FUNDAMENTALS OF SOCIAL CHOICE THEORY

by Roger B. Myerson[1]

http://home.uchicago.edu/~rmyerson/research/schch1.pdf

Abstract: This paper offers a short introduction to some of the fundamental results of social choice theory. Topics include Nash implementability, monotonic social choice correspondences, the Muller-Satterthwaite impossibility theorem, anonymous and neutral social choice correspondences, sophisticated solutions of binary agendas, the top cycle of a tournament, the bipartisan set for two-party competition, and median voter theorems. The paper begins with a simple example to illustrate the importance of multiple equilibria in game-theoretic models of political institutions.

1. An introductory model of political institutions

Mathematical models in social science are like fables or myths that we read to get insights into the social world in which we live. Our mathematical models are told in a specialized technical language that allows very precise descriptions of the motivations and choices of the various individuals in these stories. When we prove theorems in mathematical social science, we are making general statements about whole classes of such stories all at once. Here we focus on game-theoretic models of political institutions.

So let us begin our study of political institutions by a simple game-theoretic model that tells a story of how political institutions may arise from games with multiple equilibria, such as those analyzed by Schelling (1960). Consider first the simple two-person "Battle of Sexes" game shown in Table 1.

		Player 2	
		f_2	g_2
Player 1	f_1	0, 0	3, 6
	g_1	6, 3	0, 0

Table 1. The Battle of Sexes game.

[1] The author is greatly indebted to Scott Ashworth for editorial comments that helped bring this paper into the new millennium.

The two players in this game must independently choose one of two possible strategies: to defer (f_i) or to grab (g_i). If the players both grab or both defer then neither player gets anything; but if exactly one player grabs then he gets payoff 6 while the deferential player gets payoff 3.

This game has three equilibria. There is an equilibrium in which player 1 grabs while player 2 defers, giving payoffs (6,3). There is another equilibrium in which player 1 defers while player 2 grabs, giving payoffs (3,6). There is also a symmetric randomized equilibrium in which each player independently randomizes between grabbing, with probability 2/3, and deferring, with probability 1/3. In this randomized equilibrium, the expected payoffs are (2,2), which is worse for both players than either of the nonsymmetric equilibria.

Now think of an island with a large population of individuals. Every morning, these islanders assemble in the center of their island, to talk and watch the sun rise. Then the islanders scatter for the day. During the day, the islanders are randomly matched into pairs who meet at random locations around the island, and each of these matched pairs plays the simple Battle of Sexes game once. This process repeats every day. Each player's objective is to maximize a long-run discounted average of his (or her) sequence of payoffs from these daily Battle of Sexes matches.

One long-run equilibrium of this process is for everyone to play the symmetric randomized equilibrium in his match each day. But rising up from this primitive anarchy, the players could develop cultural expectations which break the symmetry among matched players, so that they will share an understanding of who should grab and who should defer.

One possibility is that the islanders might develop an understanding that each player has a special "ownership" relationship with some region of the island, such that a player is expected to grab whenever he is in the region that he owns. Notice that this system of ownership rights is a self-enforcing equilibrium, because the other player does better by deferring (getting 3 rather than 0) when he expects the "owner" to grab, and so the owner should indeed grab confidently.

But such a system of traditional grabbing rights might fail to cover many matching situations where no one has clear "ownership." To avoid the costly symmetric equilibrium in such cases, other ways of breaking the players' symmetry are needed. A system of leadership can be used to solve this problem.

That is, the islanders might appoint one of their population to serve as a leader, who will announce each morning a set of instructions that specify which one of the two players should grab in each of the daily matches. As long as the leader's instructions are clear and comprehensive, the understanding that every player will obey these instructions is a self-enforcing equilibrium. A player who grabbed when he was instructed to defer would only lower his expected payoff from 3 to 0, given that the other player is expected to follow his instruction to grab here.

To make this system of government work on our island, the islanders only need a shared understanding as to who is the leader. The leader might be the eldest among the islanders, or the tallest, or the one with the loudest voice. Or the islanders might determine their leader by some contest, such as a chess tournament, or by an annual election in which all the islanders vote. Any method of selection that the islanders understand can be used, because everyone wants to obey the selected leader's instructions as long as everyone else is expected to obey him. Thus, self-enforcing rules for a political system can be constructed arbitrarily from the equilibrium selection problem in this game.

The islanders could impose limits to a leader's authority in this political system. For example, there might be one leader whose instructions are obeyed on the northern half of the island, and another leader whose instructions are obeyed on the southern half. The islanders may even have ways to remove a leader, such as when he loses some re-election contest or when he issues instructions that violate some perceived limits. If a former leader tries to make an announcement in his former domain of authority, every player would be expected to ignore this announcement as irrelevant cheap talk.

Of course, the real world is very different from the simple island of this fable. But as in this island, coordination games with multiple equilibria are pervasive in any real society. Thus, any successful society must develop structures of law and leadership that can coordinate people's expectations in situations of multiple equilibria. So the first point of this fable is the basic social need for coordinating leadership and for political institutions that can provide it.

The second point of this fable is that the effectiveness of any political institution may be derived simply from a shared understanding that it is in effect, as Hardin (1989) has argued. Thus, any political system may be one of many possible equilibria of a more fundamental coordination game of constitutional selection. That is, the process of selecting a constitution can be viewed as an equilibrium-selection problem, but it is the equilibrium-selection problem to solve all other equilibrium-selection problems.

The arbitrariness of political structures from this game-theoretic perspective validates our treating them as exogenous explanatory parameters in political economics. A question might be posed, for example, as to whether one form of democracy might generate higher economic welfare than some other forms of government. Such a question would be untestable or even meaningless if the form of government were itself determined by the level of economic welfare. But our fable suggests that the crucial necessary condition for democracy is not wealth or literacy, but is simply a shared understanding that democracy will function in this society (so that an officer who waves his pistol in the legislative chamber should be perceived as a madman in need of psychiatric treatment, not as the new leader of the country).

2. A general impossibility theorem

There is an enormous diversity of democratic political institutions that could exist. Social choice theory is a branch of mathematical social science that tries to make general statements about all such institutions. Given the diversity of potential institutions, the power of social choice theory may be quite limited, and indeed its most famous results are negative impossibility theorems. But it is good to start with the general perspective of social choice theory and see what can be said at this level, before turning to other models of formal political theory that focus more narrowly on specific kinds of institutions that may exist in the world.

Modern social choice theory begins with the great theorem of Arrow (1951). This theorem has led to many other impossibility theorems, notably the theorem of Gibbard (1973) and Satterthwaite (1975). (See also Sen, 1970.) In this section, we focus on the theorem of Muller and Satterthwaite (1977), because this is the impossibility theorem that applies directly to Nash-equilibrium implementation. The Muller-Satterthwaite theorem was first proven as a consequence of the Gibbard-Satterthwaite and Arrow theorems, but we prove it here directly, following Moulin (1988).

Let N denote a given set of individual voters, and let Y denote a given set of alternatives or social-choice options among which the voters must select one. We assume that N and Y are both nonempty finite sets. Let L(Y) denote the set of strict transitive orderings of the alternatives in Y. Given that there are only finitely many alternatives, we may represent any individual's preference ordering in L(Y) by a utility function u_i such that $u_i(x)$ is the number of alternatives that the individual i considers to be strictly worse than x. So with strict preferences, L(Y) can be identified with the set of one-to-one functions from Y to the set {0,1,...,#Y-1}. (Here #Y denotes the number of alternatives in the set Y.)

We let $L(Y)^N$ denote the set of profiles of such preference orderings, one for each individual voter. We denote such a preference profile by a profile of utility functions $u = (u_i)_{i \in N}$, where each u_i is in L(Y). So if the voters' preference profile is u, then the inequality $u_i(x) > u_i(y)$ means that voter i prefers alternative x over alternative y. The assumption of strict preferences implies that either $u_i(x) > u_i(y)$ or $u_i(y) > u_i(x)$ must hold if $x \neq y$.

A political system creates a game that is played by the voters, with outcomes in the set of alternatives Y. The voters will play this game in a way that depends on their individual preferences over Y, and so the realized outcome may be a function of the preference profile in $L(Y)^N$. From the abstract perspective of social choice theory, an institution could be represented by the mapping that specifies these predicted outcomes as a function of the voters' preferences. So a <u>social choice function</u> is any function $F{:}L(Y)^N \rightarrow Y$, where F(u) may be interpreted as the alternative in Y that would be chosen (under some given institutional arrangement) if the voters' preferences were as in u.

If there are multiple equilibria in our political game, then we may have to talk instead about a set of possible equilibrium outcomes. So a <u>social choice correspondence</u> is any point-to-set mapping $G{:}L(Y)^N \rightarrow\rightarrow Y$. Here, for any preference profile u, G(u) is a subset of Y that may be interpreted as the set of alternatives in Y that might be chosen by society (under some institutional arrangement) if the voters' preferences were as in u.

A social choice function F is <u>monotonic</u> iff, for every pair of preference profiles u and v in $L(Y)^N$, and for every alternative x in Y, if x = F(u) and

$$\{y|\ v_i(y) > v_i(x)\} \subseteq \{y|\ u_i(y) > u_i(x)\},\ \ \forall i \in N,$$

then x = F(v). Similarly, a social choice correspondence G is <u>monotonic</u> iff, for every u and v in $L(Y)^N$, and for every x in Y, if $x \in G(u)$ and

(1) $$\{y|\ v_i(y) > v_i(x)\} \subseteq \{y|\ u_i(y) > u_i(x)\},\ \ \forall i \in N,$$

then $x \in G(v)$.

Maskin (1985) showed that any social choice correspondence that is constructed as the set of Nash equilibrium outcomes of a fixed game form must be monotonic in this sense. A game form is a function of the form $H:\times_{i\in N} S_i \to Y$ where each S_i is a nonempty strategy set for i. The pure Nash equilibrium outcomes of the game form H with preferences u is the set

$$E(H,u) = \{H(s) \mid s \in \times_{i\in N} S_i, \text{ and, } \forall i\in N, \forall r_i\in S_i, u_i(H(s)) \geq u_i(H(s_{-i},r_i))\}.$$

That is, x is a Nash equilibrium outcome in E(H,u) iff there exists a profile of strategies s such that x=H(s) and no individual i could get an outcome of H that he would strictly prefer under the preferences u_i by unilaterally deviating from s_i to another strategy r_i. Condition (1) above says that the set of outcomes that are strictly better than x for any player is the same or smaller when the preferences change from u to v, and so x=H(s) must still be an equilibrium outcome under v. So if $x\in E(H,u)$ and the preference profiles u and v satisfy condition (1) for x, then we must have $x\in E(H,v)$. Thus, for any game form H, the social choice correspondence $E(H,\bullet)$ is monotonic.

Given any social-choice function $F:L(Y)^N \to Y$, let $F(L(Y)^N)$ denote the range of the function F. That is,

$$F(L(Y)^N) = \{F(u) \mid u \in L(Y)^N\}.$$

So $\#F(L(Y)^N)$ denotes the number of elements of alternatives that would be chosen by F under at least one preference profile. The Muller-Satterthwaite theorem asserts that any monotonic social choice function that has three or more outcomes in its range must be dictatorial.

Theorem 1 (Muller and Satterthwaite, 1977). If $F:L(Y)^N \to Y$ is a monotonic social choice function and $\#F(L(Y)^N) > 2$, then there must exist some dictator h in N such that

$$F(u) = \text{argmax}_{x\in F(L(Y)^N)}\, u_h(x), \;\; \forall u \in L(Y)^N.$$

Proof. Suppose that F is a monotonic social choice function. Let X denote the range of F, $X = F(L(Y)^N)$. We now state and prove three basic facts about F, as lemmas.

Lemma 1. If $F(u) = x$, $x \neq y$, and $\{i \mid u_i(x) > u_i(y)\} \subseteq \{i \mid v_i(x) > v_i(y)\}$ then $y \neq F(v)$.

Proof of Lemma 1. Suppose that F, x, y, u, and v satisfy the assumptions of the lemma but $y = F(v)$, contrary to the lemma. Let $\hat{u}$ be derived from u by moving x and y up to the top of every individual's preferences, keeping the individual's preference among x and y unchanged.

Derive $\hat{v}$ from v in the same way. By monotonicity, we have $x = F(\hat{u})$ and $y = F(\hat{v})$. But the inclusion assumed in the lemma implies that monotonicity can also be applied to $\hat{u}$ and $\hat{v}$, with the conclusion that $F(\hat{v}) = F(\hat{u}) = x$. But $x \neq y$, and this contradiction proves Lemma 1. Q.E.D.

Lemma 2. F(v) cannot be any alternative y that is Pareto-dominated, under the preference profile v, by any other alternative x that is in X.

Proof of Lemma 2. Lemma 2 follows directly from Lemma 1, when we let u be any preference profile such that $x = F(u)$. Pareto dominance gives the inclusion needed in Lemma 2, because $\{i|\ v_i(x) > v_i(y)\}$ is the set of all voters N. Q.E.D.

Following Arrow (1951), let us say that a set of voters T is decisive for an ordered pair of distinct alternatives (x,y) in $X \times X$ under the social choice function F iff there exists some preference profile u such that

$$F(u) = x \text{ and } T = \{i|\ u_i(x) > u_i(y)\}.$$

That is, T is decisive for (x,y) iff x can be chosen by F when the individuals in T all prefer x over y but everyone else prefers y over x. Lemma 1 asserts that if T is decisive for (x,y) then y is never chosen by F when everyone in T prefers x over y.

Lemma 3. Suppose that $\#X > 2$. If the set T is decisive for some pair of distinct alternatives in $X \times X$, then T is decisive for every such pair.

Proof of Lemma 3. Suppose that T is decisive for (x,y), where $x \in X$, $y \in X$, and $x \neq y$. Choose any other alternative z such that $z \in X$ and $x \neq z \neq y$.

Consider a preference profile v such that

$$v_i(z) > v_i(x) > v_i(y), \ \forall i \in T,$$

$$v_j(y) > v_j(z) > v_j(x), \ \forall j \in N \backslash T,$$

and v has everyone preferring x, y, and z over all other alternatives. By Lemma 2, F(v) must be in {y,z}, because x and all other alternatives are Pareto dominated (by z). But F(v) cannot be y, by Lemma 1 and the fact that T is decisive for (x,y). So $F(v) = z$, and $T = \{i|\ v_i(z) > v_i(y)\}$. So T is also decisive for (z,y).

Now consider instead a preference profile w such that

$$w_i(x) > w_i(y) > w_i(z), \ \forall i \in T,$$

$$w_j(y) > w_j(z) > w_j(x), \ \forall j \in N\backslash T,$$

and w has everyone preferring x, y, and z over all other alternatives. By Lemma 2, F(w) must be in {x,y}, because z and all other alternatives are Pareto dominated (by y). But F(w) cannot be y, by Lemma 1 and the fact that T is decisive for (x,y). So F(w) = x, and $T = \{i \mid w_i(x) > w_i(z)\}$. So T is also decisive for (x,z).

So decisiveness for (x,y) implies decisiveness for (x,z) and decisiveness for (z,y). The general statement of Lemma 3 can be derived from repeated applications of this fact. Q.E.D.

To complete the proof of the Muller-Satterthwaite theorem, let T be a set of minimal size among all sets that are decisive for distinct pairs of alternatives in X. Lemma 2 tells us that T cannot be the empty set, so $\#T \neq 0$.

Suppose that $\#T > 1$. Select an individual h in T, and select alternatives x, y, and z in X, and let u be a preference profile such that

$$u_h(x) > u_h(y) > u_h(z),$$

$$u_i(z) > u_i(x) > u_i(y), \ \forall i \in T\backslash\{h\},$$

$$u_j(y) > u_j(z) > u_j(x), \ \forall j \in N\backslash T,$$

and everyone prefers x, y, and z over all other alternatives. Decisiveness of T implies that $F(u) \neq y$. If F(u) were x then {h} would be decisive for (x,z), which would contradict minimality of T. If F(u) were z then T\{h} would be decisive for (z,y), which would also contradict minimality of T. But Lemma 2 implies $F(u) \in \{x,y,z\}$. This contradiction implies that #T must equal 1.

So there is some individual h such that {h} is a decisive set for all pairs of alternatives. That is, for any pair (x,y) of distinct alternatives in X, there exists a preference profile u such that F(u) = x and $\{h\} = \{i \mid u_i(x) > u_i(y)\}$. But then Lemma 1 implies that $F(v) \neq y$ whenever $v_h(x) > v_h(y)$. Thus, F(v) cannot be any alternative in X other than the one that is most preferred by individual h. This proves the Muller-Satterthwaite theorem. Q.E.D.

The Muller-Satterthwaite theorem tells us that the only way to design a social-choice procedure that always has a unique Nash equilibrium is either to give one individual all the power or to restrict the possible outcomes to two. But an essential assumption in this theorem is that F is a social choice function, not a multi-valued social choice correspondence. Dropping this assumption just means admitting that political processes might be games that sometimes have multiple equilibria. As Schelling (1960) has emphasized, when a game has multiple equilibria, the decisions made by rational players may depend on culture and history (via the focal-point effect) as much as on their individual preferences. Thus, we can use social choice procedures which consider more than two possible outcomes at a time and which are not dictatorial, but only if we allow that these procedures might sometimes have multiple equilibria that leave some decisive role for cultural traditions and other factors that might influence voters' collective expectations.

3. Anonymity and neutrality

Having a dictatorship as a social choice function is disturbing to us because it is manifestly unfair to the other individuals. But nondictatorship is only the weakest equity requirement. In the theory of democracy, we should aspire to much higher forms of equity than nondictatorship. A natural equity condition is that a social choice function or correspondence should treat all the voters in the same way. In social choice theory, symmetric treatment of voters is called anonymity.

A permutation of any set is a one-to-one function of that set onto itself. For any preference profile u in $L(Y)^N$ and any permutation $\pi:N\to N$ of the set of voters, let $u\bullet\pi$ be the preference profile derived from u by assigning to individual i the preferences of individual $\pi(i)$ under the profile u; that is

$$(u\bullet\pi)_i(x) = u_{\pi(i)}(x).$$

A social choice function (or correspondence) F is said to be anonymous iff, for every permutation $\pi:N \to N$ and for every preference profile u in $L(Y)^N$,

$$F(u\bullet\pi) = F(u).$$

That is, anonymity means that the social choice correspondence does not ask which specific individuals have each preference ordering, so that changing the names of the individuals with each preference ordering would not change the chosen outcome. Anonymity obviously implies that there cannot be any dictator if $\#N > 1$.

There is another kind of symmetry that we might ask of a social choice function or correspondence: that it should treat the various alternatives in a neutral or unbiased way. In social choice theory, symmetric treatment of the various alternatives is called neutrality. (A bias in favor of the status quo is the most common form of non-neutrality.)

Given any permutation $\rho: Y \rightarrow Y$ of the set of alternatives, for any preference profile u, let $u \circ \rho$ be the preference profile such that each individual's ranking of alternatives x and y is the same as his ranking of alternative $\rho(x)$ and $\rho(y)$ under u. That is,

$$(u \circ \rho)_i(x) = u_i(\rho(x)).$$

Then we say that a social choice function or correspondence F is neutral iff, for every preference profile u and every permutation $\rho: Y \rightarrow Y$ on the set of alternatives,

$$\rho(F(u \circ \rho)) = F(u).$$

(When F is a correspondence, $\rho(F(u \circ \rho))$ is $\{\rho(x) | \, x \in F(u \circ \rho)\}$.) Notice that neutrality of a social choice function F implies that its range must include all possible alternatives, that is,

$$F(L(Y)^N) = Y.$$

To see the general impossibility of constructing social choice functions that are both anonymous and neutral, it suffices to consider a simple example with three alternatives $Y = \{a,b,c\}$ and three voters $N = \{1,2,3\}$. Consider the preference profile u such that

$$u_1(a) > u_1(b) > u_1(c),$$

$$u_2(b) > u_2(c) > u_2(a),$$

$$u_3(c) > u_3(a) > u_3(b).$$

We may call this example the ABC paradox (where "ABC" stands for Arrow, Black, and Condorcet, who drew attention to such examples); it is also called the Condorcet cycle. An example like this appeared at the heart of the proof of the impossibility theorem in the preceding section. Any alternative in this example can be mapped to any other alternative by a permutation of Y such that an appropriate permutation of N can then return the original preference profile. Thus, an anonymous neutral social choice correspondence must choose either the empty set or the set of all three alternatives for this ABC paradox, and so an anonymous neutral social choice function cannot be defined.

This argument could be also formulated as a statement about implementation by Nash equilibria. Under any voting procedure that treats the voters anonymously and is neutral to the various alternatives, the set of equilibrium outcomes for this example must be symmetric around the three alternatives {a,b,c}. Thus, an anonymous neutral voting game cannot have a unique pure-strategy equilibrium that selects only one out of the three alternatives for the ABC paradox.

This argument does not generalize to randomized-strategy equilibria. The symmetry of this example could be satisfied by a unique equilibrium in randomized strategies such that each alternative is selected with probability 1/3. The Muller-Satterthwaite theorem does not consider randomized social choice functions, but Gibbard (1978) has obtained related results on dominant-strategy implementation with randomization. (Gibbard characterized the dominant-strategy-implementable randomized social choice functions as probabilistic mixtures of unilateral and duple functions, which are generalizations of dictatorship and binary voting.)

Randomization confronts democratic theory with the same difficulty as multiple equilibria, however. In both cases, the social choice ultimately depends on factors that are unrelated to the individual voters' preferences: private randomizing factors in one case, public focal factors in the other. As Riker (1982) has emphasized, such dependence on extraneous factors implies that the outcome chosen by a democratic process cannot be characterized as a pure expression of the voters' will.

These impossibility results tell us that there are only two general ways to make social choices that do not depend on such extraneous factors of randomization or focal-equilibrium selection: either one individual must have all the power, or the possible outcomes to must somehow be restricted to just two alternatives. In fact, many institutions of government actually fit one of these two categories. Decision-making in the executive branch is often made by a single decision-maker, who may be the president or the minister with responsibility for a given domain of social alternatives. Many countries elect their leaders by democratic systems which are dominated by two major parties, in which case the voters may be essentially constrained to choose among the alternatives selected by these two parties. Furthermore, even in multi-party democratic systems, when a vote is called in a legislative assembly, there are usually only two possible outcomes: to approve or to reject some specific proposal that is on the floor. Of course, the current vote may be just one stage in a longer agenda, as when the assembly considers a proposal to amend another proposal that is to scheduled be considered later, and many alternatives may be ultimately considered by a sequence of votes in such a binary agenda.

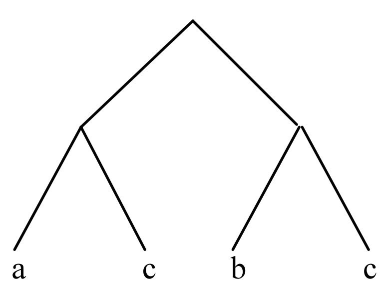
The rest of this paper develops some basic models for general analysis of sequential voting in assemblies and voting in two-party elections. Sections 4, 5, and 6 consider binary agendas for sequentially considering many alternatives. Sections 7 and 8 consider voting in elections where the alternatives are defined by two competitive parties.

<u>4. Tournaments and binary agendas</u>

When there are only two alternatives, majority rule is a simple and compelling social choice procedure. K. May (1952) showed that, when #Y = 2 and #N is odd, choosing the alternative that is preferred by a majority of the voters is the unique social choice function that satisfies anonymity, neutrality, and monotonicity.

When there are more than two alternatives, we might still try to apply the principle of majority voting by dividing the decision problem into a sequence of binary questions. For example, one simple binary agenda for choosing among three alternatives {a,b,c} is as follows. At the first stage, there is a vote on the question of whether to eliminate alternative a or alternative b from further consideration. Then, at the second stage, there is a vote between alternative c and the alternative among {a,b} that survived the first vote. The winner of this second vote is the implemented social choice.

This binary agenda is represented graphically in Figure 1. The agenda begins at the top, and at each stage the voters must choose to move down the agenda tree along the branch to the left or to the right. The labels at the bottom of the agenda tree indicate the social choice for each possible outcome at the end of the agenda. At the top of Figure 1, the left branch represents eliminating alternative b at the first vote, and the right branch represents eliminating alternative a. Then at each of the lower nodes, the right branch represents choosing c and the left branch represents choosing the other alternative that was not eliminated at the first stage.



<u>Figure 1</u>. A binary agenda for choosing among alternatives a, b, and c.

Now suppose that the voters have preferences as in the ABC paradox example (described in the preceding section). Then there is a majority (voters 1 and 3) who prefer alternative a over b, there is a majority (voters 1 and 2) who prefer alternative b over c, and there is a majority (voters 2 and 3) who prefer alternative c over a. Let us use the notation x >> y (or equivalently y << x) to denote the statement that a majority of the voters prefer x over y (so that x beats y). Then we may summarize the majority preference for this example as follows:

a >> b, b >> c, c >> a.

This cycle is what makes this example paradoxical.

Given these voters' preferences, what will be the outcome of the binary agenda in Figure 1? At the second stage, a majority would choose alternative b against c if alternative a were eliminated at the first stage, but a majority would choose alternative c against a if alternative b were eliminated at the first stage. So a majority of voters should vote to eliminate alternative a at the first stage (even though a majority prefers a over b), because they should anticipate that the ultimate result will be to implement b rather than c, and a majority prefers alternative b over c. This backwards analysis is shown in Figure 2 which displays, in parentheses above each decision node, the ultimate outcome that would be chosen by sophisticated majority voting if the process reached this node.

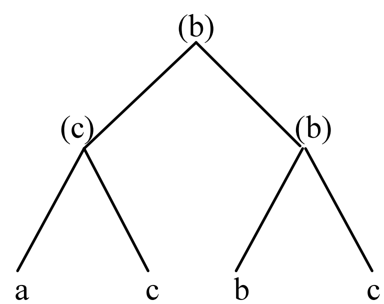


Figure 2. Sophisticated solution with a>>b, b>>c, and c>>a.

In general, given any finite set of alternatives Y, a binary agenda on Y is a rooted tree that has two branches coming out of each non-terminal node, together with a labeling that assigns an outcome in Y to every terminal node, such that each alternative in Y appears as the outcome for at least one terminal node. Given a binary agenda, a sophisticated solution extends the labeling to all nodes so that, for every non-terminal node θ, the label at θ is the alternative in Y that would be chosen by a majority vote among the alternatives listed at the two nodes that directly follow θ. A sophisticated outcome of a binary agenda is the outcome assigned to the initial node (or root) of the agenda tree in a sophisticated solution.

Given any preference profile for an odd number of voters, each binary agenda on Y has a unique sophisticated solution, which can be easily calculated by backward induction. (See Farquharson, 1969.) Sloth (1993) has shown that such sophisticated solutions correspond to subgame-perfect equilibria of voting games when individuals publicly cast their votes one at a time (as in roll-call voting) at each stage. Thus, once a binary agenda has been specified, it is straightforward to predict the outcome that will be chosen under majority rule, assuming that the voters have a sophisticated understanding of each others' preferences and of the agenda.

But different agendas may lead to different majority-rule outcomes for the same voters' preferences. Thus, the chairman who sets the agenda may have substantial power to influence the sophisticated majority-rule outcome. To quantify the extent of such agenda-setting power, we want to characterize the set of alternatives that can be achieved under binary agendas, for any given preference profile.

To compute sophisticated solutions, it is only necessary to know, in each pair of alternatives, which one would be preferred by a majority of the voters. That is, we only need to know, for each pair of distinct alternatives x and y in Y, whether x >> y or y >> x. (Read ">>" here as "would be preferred by a majority over" or simply "beats".)

The ABC paradox shows that the majority-preference relation >> is not necessarily transitive, even though each individual voter's preferences are assumed to be transitive. In fact, McGarvey (1953) showed that a relation >> can be generated as the majority-preference relation for an odd number of voters whose individual preferences are transitive if and only if it satisfies the following completeness and antisymmetry condition:

$$x \gg y \text{ or } y \gg x, \text{ but not both, } \forall x \in Y, \forall y \in Y \backslash \{x\}.$$

Any such relation >> on Y may be called a <u>tournament</u>.

<u>5. The top cycle</u>

Let >> be any fixed tournament (complete and antisymmetric) on the given set of alternatives Y. We now consider three definitions that each characterize a subset of Y.

Let $Y^*(1)$ denote the set of all alternatives x such that there exists a binary agenda on Y for which x is the sophisticated outcome. That is, $Y^*(1)$ is the set of outcomes that could be achieved by agenda-manipulation, when the agenda-setter can plan any series of binary questions, subject only to the constraint that every alternative in Y must be admitted as a possibility under the agenda, and all questions will be resolved by sophisticated (forward-looking) majority votes.

Let Y*(2) denote the set of alternatives y such that, for every alternative x in Y\{y}, there must exist some chain $(z_0,z_1,...,z_m)$ such that $x = z_0$, $z_m = y$, and $z_{k-1} << z_k$ for every k = 1,...,m. That is, an alternative y is in Y*(2) iff, starting from any given status quo x, the voters could be manipulated to give up x for y through a sequence of replacements such that, at each stage, the majority would always prefer to give up the previously chosen alternative for the manipulator's proposed replacement if they believed that this replacement would be the last. In contrast to Y*(1) which assumed sophisticated forward-looking voters, Y*(2) is based on an assumption that voters are very naive or myopic.

Let Y*(3) be defined as the smallest (in the set-inclusion sense) nonempty subset of Y that has the following property:

> for any pair of alternatives x and y, if y is in the subset and x is not in the subset then y >> x.

An argument is needed to verify that this set Y*(3) is well defined. Notice first that Y is itself a "subset" that has this property (because the property is trivially satisfied if no x outside of the subset can be found). Notice next that, if W and Z are any two subsets that have this property then either $W \subseteq Z$ or $Z \subseteq W$. (Otherwise we could find w and z such that $w \in W$, $w \notin Z$, $z \in Z$, and $z \notin W$; but then we would get w >> z and z >> w, which is impossible in a tournament.) Thus, assuming that Y is finite, there exists a smallest nonempty set Y*(3) that has this property, and which is a subset of all other sets that have this property.

A fundamental result in tournament theory, due to Miller (1977), is that the above three definitions all characterize the same set Y*. This set Y* is called the <u>top cycle</u>.

<u>Theorem 2.</u> Y*(1) = Y*(2) = Y*(3).

<u>Proof</u>. We first show that $Y^*(1) \subseteq Y^*(2)$. Suppose that y is in Y*(1). Then there is some binary agenda tree such that the sophisticated solution is y. Given any x, we can find a terminal node in the tree where the outcome is x. Now trace the path from this terminal node back up through the tree to the initial node. A chain satisfying the definition of Y*(2) for x and y

can be constructed simply by taking the sophisticated solution at each node on this path, ignoring repetitions. (This chain begins at x, ends at y, and only changes from one alternative to another that beats it in the tournament. For example, Figure 2 shows that alternative b is in Y*(2), with chains a << c << b and c << b.)

We next show that $Y^*(2) \subseteq Y^*(3)$. If not, then there would be some y such that $y \in Y^*(2)$ but $y \notin Y^*(3)$. Let x be in Y*(3). To satisfy the definition of Y*(2), there must be some chain such that $x = z_0 << z_1 << ... << z_m = y$. This chain begins in Y*(3) and ends outside of Y*(3), and so there must exist some k such that $z_{k-1} \in Y^*(3)$ and $z_k \notin Y^*(3)$, but then $z_{k-1} << z_k$ contradicts the definition of Y*(3).

Finally we show that $Y^*(3) \subseteq Y^*(1)$. Notice first that Y*(1) is nonempty (because binary agendas and their sophisticated outcome always exist). Now let y be any alternative in Y*(1) and let x be any alternative not in Y*(1). We claim that y >> x. If not then we would have x >> y; but then x would be the sophisticated outcome for a binary agenda, in which the first choice is between x and a sub-tree that is itself a binary agenda for which the sophisticated outcome would be y, and this conclusion would contradict the assumption $x \notin Y^*(1)$. So every y in Y*(1) beats every x outside of Y*(1); and so Y*(1) includes Y*(3), which is the smallest nonempty set that has this property.

Thus we have $Y^*(1) \subseteq Y^*(2) \subseteq Y^*(3) \subseteq Y^*(1)$, and so all three are the same. Q.E.D.

In the ABC paradox from the previous section, the top cycle includes all three alternatives {a,b,c}. If we add a fourth alternative d that appears immediately below c in each individual's preference ranking (so that b >> d and c >> d but d >> a), then d is also included in the top cycle for this example, even though d is Pareto-dominated by c.

When the top cycle consists of a single alternative, this unique alternative is called a Condorcet winner. That is, a Condorcet winner is an alternative y such that y >> x for every other alternative x in Y\{y}. The existence of a Condorcet winner requires very special configurations of individual preferences. For example, suppose that each voter's preference ordering is selected at random from the (#Y)! possible rank orderings in L(Y), independently of all other voters' preferences. R. May (1971) has proven that, if the number of voters is odd and more than 2, then the probability of a Condorcet winner existing among the alternatives in Y goes to zero as #Y goes to infinity. (See also Fishburn, 1973.)

McKelvey (1976, 1979) has shown that, under some common assumptions about voters' preferences, if a Condorcet winner does not exist then the top cycle is generally very large. Let us now state and prove, as Theorem 3 below, a simple result that is similar to McKelvey's.

We assume a given finite set of alternatives Y, and a given odd finite set of voters N, each of whom has strict preferences over Y. Let $\Delta(Y)$ denote the set of probability distributions over the set Y. We may identify $\Delta(Y)$ with the set of lotteries or randomized procedures for choosing among the pure alternatives in Y. Suppose that each individual i has a von Neumann Morgenstern utility function $U_i{:}Y \rightarrow \mathbb{R}$ such that, for any pair of lotteries, individual i always prefers the lottery that gives him higher expected utility. So if we extend the set of alternatives by adding some lotteries from $\Delta(Y)$, then U_i defines individual i's preferences on this extended alternative set. With this framework, we can prove the following theorem.

Theorem 3. If the top cycle contains more than one alternative then, for any alternative z and any positive number ε, then we can construct an extended alternative set, composed of Y and a finite subset of $\Delta(Y)$, such that the extended top cycle includes a lottery in which the probability of z is at least $1\text{-}\varepsilon$.

Proof. If the top cycle is not a single alternative, then the top cycle must include a set of three or more alternatives $\{w_1, w_2, ..., w_K\}$ such that $w_1 << w_2 << .. << w_K << w_{K+1} = w_1$.

VonNeumann Morgenstern utility theory guarantees that every individual who strictly prefers w_{j+1} over w_j, will also strictly prefer $(1\text{-}\varepsilon)q + \varepsilon[w_{j+1}]$ over $(1\text{-}\varepsilon)q + \varepsilon[w_j]$ for any lottery q. (Here $(1\text{-}\varepsilon)q + \varepsilon[w_j]$ denotes the lottery that gives outcome w_j with probability $\varepsilon>0$, and otherwise implements the outcome randomly selected by lottery q.) Thus, by continuity, there must exist some large integer M such that every individual who strictly prefers w_{j+1} over w_j, will also strictly prefer

$$(1\text{-}\varepsilon)\big((1 - (m+1)/M)[w_1] + ((m+1)/M)[z]\big) + \varepsilon[w_{j+1}]$$

over

$$(1\text{-}\varepsilon)\big((1 - m/M)[w_1] + (m/M)[z]\big) + \varepsilon[w_j]$$

for any m between 0 and M-1. That is, for some sufficiently large M, the same majority that would vote to change from w_j to w_{j+1} would also vote to change from w_j to w_{j+1} with probability

ε even when this decision also entails a probability $(1-\varepsilon)/M$ of changing from w_1 to z.

We now prove the theorem using the Y*(2) characterization of the top cycle. Because w_1 is in the top cycle, we can construct a naive chain from any alternative x to w_1 ($x \ll \ldots \ll w_1$). This naive chain can be continued from w_1 to z through M lotteries as follows:

$$w_1 = (1-\varepsilon)[w_1] + \varepsilon[w_1]$$

$$\ll (1-\varepsilon)\big((1-1/M)[w_1] + (1/M)[z]\big) + \varepsilon[w_2]$$

$$\ll (1-\varepsilon)\big((1-2/M)[w_1] + (2/M)[z]\big) + \varepsilon[w_3]$$

$$\ldots \ll (1-\varepsilon)\big((1-M/M)[w_1] + (M/M)[z]\big) + \varepsilon[w_j] = (1-\varepsilon)[z] + \varepsilon[w_j] \quad \text{(for some j).}$$

So including all the lotteries of this chain as alternatives gives us an extension of Y in which a lottery $(1-\varepsilon)[z] + \varepsilon[w_j]$ can be reached by a naive chain from any alternative x. Q.E.D.

The proof of Theorem 3 uses the naive-chain characterization of the top cycle (Y*(2)), but the equivalence theorem tells us that this result also applies to agenda manipulation with sophisticated voters. That is, if the chairman can include randomized social-choice plans among the possible outcomes of an agenda then, either a Condorcet winner exists, or else the chairman can design a binary agenda that selects any arbitrary alternative (even one that may be worst for all voters) with arbitrarily high probability in the majority-rule sophisticated outcome. (See also Abreu and Matsushima, 1992, for other results about general implementability of social choice functions with randomization when voters have von Neumann Morgenstern utility functions.)

If more restrictions are imposed on the form of the agenda that can be used, then the set of alternatives that can be achieved by agenda manipulation may be substantially smaller. For example, Banks (1985) has characterized the set of alternatives that can be achieved as sophisticated outcomes of successive-elimination agendas of the following form: The alternatives must be put into an ordered list; the first question must be whether to eliminate the first or second alternative in this list; and thereafter the next question is always whether to eliminate the previous winner or the next alternative on the list, until all but one of the alternatives have been eliminated. For any given tournament (Y,>>), the set of alternatives that can be sophisticated outcomes of such successive-elimination agendas is called the Banks set.

Given a tournament (Y,>>), an alternative x is covered iff there exists some other alternative y such that y >> x and

$\{z|\ x >> z\} \subseteq \{z|\ y >> z\}$.

If x is Pareto-dominated by some other alternative then x must be covered. The uncovered set is the set of alternatives that are not covered. (See Miller, 1980.) The uncovered set is always a subset of the top cycle, because any alternative not in Y*(3) is covered by any alternative in Y*(3). On the other hand, the Banks set is always a subset of the uncovered set. (See Shepsle and Weingast, 1984; Banks, 1985; McKelvey, 1986; or Moulin, 1986.) Thus, Pareto-dominated alternatives cannot be sophisticated outcomes of successive-elimination agendas.

The power of the chairman who sets the agenda would be increased if we dropped the assumption that all alternatives in the given set Y must be included as possible outcomes of the agenda. Romer and Rosenthal (1978) have considered the assumption that the agenda-setter only needs to include one given alternative that represents the status quo. In the analysis of their setter model, the agenda-setter could generally get a better outcome if the status quo became worse for all individuals (even including the agenda-setter), because such a change would enlarge the set of feasible alternatives that a majority would prefer to the status quo.

There are other alternatives to assuming agenda control by a powerful chairman. Baron and Ferejohn (1989) assume that randomly selected voters will be given sequential opportunities to propose a social-choice alternative for an up-or-down vote until one such proposal is approved by a majority. Their basic model has a wide multiplicity of subgame-perfect equilibria, but they find unique expected values with stationary equilibria. (See also Banks and Duggan (2000).)

6. Monotonicity and learning from sequential votes

With an odd number of voters in N, any binary agenda with outcomes in Y defines a social choice function $F:L(Y)^N \rightarrow Y$ such that F(u) is the unique sophisticated outcome of the binary agenda when u is the profile of individuals' preferences over Y. For the binary agenda shown in Figure 1, the social choice function defined by its sophisticated outcomes is non-dictatorial and has a range that includes all of Y={a,b,c}. This result does not contradict the Muller-Satterthwaite theorem, because this social choice function is not monotonic.

For example, let u and v be preference profiles for N={1,2,3} such that

$u_1(a)>u_1(b)>u_1(c),\ \ u_2(b)>u_2(c)>u_2(a),\ \ u_3(c)>u_3(a)>u_3(b),$

$v_1(a)>v_1(b)>v_1(c),\ \ v_2(b)>v_2(a)>v_2(c),\ \ v_3(c)>v_3(a)>v_3(b).$

Here u is taken from the basic ABC paradox, and v differs from u only in that individual 2 prefers a over c in v. The sophisticated outcomes of this binary agenda for these preference profiles are F(u) = b and F(v) = a, but these social choices violate monotonicity, as

$$\{y|\ v_i(y) > v_i(b)\} = \{y|\ u_i(y) > u_i(b)\},\ \forall i \in N.$$

With b chosen for u, monotonicity would require us to include b as a social choice with preferences v, but alternative a is the Condorcet winner under majority rule with preferences v.

So in our consideration of sequential voting in binary agendas, we have been using a concept of rational behavior that does not entail monotonicity, even though we have seen that the set of Nash equilibrium outcomes of any game form must be a monotonic correspondence. There is no contradiction here because the sophisticated solutions of binary agendas correspond to subgame-perfect equilibria of sequential voting games (by Sloth, 1993), and subgame perfection is an equilibrium refinement that admits only a subset of all Nash equilibria.

In our example, the difference between the sophisticated outcomes of u and v under the given binary agenda depends on voter 1's beliefs about how voter 2 would behave if they chose to eliminate alternative b at the first stage. In the sophisticated solution with preferences v, 1 believes that eliminating b at the first stage would lead to a, the best possible outcome for 1, as 2 would vote with 1 at the second stage to choose a against c. But in the sophisticated solution with preferences u, 1 believes that eliminating b at the first stage would lead to c, the worst possible outcome for 1, as 2 would then vote with 3 to choose c against a. This analysis uses an assumption that that voters' behavior should be sequentially rational in all events. Nash equilibrium does not require sequentially rational behavior after events that have that have probability 0, however, and so the sophisticated solution with preferences u would also be a Nash equilibrium of this binary agenda with the preferences v. Individual 2's voting for c against a would not be sequentially rational with preferences v, but the second-stage vote between c and a has probability 0 in this scenario.

Thus, in using sequential rationality to refine the rational equilibrium analysis of sequential voting games, we have found that voters could rationally implement non-monotonic social choice functions. (See Moore and Repullo, 1988, for more on implementation of general non-monotonic social choice functions by subgame-perfect equilibria of multistage game forms.) But this analysis is based on an assumption that the voters' preferences in $L(Y)^N$ are common knowledge, so that nobody could actually communicate anything or learn anything about their preferences in the voting process. Aghion, Fudenberg, Holden, Kunimoto, and Tercieux (2012) have shown that, when we admit that voters might have different information about unobserved factors that affect their preferences, such implementation of non-monotonic social choice functions may not be robust to small changes in our assumptions about the voters' information.

In our example, even if 2's preference were actually v_2, individual 2 would be better off having 1 believe that 2 would behave according to the preference u_2, even if this required 2 to make a binding commitment to do so, as $v_2(F(v)) = v_2(a) < v_2(b) = v_2(F(u))$. (Recall, the only difference between u and v is that $u_2 \neq v_2$.) Thus, if 1 is uncertain about 2's preferences, 2 would want to manipulate 1's beliefs.

Remarkably, if 2's preferences depend on uncertain factors about which other individuals might have better information, then even a small degree of uncertainty can admit alternative b as a sequentially rational outcome for the given binary agenda with voters' preferences v. Suppose, for example, that the difference between the preference profiles u and v depends on an unobserved state that affects the quality of alternative c and thus determines whether c would be better or worse than alternative a for individual 2. For simplicity, we may assume that 2's utility differences between alternatives a and c have the same magnitude in each state ($u_2(c)-u_2(a) = v_2(a)-v_2(c)$) so that, in choosing between a and c, 2 would prefer the alternative that she believes is more likely to be better for her. We may also assume that, before getting any information, everyone would believe that the true state was equally likely to correspond to the preferences u or v. Then suppose that, given either state u or v, each individual would get a conditionally independent signal "u" or "v" that has a high probability of being equal to the true state, but 1's signal is the most reliable. To be specific, for some small number ε between 0 and 1/2, suppose that, given either state, individuals 2 and 3 get a signal that has probability ε of being wrong, but individual 1 gets a signal that has the smaller probability ε^2 of being wrong. In the voting process, each individual's vote can depend only on the signal that the individual has received and, at the second stage, on how the others voted at the first stage.

With this information structure, the binary agenda from Figure 1 has a sequential equilibrium in which, regardless of whether the true state is u or v, the voters are always expected to behave as in the sophisticated solution for preferences u, and so alternative b is the outcome with probability one in both states. In this equilibrium, the event of individual 1 voting (with 3) to eliminate b at the first stage has probability 0, but Bayes-consistent inference allows that, in this zero-probability event, individual 2 could take 1's surprising behavior as an indication that 1 got the signal "u". Then even if 2 got the signal "v", 2 would rationally vote (with 3) for c against a at the second stage, because 2 knows that 1's signal is a more reliable indicator of the true state. As 1 prefers b over c in both states, it would then always be a mistake for 1 to vote for eliminating b at the first stage, but 2 could believe that 1's voting hand would be much less likely to tremble into this mistake after observing "v" than "u". So the threat of 2 voting for c at the second stage if 1 votes to eliminate b at the first stage can be made credible in both states by the possibility of 2 deriving information about the state from 1's previous vote.

Thus, we can find a sequential equilibrium in which alternative b is always chosen in both states u and v, even though alternative a would be the Condorcet winner if the state v were common knowledge, and even though (with ε small) all voters are very likely to have accurate information about the state. Aghion, Fudenberg, Holden, Kunimoto, and Tercieux (2012) have shown that such deviations from full-information sequential-equilibrium analysis can always be constructed whenever full-information sequential-equilibrium analysis predicts a social choice function that is not monotonic. (See also related results of Chung and Ely, 2003.) In this sense, monotonicity may be seen as a necessary condition for taking account of the possibility that voters might learn preference-relevant information from each others' previous votes in a sequential voting process.

7. Two-party competition

We have studied binary agendas because they allow us to reduce the problem of choosing among many alternatives to a sequence of votes each of which is binary, in the sense that it has only two possible outcomes. However, the number of binary votes that are needed to work through a large number of alternatives goes up at least as the log (base 2) of the number of alternatives. When we move from voting in small committees to voting in large democratic nations, the increased cost of each round of voting makes it impractical to work through a long sequence of votes. So democracies generally rely on political leaders to select a small subset of the potential social alternatives, and then only this small selected set of social alternatives will be considered by the voters in the general election. The hope for a successful democracy is that competition among political leaders should ensure that they will try to select alternatives that are highly preferred by a large fraction of the voting population.

So let us consider a simple model of how political leaders might select alternatives to put before the voters in a general election. We assume that the set of all possible social alternatives Y is a nonempty finite set. To keep voting binary, we assume here that there are only two political leaders, each of whom must select an alternative in Y, which we may call the leader's policy position. Making the simplest assumption about timing, let us suppose that the two political leaders must choose their policy positions simultaneously and independently.

Let >> denote the majority preference relation, satisfying the completeness and antisymmetry properties of a tournament. We assume that the leader whose policy position is preferred by a majority of the voters will win the election if they choose different positions, but each leader has a probability 1/2 of winning if both leaders choose the same policy position. Assuming that each leader is motivated only by the desire to win, we get a simple two-person zero-sum game. In this game, when leader 1 chooses position x_1 and leader 2 chooses position x_2, the payoffs are

+1 for leader 1 and -1 for leader 2 if $x_1 >> x_2$,

-1 for leader 1 and +1 for leader 2 if $x_2 >> x_1$,

and 0 for both leaders if $x_1 = x_2$.

If Y contains a Condorcet winner that beats every other alternative in Y, then the unique equilibrium of this game is for both political leaders to choose this Condorcet winner as their policy position. But if no alternative is a Condorcet-winner then this game cannot have any equilibria in pure strategies, because any position could be beaten by at least one other alternative, and so each leader could make himself the sure winner if he knew which position would be chosen by his opponent.

The general existence theorems of von Neumann (1928) and Nash (1951) assure us that this game must have at least one equilibrium in mixed strategies, even if a Condorcet winner does not exist. We know that all equilibria must give the same expected payoff allocation, because this game is two-person zero-sum. The ex-ante symmetry of the leaders who are playing this game makes it obvious that each player must have the same set of equilibrium strategies, and the expected payoff allocation in equilibrium must be (0,0). That is, each leader's probability of winning the election must be 1/2 at the beginning of the game (before the randomized strategies are implemented).

For the example of the ABC paradox from Section 3, the unique equilibrium strategy for each leader is to randomize uniformly over the three alternatives, choosing each with probability 1/3. Then there is a 1/3 probability of leader 1 choosing a position that beats leader 2's position (in the sense that $x_1 >> x_2$); there is a 1/3 probability of leader 1 choosing a position that is beaten by leader 2's position; and there is a 1/3 probability of leader 1 choosing the same position as leader 2, in which case each has an equal probability of winning the election.

If we add a fourth alternative d such that every voter ranks d immediately below c, then the unique equilibrium strategy remains the same. That is, the alternative d would not chosen by either leader, even though d is in the top cycle. Notice that d is covered by c in this example. In fact, the covered alternatives in any tournament are precisely the dominated pure strategies for the leaders in this policy-positioning game.

Remarkably, Fisher and Ryan (1992) showed that there is always a unique equilibrium strategy in this game. Our formulation and proof of this uniqueness theorem here is based on Laffond, Laslier, and Le Breton (1993).

Theorem 4. The two-person game of choosing positions in a finite tournament (Y,>>) has a unique Nash equilibrium. In this equilibrium, every alternative that is a best response has positive probability.

Proof. Let p and q be randomized strategies in $\Delta(Y)$ in two Nash equilibria of this game. The two players in this game are symmetric, and Nash equilibria of two-person zero-sum games are always interchangeable, and so (p,p), (q,q), and (p,q) must all be Nash equilibria of this game. Let

$$B = \{y \in Y \mid p(y) > 0 \text{ or } q(y) > 0\}.$$

Because randomizing according to p or q is optimal for a player against p or q, all of the alternatives in B must offer the equilibrium expected payoff against both p and q. But we know that the equilibrium expected payoff is 0 in this game. So choosing any alternative in B must give a player a probability of winning that equals his probability of losing, when the other player randomizes according to p or q. That is,

$\sum_{x>>y} p(x) = \sum_{x<<y} p(x), \ \forall y \in B$

$\sum_{x>>y} q(x) = \sum_{x<<y} q(x), \ \forall y \in B.$

Now let $d(x) = p(x) - q(x)$ for every x in B. So we have

$\sum_{x>>y} d(x) = \sum_{x<<y} d(x), \ \forall y \in B,$

$\sum_{x\in B} d(x) = 0.$

(The latter equation holds because the p(x) and q(x) both sum to 1.) To show uniqueness, we need only to show that this system of equations for d has no nonzero solutions.

So suppose to the contrary that this system of equations has some nonzero solution for d. Then it must have at least one nonzero solution such that all d(x) numbers are rational, because all coefficients are rational in these linear equations. Furthermore, multiplying through by the lowest common denominator, this system of equations must have at least one solution such that all d(x) are integers, and (dividing by 2 as necessary) we can guarantee that at least one integer d(y) must be odd. Then for this alternative y, we get

$0 = d(y) + \sum_{x<<y} d(x) + \sum_{z>>y} d(z) = d(y) + 2(\sum_{x<<y} d(x)).$

But $d(y) + 2(\sum_{x<<y} d(x))$ is an odd integer, and 0 is even. This contradiction proves that there cannot be any nonzero solutions for d. Thus p = q, and so the equilibrium is unique.

The components of p must be rational numbers, because p is the unique equilibrium strategy for a two-person zero-sum game that has payoffs in the rational numbers. Now let p* denote the smallest positive multiple of p that has all integer components. This vector p* satisfies

$\sum_{x<<y} p^*(x) = \sum_{x>>y} p^*(x),$

for every pure strategy y that is a best response to the equilibrium strategy p (which includes all y in its support B), because all best responses give expected payoff zero. At least one p*(z) must be odd (or else we could divide them all by 2). So the sum of the components

$\sum_{x\in B} p^*(x) = p^*(z)+ \sum_{x<<z} p^*(x) + \sum_{x>>z} p^*(x)$

$= p^*(z) + 2(\sum_{x<<z} p^*(x))$

is an odd integer. For every y that is a best response to the equilibrium strategy p, we have

$p^*(y) + 2(\sum_{x<<y} p^*(x)) = \sum_{x\in B} p^*(x),$

and so p*(y) is also an odd integer. But 0 is not odd. Thus, if y is any best response to the equilibrium strategy p then

$p(y) = p^*(y)/\left((\sum_{x \in B} p^*(x)\right) \neq 0.$ Q.E.D.

The set of alternatives that may be chosen with positive probability by the party leaders in the equilibrium of this policy-positioning game is called the bipartisan set of the tournament (Y,>>). The bipartisan set is a subset of the uncovered set, because the covered alternatives are the dominated strategies in this policy-positioning game, and so the bipartisan set is always contained in the top cycle. Laffond, Laslier, and Le Breton (1993) have shown, however, that the bipartisan set may contain alternatives that are not in the Banks set, and the Banks set may contain alternatives that are not in the bipartisan set.

8. Median voter theorems

We have seen that, if a Condorcet winner exists, then we can expect it as the outcome of rational voting in any binary agenda, or as the unique policy position that would be chosen by party leaders in two-party competition. But when a Condorcet winner does not exist, then agenda manipulation with randomized alternatives can achieve virtually any outcome, and the outcome of two-party policy positioning must have some unpredictability. So we should be interested in economic conditions that imply the existence of a Condorcet winner in Y. The most natural such condition is expressed by the median voter theorems.

There are two basic versions of the median voter theorem. One version (from Black, 1958) assumes single-peaked preferences, and another version (from Gans and Smart, 1994, Rothstein, 1990, 1991, and Roberts, 1977) assumes a single-crossing property.

To develop the single-crossing property, we begin by assuming that the policy alternatives in Y are ordered completely and transitively, say from "left" to "right" in some sense. We may write "$x < y$" to mean that the alternative x is to the left of alternative y in the space of policy alternatives. We also assume that the voters (or their political preferences) are transitively ordered in some political spectrum, say from "leftist" to "rightist," and we may write "$i < j$" to mean that voter i is to the left of voter j in this political spectrum.

The meaning of this ordering of voters is only that leftist voters tend to favor left policies more than voters who are rightist in political preference. Formally, we assume that, for any two voters i and j such that $i < j$, and for any two policy alternatives x and y such that $x < y$,

$$\text{if } u_i(x) < u_i(y) \text{ then } u_j(x) < u_j(y),$$

but if $u_j(x) > u_j(y)$ then $u_i(x) > u_i(y)$.

This assumption is called the single-crossing property.

Let us assume that the number of voters is odd and their ordering is complete and transitive. Then there is some median voter h, such that $\#\{i \in N \mid i < h\} = \#\{j \in N \mid h < j\}$. For any pair of alternatives x and y such that $x < y$, if the median voter prefers x then all voters to the left of the median agree with him, but if the median voter prefers y then all voters to the right of the median agree with him. Either way, there is a majority of voters who agree with the median voter. So the majority preference relation (>>) is the same as the preference of the median voter. Thus, the alternative that is most preferred by the median voter must be a Condorcet winner. That is, we have proven the following theorem.

Theorem 5. Suppose that there is an odd number of voters. If the alternatives in Y have a complete transitive ordering and the voters in N have a complete transitive ordering which together satisfy the single-crossing property, then the ideal point of the median voter is a Condorcet winner in Y.

In the single-peakedness version of the median-voter theorem, a complete transitive ordering (<) is assumed on the set of alternatives Y only. For each voter i, it is assumed that there is some ideal point θ_i in Y such that, for every x and y in Y,

if $\theta_i \leq x < y$ or $y < x \leq \theta_i$ then $u_i(x) > u_i(y)$.

That is, on either side of θ_i, voter i always prefers alternatives that are closer to θ_i. This property is called the single-peakedness assumption. Assuming that the number of voters is odd, the median voter's ideal point is the alternative θ^* such that

$\#N/2 \geq \#\{i \mid \theta_i < \theta^*\}$ and $\#N/2 \geq \{i \mid \theta^* < \theta_i\}$.

The voters who have ideal points at θ^* and to its left form a majority that prefers θ^* over any alternative to the right of θ^*, while the voters who have ideal points at θ^* and to its right form a majority that prefers θ^* over any alternative to the left of θ^*. Thus, this median voter's ideal point θ^* is a Condorcet winner in Y.

Single-crossing and single-peakedness are different assumptions, and neither is logically implied by the other. Both assumptions give us a result that says "the median voter's ideal point is a Condorcet winner," but there is a subtle difference in the meaning of these results. With the single-crossing property we are speaking about the ideal point of the median voter, but with the single-peakedness property we are speaking about the median of the voters' ideal points. Notice also that the majority preference relation can be guaranteed to be a full transitive ordering under the single-crossing assumption, but not under the single-peakedness assumption.

In both versions of the median voter theorem, the set of policy alternatives must be essentially one-dimensional, because otherwise we cannot put the alternatives in a transitive order. In general applications that do not have this simple one-dimensional structure, we do not generally expect to find a Condorcet winner. (For more on generalizations of the median-voter rule, see also Moulin, 1980, Saporiti, 2009, and Penn, Patty, and Gailmard, 2011.)

9. Conclusions

We have considered binary agendas and two-party competition, because they are procedures for reducing general social choice problems with many alternatives into a simple framework of majority voting on pairs of alternatives. This reduction requires some decision-making by political leaders: the chairman who sets the agenda, or the leaders who formulate policy for the two major parties. So it is natural ask, to what extent do the outcomes of binary agendas or two-party competition depend on the decision-making by such political leaders, rather than on the preferences of the voters. The answer, we have seen, is that manipulations of an agenda-setter or arbitrary and unpredictable positioning decisions of political leaders can substantially affect the outcome of majority voting, except in the special case where a Condorcet winner happens to exist.

To find ways of avoiding such dependence on an agenda setter or a couple of party leaders, we must go on to study more general voting systems that allow voters to consider more than two alternatives at once. K. May's theorem (1952) assured us that majority rule is the unique obvious way to implement the principles of democracy (anonymity, neutrality) in social decision-making when only two alternatives are considered at a time. In contrast, there is a wide variety of anonymous neutral voting systems that have been proposed for choosing among more than two alternatives (plurality voting, Borda voting, approval voting, single transferable vote, etc.), and all of these deserve to be called democratic. Furthermore, the impossibility theorems of social choice theory tell us that no such voting system can guarantee a unique pure-strategy equilibrum for all profiles of voters' preferences. Multiplicity of equilibria means that the social outcome can depend on any factor that focuses public attention on one equilibrium. As discussed by Schelling (1960), these focal factors may include history, cultural tradition, and public speeches of political leaders. (Comparative analysis of focal-factor dependence in different electoral systems has been considered at length by Myerson and Weber, 1993, although the formal solution concept in that paper has been found to have some problems of consistency. See also Myerson, 1999.)

Our initial fable suggested that political institutions may arise out of a need to coordinate on better equilibria in social and economic arenas, and we have found that some of this multiplicity of equilibria may inevitably remain in any democratic political system. But having multiple equilibrium outcomes for some preference profiles does not imply that everything must be an equilibrium outcome for all preference profiles. Game-theoretic analysis of political institutions can show substantial differences in the equilibrium outcomes under different political institutions. If social choice theory has not given us one perfect voting system, then it has left us the important task of characterizing the properties and performance of the many voting systems that we do have.

REFERENCES

Dilip Abreu and Hitoshi Matsushima, "Virtual implementation in iteratively undominated strategies," Econometrica 60(5):993-1008 (1992).

Philippe Aghion, Drew Fudenberg, Richard Holden, Takashi Kunimoto, and Olivier Tercieux, "Subgame perfect implementation under information perturbations," Quarterly Journal of Economics 127(4):1843-1881 (2012).

Kenneth J. Arrow, Social Choice and Individual Values, Wiley (1951).

Jeffrey Banks and John Duggan, "A bargaining model of collective choice," American Political Science Review 94 (1):73-88 (2000).

David Baron and John Ferejohn, "Bargaining in legislatures." American Political Science Review 83:1181-1206 (1989).

Duncan Black, Theory of Committees and Elections, Cambridge (1958).

Jeffrey S. Banks, "Sophisticated voting outcomes and agenda control," Social Choice and

Welfare 1:295-306 (1985).

Kim-sau Chung and Jeffrey C. Ely, "Implementation with near-complete information," Econometrica 71(3):857-871 (2003).

Robin Farquharson, Theory of Voting, Yale, 1969.

Peter C. Fishburn, The Theory of Social Choice, Princeton, 1973.

David C. Fisher and Jennifer Ryan, "Optimal strategies for a generalized 'scissors, paper, and stone' game," American Mathematical Monthly 99:935-942 (1992).

Joshua S. Gans and Michael Smart, "Majority voting with Single-Crossing Preferences," Journal of Public Economics 59:219-237 (1996).

Allan Gibbard, "Manipulation of voting rules: a general result," Econometrica 41:587-601 (1973).

Allan Gibbard, "Straightforwardness of game forms with lotteries as outcomes," Econometrica 46 (1978), 595-614.

Russell Hardin, "Why a Constitution," in The Federalist Papers and the New Institutionalism, Bernard Grofman and Donald Wittman, eds., NY: Agathon Press (1989).

Gilbert Laffond, Jean-Francois Laslier, and Michel Le Breton, "The bipartisan set of a tournament game," Games and Economic Behavior 5:182-201 (1993).

Eric Maskin, "The theory of implementation in Nash equilibrium: a survey," in L. Hurwicz, D. Schmeidler, and H. Sonnenschein eds., Social Goals and Social Organization, Cambridge U. Press (1985), pages 173-204.

Kenneth O. May, "A set of independent necessary and sufficient conditions for simple majority decision," Econometrica 20:680-684 (1952),.

Robert M. May, "Some mathematical remarks on the paradox of voting," Behavioral Science 16:143-151 (1971).

David C. McGarvey, "A theorem in the construction of voting paradoxes," Econometrica 21:608-610 (1953).

Richard D. McKelvey, "Intransitivities in multidimensional voting models and some implications for agenda control," Journal of Economic Theory 12:472-482 (1976),.

Richard D. McKelvey, "General conditions for global intransitivities in formal voting models," Econometrica 47:1085-1112 (1979).

Richard D. McKelvey, "Covering, dominance, and institution-free properties of social choice," American Journal of Political Science 30:283-314 (1986).

Nicholas R. Miller, "Graph theoretical approaches to the theory of voting," American Journal of Political Science 21:769-803 (1977).

Nicholas R. Miller, "A new solution set for tournaments and majority voting," American Journal of Political Science 24:68-96 (1980). (Erratum 1983.)

John Moore and Raphael Repullo, "Subgame perfect implementation," Econometrica 56:1191-1220 (1988).

Herve Moulin, "On strategy-proofness and single-peakedness," Public Choice 35:437-455 (1980).

Herve Moulin, "Choosing from a tournament," Social Choice and Welfare 3:271-291 (1986).

Herve Moulin, Axioms of Cooperative Decision Making, Cambridge (1988).

Eitan Muller and Mark Satterthwaite, "The equivalence of strong positive association and strategy-proofness," Journal of Economic Theory 14:412-418 (1977).

Roger B. Myerson and Robert J. Weber, "A theory of voting equilibria," American Political Science Review 87:102-114 (1993).

Roger B. Myerson, "Theoretical comparison of electoral systems," European Economic Review 43:671-697 (1999).

John F. Nash, "Noncooperative Games," Annals of Mathematics 54:289-295 (1951).

John von Neumann, "Zur Theories der Gesellschaftsspiele." Mathematische Annalen 100:295-320 (1928). English translation by S. Bergmann in R. D. Luce and A. W. Tucker, eds., Contributions to the Theory of Games IV (1959), pp. 13-42, Princeton University Press.

Elizabeth Maggie Penn, John W. Patty, and Sean Gailmard, "Manipulation and single-peakedness: a general result," American Journal of Political Science 55(2):436-449 (2011).

William H. Riker, Liberalism against Populism, San Francisco, Freeman (1982).

Kevin W. S. Roberts, "Voting over income tax schedules," Journal of Public Economics 8:329-340 (1977).

Thomas Romer and Howard Rosenthal, "Political resource allocation, controlled agendas, and the status quo," Public Choice 33:27-43 (1978).

Paul Rothstein, "Order-restricted preferences and majority rule" Social Choice and Welfare 7:331-342 (1990).

Paul Rothstein,"Representative voter theorems" Public Choice 72:193-212 (1991).

Alejandro Saporiti, "Strategy-proofness and single-crossing," Theoretical Economics 4:127-163 (2009).

Mark A. Satterthwaite, "Strategy-proofness and Arrow's conditions," Journal of Economic Theory 10:198-217 (1975).

Amartya K. Sen, Collective Choice and Social Welfare, Holden-Day, (1970).

Thomas C. Schelling, Strategy of Conflict, Harvard University Press (1960).

Kenneth Shepsle and Barry Weingast, "Uncovered sets and sophisticated voting outcomes, with implications for agenda institutions," American Journal of Political Science 28:49-74 (1984).

Birgitte Sloth, "The theory of voting and equilibria in noncooperative games," Games and Economic Behavior 5:152-169 (1993).

Author's address: Economics Dept., University of Chicago, 1126 East 59th Street, Chicago, IL 60637.
Phone: 1-773-834-9071. Fax: 1-773-702-8490.
Email: myerson@uchicago.edu. URL: http://home.uchicago.edu/~rmyerson/
Original version: September 1996. Date of this version: January 8, 2013.