists of deliberative democracy believe it to have superior normative properties to democracy as understood in the purely aggregative terms familiar to rational choice theory, emphasizing selfish preferences and strategic action.

Rational choice theorists, newly equipped with formal models of information transmission and imperfectly rational choice, have in several instances taken up the idea of deliberation for formal treatment in the context of preferences and choice. Some, beginning with the work of Austen-Smith and Riker (1987) on political debate and Gilligan and Krehbiel (1987) on informative legislative committees, treat pre-choice communication processes as occasions for the strategic revelation of information about the state of the world, such as the likely effects of contemplated laws. (Such theories were first brought to bear explicitly on deliberation by Austen-Smith (1990).) Other formal theorists, such as Hafer and Landa (2004), have focused on how deliberation, as the offering of “reasons,” especially the stating of logical connections or implications, might affect the actions of rational actors who are less than omniscient in their use of information that is in principle already known, such as logical implications. Such models appeal to the direct modeling of gaps in logical reasoning and how those gaps are filled by the statement of such reasoning. These models have in common the making of a fairly simple choice—whether explicitly by voting, as in Gilligan and Krehbiel (1987), or implicitly as choosing or casting a vote—which is preceded by a relatively complex model of communication.

The present paper departs from the pattern of modeling deliberation as the preliminary stage to a simple choice problem. Instead it envisions a potentially complex process of “choice” preceded by a simpler process of communication. The goals of this approach are to (1) establish an important alternative to seeing deliberation as either a matter of determining preferences or of pooling objective information; and (2) to relate deliberation to important features of social interaction abstracted away in previous models.

Specifically, this paper assumes that political action takes place in the context of a complex social interaction that is itself a game, and follows one of many possible equilibria to that game. This “game of life” is preceded by a deliberation stage in which individuals propose or announce plans, rules, or strategies according to which they will act in the action stage. Participants have common expectations about the patterns of action that will ensue depending on the pattern of proposals announced. The idea is to capture the extensive arbitrariness of how principles may be selected to motivate different desired patterns of social action. At the same time, agreement or coordination on the actions, and hence on the application of principles, yields outcomes potentially better for all participants, so there is an incentive to control arbitrary claims in order to promote efficient coordination.
Formally, the model here explores a many-player extension of the work of Farrell (1987) and others on coordination via pre-play, cheap-talk communication in battle-of-the-sexes games. With more than two players, the patterns by which proposals lead to coordination vary along dimensions that do not apply in the standard two-player model. In particular, I analyze the illustrative cases of coordination based on “consensus” in deliberation and on “plurality” in deliberation, comparing their welfare properties and the strategic deliberation that takes place under them.

The next section describes the conception of the political or social action game that is at issue following deliberation. The overall strategy of this paper is to model the choice among equilibria in that action game as a problem of full or partial solution of a simple coordination dilemma. Deliberation, then, occurs through a process of communication prior to the play of that game. Section 3 examines the implications of previous research on pre-play communication and coordination. The main body of the paper presents a generalized mixed-motive coordination problem among many players (Section 4), and compares the behavioral and welfare properties of two simple equilibria in the overall game of deliberation and action (Section 5). A final section discusses general conclusions, possible extensions, and alternative applications.

2 A Political “Game of Life” and its Attendant Coordination Problems

This paper views deliberation not as a process for adjusting preferences or for pooling information about the state of the world, but rather as a process that prepares the way for compliance with group decisions. That is, deliberation enhances the authority and legitimacy of group decisions. A legislative vote is just words spoken in a certain venue; a law is just a piece of paper issued by a group of people; and a government is just a designated collection of people who react to these actions, and to each other’s actions. Common expectations (and sufficient incentives) are what make votes, laws, and governments behave as such. Deliberation helps maintain those expectations.

The collective activities of politics are best seen not as just one-off acts of social choice, but as complex sequences of actions, each of which may be contingent on previous actions: taxation, administration, enforcement, public good production and distribution, etc.; and attendant opportunities for evasion, corruption, theft, coercion, and shirking. Even if information about preferences and other aspects of the state of the world were fully pooled, any reasonably rich game-theoretic representation of these events is likely to have many equilibria. And to the extent such problems are faced repeatedly over an indefinite horizon, of course, the folk theorem applies.

A classic theme of political theory is the problem of security or order, such as famously explored by Hobbes in *Leviathan*. Game theorists ought to understand security as simply one aspect of the wider problem of implementing relatively efficient, and avoiding relatively inefficient, equilibria. Hobbes proposed that this required putting all power in the hands of a single ruler. Subsequent democratic theorists have seen this prescription as unnecessarily harsh, but to a game theorist the creation of a Leviathan is actually insufficient to ensure
security. Power must be organized, internal incentives maintained, and the King’s enforcers must in turn be supervised and sanctioned, if the Leviathan is to retain and wield power. A democratic polity, of course, faces even more complicated versions of those problems.

To achieve equilibria that realize at least some measure of security, not to mention prosperity, requires the alignment of expectations. Participants’ expectations about what actions are to be taken and demanded, and what consequences are to follow from nonperformance, must be (1) established by communication, (2) consistent with incentives, and (3) reinforced by experience. Actions taken and words spoken at one time determine individuals’ subsequent choices about compliance, cooperation, political opposition, opportunism, and rebellion.

The broad problem of achieving and maintaining a good equilibrium is fundamentally a problem of coordination. If most people believe that most other people will adhere to constitutional and legal prescriptions, will not tolerate those who fail to do so, and will in fact participate in the sanctioning of violators, then (assuming those forms of monitoring and enforcement are sufficient) most people will so behave. If they believe that officials systematically demand bribes and tolerate others who do so, they will demand, pay, and tolerate. If they believe that sufficiently many others in their group will cooperate to perpetrate violence and theft against another group, and will not tolerate those who do not, there will be group exploitation and violence. And wherever common expectations about social structure and compliance are absent, disorder results—not, game theoretically speaking, a failure to reach equilibrium, but rather something more like either (1) a symmetric mixed-strategy solution to problems of social coordination or (2) coordination within smaller groups, resulting in correspondingly lower average coordination gains.

Deliberation can be important for achieving such coordination. The reasons people enunciate for a decision often carry implications for the criteria by which future decisions will be made. To the extent these reasons are taken seriously, losers in the current decision might see them as an assurance that future decisions will tend to be more agreeable. And even without giving full credibility to the reasons offered, opponents on the losing side might expect those giving the reasons to have trouble contradicting those reasons later, or might be able to use the same reasons for later rhetorical leverage in their own favor in winning coalition or constituent support. (Rhetorical leverage, in this context, means simply the ability to create focal points by appealing to the right verbal formulation.)

Coordination models, generally speaking, focus on the creation of expectations that bring about a focus on one equilibrium rather than another. This can be an equilibrium in a single-play political choice game, in which case the payoffs involve expectations about future expected payoffs; or equilibrium in a subgame during an ongoing process of successive political decisions. In the latter case, successive decisions might be preceded by the revelation of new situational details that create successive, new problems of coordination on subgame equilibria.

To theorize about the phenomenon of deliberation in political decision making, then, requires theorizing about how communication solves such coordination problems and establishes common, equilibrium expectations about how people will act and react. In what follows, I approach this problem by extending a common game-theoretic model of coordination to the multi-player setting, and then examining the role of strategic pre-play communication in solving the coordination problem.
3 Pre-play Communication and Coordination

3.1 Pre-play Communication in the Two-Player Battle of the Sexes

Consider a standard two-player battle-of-the-sexes (henceforth BoS) game $G$, capturing as simply as possible the problem of symmetrical, mixed-motive coordination:

\[
\begin{array}{c|cc}
\text{Row Player} & A & B \\
\hline
A & x,1 & 0,0 \\
B & 0,0 & 1,x \\
\end{array}
\]

assuming, without loss of generality, that $x \geq 1$. The game’s Nash equilibria are $(A, A)$, $(B, B)$, and a mixed-strategy equilibrium in which the row player uses $A$ with probability $\frac{x}{x+1}$ and the column player uses $B$ with that same probability. I will refer throughout the remainder of the paper to $(A, A)$ as the “favorite outcome” or “favorite equilibrium” of the row player, and similarly for the column player; and, more loosely, to $A$ as the “favorite action” of Row. As Farrell (1987) argues, the mixed-strategy equilibrium is the natural prediction for this essentially symmetric game, at least in the absence of any contextual information that would justify the choice of one of the pure strategy equilibria in contradiction to that symmetry. Moreover, the mixed-strategy equilibrium results in a probability $\frac{x^2+1}{(x+1)^2}$ of failing to coordinate, nicely capturing the inefficiency inherent in the coordination problem. At the opposite extreme of coordination ability, however, the game has a symmetric, efficient correlated equilibrium in which $(A, A)$ and $(B, B)$ are each chosen with probability $1/2$.

Farrell (1987) analyzes a model in which $G$ is preceded by a multistage process of communication; I refer to these as the “action stage” preceded by a “communication stage,” which may itself consist of multiple rounds of simultaneous broadcasting of messages. In principle, the opportunity for the players to exchange even a single round of simultaneous messages, chosen prior to play from a set of at least two available messages, would be sufficient to implement the symmetric correlated equilibrium. This is done via the device introduced by Aumann, Maschler, and Stearns (1968; henceforth AMS): each player uses the mixed communication strategy of sending each of two messages with probability $1/2$; they then take actions consistent with Player 1’s favorite if the messages match, or with Player 2’s if the messages differ. In this Nash-equilibrium implementation of the correlated equilibrium, the meaning of the communication can be captured by labeling the two messages as, say,

\[1\] This is the only correlated equilibrium that is both symmetric and efficient, but there are many others that are symmetric, or are efficient, or are neither. The efficient ones are all those in which $(A, A)$ is chosen with any probability $p$ and $(B, B)$ with probability $1-p$. The symmetric ones are all those in which $(A, A)$ and $(B, B)$ are each chosen with some probability $p \geq \frac{x^2}{(x+1)^2}$, and $(A, B)$ and $(B, A)$ are each chosen with probability $\frac{1-p}{2}$.

\[2\] Banks and Calvert (1992) spell out this procedure for the BoS game. Aumann and Hart (2003) lay out a more general explanation of how mixed, independent communication strategies can be used to implement correlated equilibria.
“heads” and “tails,” and describing the rule for prescribing subsequent actions as “matching pennies.”

Farrell, however, argues the relevance of a less efficient “natural language” equilibrium in the combined game of communicating and acting. In effect, he suggests that prior experience with communication imparts strong, mutual equilibrium expectations in which the messages’ meanings will be derived from previous usage: one message will mean, in effect, “I intend to take the action consistent with achieving my favorite equilibrium” while the other means “I intend to take the action consistent with your favorite,” so that the exchange of messages amounts to a crude form of bargaining. This previous experience with the language creates a focal point such that, in the absence of any sort of pre-pre-play communication for agreeing on an AMS-style procedure for correlated play, the players are likely to settle on this equilibrium.\(^3\)

Under this assumption, Farrell shows that (1) even with only one round of such communication, the probability of coordination in equilibrium is increased; (2) more rounds of communication increases the probability of coordination; (3) even if the number of rounds is made arbitrarily large, the probability of coordination is bounded away from 1. In the Farrell equilibrium, the first communication period in which the players signal coordinated actions determines the outcome of \(G\); if they do not signal coordinated strategies by the end of the communication phase, they play the symmetric mixed-strategy equilibrium.

### 3.2 Issues in Multi-Player Extension

Applying similar reasoning to a problem of reaching political agreement among a large number of participants requires that we address several new complications. First, a large number of players might have an equally large variety of preferences concerning what outcomes are most desirable, so a suitable generalization of the BoS should have at least as many actions as players. Second, with more than two participants there is the possibility of incomplete coordination, in which some players, but not all, succeed in choosing the same action. The BoS game representing the action stage could be extended in various ways to reflect the value of incomplete coordination. Then, the model must specify how payoffs would respond to incomplete coordination over those actions. At low levels of partial coordination, a player might prefer to “coordinate” alone on her own favorite outcome rather than to join the small crowd at another outcome, so this tradeoff too must be specified.

Other issues concern the nature of the communication process and its equilibria. Suppose we maintain the meanings suggested by Farrell: each message used in equilibrium corresponds to an outcome; and that message amounts to a proposal that all players take the corresponding action. Even so, there is still the problem of how such messages are aggregated to form players’ expectations about eventual actions. At one extreme in terms of simplicity would be to imagine that if the players unanimously name the same action during the communication process, they will all play that action, ignoring any further communication; and otherwise their further communication strategies will be unaffected, and if they reach the last stage without having achieved such unanimity, they will choose the symmetric mixed equilibrium in the action stage. An alternative just as reasonable, and only slightly

\(^3\)The idea of natural language equilibria receives more general discussion in Farrell and Rabin (1996).
more complex, would "stop" the process whenever one action is named by a clear plurality of players. I examine both these possibilities below. Another alternative would be for the players to move in stages toward consensus, perhaps requiring an increasing plurality in successive stages and reverting to the starting condition if that threshold is not met at some stage. I leave this possibility for future research.\footnote{Rabin (1994) envisions a process in which a suitable combination of proposals at one stage might advance a "provisional agreement" from a status quo toward some eventual outcome, showing how conditions on this abstract process of advancement can yield guarantees about the welfare qualities of the final outcome. The present paper examines the microfoundations of such an advancement process, although such multi-stage agreement processes remain for future research. However, the basic processes studied here (the plurality and consensus criteria), if extended to such a process of successive advancement of a provisional agreement, would not satisfy any of Rabin’s basic conditions guaranteeing welfare enhancement.}

In the many-player setting, the AMS procedure, generalized to a multi-player setting, would correspond roughly to having participants agree to choose the outcome by lottery, where each alternative involves a prescribed pattern of behavior, rewards, and penalties whose carrying out would constitute an equilibrium in the political action stage. If participants did this, their political arrangements would always have, at least, the virtue of efficiency, forgoing none of the available gains from political organization and voluntary order. Such choice by lottery is, of course, rare in politics. Even in a smoothly working democracy, by the time an issue is cognized and general policy alternatives formulated, participants argue for their preferred alternatives rather than submitting all to the drawing of numbered ping pong balls on the Eleven PM News. In this setting, something akin to the Farrell-Rabin natural language reasoning seems compelling.

4 A Many-player Battle of the Sexes Game

As a concrete, and generalizable, model of the BoS game with more than two players, consider a version with a set of players $N = \{1, 2, \ldots, n\}$, each having a different, favorite action on which she would like to see the others coordinate; for simplicity assume these favorites are the only actions available, so $N$ is the strategy set of each player as well, and $i$ is the favorite action of player $i$. Let $a_i$ denote the action chosen by player $i$, and let $m_i$ denote the number of players choosing strategy $i$:

$$m_i = \#\{j \in N | a_j = i\}.$$

The payoff to player $i$ from action profile $(a_1, \ldots, a_n)$ is

$$u_i(a_1, \ldots, a_n) = Am_i^B + \sum_{j \neq i} m_j^B$$

where $A$ and $B$ are parameters greater than 1. These parameters accomplish the spirit of the BoS game: the greater the extent to which players congregate on the same actions, the better; and it’s better if that action is my favorite one. If $n = B = 2$ and $A = 1.4$, this is the textbook BoS game with payoffs, once normalized, of 0, 1, and 2.\footnote{More generally, the case $n = 2$ produces the textbook BoS provided that, if we normalize the payoff from not coordinating to be zero and normalize the payoff to coordinating on a non-favorite action to be}
The worst outcome, from every player’s point of view, is for each player to choose a
different action; in that case, the payoff to each player is just \( n \). Notice that it doesn’t matter
who chooses which action; and in particular, a player doesn’t care which players (including
herself) choose her favorite action, only about how many do so. The game incorporates no
conception of “similar” actions or any other idea of near-coordination, other than partial
agglomeration. The best outcome, from player \( i \)’s standpoint, is when every other player
chooses \( i \)’s favorite action, giving \( i \) a payoff of \( An^B \). If everyone coordinates on some other
player’s action, the payoff to \( i \) is \( n^B \). The parameters \( A \) and \( B \) determine the tradeoff
between adding a player to the group who choose \( i \)’s favorite, and adding a player to the
group who choose the most popular action. Specifically, \( A \) measures the extent of competition
in the game, that is, the advantage to \( i \) of having players choose \( i \)’s favorite action rather
than another action; while \( B \) measures the value of coordination, that is, the advantage
of having players congregate on the same action rather than choosing different actions. It
is immediately apparent that, to capture the spirit of the BoS game, it must at least be
preferable for player \( i \) to join the other \( n - 1 \) players when they all choose the favorite action
of player \( j \neq i \), rather than to choose her own favorite alone; so assume \( A < n^B - (n - 1)^B \).

In the same manner as Farrell, the focal outcome of this symmetric game, in the absence of
any other pre-determined focal point or any helpful pre-play communication, is a symmetric
equilibrium in which every player \( i \) chooses \( a_i \) with the same probability \( p \) and chooses each
of the other strategies with equal probability, hence with probability \( \frac{1-p}{n-1} \). Equilibrium values
of \( p \) are the solutions to a combinatoric polynomial equation.

The tractable three-player case is sufficient to illustrate some features of this model of
deliberation not apparent in the two-player case, so we begin by identifying the symmetric
equilibria of the three-player Battle-of-the-Sexes game (henceforth “3-BoS”). An equilibrium
value of \( p \) occurs when, given that players 2 and 3 use \( p \), player 1 is indifferent between
\( a_1 = 1 \) and \( a_1 = 2 \) (or 3, of course). The details needed for this calculation are given in Table 1.

Let \( E u_i(p_1, p_2, p_3) \) represent player \( i \)’s expected payoff for the profile of mixed strategies in
which each player \( j \) uses her favorite action with probability \( p_j \) and each of the other actions
with probability \( \frac{1-p_j}{2} \); in the usual abuse of notation, I write \( (a_1, p_2, p_3) \) for the profile in
which player 1 uses pure strategy \( a_1 \), etc. The requirement for all three players to mix at \( p \)
is that

\[ E u_1(1, p, p) = E u_1(2, p, p) \]

one, the resulting favorite-coordination payoff is 2:

\[ \frac{u(\text{coord on favorite}) - u(\text{no coord})}{u(\text{coord on non-favorite}) - u(\text{no coord})} = 2 \]

that is,

\[ \frac{A2^B - (A + 1)}{2^B - (A + 1)} = 2 \]

which simplifies to

\[ A = \frac{2^{B+1} - 1}{2^B + 1} \]

Thus the same textbook BoS can be obtained for other combinations of \( A \) and \( B \)—for example, \( B = 3 \) and
\( A = 5/3 \).
Table 1: Elements for calculating payoffs to player 1’s pure strategies when players 2 and 3 mix.

( and similarly for players 2 and 3). The resulting quadratic equation in $p$ is


Further traction can be gained by focusing on the case when $B = 2$. Then the quadratic term is zero, and the equilibrium value of $p$ becomes

$$p_0 \equiv \frac{3B - 2}{2A + 1}.$$  

If the value of $A$ is too large, this solution may be greater than 1, indicating that the only symmetric equilibrium is $(a_1, a_2, a_3) = (1, 2, 3)$, yielding each player a payoff of $A+2$. The upper limit on $A$ at which this happens may be tighter than the $A < 3B - 2B = 5$ restriction for coordination games mentioned above, due to the additional consideration that, in a symmetric equilibrium, if a player does choose to try to coordinate on one of the other players’ preferred actions, she may choose the wrong one. Hence the upper limit condition on $A$ for the case $B = 2$ is actually $A < 3$ rather than $A < 5$.

If the players could coordinate perfectly to achieve a symmetric correlated equilibrium, coordinating fully on each action with probability 1/3, the expected payoff would be $\left(\frac{1}{3}\right)(A \cdot 3^2) + (\frac{2}{3})3^2 = 3A + 6$. This ranges, of course, linearly from 9 to 15 as $A$ increases from 1 to 3.

In the symmetric Nash equilibrium, the probability that the players will successfully coordinate on the same action is $3p_0 \left(\frac{1-p_0}{2}\right)^2$, which declines from .11 at $A = 1$ to zero at $A = 3$. The probability that they will choose three different actions simplifies to $p_0^3 + \frac{1}{4}(1-p_0)^3 + \frac{3}{4}p_0(1-p_0)^2$. This probability rises from .22 at $A = 1$ to 1 at $A = 3$.

The expected payoff from this equilibrium is (laboriously) calculated using the profile-by-profile payoffs and probabilities shown in Table 2. (One can then check to make sure that the results match the expected payoffs from either of player 1’s pure strategies as shown in Table 1.) Summing and rearranging these terms gives the following expected payoff to each player in the symmetric equilibrium of the 3-player BoS with $B = 2$:

$$U_0 = \frac{1}{2}(A+2)(-3p_0^2 + 2p_0 + 3).$$
Or, since $p_0$ is itself a function of $A$, 
\[ U_0 = \frac{(A + 2)(-3A^2 + 46A - 13)}{2(2A + 1)^2} \]
As $A$ ranges from 1 to 3, this expected payoff increases from 5.00 at $A = 1$ to about 5.40 at about $A = 5/3$, above which it again decreases to 5 at $A = 3$. At all values of $A$ below 3, then, the equilibrium expected payoff is better than the payoff of $A + 2$ when each player uses her own favored action, but far below the payoff of 9 for player $i$ when all players coordinate on another player $j$’s preferred action.

5 Pre-play communication in the 3-player BoS

Suppose, as in Farrell, that before playing the BoS players 1, 2, and 3 first engage in a round of cheap-talk communication. In particular, they simultaneously choose costless messages from a message set equal to the strategy set. A natural interpretation of these messages, in the equilibria examined below, is that message $j$ means “I propose we all use action $j$.” However, sending message $j$ does not necessarily entail any commitment to actually choose $j$; rather, action is still a fully strategic decision. Strategies in the two-stage game now consist of a (mixed) strategy for choosing messages along with a (mixed) strategy for action, the latter possibly depending on the triple of messages actually announced.

5.1 The consensus rule

First, suppose that the common expectation of the players is that, if all three announce the same action, then all three will play that action in the second stage; but otherwise, they will play the symmetric mixed strategy equilibrium $p_0$ at the action stage. I refer to this as the consensus rule, or more properly the consensus expectation, and seek a symmetric equilibrium in this two-stage game that conforms to this pattern. A symmetric strategy would begin with each player $i$ announcing message $i$ with some common probability $p$, and announcing each of the other two messages with probability $1 - p$. Such an equilibrium requires the usual condition of indifference for each player $i$ between announcing message $i$ and announcing another message $j$. As shorthand, let $\Pr(\text{all } k|j)$ represent the probability that all players announce $k$ given that player $i$ announces $j$. Then symmetric equilibrium must satisfy
\[
9A \cdot \Pr(\text{all } i|i) + U_0 \cdot [1 - \Pr(\text{all } i|i)] = 9 \cdot \Pr(\text{all } j|j) + U_0 \cdot [1 - \Pr(\text{all } j|j)].
\]
Since $\Pr(\text{all } i|i) = (\frac{1-p}{2})^2$ and $\Pr(\text{all } j|j) = p(1 - p)/2$, this condition is
\[
9A \left( \frac{1-p}{2} \right)^2 + U_0 \cdot \left[ 1 - \left( \frac{1-p}{2} \right)^2 \right] = 9p \frac{1-p}{2} + U_0 \cdot \left[ 1 - p \left( \frac{1-p}{2} \right) \right]
\]
which simplifies to give the symmetric equilibrium value of $p$ for the single communication stage,
\[
p_1 = \frac{9A - U_0}{9A - 3U_0 + 18} = \frac{9A - U_0}{9A - U_0 + 18 - 2U_0}.
\]
Table 2: Payoffs to player 1 to each pure strategy profile and their probabilities in symmetric mixed strategy $p$.

<table>
<thead>
<tr>
<th>$(a_1, a_2, a_3)$</th>
<th>$\Pr(a_1, a_2, a_3)$</th>
<th>$u_1(a_1, a_2, a_3)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,1,1</td>
<td>$p \left( \frac{1-p}{2} \right)^2$</td>
<td>$9A$</td>
</tr>
<tr>
<td>1,1,2</td>
<td>$p \left( \frac{1-p}{2} \right)^2$</td>
<td>$4A + 1$</td>
</tr>
<tr>
<td>1,1,3</td>
<td>$p^2 \left( \frac{1-p}{2} \right)$</td>
<td>$4A + 1$</td>
</tr>
<tr>
<td>1,2,1</td>
<td>$p^2 \left( \frac{1-p}{2} \right)$</td>
<td>$4A + 1$</td>
</tr>
<tr>
<td>1,2,2</td>
<td>$p^2 \left( \frac{1-p}{2} \right)$</td>
<td>$A + 4$</td>
</tr>
<tr>
<td>1,2,3</td>
<td>$p^3$</td>
<td>$A + 2$</td>
</tr>
<tr>
<td>1,3,1</td>
<td>$p \left( \frac{1-p}{2} \right)^2$</td>
<td>$4A + 1$</td>
</tr>
<tr>
<td>1,3,2</td>
<td>$p \left( \frac{1-p}{2} \right)^2$</td>
<td>$A + 2$</td>
</tr>
<tr>
<td>1,3,3</td>
<td>$p^2 \left( \frac{1-p}{2} \right)$</td>
<td>$A + 4$</td>
</tr>
<tr>
<td>2,1,1</td>
<td>$\left( \frac{1-p}{2} \right)^3$</td>
<td>$4A + 1$</td>
</tr>
<tr>
<td>2,1,2</td>
<td>$\left( \frac{1-p}{2} \right)^3$</td>
<td>$A + 4$</td>
</tr>
<tr>
<td>2,1,3</td>
<td>$p \left( \frac{1-p}{2} \right)^2$</td>
<td>$A + 2$</td>
</tr>
<tr>
<td>2,2,1</td>
<td>$p \left( \frac{1-p}{2} \right)^2$</td>
<td>$A + 4$</td>
</tr>
<tr>
<td>2,2,2</td>
<td>$p \left( \frac{1-p}{2} \right)^2$</td>
<td>$9$</td>
</tr>
<tr>
<td>2,2,3</td>
<td>$p^2 \left( \frac{1-p}{2} \right)$</td>
<td>$5$</td>
</tr>
<tr>
<td>2,3,1</td>
<td>$\left( \frac{1-p}{2} \right)^3$</td>
<td>$A + 2$</td>
</tr>
<tr>
<td>2,3,2</td>
<td>$\left( \frac{1-p}{2} \right)^3$</td>
<td>$5$</td>
</tr>
<tr>
<td>2,3,3</td>
<td>$p \left( \frac{1-p}{2} \right)^2$</td>
<td>$5$</td>
</tr>
<tr>
<td>3,1,1</td>
<td>$\left( \frac{1-p}{2} \right)^3$</td>
<td>$4A + 1$</td>
</tr>
<tr>
<td>3,1,2</td>
<td>$\left( \frac{1-p}{2} \right)^3$</td>
<td>$A + 2$</td>
</tr>
<tr>
<td>3,1,3</td>
<td>$p \left( \frac{1-p}{2} \right)^2$</td>
<td>$A + 4$</td>
</tr>
<tr>
<td>3,2,1</td>
<td>$p \left( \frac{1-p}{2} \right)^2$</td>
<td>$A + 4$</td>
</tr>
<tr>
<td>3,2,2</td>
<td>$p \left( \frac{1-p}{2} \right)^2$</td>
<td>$5$</td>
</tr>
<tr>
<td>3,2,3</td>
<td>$p^2 \left( \frac{1-p}{2} \right)$</td>
<td>$5$</td>
</tr>
<tr>
<td>3,3,1</td>
<td>$\left( \frac{1-p}{2} \right)^3$</td>
<td>$A + 4$</td>
</tr>
<tr>
<td>3,3,2</td>
<td>$\left( \frac{1-p}{2} \right)^3$</td>
<td>$5$</td>
</tr>
<tr>
<td>3,3,3</td>
<td>$p \left( \frac{1-p}{2} \right)^2$</td>
<td>$9$</td>
</tr>
</tbody>
</table>
This probability increases monotonically from $1/3$ at $A = 1$ to $11/15$ at $A = 3$, so it is a valid probability—and therefore the desired symmetric communication equilibrium exists—throughout the allowable range of $A$. These probabilities, as functions of $A$, are illustrated in Figure 1, where $p_0$ is the solid line and $p_1$ the dotted line.

![Figure 1: Symmetric equilibrium mixed-strategy probabilities $p_0$ (solid line), $p_1$ (dotted line), and $q_1$ (dashed line), as functions of $A$.](image)

For $A \in [1, 3]$, it is always true that $p_1 \leq p_0$, with equality occurring only at $A = 1$. This contrasts with Farrell’s (1987) result for the two-player BoS—there, the equilibrium probability that a player “insists” on her own preferred action in the communication stage is *higher* than her probability of playing her own preferred action in the action stage. This contrast is due to the difference between the two models’ payoffs for partial, or imperfect, coordination. With only two players, of course, that issue does not arise. But consider a different kind of $n$-player BoS, $n > 2$, in which only *perfect* coordination at the action stage yields positive payoffs, and otherwise all players receive zero. Such a model would exhibit the same pattern as does the Farrell model: with one stage of pre-play communication, the symmetric equilibrium probability of announcing one’s preferred action would be higher than the symmetric equilibrium probability of playing one’s preferred action in the action stage. The present model, however, provides *intermediate* levels of payoff for partial coordination, which has the effect of increasing the symmetric equilibrium $p_0$ in the action stage relative to the Farrell-type model. At the communication stage, however, any non-consensus profile of announcements yields the same, low expected payoff, $U_0$—in effect putting the players back in the Farrell-type situation. This shift of payoff patterns between partial coordination in the communication stage and partial coordination in the action stage is what makes the present model behave so differently from the Farrell model.

Obviously, having one stage of communication increases the probability that the players will ultimately coordinate on the same action, since it gives them an extra opportunity to coordinate, while leaving the strategies in the action stage unchanged if they fail at the communication stage. Under the consensus rule, the probability of coordinating in the communication stage is $3p_1 \left( \frac{1-p_1}{2} \right)^2$, which lies below the probability of coordinating in mixed strategies in the final stage since $p_1 < p_0$. However, the overall probability of coordinating
is now
\[3p_1 \left( \frac{1 - p_1}{2} \right)^2 + \left[ 1 - 3p_1 \left( \frac{1 - p_1}{2} \right)^2 \right] 3p_0 \left( \frac{1 - p_0}{2} \right)^2\]
which can be rewritten as
\[3p_0 \left( \frac{1 - p_0}{2} \right)^2 + \left[ 1 - 3p_0 \left( \frac{1 - p_0}{2} \right)^2 \right] 3p_1 \left( \frac{1 - p_1}{2} \right)^2,
\]
which is the no-communication probability of complete coordination, plus a positive term.

The probability that the three players will eventually choose three different actions when there is communication under the consensus rule also obviously decreases compared with the no-communication game. This probability equals the probability of doing so in the absence of communication, times the probability of failing to reach consensus in the communication stage:
\[\left[ p_0^3 + \frac{1}{4}(1 - p_0)^3 + \frac{3}{4}p_0(1 - p_0)^2 \right] \left[ 1 - 3p_1 \left( \frac{1 - p_1}{2} \right)^2 \right]\]

The payoff to each player in the symmetric equilibrium with pre-play communication under the consensus rule is the probability of consensus in announcements, times the average payoff from full coordination, \((9A + 9 + 9)/3 = 3A + 6\), plus the probability of no consensus in announcement times \(U_0\). The average payoff from full coordination ranges from 9 to 15 as \(A\) ranges from 1 to 3, so it is always at least twice as large as \(U_0\); it is immediately clear that equilibrium communication with the consensus rule greatly increases expected payoff. Specifically, the expected payoff from one stage of pre-play communication is
\[3p_1 \left( \frac{1 - p_1}{2} \right)^2 (3A + 6) + \left[ 1 - 3p_1 \left( \frac{1 - p_1}{2} \right)^2 \right] U_0\]
that is,
\[U_{1C}^C = U_0 + 3p_1 \left( \frac{1 - p_1}{2} \right)^2 (3A + 6 - U_0)\]

Notice that, just treating \(p_1\) as a variable, \(U_{1C}^C\) is increasing in \(p_1\):
\[\frac{\partial U_{1C}^C}{\partial p_1} = (3A - 6 + U_0) \left[ 3 \left( \frac{1 - p_1}{2} \right)^2 - 3p_1^2 \right] = \frac{3}{4}(3A - 6 + U_0)(3p_1^2 + 2p_1 - 1)\]
whose sign is opposite the sign of the last parenthesized factor. Since \(p_1\) ranges upwards from 1/3 as \(A\) increases from 1 to 3, that factor is nonnegative, so the derivative itself is nonpositive. Thus as the equilibrium probability of announcing one’s own favorite action increases, the expected utility of the overall BoS with one stage of consensus communication decreases (strictly, except at its upper and lower limits).
5.2 The plurality rule

As an alternative to the consensus rule, suppose the common expectation of the players is that if any two or more players announce the same action, all will play that action with certainty in the final stage. On the one hand, we might expect a “plurality” equilibrium to yield coordination with higher probability in any particular period, and therefore to yield a higher expected payoff. On the other hand, this very tendency should lead a strategic player to assign her own favorite action a higher probability in the mixed-strategy equilibrium. Which tendency will predominate?

The principle for identifying a symmetric equilibrium is the same as before. Let $\Pr(\text{two } k|j)$ represent the probability that at least two players announce $k$ given that player 1 announces $j$. These probabilities are

\[
\Pr(\text{two } 1|1) = p(1 - p) + 3 \left( \frac{1 - p}{2} \right)^2 = \frac{3 - 2p - p^2}{4}
\]

\[
\Pr(\text{two } 2|1) = \Pr(\text{two } 3|1) = p \frac{1 - p}{2}
\]

\[
\Pr(\text{two } 1|2) = \Pr(\text{two } 1|3) = \left( \frac{1 - p}{2} \right)^2
\]

\[
\Pr(\text{two } 2|2) = \Pr(\text{two } 3|3) = 2 \left( \frac{1 - p}{2} \right)^2 + p(1 - p) + p^2 = \frac{1 + p^2}{2}
\]

\[
\Pr(\text{two } 2|2) = \Pr(\text{two } 3|3) = \frac{1 - p}{2}.
\]

Thus the total probability of a plurality for any other player’s favorite action if player 1 announces another player’s favorite action is

\[
\Pr(\text{two } 2|2) + \Pr(\text{two } 3|2) = \frac{1 + p}{2}.
\]

Finally, the probability of no plurality, that is, three different announcements from the three players, depending on 1’s announcement, is

\[
\Pr(\text{three different}|1) = p^2 + \left( \frac{1 - p}{2} \right)^2 = \frac{5p^2 - 2p + 1}{4}
\]

\[
\Pr(\text{three different}|2) = \Pr(\text{three different}|3) = p \frac{1 - p}{2} + \left( \frac{1 - p}{2} \right)^2 = \frac{1 - p^2}{4}.
\]

Thus if player 1 announces 1, her expected payoff is

\[
9A \cdot \Pr(\text{two } 1|1) + 9 \cdot \Pr(\text{two } 2|1) + 9 \cdot \Pr(\text{two } 3|1) + U_0 \cdot \Pr(\text{three different}|1))
\]

or, formally,

\[
9A \frac{3 - 2p - p^2}{4} + 9p(1 - p) + U_0 \frac{5p^2 - 2p + 1}{4}.
\]
If player 1 announces 2 or 3, her expected payoff is
\[ 9A \cdot \Pr(\text{two 1}|2) + 9 \Pr(\text{two 2}|2) + 9 \Pr(\text{two 3}|2) + U_0 \cdot \Pr(\text{three different}|2) \]
that is,
\[ 9A \left( \frac{1-p}{2} \right)^2 + 9 \frac{1+p}{2} + U_0 \frac{1-p^2}{4} \]
Setting these two expected payoffs equal and solving for a positive root \( p \) gives the symmetric equilibrium mixed-strategy probability, \( q_1 \), of announcing one’s own favorite action, under the “plurality rule.” Leaving the solution partially in terms of \( U_0 \), which is itself a function of \( A \), we have
\[ p = q_1 = \frac{9 - U_0 + \sqrt{-567 + 324A + 324A^2 + 90U_0 - 108AU_0 + U_0^2}}{6(6 + 3A - U_0)} \]
Given entirely in terms of \( A \), the solution is one of the roots of a really ugly quadratic:
\[ q_1 = \frac{-44 + 7A - 32A^2 - 3A^3 - r}{2(-114 + 75A - 96A^2 - 81A^3)} \]
where
\[ r = \sqrt{-6272 - 19840A + 12153A^2 + 6080A^3 - 38114A^4 + 27840A^5 + 23337A^6} \]
(These calculations are pure Mathematica, checked to make sure I’ve adopted the correct root and that it is a valid probability throughout the allowable range of \( A \).)
Comparing \( q_1 \) with \( p_0 \) shows that the two are equal at approximately \( A = 1.8 \). For values of \( A \) between 1 and 1.8, \( q_1 > p_0 \), while for larger \( A \), \( p_0 > q_1 \) by an increasing margin. Communication under the plurality criterion, unlike the consensus rule, rewards partial coordination in the announcement stage as well as at the action stage. Because of this, the question of whether, as in Farrell (1987), the probability of sticking to one’s preferred action is higher in the communication stage than in the action stage is more complicated. For lower values of \( A \), the Farrell intuition predominates, and players “talk tougher” than they finally act. For higher \( A \)-values, however, the prospect of a positive payoff from partial cooperation at the action stage evidently acts as more of a safety net for selfish choice than it does at the communication stage. This switch-over offers a nice opportunity for an experimental investigation of a sort of equilibrium bargaining rationality.
As under the consensus rule, the equilibrium expected payoff to each player from the game with one stage of communication under the plurality rule can be expressed as the expected payoff from the action stage, plus the expected gain from the communication stage: informally,
\[ U_1^P = U_0 + \Pr(\text{coordinated announcements})(3A + 6 - U_0) \]
For the plurality rule, the probability in the last term is most easily calculated as one minus the probability of getting three different announcements. There are six different ways to get three different announcements; using the probabilities from Table 2 and simplifying, the total probability of not doing so is
\[ \frac{3}{4}(-2q_1^3 + q_1^2 + 1) \]
thus,
\[ U_1^P = U_0 + \frac{3}{4}(-2q_1^3 + q_1^2 + 1)(3A + 6 - U_0). \]

5.3 Comparison between consensus and plurality communication

Like \( p_0 \) and \( p_1 \), \( q_1 \) takes on the value 1/3 when \( A = 1 \) and is monotonically increasing and concave, staying above \( p_1 \) and reaching about 0.844 at \( A = 3 \); see Figure 1 for illustration. In other words, each player is more likely to insist on her own preferred action in the communication stage under the plurality rule than under the consensus rule. This makes some intuitive sense, since, with more ways to coordinate, there is perhaps less need to give up on one’s preferred action in the communication stage.

Under the plurality rule, the probability of coordinating at the communication stage is much higher than under the consensus rule, despite the larger probability that each player insists on her own preferred action. Consequently, the expected payoff under plurality is much higher: depending on \( A \), expected payoff under consensus-rule communication ranges between about 5.39 and 5.85, only about 0.4 above the expected payoff with no communication. Under plurality, expected payoff is higher by around 3.0, ranging between about 8.12 and 8.82. Figure 2 graphs \( U_0 \), \( U_1^C \), and \( U_1^P \) as functions of \( A \).

Another notable difference between plurality and the no-communication and consensus procedures is that plurality yields ever higher expected payoffs as \( A \), the measure of competition, increases toward its maximum value. Under consensus, or with no communication, expected payoff peaks at intermediate values of \( A \), after which the temptation to pursue one’s own favorite outcome begins to outweigh the attraction of coordinating. Under the plurality procedure, despite the higher probability of insisting on one’s favorite, the lower threshold for reaching agreement yields an ever higher expected payoff.

Figure 2: Symmetric equilibrium expected payoffs from no communication (\( U_0 \), solid line), one stage of communication under the consensus rule (\( U_1^C \), dotted line), and one stage of communication under the plurality rule (\( U_1^P \), dashed line), as functions of \( A \).
Multiple stages of pre-play communication

Rather than a single stage of pre-play communication, suppose there are $T$ identical stages, with the players hearing the announcements of one stage before going on to the next. As does Farrell (1987), I focus on the following simple type of equilibrium treatment of the repeated announcements: If at any stage the three players reach sufficient agreement, then all future communication is ignored and, in the final stage, they play that action. Below I examine the case in which “sufficient” agreement at each stage is “consensus” and the case of “plurality.” In each stage, If sufficient agreement has not previously been reached, the players use a symmetric mixed-strategy equilibrium that is completely independent of previous messages.

6.1 The consensus rule

In the Farrell (1987) model, increasing the number of communication stages increases the expected payoff, but even in the limit as $T \rightarrow \infty$ the total payoff is bounded away from the total payoff from perfect symmetric coordination. (In particular, in the limit both players receive the same expected payoff that they would get from coordinating on the other player’s favored equilibrium.) That is, more communication is always better, but it can never approximate the expected payoff from a correlated strategy profile that puts equal probability on each of the (fully) coordinated outcomes and no probability on any other outcome. The same holds true for communication in the multi-player BoS.

Let $t$ be the number of remaining communication stages, counting the current stage. The calculation for $p_t$, the symmetric mixed-strategy equilibrium probability of announcing one’s own favorite action, is very similar to that for the case of one-stage communication. Let $U_t^C$ represent the expected payoff from this equilibrium when

- there are $t$ communication stages remaining,
- the stage-$t$ announcements have not yet been made, and
- no previous stage has resulted in a consensus of announcements.

Then for each $t = 1, 2, \ldots, T$, mixed-strategy equilibrium requires

$$9A \cdot \Pr(\text{all } i|i) + U_{t-1}^C \cdot [1 - \Pr(\text{all } i|i)] = 9 \cdot \Pr(\text{all } j|j) + U_{t-1}^C \cdot [1 - \Pr(\text{all } j|j)]$$

just as before, except that now the probabilities of consensus are calculated with $p_t$ instead of $p_1$. Formally, this is

$$9A \left( \frac{1 - p_t}{2} \right)^2 + U_{t-1}^C \cdot \left[ 1 - \left( \frac{1 - p_t}{2} \right)^2 \right] = 9 \cdot \frac{1 - p_t}{2} + U_{t-1}^C \cdot \left[ 1 - \frac{1 - p_t}{2} \right].$$

In the same way that $p_1$ and $U_1^C$ were solved in terms of $p_0$ and $U_0$ previously, then, one can derive the following recursive formulas for $p_t$ and $U_t$, for $t = 1, 2, \ldots$:
\[ p_t = \frac{9A - U_{t-1}^C}{9A - 3U_{t-1}^C + 18} \quad (1) \]
\[ U_t^C = 3p_t \left( \frac{1 - p_t}{2} \right)^2 (3A + 6) + \left[ 1 - 3p_t \left( \frac{1 - p_t}{2} \right)^2 \right] U_{t-1}^C. \quad (2) \]

Thus \( U_t^C \) is a weighted average of \( U_{t-1}^C \) and \( 3A + 6 \), the latter being the expected payoff from the correlated equilibrium that uses each fully coordinated outcome with probability \( 1/3 \). Substituting equation 1 for \( p_t \) into equation 2 and simplifying gives the following recursive expression in \( U_t^C \) alone:

\[ U_t^C = U_{t-1}^C + \frac{(9A - U_{t-1}^C)(9 - U_{t-1}^C)^2}{(9A - 3U_{t-1}^C + 18)^2}. \quad (3) \]

Recall that \( U_0^C \equiv U_0 \) lies between 5 and 5.4 for all \( A \in [1, 3] \). Throughout the range \( U_{t-1}^C \in [5, 9] \), equation 3 for \( U_t^C \) is an increasing function of \( U_{t-1}^C \) that lies above \( U_{t-1}^C \) and converges to it monotonically at \( U_{t-1}^C = 9 \), as shown in Figure 3. The figure is plotted based on a value of \( A = 2 \), but varying \( A \) across its whole range \([1, 3]\) makes no perceptible difference in the graph at this level of resolution. Thus in general, \( (U_t^C)_{t=1}^\infty \) is a monotone increasing sequence converging to 9, which is the payoff from coordinating with certainty on another player’s favorite action. This sequence can be traced out in Figure 3 by beginning at any point on the solid curve above a \( U_{t-1}^C \)-value between 5 and 5.4; and then moving horizontally to the dashed 45-degree line, vertically to the solid curve, horizontally again to the dashed line, and so forth. The values thus traced out on either axis are the elements of the sequence.

Aside from the extreme case of \( A = 1 \), then, even with many stages of communication, the expected payoff for all players is less than, and bounded away from, the efficient, symmetric ideal, just as in the Farrell model. This convergence of \( U_t^C \) also implies, by equation 1, that \( p_t \to 1 \) as \( t \to \infty \), again the same as with the two-player BoS.

### 6.2 The plurality rule

The corresponding recursion equations for communication under the plurality rule are

\[ q_t = \frac{9 - U_{t-1}^P + \sqrt{-567 + 324A + 324A^2 + (90 - 108A)U_{t-1}^P + (U_{t-1}^P)^2}}{6(6 + 3A - U_{t-1}^P)} \quad (4) \]

and

\[ U_t^P = U_{t-1}^P + \frac{3}{4}(-2q_t^3 + q_t^2 + 1)(3A + 6 - U_{t-1}^P). \quad (5) \]

Numerical analysis indicates the same result as for consensus communication, although the approach to the limits—again \( q_t = 1 \) and \( U_t^P = 9 \)—is much faster under the plurality rule.
Figure 3: $U_t^C$ as a function of $U_{t-1}^C$, for the case $A = 2$. The dashed line marks 45 degrees from the origin.

Writing the radical in equation 4 as $R$, and then using equation 4 to substitute for $q_t$ in equation 5, gives

$$U_P^t = U_{t-1}^P + \frac{3}{4}(6 + 3A - U_{t-1}^P) + \frac{(9 - U_{t-1}^P + R)(9 + 9A - 2U_{t-1}^P - R)}{144(6 + 3A - U_{t-1}^P)^3}. \quad (6)$$

Figure 4 shows the graph of $U_P^t$ against $U_{t-1}^P$. Once again, the expected payoff monotonically approaches 9, the payoff from full coordination off one’s favorite outcome, as $t$ becomes large. This convergence is more rapid the higher the value of $A$.

### 6.3 Comparing consensus and plurality communication with many rounds

Comparing Figures 3 and 4, it is obvious that, although the two procedures approach the same expected payoff in the limit as $t \to \infty$, the convergence is much more rapid under plurality. Numerical analysis shows that a half-dozen plurality-rule communication rounds is sufficient to reach an expected payoff over 8.999 regardless of the value of $A$. Under the consensus rule, by contrast, this payoff level is not reached until $t > 70$ even for the most favorable case of $A = 1$.

Nevertheless, under our assumption that the messages are interpreted as “proposals,”
both methods of communication bump up against the same ceiling. They approach the expected payoff of the efficient, symmetric correlated equilibrium only in the trivial case of $A = 1$.

Although equations 1 and 4 are intractable, this result concerning the limit payoffs proves that neither method of communication can eliminate coordination failure. If coordination failure disappeared in the limit, each player’s payoff would be $9A$ in the $1/3$ of cases when they coordinate on that player’s favorite outcome, and $9$ in all other cases, for an expected payoff of $3A + 6 \geq 9$. The entire shortfall is attributable to coordination failure. The limit probability of coordination failure must be a value $\phi$ such that

$$(1 - \phi) \left[ \frac{1}{3} \cdot 9A + \frac{2}{3} \cdot 9 \right] + \phi U_0 = 9$$

Plugging in the formula for $U_0$ in terms of $A$ and solving gives $\phi$ as a function of $A$:

$$\phi = \frac{6(A - 1)(2A + 1)^2}{27A^3 + 32A^2 - 25A + 38},$$

a concave, increasing function ranging from zero when $A = 1$ to $3/5$ when $A = 3$. Thus for any particular $A > 1$, the probability of failing to coordinate remains bounded away from 0 no matter how many rounds of communication are allowed.

Figure 4: $U_P^n$ as a function of $U_P^{n-1}$, for the case $A = 1.1$ (dotted curve) and $A = 2.5$ (solid curve). The dashed line marks 45 degrees from the origin.
7 Conclusions

I have argued here that political deliberation should be viewed as a process for determining an equilibrium of behavior in the ensuing “subgame” of social interaction. That subgame has multiple equilibria, presenting the participants with a mixed-motive coordination problem. Deliberators propose, in effect, particular equilibria by stating what they themselves would do. In real life, of course, they do so by appealing to matters of principle and precedent that come to have shared meaning. Evolutionary game theory models indicate that such communication games tend to evolve meaning; that is, equilibria in which messages influence future behavior have stronger stability properties than equilibria in which the messages are essentially meaningless (for example, Blume et al. 2001). These results from the theoretical literature lend credence to the intuition of Farrell and Rabin (1996) that participants in many communication games are likely to develop a system of “natural language” that creates focal points. In the setting of deliberation, I suggest that this insight about “natural language” can be applied similarly to the development of a shared system of political values or principles, language elements that can be invoked to influence the formation of focal points in social behavior. The most basic form of such principles or values, which shows up in a literal way in the coordination game, is the idea that we can all gain if we can just agree on coordinating, even if we can’t all get our most-preferred outcome. The temptation for deliberators is to advocate, on general-interest grounds, one’s own preferred outcome as the place where the coordination should focus.

Using Farrell’s (1987) model of coordination through pre-play communication as a template, then, I have examined how pre-play communication might influence such coordination among a larger group of actors. My basic model of the extended battle-of-the-sexes coordination problem offers a picture of such a coordination game, and the “plurality” and “consensus” criteria for establishing mutual expectations represent two simple kinds of equilibrium in the extended game of pre-play communication followed by action. Both aspects of the model, the game and the communication equilibria examined, obviously leave plenty of room for elaboration. For example—especially for purposes of theorizing about politics—the model ideally should include the possibility of choosing actions that are relatively near or distant from each other, allowing for another form of partial coordination, namely the agglomeration of players on “nearby,” but not identical, actions. And of course, the plurality and consensus criteria hardly do justice to most real-life processes of reaching general agreement and compliance.

(The use of the plurality and consensus criteria here, however, does suggest another interesting application of the model to our understanding of self-enforcing political institutions. In most theories of political choice, legislative votes are assumed to determine outcomes with no possibility of slippage—separation-of-powers models notwithstanding, since they too involve voting or veto acts that are assumed to be determinative. But votes, really, are just communicative acts; in an established political system, saying “Yea” or “Nay” by certain people in certain circumstances has, by mutual expectation of officials and citizens alike, a hugely disproportionate impact upon everyone’s future behavior. But at root, that influence is either just an equilibrium phenomenon, or is soon to disintegrate. So the “plurality” and “consensus” analyses pursued in this paper are the beginnings of a theory of endogenous political decision making institutions, in the spirit of Schotter (1981) and Calvert (1998).)
In the model, the relative difficulty of achieving coordination under the consensus criterion is potentially offset by its greater demands on strategic players to “give in” and agree to some other actor’s favored coordination solution. The analysis of the three-player example showed that, in fact, the difficulty was largely not offset: under the plurality criterion, despite the greater tendency of participants to insist on their own favorite outcomes, the likelihood of successful coordination after just one round of communication is in fact much higher than under consensus.

If an extended sequence of back-and-forth communication is allowed, however, the likelihood of coordination under both criteria approaches the same upper bound as the length of communication increases, the plurality rule approaching this limit much more slowly than the consensus rule. That upper limit, exactly as in the original Farrell (1987) analysis, yields an expected payoff to each player equal to that achieved by coordinating with certainty on one of her un-favored outcomes. Strategic uncertainty, mixed motives, and the “natural language” combine to leave an unavoidable amount of loss relative to the expected payoff that could be achieved in a correlated equilibrium. Deliberation, viewed in this light, improves political stability and the achievement of jointly desired political ends; but it cannot completely reverse the effects of differing interests and the temptation to promote one’s private interest based on general-interest arguments.
References


