

## The constitution as a utilitarian contract<sup>1</sup>

The individuals themselves, each in his own personal and sovereign right, entered into a compact with each other to produce a government; and this is the only mode in which governments have a right to arise and the only principle on which they have a right to exist.

Thomas Paine

The ideally perfect constitution of a public office is that in which the interest of the functionary is entirely coincident with his duty. No mere system will make it so, but still less can it be made so without a system, aptly devised for the purpose.

John Stuart Mill

We have already discussed several works that have assumed uncertainty over future position to derive a normative theory of social choice. Rawls's (1971) theory discussed in Chapter 25 uses uncertainty over future position to derive principles of justice to be included in a social contract; Harsanyi (1953, 1955, 1977) uses it to derive an additive SWF (see Chapter 23).

Buchanan and Tullock (1962) develop a theory of constitutional government in which the constitution is written in a setting resembling that depicted by Harsanyi and Rawls. Individuals are uncertain about their future positions and thus are led out of self-interest to select rules that weigh the positions of all other individuals (Buchanan and Tullock, 1962, pp. 77–80).<sup>2</sup> Buchanan and Tullock's theory is at once positive and normative. Its authors state: "The uncertainty that is required in order for the individual to be led by his own interest to support constitutional provisions that are generally advantageous to all individuals and to all groups seems likely to be present at any constitutional stage of discussion" (Buchanan and Tullock, 1962, p. 78). And the tone of their entire manuscript is strongly positivist in contrast to, say, the works of Rawls and Harsanyi. But they also recognize the normative antecedents to their approach in the work of Kant and the contractarians (see, especially, Buchanan and Tullock, 1962, Appendix 1). Indeed, they state that the normative content of their theory lies precisely in the unanimity achieved at the constitutional stage (p. 14).

<sup>1</sup> This chapter draws heavily from Mueller (2001).

<sup>2</sup> Leibenstein (1965) achieves the same effect by envisaging collective decisions being made by a group of aging individuals for their descendants. Vickrey (1960) assumes people are moving to an island and are uncertain of their positions on the island.

One of the important contributions of Buchanan and Tullock's book is that it demonstrates the conceptual usefulness of the distinction between the constitutional and parliamentary stages of democratic decision making. If unanimous agreement can be achieved behind the veil of uncertainty that shrouds the constitutional stage, then a set of rules can be written at this stage that will allow individuals to pursue their own self-interests at the parliamentary stage in full possession of knowledge of their own tastes and positions. This obviously requires that any redistribution which is to take place be undertaken at the constitutional stage, where uncertainty over future positions holds (Buchanan and Tullock, 1962, ch. 13). Here the similarity to Rawls is striking. Unlike Rawls, however, Buchanan and Tullock allow individuals not just *more* information about themselves at the parliamentary stage, but full information.

The differences in the degrees of uncertainty assumed by Harsanyi, Rawls, and Buchanan and Tullock lead them in quite different directions in describing the principles and institutions that are optimal for making social choices. In this chapter we spell out these differences and draw out their implications. In so doing we outline a general theory of constitutional choice that builds on the Buchanan and Tullock mode of analysis.

### 26.1     **The constitutional context**

Each individual  $R$  can undertake one of  $n$  possible actions,  $a_{rj}$ ,  $j = 1, n$ . These can range from very private actions like scratching one's ear, to very public ones like bombing the local pub. Among the set of actions might be paying a tax to provide a pure public good. Thus, all *collective* action questions can be viewed as decisions about *individual* actions. A law against driving above 65 mph restricts one's freedom to drive fast. A tax on gasoline to finance highway construction both restricts one's ability to purchase gasoline, and expands one's driving opportunities. All *collective choices* are decisions about individual actions.

All actions fall into one of three categories: *neutral* actions that affect only the welfare of the actor; *negative externalities*, actions that make other parties worse off; and *positive externalities*, actions that make other parties better off. Since we deal with situations involving risk and uncertainty, we assume that individual utility functions satisfy the von Neumann-Morgenstern utility axioms, and thus that the utilities of each individual can be regarded as *cardinal* indices (Ng, 1984a; Binmore, 1994, ch. 4).

The community can make three mutually exclusive decisions with respect to any individual  $R$  and her action  $a_{rj}$ : (1) it can *allow*  $R$  the freedom to make the action or not, (2) it can *ban*  $R$  from undertaking the action, or (3) it can *obligate*  $R$  to undertake it. A ban of an action can be regarded as setting an infinite price on the action.

Any action that creates an externality can lead to conflict in the post-constitutional stage over whether to ban or compel this action, and at the constitutional stage over the political institutions to be used to resolve this postconstitutional conflict. This sort of conflict at the constitutional stage can prevent unanimity over the constitutional contract. Following Harsanyi (1955), Rawls (1971), and Buchanan

and Tullock (1962) unanimity can be obtained by assuming uncertainty over future positions at the constitutional stage. Each of these authors defended this assumption in different ways and, at least insofar as Harsanyi and Rawls are concerned, assumed different thicknesses in the *veil of ignorance*, which screens out information about the future. The assumptions one makes about the “thickness” of the veil of ignorance, that is, the information citizens have at the constitutional stage, have important consequences for the types of institutions that are placed in the constitution.

At the constitutional stage individuals choose bans, obligations, and voting rules to maximize their *expected* utilities. The agreement on the constitution must be unanimous, the existence of uncertainty ensures that this unanimity is obtained. At the post constitutional stage individuals know who they are, what their preferences are, and so forth. All private actions have the goal of maximizing utility, as do all collective actions under the voting rules established in the constitution. It is of course possible that an individual will vote to ban an action at the constitutional stage, when she is uncertain about her future preferences, and then in the postconstitutional stage, when she knows her preferences, try to violate the ban. Thus, the community obviously must include in the constitution institutions to ensure *compliance* with it. The compliance issue is taken up in Section 26.8.

## 26.2 The two-action case

Assume that there are only two groups of individuals, Rows ( $R$ ) and Columns ( $C$ ). Each can undertake any one of  $n$  possible actions,  $a_{rj}$ ,  $j = 1, n$ , and  $a_{cj}$ ,  $j = 1, n$ . Each individual in a group has the identical utility function defined over his own action and the action of the players in the other group,  $U_i(a_{rj}, a_{cj})$ ,  $i = R, C$ . Since all  $R$ s have identical utility functions, if one  $R$  experiences a higher utility from undertaking action  $a_{rj}$ , then all  $R$ s do, and so we can think of  $U_i(a_{rj}, a_{ck})$  as the utility an  $i$  experiences when all  $R$ s undertake  $a_{rj}$ , and all  $C$ s undertake  $a_{ck}$ . Each individual undertakes only one action at a time. Actions  $a_{rn}$  and  $a_{cn}$  are defined as no action and are assumed to produce no externalities.

Now consider the possible consequences of  $R$ s and  $C$ s undertaking the actions  $a_{rj}$  and  $a_{cj}$  versus the nonactions  $a_{rn}$  and  $a_{cn}$ . Action  $a_{rj}$  has three possible consequences for an  $R$ : (1) it raises his utility relative to when he undertakes  $a_{rn}$  – we represent this situation as  $u_{rj} > 0$ ; (2) action  $a_{rj}$  does not alter  $R$ 's utility,  $u_{rj} = 0$ ; or (3) action  $a_{rj}$  reduces  $R$ 's utility,  $u_{rj} < 0$ . These utility changes can be thought of as the combined effect on  $R$  of his own gain or loss from action  $a_{rj}$ , and any gain or loss he experiences from contemplating the effect of this action on the  $C$ s. For example, suppose  $a_{rj}$  is  $R$ 's smoking a cigar, although he knows this makes all  $C$ s worse off. If  $R$  suffers sufficient disutility from the knowledge that  $C$ s suffer from his smoking cigars, then  $u_{rj} < 0$  for this action even though in the absence of any  $C$ s, an  $R$  would get positive utility from smoking. The same three possible utility payoffs exist for the action  $a_{cj}$  by the  $C$ s.

Each action by an  $R$  or a  $C$  can have no effect on the other group, or a positive or negative externality. Let us call  $e_{rj}$  the utility change a  $C$  experiences from the action  $a_{rj}$  by the  $R$ s. A positive externality thus implies  $e_{rj} > 0$ , with  $e_{rj} = 0$ ,

Matrix 26.1. Collective action options when external effects are separable

		Column						
		1	2	3	4	5	6	
		$u_{cj} > 0$ $e_{cj} > 0$	$u_{cj} > 0$ $e_{cj} = 0$	$u_{cj} \leq 0$ $e_{cj} < 0$	$u_{cj} \leq 0$ $e_{cj} = 0$	$u_{cj} > 0$ $e_{cj} < 0$	$u_{cj} \leq 0$ $e_{cj} > 0$	
R o w	1	$u_{rj} > 0$ $e_{rj} > 0$	NN	NN	NN	NN	NB	NO
	2	$u_{rj} > 0$ $e_{rj} = 0$	NN	NN	NN	NN	NB	NO
	3	$u_{rj} \leq 0$ $e_{rj} < 0$	NN	NN	NN	NN	NB	NO
	4	$u_{rj} \leq 0$ $e_{rj} = 0$	NN	NN	NN	NN	NB	NO
	5	$u_{rj} > 0$ $e_{rj} < 0$	BN	BN	BN	BN	BB	BO
	6	$u_{rj} \leq 0$ $e_{rj} > 0$	ON	ON	ON	ON	OB	OO

Notes: N = No action required; B = ban of the action; O = obligation to act.  
First letter applies to Rows, second to Columns

and  $e_{rj} < 0$  representing neutral actions and negative externalities. To begin we make the simplifying assumption that the utility functions are separable. Under this assumption the effects of  $R$ 's action,  $a_j$ , and the external effects of  $C$ 's action,  $a_j$ , are both constants, and their combined effect on  $R$ 's utility is simply the sum of the two effects,  $u_{rj} + e_{cj}$ .

Action  $a_{rj}$  has three possible utility consequences for each  $R$ , and three possible external effects producing nine combinations of own effect and externality. The same holds for the  $C$ s, giving 81 combinations of utility payoffs taking into account the possible actions and interactions of the two groups. The number of combinations can be reduced to 36, however, if we assume that an  $R$  does not voluntarily undertake  $a_{rj}$  when  $u_{rj} = 0$ , and likewise for  $C$  when  $u_{cj} = 0$ . The remaining 36 combinations are depicted in Matrix 26.1.

Of the 36 possible situations, 16 require no collective action. The matrix has been constructed so that these cases appear in the upper left-hand portion of the matrix, and are indicated by an NN. The first N indicates that *no* collective decision need be taken with respect to a Row's action  $a_{rj}$ ; the second N has the same implication with respect to a Column's action. In the row 3, column 1 entry, for example,  $R$ 's undertaking  $a_{rj}$  would create a negative externality for the  $C$ s, while a  $C$ 's undertaking  $a_{cj}$  creates a positive externality for the  $R$ s. Since  $u_{rj} \leq 0$ , however, and  $u_{cj} > 0$ , the  $R$ s find it in their own interest not to undertake the action, while the  $C$ s find it in their interest to do so, and the optimal outcome occurs without the need for any collective decision.<sup>3</sup>

<sup>3</sup> Recall, however, that the reason why  $u_{rj} \leq 0$  may be that the  $R$ s suffer disutility if they create a negative externality, that is, because  $e_{rj} < 0$ .

For all entries containing a  $B$ , a ban on a group's undertaking the action *may be optimal*. In row 5, column 1, for example,  $R$ s obtain positive utility from undertaking the action,  $u_{rj} > 0$ , but the action also produces a negative externality,  $e_{rj} < 0$ . If  $e_{rj}$  is large enough relative to  $u_{rj}$ , a ban on the  $R$ 's undertaking the action may be socially optimal. Note that when the  $C$ s commit the same action it produces a positive externality, so that *if* a ban in this situation were optimal, it would be an *asymmetric* ban against only the  $R$ s. Entries containing an  $O$  designate situations in which an obligation *might be optimal* because of the existence of positive externalities, with the two squares labeled  $OB$  and  $BO$  representing the unusual cases of a simultaneous asymmetric ban and obligation being optimal. We return to these and the other asymmetric cases below.

The 16 entries with an  $NN$  designate situations in which collective action is *never* required, because each group acting independently of the other produces the optimal outcome. The 20 additional entries designate situations in which bans or obligations *may be optimal*. It is, of course, conceivable that no collective action of any kind is necessary. A single  $R$ , Robinson, and  $C$ , Crusoe, inhabit an island that is so bountiful that no collective action produces benefits that exceed its costs, and the island is big enough so that all negative externalities are small in comparison with the gains to the perpetrator of the externality. Blissful anarchy is a logical possibility.

In more populous communities and harsher environments, one expects potential gains from collective action. We now explore how optimal collective agreements might emerge out of a two-stage constitutional process in which individuals in the first stage are uncertain over future positions.

### 26.3 The constitutional contract

In the context of a two-stage democratic process, uncertainty can take several forms. The minimum uncertainty needed to produce unanimous agreement on a constitution covering the full spectrum of possible actions is over future identities. Assume that each individual at the constitutional stage can forecast all possible future collective actions and their consequences, that is to say, the entries in Matrix 26.1 and all similar matrices for all other pairs of future actions, including the utility payoffs to the different players. Since each possible "state of the world" is a pair of actions by Rows and Columns, this assumption is equivalent to assuming that each individual at the constitutional stage can envisage all possible future states of the world. Each individual at the constitutional stage knows the  $u_{rj}$ ,  $u_{cj}$ ,  $e_{rj}$ , and  $e_{cj}$  in Matrix 26.1 for every possible pair of actions, and the numbers of Row and Column players,  $n_r$  and  $n_c$ . Each individual at the constitutional stage knows *everything* about the future *except* whether she will be an  $R$  or a  $C$  player. We refer to this situation as one of *identity uncertainty*. One way to think of identity uncertainty arising is to think of individuals choosing a constitution for their future children. Let  $R$  stand for female and  $C$  for male. It may be possible to envisage the utilities men and women will experience from a given pair of actions, and the numbers of men and women in the future. But it may not be possible at the constitutional stage to predict the sex of one's unborn children. If so, then identity uncertainty exists.

If individuals at the constitutional stage know the numbers of Row and Column players, the  $n_r$  and  $n_c$ , then they can calculate the probabilities that they are an  $R$  or a  $C$ . A further degree of uncertainty is added by assuming that these numbers are unknown.  $R$  and  $C$  now represent ethnic groups and the future population growth of each group is unknown. We refer to this as *numbers uncertainty*.

The degree of uncertainty is increased still further by assuming that individuals at the constitutional stage are uncertain about the future utility payoffs – the  $u_{rj}$ ,  $u_{cj}$ ,  $e_{rj}$ , and  $e_{cj}$  – in different situations. We refer to this situation as one of *payoff uncertainty*. A person at the constitutional stage can make no judgment about the likely suffering of a future slave or the benefits to her master.

Each of these types of uncertainty leads to a different institutional solution to the collective action problem.

### 26.3.1 Optimal collective action with only identity uncertainty

Every individual at the constitutional stage can envisage the kinds of issues that will come up in the future, the numbers of individuals in each group, and their utility payoffs. They are uncertain over only whether they will be an  $R$  or  $C$ . Thus, each individual at the constitutional stage can predict for every possible pair of future actions  $(a_{rj}, a_{cj})$  the box in Matrix 26.1 in which the community will be located. If the box is one of those containing an NN, no collective decision is necessary. Many actions are likely to fall into these 16 boxes, so many that the constitution framers are likely to include a clause that allows everyone to do anything he chooses *unless the constitution or a law passed in accordance with the constitution specifically forbids or requires a certain action*, thereby handling all of the possible actions an individual can undertake that affect no one's welfare other than the actor, or have positive external effects on others.

Now consider an action in one of the remaining 20 boxes, say row 5, column 1. Column's action creates a positive externality and gives Column positive utility. Thus,  $C$  need not be compelled to undertake the action and should not be prevented from doing so. Row's action, on the other hand, creates a negative externality while giving Row positive utility. The rational individual at the constitutional stage, uncertain over whether she will be a future  $R$  or  $C$  player, chooses to ban future  $R$ s from undertaking the action if the expected utility from such a ban is positive. The probability that an individual is an  $R$  is  $\pi_r = n_r / (n_r + n_c)$ , while the probability that she is a  $C$  is  $\pi_c = n_c / (n_r + n_c)$ . Her expected utility from the action is then

$$\mathcal{E}(U) = \pi_r u_{rj} + \pi_c e_{rj}. \quad (26.1)$$

If (26.1) is negative for an action leading to a box in row 5, the constitution should ban  $R$ 's undertaking this action. If (26.1) is negative, then so too is (26.2), which is just (26.1) multiplied by  $(n_r + n_c)$ .

$$n_r u_{rj} + n_c e_{rj} < 0. \quad (26.2)$$

Condition (26.2) reveals the close link between the expected utility maximizing choices of an individual at the constitutional stage and the Benthamite SWF; the

optimal collective decision regarding action  $a_{rj}$  maximizes the sum of the utility changes caused by this action.

If (26.3) holds for an action leading to any box in row 6, the constitution framers should agree to obligate  $R$  to undertake the action.

$$n_r u_{rj} + n_c e_{rj} > 0. \quad (26.3)$$

Analogous inequalities with respect to entries in columns 5 and 6 define the conditions under which actions by  $C$  should be banned or compelled. Notice that only the boxes in (row 5, column 5) and (row 6, column 6) can *possibly* lead to *symmetric* bans or obligations on all citizens. We discuss symmetric and asymmetric bans and obligations in the next section.

If the only information individuals at the constitutional stage lacked was knowledge of which future citizen they would be, then the constitution could contain all of the bans and obligations that would ever be needed. Strictly speaking, such a situation involves only Knightian risk, rather than true uncertainty, and individuals at the constitutional stage have all of the information they need to calculate their expected utilities for every pair of actions by Rows and Columns (Knight, 1921). If in 20 or 100 years time, the threat of a flood would require the construction of a dike, the constitution framers could forecast this event, the future preferences of citizens, and determine their tax and effort obligations. These could then be written into the constitution. No second stage in the democratic process would be needed. From the point of view of individuals at the constitutional stage, the constitution could optimally resolve all issues for all time.

**Proposition 1:** *Identity uncertainty combined with full knowledge of preferences and numbers of all future citizens allows individuals at the constitutional stage to specify all future bans and obligations so as to maximize their expected utility in the postconstitutional stage. No second stage of collective decision making is required.*

The assumptions in Proposition 1 are essentially those that Harsanyi (1955, 1977) made in determining principles for moral choices. Each individual can envisage the utility of every individual in every possible future state of the world, and the probabilities that she will be any of those individuals. She chooses that social state, that is, a combination of actions for Rows and Columns, that maximizes her expected utility. This choice maximizes the sum of the future utilities of the community, and thus can be viewed as maximizing a Benthamite SWF.<sup>4</sup> If only identity uncertainty is present at the constitutional stage, then the constitution specifies all actions for all future citizens so as to maximize the Benthamite sum of individual utilities. The social contract/constitution specifies all of the necessary actions of those who are a party to it. No second stage of the political process is needed.

<sup>4</sup> With a few additional axioms Harsanyi (1955) proves that the ethical choices of individuals, which consist of maximizing their expected utilities under the assumption that they have an equal probability of being any future citizen, are equivalent to maximizing a Benthamite SWF. See ch. 23.

### 26.3.2 Optimal collective action with identity and numbers uncertainty

We continue to assume that individuals at the constitutional stage know and can compare the  $u_{rj}$ ,  $u_{cj}$ ,  $e_{rj}$ , and  $e_{cj}$  associated with all future actions by members of the two groups. Thus, the optimal collective decision with respect to an  $R$ 's action that leads to a ban in rows 5 or 6 in Matrix 26.1 must still satisfy equations (26.2) or (26.3). Equation (26.2) requires that the following condition be satisfied:

$$n_r/n_c < -e_{rj}/u_{rj} \quad (26.4)$$

and with respect to a ban of  $a_{cj}$ ,

$$n_c/n_r < -e_{cj}/u_{cj}. \quad (26.5)$$

Since the right-hand sides of (26.4) and (26.5) are assumed to be known, the optimal collective choices can be made once the numbers of individuals in the two groups are established. This information can be obtained simply by citizens voting in the second stage of the political process on the bans. It is in an  $R$ 's interest to vote against a ban of  $a_{rj}$ , and in a  $C$ 's interest to vote for it. The constitution framers can ensure that the optimal collective choice is made with respect to the ban on  $R$ 's action by requiring a referendum with a majority of votes in favor of a ban satisfying (26.4). For example, if the utility gain to an  $R$  from  $a_{rj}$  is known to be three times the loss imposed on a  $C$  from the action ( $u_{rj} = -3e_{rj}$ ), then the expected utility of the constitution framers is maximized by requiring that a future ban against Rows undertaking this action obtain a three-fourths majority or more.

**Proposition 2:** *With  $u_{rj} > 0$ ,  $u_{cj} > 0$ ,  $e_{rj} < 0$ , and  $e_{cj} < 0$ , identity and numbers uncertainty combined with full knowledge of the preferences of all future citizens allows individuals at the constitutional stage to maximize their expected utility by specifying a voting rule for the second stage of collective decision making to decide all future bans against  $a_{rj}(a_{cj})$  such that condition (26.4) [(26.5)] is satisfied.*

From (26.3) we can analogously derive the conditions for obligating  $a_{rj}$  and  $a_{cj}$ :

$$n_r/n_c > -e_{rj}/u_{rj} \quad (26.6)$$

$$n_c/n_r > -e_{cj}/u_{cj}, \quad (26.7)$$

from which we obtain

**Proposition 3:** *With  $u_{rj} < 0$ ,  $u_{cj} < 0$ ,  $e_{rj} > 0$ , and  $e_{cj} > 0$ , identity and numbers uncertainty combined with full knowledge of the preferences of all future citizens allows individuals at the constitutional stage to maximize their expected utility by specifying a voting rule for the second stage of collective decision making to decide all future obligations of  $a_{rj}(a_{cj})$  such that condition (26.6) [(26.7)] is satisfied.*

In the special case that  $u_{rj} = -e_{rj} > 0$  the expected utility of an individual at the constitutional stage is maximized if the ban against  $a_{rj}$  is decided using the simple



majority rule. This is essentially the Rae-Taylor theorem in favor of the simple majority rule, which we discussed in Chapter 6, and rests clearly on the assumption of equal intensities on both sides of the issue.<sup>5</sup>

When the equal intensity condition holds with respect to symmetric negative externalities, that is,  $u_{rj} = -e_{rj} > 0$  and  $u_{cj} = -e_{cj} > 0$ , then the simple majority rule is the optimal voting rule to decide whether to ban action  $a_{rj}$  by Row players, and  $a_{cj}$  by Column players. If Rows are in the majority they will vote to ban  $a_{cj}$  and to allow themselves to undertake  $a_{rj}$ . The relentless logic of expected utility maximization coupled with the equal intensity assumption leads to “a tyranny of the majority” as the optimal outcome of the process of choosing a voting rule that maximizes the expected utility of a citizen at the constitutional stage. The majority votes to allow themselves to do that which it forbids the minority from doing.

**Proposition 4:** *With symmetric negative (positive) externalities and equal intensities on the two sides of the issue (that is,  $u_{rj} = -e_{rj}$ , and  $u_{cj} = -e_{cj}$ ), identity and numbers uncertainty combined with full knowledge of the preferences of all future citizens implies that the simple majority rule is the optimal voting rule to decide whether to ban (obligate) actions  $a_{rj}$  and  $a_{cj}$  by future Row and Column players. The application of this voting rule in the second stage of collective decision making under these assumptions must lead to an asymmetric ban (obligation) of the actions  $a_{rj}$  and  $a_{cj}$ . (Note that the equal intensities assumptions imply that the right-hand sides of both (26.4) and (26.5) equal 1. For a symmetric ban to be optimal,  $n_r/n_c < 1$  and  $n_c/n_r < 1$  would both need to hold, which is impossible.)*

Conversely, we can see that a symmetric ban can be optimal with identity and numbers uncertainty, only when the payoffs are *known* and are such as to make *different* voting rules optimal for the respective bans. For example, if  $u_{rj} > 0$ ,  $u_{cj} > 0$ ,  $-e_{rj}/u_{rj} = 1$ , and  $-e_{cj}/u_{cj} = 2$ , then the simple majority rule would be optimal for banning  $a_{rj}$ , while  $a_{cj}$  should be banned if even a third of the community chooses to do so. If  $1 < n_c/n_r < 2$ , Columns are able to ban  $a_{rj}$  but are not able to block Rows from banning  $a_{cj}$ .

Conditions (26.4) and (26.5) require that the majority required to ban an action be higher, the smaller the gain in utility to an individual in favor of a ban relative to the gain in utility for the person who is allowed to act. In the limit, as the right-hand sides of (26.4) and (26.5) approach infinity, the constitution framers would allow a future ban only if the community unanimously voted in favor of it.

Conversely, as  $-e_{rj}$  grows large relative to  $u_{rj}$  the constitutional convention will wish to establish a presumption against action  $a_{rj}$ . This could be accomplished through a constitutional ban on  $a_{rj}$  with a provision that it could be lifted with a majority of  $m_j \geq -e_{rj}/(-e_{rj} + u_{rj})$ . In the limit, as the utility loss to a Column becomes very large relative to the gain to a Row from the action, its constitutional ban could be lifted only by a unanimous vote of the community.

Analogous considerations once again apply with respect to obligations.

<sup>5</sup> See Rae (1969), Taylor (1969), and Rae and Schickler (1997). Buchanan and Tullock (1962, pp. 128–30) also stress the importance of assuming equal intensities in choosing the simple majority rule.

**26.3.3** *Optimal collective action with identity, numbers, and payoff uncertainty*

For many sorts of actions the most realistic assumption to make is that an individual at the constitutional stage is uncertain over identities, numbers, and future utility payoffs from these actions. For example, it might be reasonable to assume that in 1787 an individual could compare the utility he perceived a smoker obtained from smoking, and the negative externality this action caused at that time, but he would not have been able to envisage very accurately future citizens' utilities and disutilities from smoking, or the numbers of smokers and nonsmokers. More generally, he could not anticipate whether other stimulants similar to tobacco would be discovered, their positive and negative effects, and so on. Both the  $e_{js}$  and the  $u_{js}$  in (26.2) and (26.3) are in these situations unknown.

If the constitution framers can envisage the distribution of utility changes associated with a particular action, then we can simply substitute the expected values of the  $e_{js}$  and  $u_{js}$  into our optimality conditions, and proceed as above. If we think of the constitution as governing the collective decisions of the community over a very long period, however, even this assumption may be questionable. On the other hand, if all elements in the equations defining the optimality conditions are unknowns, no voting rule specifying a qualified majority for making future collective choices can be written into the constitution that maximizes the expected utility of someone at the constitutional stage.

Thus, when reasonable predictions of the utility gains and losses from particular actions cannot be made, the constitution might simply be silent on how future generations should decide them. Although this approach would be intellectually honest, it would impose on future generations the difficult task of both choosing and applying voting rules to deal with many potentially divisive issues, once their preferences were fully known.

Rather than saddle future generations with such choices, the constitution framers might make "an educated guess" as to the magnitudes of the  $-e_j$  and  $u_j$  and define a voting rule accordingly. But what is a reasonable guess?  $-e_{rj}$  is half of  $u_{rj}$ , three times as great? Assuming they are of equal magnitude constitutes a form Schelling point, or alternatively might be interpreted as an application of the principle of insufficient reason to this problem. With  $-e_{rj}$  and  $u_{rj}$  equal, condition (26.4) requires that any ban of an action that fits entries in row 5 be resolved using the simple majority rule. Condition (26.6) demands the simple majority rule for obligations in situations that fit entries in row 6. We have then a normative justification for the ubiquitous use of this voting rule. Unable to estimate the future gains and losses from many collective decisions, the constitution writers assume that they are equal and opt for the voting rule that maximizes their expected utility under this assumption.

**26.4**     **Symmetric and asymmetric bans and obligations**

Although asymmetric bans or obligations are likely to be optimal from the point of view of an individual at the constitutional stage who is uncertain of her future

identity, they may often be infeasible. Suppose that both Rows and Columns get utility out of being free to drive faster than 65 mph when they so choose ( $u_{rj} > 0$  and  $u_{cj} > 0$ ). Rows are skillful and prudent drivers and only drive at these speeds when there is no danger of their harming anyone ( $e_{rj} = 0$ ). Columns, on the other hand, are poor and somewhat reckless drivers ( $e_{cj} < 0$ ). From behind the veil of ignorance, the community could unanimously agree to ban Columns from driving at more than 65 mph, while allowing Rows to drive at whatever speed they choose. But unless Rows and Columns can be identified prior to their stepping behind the wheel, such a ban will be unenforceable. Since both Rows and Columns prefer having the freedom to drive above 65, Columns will simply pretend to be Rows. Given the infeasibility of enforcing an asymmetric ban, a symmetric ban may be optimal. This will be the case when the expected utility of someone at the constitutional stage from a total ban is positive, that is, when (26.8) is satisfied:

$$n_r u_{rj} + n_c e_{rj} + n_c u_{cj} + n_r e_{cj} < 0. \quad (26.8)$$

When (26.8) does not hold the optimal rule will be a symmetric freedom to drive above 65. An analogous condition with the inequality reversed applies to symmetric obligations in the presence of positive externalities. Thus, owing to the transaction costs of enforcing asymmetric bans and obligations, more rules must be applied symmetrically than is suggested by Matrix 26.1.

Identifying those who have different preferences and/or who generate different externalities is, on the other hand, often feasible, thus so too are asymmetric bans, for example, a ban against those who are under 21 consuming alcohol. Thus, an expected utility-maximizing constitution would impose asymmetric bans whenever differences in utility payoffs and external effects from actions can be readily identified.

Row 6, column 5, and row 5, column 6 in Matrix 26.1 contain entries that may seem highly unlikely – a simultaneous ban and obligation for the two groups with respect to the same action. Nevertheless, such asymmetric treatments of different groups are both logically possible and observable in practice. A somewhat archaic and sexist example of this sort of asymmetry would be a constitutionally defined obligation for men to serve in the army, and a ban against women serving. Such asymmetric treatment of these two groups could arise if men got negative utility from being in the army but their service generated positive externalities, while women fancied being in the army but their service generated negative externalities. Under these conditions, citizens who were uncertain of their future sex could unanimously agree on an asymmetric ban and obligation regarding military service.

## 26.5 Continuous actions with interdependent utilities

The assumptions of binary actions and separable external effects have allowed us to illustrate rather simply several important features of the optimal political institutions in a utility-maximizing constitution. Moreover, these assumptions are realistic with respect to many collective choices. Slavery, abortion, and legalized drugs are just three examples of issues that many people view as binary choices. The loss Column experiences when a Row steals from him may reasonably be assumed to be independent of whether Column is also a thief.

In other situations more complex relationships must be assumed to exist, however. The risk of harm that *Rs* experience from *Cs*' driving depends on whether the *Rs* are also driving. Cars can be driven at any one of a continuous range of speeds. Money to provide a pure public good can be contributed in various amounts. To handle such cases, we need to think of  $a_j$  as a continuous variable. To see what is involved, let us assume that Rows and Columns have twice differentiable utility functions defined over the two scalars  $a_{rj}$  and  $a_{cj}$  of the following forms:

$$U_R = U_R(a_{rj}, a_{cj}) \text{ and } U_C = U_C(a_{cj}, a_{rj}). \quad (26.9)$$

An individual at the constitutional convention wishes to maximize his expected utility, which again amounts to maximizing the Benthamite function

$$W = n_r U_R(a_{rj}, a_{cj}) + n_c U_C(a_{cj}, a_{rj}), \quad (26.10)$$

which yields the first-order conditions

$$\begin{aligned} \frac{\partial W}{\partial a_{rj}} &= n_r \frac{\partial U_R}{\partial a_{rj}} + n_c \frac{\partial U_C}{\partial a_{rj}} = 0 \\ \frac{\partial W}{\partial a_{cj}} &= n_r \frac{\partial U_R}{\partial a_{cj}} + n_c \frac{\partial U_C}{\partial a_{cj}} = 0. \end{aligned} \quad (26.11)$$

If both the utility functions and the numbers of Row and Column players are known, we again have essentially the situation first analyzed by Harsanyi (1955), and the constitution framers stipulate the levels of each action ( $a_{rj}, a_{cj}$ ) so as to maximize the SWF in (26.10).

When the utility functions  $U_R$  and  $U_C$  are known, but the  $n_r$  and  $n_c$  are not, one might wish to define a voting rule to reveal the  $n_r$  and  $n_c$ . When  $U_R$  and  $U_C$  are continuous functions of  $a_{rj}$  and  $a_{cj}$ , however, such an option no longer exists. From (26.11) we can solve for the optimal relationships between the numbers of individuals in each group and the marginal utilities from each action.

$$\begin{aligned} \frac{n_r}{n_c} &= - \frac{\partial U_C / \partial a_{rj}}{\partial U_R / \partial a_{rj}} \\ \frac{n_r}{n_c} &= - \frac{\partial U_C / \partial a_{cj}}{\partial U_R / \partial a_{cj}}. \end{aligned} \quad (26.12)$$

If both marginal utilities from  $a_j$  are positive ( $\partial U_R / \partial a_{rj} > 0$  and  $\partial U_C / \partial a_{cj} > 0$ ) and the actions cause negative externalities, then (26.12) defines conditions that determine the optimal levels of both actions. But no voting rule leads to this outcome. If the simple majority rule is used to decide the levels of  $a_{rj}$  and  $a_{cj}$ , and the Rows are in the majority, they will not choose to require levels of  $a_{rj}$  and  $a_{cj}$  that satisfy (26.12). Instead they will allow themselves full freedom to act, so that  $\partial U_R / \partial a_{rj} = 0$ , and the right-hand side of the first equation in (26.12) goes to infinity, while totally banning  $a_{cj}$ . When multiple degrees of an action are possible and utility varies with the level of the action, no qualified majority rule *alone* can be relied upon to determine the optimal level of the action.

The potential scope for a tyranny of the majority is obviously great when the levels of action vary over a wide range. Moreover, unlike the situation when only two choices exist – action or no action – with multiple actions the simple majority rule is likely to produce an outcome that deviates greatly from that which would maximize the expected utility of an individual at the constitutional stage. A closer approximation to the levels of actions that are optimal might be achieved in this situation if the constitution coupled the choice of a qualified majority to decide the level of an action with a symmetry constraint. Whatever level of the action that is allowed (required) of one group must pertain to the other. With this symmetry condition, the simple majority rule in use, and Rows, say, in the majority, they would choose a level of  $a_j$  such that  $\partial U_R/\partial a_{rj} = -\partial U_R/\partial a_{cj}$ , that is, a level that equates the denominators of the right-hand sides of the two equations in (26.12). If the utility functions of the Rows and Columns were similar, then this level of activity would also equate the numerators, and the right-hand sides of (26.12) would both equal 1. Although the levels of  $a_{rj}$  and  $a_{cj}$  would not maximize (26.10), given  $n_r$  and  $n_c$ , they would most likely come much closer to achieving this outcome than allowing one group to set different levels of  $a_j$  for each group so as to maximize its utility.<sup>6</sup>

We conclude that a constitutional convention that expected future members of the community to have similar utility functions defined over continuous levels of different activities could achieve a higher level of expected utility at the constitutional stage, if it coupled the use of the simple majority rule to the requirement that decisions made with this rule apply uniformly to all members of the community.

## 26.6 Decision-making costs

Consider again the entries in row 5 of Matrix 26.1. Rows obtain positive utility from an action that causes a negative externality. It is tempting to argue that no collective action is necessary in these cases, and to rely on the Coase theorem to ensure that a Pareto-optimal outcome is obtained.<sup>7</sup> Columns can simply bribe Rows not to act.

In thinking about the resolution of these conflicts at the constitutional stage, however, such a way around these sorts of difficulties seems illegitimate, at least with respect to the first four entries in row 5. To prevent Rows from acting Columns must offer them a sufficiently large bribe. But with what can Columns bribe Rows if, at the constitutional stage, property rights are not yet secure? Thus, with respect to the kinds of *conflict* issues that are represented in the first four boxes of rows 5 and 6, it seems reasonable to assume that Coasian solutions are not feasible, and provision in the constitution must be made for their optimal resolution.

<sup>6</sup> With diminishing marginal utility from undertaking  $a_{rj}$ , the gain to a Row from going from the constrained level of  $a_{rj}$ , where  $\partial U_R/\partial a_{rj} = -\partial U_R/\partial a_{cj}$ , to the level of  $a_{rj}$ , where Rs are unconstrained ( $\partial U_R/\partial a_{rj} = 0$ ), will tend to be less than the loss in utility if Columns are constrained and  $a_{cj} = 0$ .

Buchanan and Congleton (1998) present examples of situations in which the imposition of a symmetry constraint can improve the realized aggregate utilities of a community.

<sup>7</sup> See Coase (1960), Bernholz (1997a), and the discussion in Chapter 2.

This argument does not hold for the four entries in the bottom right-hand corner of Matrix 26.1, where rows 5 and 6, and columns 5 and 6 intersect. Now each person does have something to trade – her freedom to undertake action  $a_j$ . These four cases can give rise to different forms of prisoners' dilemmas, and the optimal outcome could conceivably be reached by requiring that bans and obligations of these actions be made jointly using the unanimity rule. In discussing the possible problems raised by decision-making costs, therefore, we distinguish between the four boxes in Matrix 26.1 which potentially give rise to prisoners' dilemmas (the intersections of rows 5 and 6 with columns 5 and 6), and the other 16 entries in these two rows and columns, which we refer to as *conflict* issues.

### 26.6.1 Prisoners' dilemmas

In a prisoners' dilemma, a unanimous agreement to adopt the cooperative strategies is possible *without* any uncertainty over who the players are or their utility payoffs (Müller, 1998). Thus, *even* when none of the three forms of uncertainty is present at the constitutional stage, for actions giving rise to prisoners' dilemmas, the players have the incentive to agree to the jointly cooperative actions, and this agreement can, in principle, be written directly into the constitution.

Unfortunately, of course, in prisoners' dilemma situations each individual has an incentive to break the agreement in the postconstitutional stage. To achieve the gains from cooperation in prisoners' dilemmas, agreements must also include incentives to cooperate as, for example, penalties for noncooperation. An effective ban of stealing must stipulate the penalty to be imposed if the ban is violated. The optimal penalties to deter stealing a loaf of bread may differ from those to deter robbing a bank. Thus, collective decisions in many prisoners' dilemma situations do not simply involve the specification of the desired actions by each party – do not steal – they also involve multiple, possible retaliatory actions by the community.

Contributions to the provision of a pure public good also have the characteristics of a prisoners' dilemma, but in this case the *action* involved – how much each person contributes – is essentially a continuous variable. The optimal contribution of each citizen depends on her preferences and income, and the number of groups with different preferences for the public good is likely to exceed two. In communities with large numbers of individuals with different preferences and incomes, the *decision-making costs* of determining each individual's contribution, the penalty for failing to contribute, and so on will be large. When these costs are taken into account, some less-than-unanimity rule may prove optimal.

Once collective decisions are made with a qualified majority rule, however, an individual loses the protection afforded by the unanimity rule against decisions that make her worse off. She becomes exposed to the *external costs of collective decision making*.<sup>8</sup> Thus, the decision-making costs associated with the unanimity rule convert a potentially cooperative game to find a Pareto-preferred set of actions into a conflict between those in the winning coalition who obtain net benefits from the collective action, and those in the losing coalition who do not. Uncertainty

<sup>8</sup> See Buchanan and Tullock (1962, pp. 63–91) and discussion in Chapter 4.

reappears at the constitutional stage over whether a given individual will be in a future winning or losing coalition.

The impact of decision-making costs on the choice of collective decision rule can be studied under the assumption that there are again only two groups in the community, the winners and the losers under a given collective decision. Let  $w$  be the gain in utility an individual at the constitutional stage expects from a particular collective action should he be on the winning side on this issue, and  $s$  the loss if he is on the losing side. The probability that the individual is on the winning side of issue  $j$ ,  $p(m_j)$  is a function of the majority required to pass it,  $m_j$ , where  $p'(m_j) > 0$ , and  $p''(m_j) < 0$  up to  $m_j = 1$ . In choosing a voting rule to decide this issue, an individual at the constitutional stage must weigh the gain in utility he expects from increasing the majority required to pass an issue, and thus his chances of being on the winning side, against the decision-making costs of finding a set of actions that can win a higher majority. Let us call these  $d(m_j)$ , where it is reasonable to assume  $d'(m_j) > 0$ , and  $d''(m_j) > 0$  up to  $m_j = 1$ . A member of the constitutional convention must thus choose  $m_j$  to maximize

$$\mathcal{E}(U) = p(m_j)w - [1 - p(m_j)]s - d(m_j), \quad (26.13)$$

which yields the  $m_j$  satisfying

$$p'(m_j)(w + s) = d'(m_j). \quad (26.14)$$

The left-hand side of (26.14) is the marginal gain in utility expected from increasing the required majority; the right-hand side is the marginal increase in decision-making costs. The voting rule that maximizes the expected utility of someone at the constitutional stage balances these marginal gains and costs of alternative required majorities.

If we think of the voting process as a search for information about individual preferences, for example, the willingness of each individual to contribute to the provision of a public good, it seems reasonable to think of marginal decision-making costs rising continuously with the majority required to pass an issue, as it becomes more and more difficult to discover a contribution that makes an individual with outlier preferences better off, and the incentive to engage in strategic holdouts increases. An alternative way to envisage the process, however, is as a search for winning coalitions. Each new proposal may be quite different from its predecessor and win support from a quite different set of voters. When the voting process is of this form, the possibility of cycling must be entertained. Decision-making costs might then actually *fall* as the required majority is increased over some range of  $m_j$ , because increasing  $m_j$  lowers the probability of a cycle. This is particularly likely if the issues to be decided resemble the determination of the quantities of pure public goods, and thus it is reasonable to assume that the conditions needed to invoke Caplin and Nalebuff's (1988) theorem are satisfied. In this case the probability of cycles can be expected to fall as the required majority increases, reaching zero at an  $m_j$  of 0.64. This implies that marginal decision-making costs are U-shaped with the bottom of the  $U$  perhaps somewhere around 0.64 (see Figure 26.1). The marginal benefits from increasing  $m_j$ ,  $p'(m_j)(w + s)$ , are then likely to cut  $d'(m_j)$

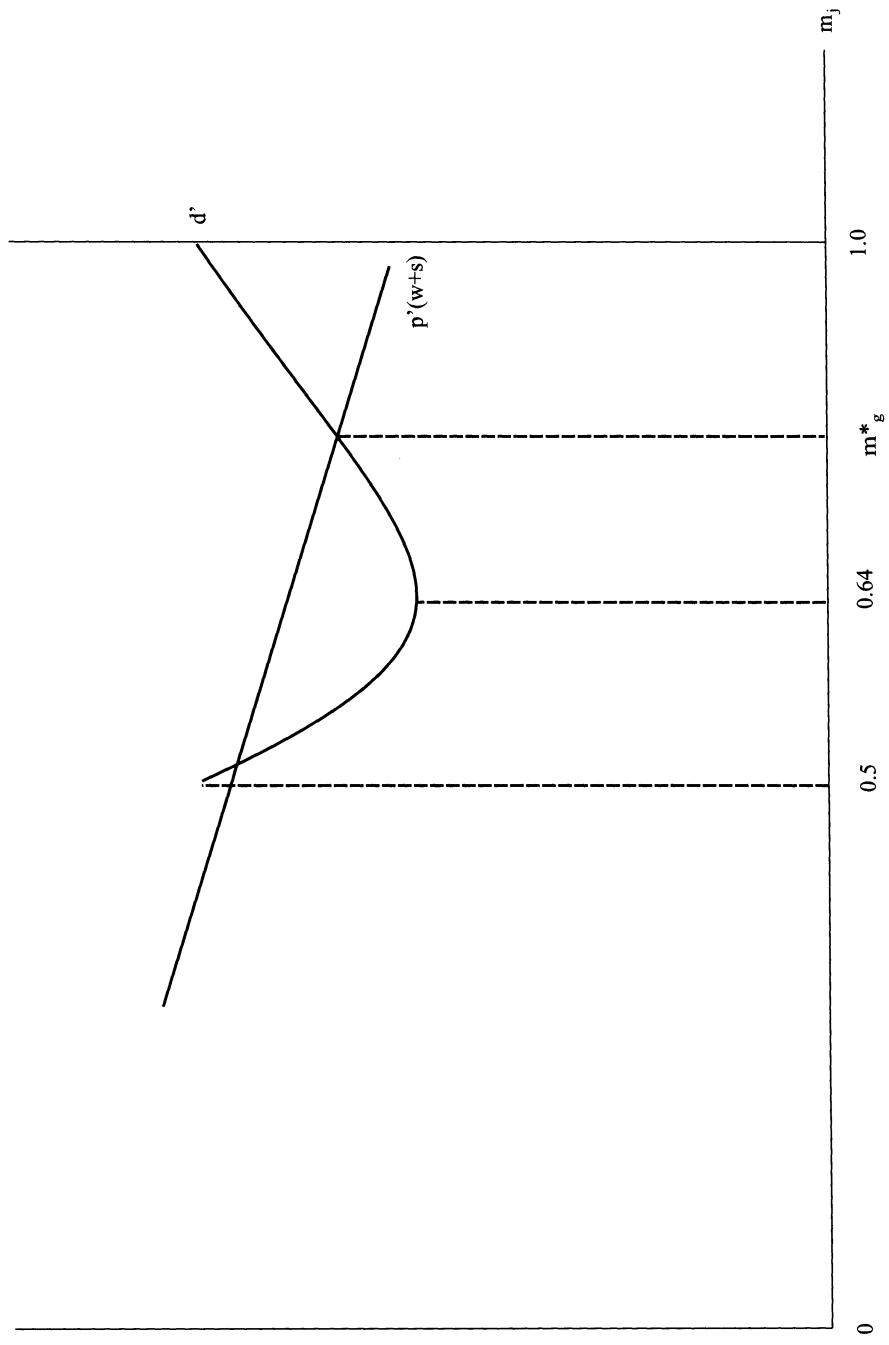


Figure 26.1. Choosing the optimal majority.



twice with the optimal  $m_j^*$  being somewhere around or above 0.64. Taking into account the possibility of cycling and the decision-making costs it causes would lead a constitutional convention to reject the simple majority rule for public good–prisoners’ dilemma issues in favor of a higher qualified majority rule.<sup>9</sup>

### 26.6.2 Direct conflicts

The other 16 entries in rows and columns 5 or 6 of Matrix 26.1 involve one-way externalities. Entries in row or column 5 involve negative externality issues like smoking in public places and driving at high speeds; entries in row or column 6 positive externalities. These issues can be thought of as single-dimensional, ranging from out right prohibitions and obligations to blanket freedoms. The collective action involves the resolution of a conflict among the citizens over the optimal severity of a ban or obligation. It is reasonable to assume that individuals have single-peaked preferences with respect to these sorts of issues; that is, each person favors a ban or obligation of a particular degree of severity with utility falling off as the severity chosen deviates from this ideal level. The unanimity rule is not an option for resolving such conflicts unless side payments are also allowed as a way of securing a Coasian exchange.

With a single-dimensional issue an individual has an incentive to vote sincerely. Proposals to restrict an action can be made in increasing degrees of severity. The winning proposal under an  $m_j$ -qualified majority rule will impose a restriction corresponding to the ideal point of the voter at the  $m_j$ th percentile of the distribution of voter ideal points. The choice of  $m_j$  amounts to the choice of the percentile of the distribution of ideal points where the restriction will lie. The time required to select one percentile should not differ much from the time to select another;  $d'(m_j)$  might reasonably be assumed to be zero. The constitutional convention can treat decision-making costs as a constant deadweight loss. When choosing the optimal majority to resolve single-dimensional conflict issues, only the effects of this choice on the expected utility payoffs need be weighed. Future decision-making costs should not be a factor.

## 26.7 Rights and obligations

Consider equation (26.4) once again. As the right-hand side approaches zero, the majority required to prohibit  $a_{rj}$  approaches unanimity. Now there are two ways in which the right-hand side of (26.4) might approach zero. First, of course, it equals zero if  $e_{rj} = 0$ . If  $C$ ’s utility is unaffected by  $a_{rj}$ , then  $R$  should be free to act, and (26.4) calls for a unanimous vote of the community in the second stage of the political process to prohibit her from doing so. But there are a myriad of actions,  $a_{rj}$ , that benefit  $R$  and have no impact on others. It would be impossible for the constitution to list all of these and specify that they could be abridged only through a unanimous vote of the community. As noted above, such actions seem

<sup>9</sup> See discussion in Chapter 5.

most efficiently handled through a blanket provision that allows all actions that have not been specifically prohibited.

The ratio  $-e_{rj}/u_{rj}$  also approaches zero even when  $-e_{rj} > 0$ , as  $u_{rj}$  becomes very large. In this case individuals at the constitutional stage who thought that they might be a future  $R$  could not simply count on a broadly defined freedom to do what one chooses to protect their freedom to do  $a_{rj}$ . Because  $C$ s experience a utility decline from  $a_{rj}$ , future  $C$ s may try to prohibit  $R$ s from doing  $a_j$ . Such restrictions might be imposed by a simple majority vote and result in a loss in net utility if an  $R$ 's freedom to do  $a_j$  were not explicitly protected. Individuals who are uncertain of whether they would be a future  $R$  or  $C$  would maximize their expected utility at the constitutional stage by explicitly requiring that a proposal to restrict the freedom to do  $a_j$  must pass by a supramajority, which could range up to unanimity (see Figure 26.2).

If  $R$ s experience a great loss from not doing  $a_j$ , they would only vote for a proposal to restrict their freedom to do  $a_j$  if they were compensated for this loss or cajoled into accepting it. Although one can imagine groups being somehow convinced to give up their veto powers in such situations, one expects this to be rare if the constitutional convention correctly anticipated the relative payoffs from the action when it chose to protect it by invoking the unanimity rule.  $R$ s would nearly always vote down proposed restrictions. Time spent debating and voting on such restrictions would be wasted. Anticipating that most future proposals to restrict this action would lose under the unanimity rule, future decision-making costs could be economized by defining a constitutional *right* guaranteeing  $R$ s the freedom to do  $a_j$ . This guarantee would prohibit any future political or private attempts to infringe on an  $R$ 's freedom to commit the defined action, or if the analogous condition holds for  $C$ s, on anyone's freedom. Since a right always carries with it the freedom *not* to undertake the action, the community could still try and bribe or persuade a group to refrain from a particular action, and so both outcomes possible under the unanimity rule are still open to the community after it defines a right.

Several features of constitutional rights under this theory are to be noted. First, explicit rights are defined only for actions capable of generating sufficiently strong negative externalities to elicit efforts by some members of the community to restrict the actions. In the absence of possible negative externalities, even actions that provide considerable benefits for the actor will not be challenged and need not be protected. Second, there is an inherent tension between constitutional rights and the principle of majoritarian democracy. When the institutions of explicitly defined rights and the simple majority rule are both found in the constitution to deal with situations where individual interests conflict, these situations will differ dramatically in the perceived losses imposed on the different sides from curtailing the action. The simple majority rule is optimal for resolving a negative externality when an individual at the constitutional stage expects the utility gain from undertaking the action to equal the loss it causes. Rights are defined precisely where the simple majority rule is not optimal, because the expected gains and losses from a ban are dramatically different and the constitution framers wish to preclude its use. Because rights will be defined only when significant losses are expected for those prevented from acting relative to the losses imposed on others, disputes over rights are likely to

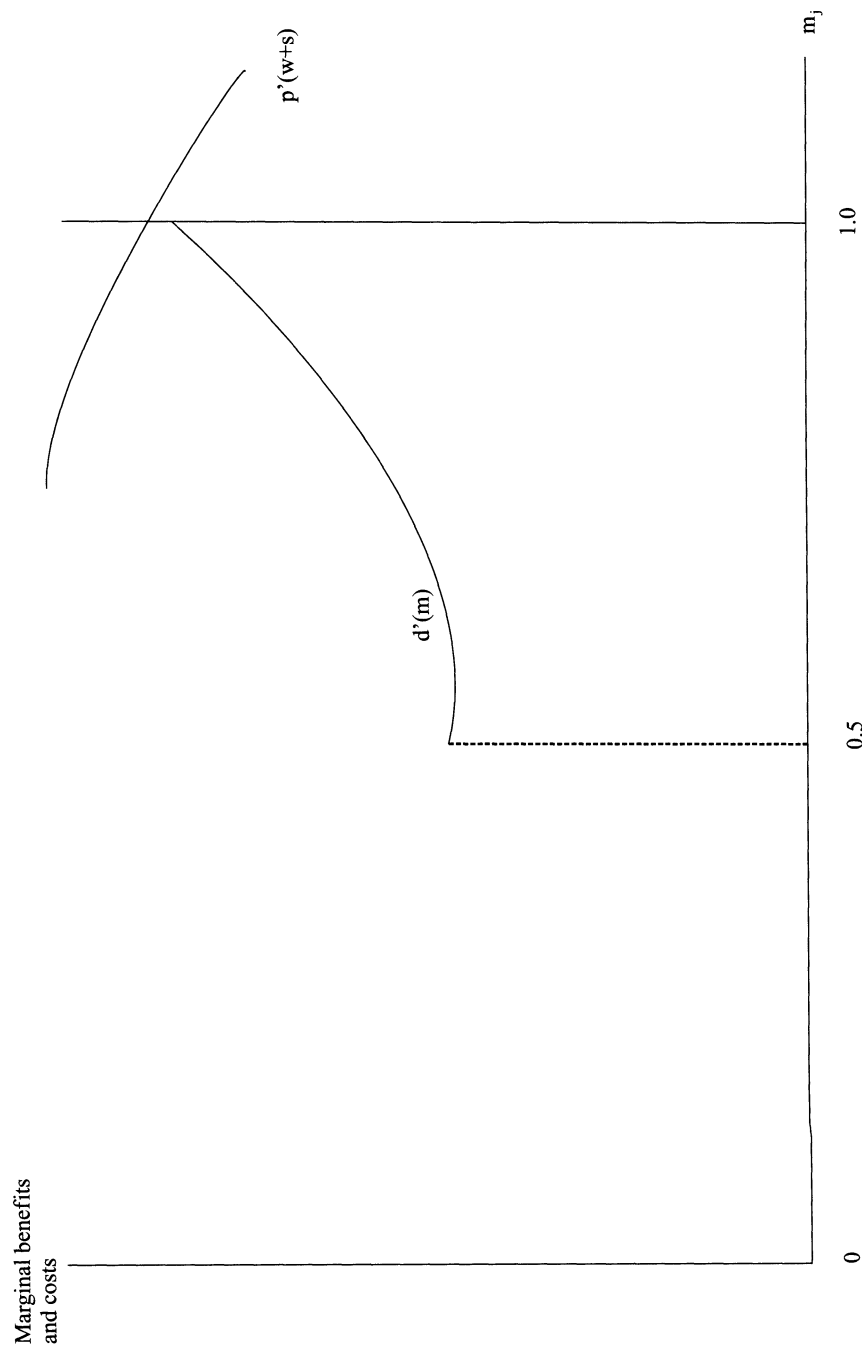


Figure 26.2. The unanimity rules as the optimal majority.

be emotionally charged, as they pit a perhaps substantial majority that feel harmed by the action against an intense minority that benefit from it.<sup>10</sup>

A right protects an individual's freedom *to act*. Therefore, all rights explicitly defined in a constitution contain an *implicit obligation* on all individuals *not* to interfere with an individual who undertakes a constitutionally protected action.<sup>11</sup> *Explicit obligations*, on the other hand, *compel* certain actions. Just as a constitutional right to do  $a_j$  can be thought of as a substitute for a provision requiring that any future restrictions on  $a_j$  imposed by the community be unanimously approved, a constitutionally mandated obligation to do  $a_j$  is a substitute for a provision requiring that all future exemptions to this obligation be unanimously approved. Both need to be defined only in situations of conflict. There is, however, a very important difference between a right and an obligation. A right *allows* an individual the freedom to do  $a_j$ , but does not compel this *choice*. The individual remains free to do  $a_j$  or not. Thus a right extends and strengthens the blanket freedom to do as one chooses that a constitution should provide.

An explicit obligation compels  $a_j$ . The individual is left no choice. Such compulsion is needed because the actor is made worse off to benefit the rest of the community, that is, only in a situation of conflict. Thus, an obligation is a form of slavery to the community. One expects, therefore, in communities where individuals perceive significant gains from allowing people the freedom to make their own choices, the number of rights *to act* defined in the constitution greatly exceeds the explicit obligations *to act*.

## 26.8     **Constitutions: contracts or conventions?**

The idea that a constitution is a kind of contract among the members of a community to establish institutions of government can be traced back at least as far as Thomas Hobbes (1651) and, as already noted, has featured prominently in the public choice literature beginning with Buchanan and Tullock's *The Calculus of Consent*. This view of constitutions as contracts has been seriously challenged by several writers in the last decade, who prefer to think of it as a *convention* or *device* for coordinating the actions of members of society.<sup>12</sup> Although the issue is partly one of semantics, more is at stake than just the proper use of words, since underlying the two perspectives are somewhat different perceptions of what constitutions are and what they accomplish. We shall pause, therefore, to examine the reasoning behind the two perspectives.

### 26.8.1     *Constitutions as contracts*

Following Hobbes constitutional contracts are often seen as agreed to in a state of anarchy (e.g., Buchanan, 1975a). Let us imagine, therefore, a small community

<sup>10</sup> For further discussion of these issues, see Mueller (1991, 1996a, ch. 14).

<sup>11</sup> The word "rights" is often used today in reference to entitlements. Such "economic rights" can also be defended as constitutional provisions. Here the definition of an action must encompass obtaining needed medical care, an adequate diet, and so forth. See Mueller (1991, 1996a, ch. 16), and discussion in Chapter 3 of this book.

<sup>12</sup> See Hardin (1989, 1990); Ordeshook (1992); Binmore (1994, pp. 28–31); Kolmar (2000); and Filippov, Ordeshook, and Shvestova (2001).

living in anarchy that considers creating political institutions to facilitate making future collective decisions. The community is small enough so that all members can meet in assembly, and it meets to draft a constitution. It is soon decided that future collective decisions will also be made in an assembly composed of all members of the community. The question of the voting rule to be used in the future takes more time to resolve. Some favor the simple majority rule, others a three-fourths majority, a few full unanimity. After much debate the assembly agrees – unanimously – to require a two-thirds majority for all future collective decisions.

Why might the community *require* that the choice over the future voting rule be unanimous at this first meeting? There are at least two reasons. The first is to solve the infinite regress problem. What voting rule should be used to choose a voting rule? *If* the community can unanimously agree on a voting rule for making future collective decisions the infinite regress is broken.<sup>13</sup> The second reason for requiring unanimous agreement at the constitutional stage is to increase the likelihood of future compliance with the constitution's provisions. If the two-thirds majority rule is chosen, then there will be some people who are harmed by future collective decisions. How can one be sure that they will go along with the community's decision? The answer, of course, is that one can never be sure, but the likelihood is higher if the losers on a future decision agreed to this choice of voting rule, because in agreeing they endorsed a procedure that they knew would allow some people to be harmed by collective decisions and at least implicitly agreed to be among those harmed.<sup>14</sup>

To further enhance the likelihood of compliance, one could well imagine all members of the community *signing* the constitution after it was voted upon, just as they would sign a private contract. In engaging in this symbolic act each citizen would further commit herself to abide by the constitution.

This point might be buttressed by drawing an analogy between constitutions and marriages.<sup>15</sup> Today a man and woman often live together for several years without being married, and then decide to marry even though this act will not alter their day-to-day life-style. Why do they go through the bother and expense of a formal marriage? One explanation is that they have decided to make a deeper commitment to the relationship, and to symbolize this deeper commitment by formally wedding. By so doing they both signal to one another a willingness to live together for a long period of time, "for better or worse," and so on. Signing the marriage contract may strengthen each party's commitment to the relationship, just as signing a constitution – or voting to ratify one – might strengthen each citizen's commitment to it. For some people such symbolic acts have meaning.

In communities that are too large to function solely as direct democracies, a second form of compliance problem arises. How can one ensure that the future representatives of the citizens will make decisions that advance the citizens' interests

<sup>13</sup> See Buchanan and Tullock (1962, pp. 6–8).

<sup>14</sup> Rawls (1971) is much concerned about the question of ensuring compliance in his theory of the *social contract*, as is Buchanan (1975a) in his Hobbesian theory of constitutions.

<sup>15</sup> Although highly critical of the contractarian approach to constitutions, Filippov, Ordeshook, and Shvestova (2001) draw the same analogy.

and not simply their own? Here again we can think of a constitution as a contract, but now as a *principal–agent contract*. As in all principal–agent contracts the question of creating the proper incentives for the agents is salient. Competitive elections are one obvious device – constitutional checks and balances another.

Thus, the constitution-as-contract approach to constitutional analysis can be seen to have three possible conceptual advantages: (1) it solves the infinite regress problem of the choice of voting rule, (2) it gives a motive for why citizens might comply with the constitution, and (3) it highlights the principal–agent nature of representative government and the need, therefore, to design institutions that align representatives’ interests with those of the citizens.

### 26.8.2 *Constitutions as conventions*

Ordeshook is concerned with a different sort of infinite regress problem from the one described above.

But if contracts ensure that people do things that they would not otherwise do, it is difficult to isolate the ultimate source of a constitution’s durability. Are its provisions enforced by yet a second contract, that is enforced by a third, and so on? Are they enforced from within, by the police, the courts and the military? Or must they be enforced by force to be administered by an oligarchy that stands removed from constitutional limits? The answer to the first question is obviously “No,” the second question merely pushes the problem back a step so that we must ask, “How are the provisions enforcing those enforcement mechanisms enforced?”

Ordeshook (1992, p. 144)

According to Ordeshook the only solution to this infinite regress problem is for the constitution to be *self-enforcing*. The constitution must consist of a set of *devices* or *conventions*, which provide the proper incentives for their self-enforcement.<sup>16</sup>

All of those who reject the constitution-as-contract perspectives emphasize the game-theoretic nature of the problem of constitutional design. Society confronts a plethora of recurring social dilemmas, and it must somehow settle in on one of the many possible equilibria to these supergames. A constitution is a mechanism, a set of conventions, that selects the equilibria.<sup>17</sup>

A frequently used example of a convention is whether a community drives on the right or left side of the road. Young (1993) has employed evolutionary game theory to demonstrate how a community would converge on one of the two possible equilibria in a supergame, even without any communication among the citizens. Such convergence is likely to be much quicker, however, if the citizens do communicate. Imagine, therefore, that the automobile is about to be introduced into our small community. It has observed what has happened in other communities when

<sup>16</sup> See also Filippov, Ordeshook, and Shvestova (2001, especially ch. 5).

<sup>17</sup> Given his emphasis on game theory as a tool for analyzing constitutions, one might expect Cooter (2000) to commit himself to the constitutions-are-conventions position. But he also recognizes the advantage of the constitutions-are-contracts idea in fostering compliance (pp. 273–6).

the automobile was introduced, and wishes to avoid the many accidents that occur through its anarchic use. A meeting is called to decide which side of the road cars should drive on.

As each person enters the meeting she has a preference for either the right side or the left. Each also knows, however, that if her favored side is not chosen her loss will not be great. The first decision that the meeting must make is to choose a voting rule for making the choice of road side. The community unanimously agrees to use the simple majority rule for this choice, because no one wants to spend a lot of time deciding this issue. A motion is then made, a vote taken, and a side of the road is chosen. The meeting is over quickly.

This simple example illustrates the main characteristics of the constitutions-are-conventions perspective. There are multiple (two) equilibria from which to choose. Once a selection is made, the rule becomes self-enforcing. No one has an incentive to break the convention. Note also that there would be little gained in thinking of this decision as being some sort of contract. It would be far less likely than in the first example that anyone would suggest that everyone should *sign* a proclamation that all citizens should drive on the left side of the road. No symbolic acts of commitment are needed because of the self-enforcing nature of the convention.

### 26.8.3 Discussion

Constitutions are contracts. Constitutions are conventions. Both statements are metaphors, and like all metaphors neither one is literally true. On the other hand, each does connote a set of ideas that helps reveal important underlying characteristics of constitutions. Constitutions resemble *both* contracts and conventions.

A provision in a constitution that the head of state should be elected every four years fits the metaphor that constitutions are conventions nicely. There are a variety of terms of office that might be chosen – three years, four years, life. Each can be thought of as resulting in a different long-run equilibrium. By establishing four years as the convention, the constitution effectively selects one of these equilibria. Once it becomes established, it is likely to be self-enforcing. There are many elements in most constitutions that resemble this one.

Now consider, however, a constitutional ban of divorces. Such a provision does not seem to be reasonably characterized as a convention. Certainly it is not likely to be self-enforcing. At some time, some people are likely to want to get divorced, and if this provision of the constitution is to be enforced it will have to be by the police, the courts, and/or the military. If the constitution includes provisions like this one it will also have to create agencies to enforce them, and the infinite regress problem described by Ordeshook raises its ugly head. Once the police, courts, and military can prevent divorces, what stops them from preventing the use of birth control devices, sexual relationships out of marriage, and many other actions that are *not* banned in the constitution?

All constitutions contain provisions that require enforcement agencies for them to be effective. Their inclusion in the constitution immediately thrusts the citizenry into a principal-agent relationship with the state, and the metaphor of constitutions

as contracts begins to have value. Why would a community ever ban divorces? Why would a community ever ban slavery? The two-stage theory of constitutions presented in this chapter gives some insight into possible answers to these questions. A community might ban divorces or slavery if it believed that the future gains to anyone seeking a divorce or wishing to own slaves would be very small relative to the losses imposed on others by these actions. When placing such provisions into the constitution, the citizens must also create agencies to enforce these bans, and thus also include provisions that give the agents incentives to abide by the constitutional contract. The principal-agent nature of the constitutional contract must be faced head on.<sup>18</sup>

One danger in thinking of constitutions as self-enforcing conventions is that one obtains the impression that these “mechanisms” or “devices” for coordinating actions will – if properly designed – run forever. The value of thinking of them as contracts is that one recognizes that contracts often need to be rewritten to reflect changing situations, *and* one immediately identifies who it is that should do the rewriting. If the community thought when it first wrote the constitution 200 years ago that a ban of divorces should be included, and today it feels differently, then it should meet again and change the constitution. This in turn implies that when the constitution is first written, provisions should be made that allow the community to rewrite the constitution as conditions change.<sup>19</sup> In contrast, the metaphor of constitutions as conventions connotes an evolutionary process of selecting conventions and equilibria, which is somehow out of the hands of the citizens.

Although all constitutions have the attributes of both contracts and conventions, they differ in the extent to which they exhibit these attributes. The British Constitution comes closest to resembling a set of conventions that serves to coordinate the political activity of the nation. Except for the signing of the Magna Carta at Runnymede in 1215, there are no “constitutional moments” in British history which might be identified as instances of contractual agreement among the citizens.<sup>20</sup> The unwritten nature of the British Constitution gives it a great deal of flexibility in responding to changes in the environment. Over time the convention has evolved that a national election must be held at least once every five years, but in times of war or national crisis national elections have occasionally been suspended. The British Constitution is a flexible and evolving set of conventions.

In contrast U.S. history does contain that great “constitutional moment” at the end of the eighteenth century, when the U.S. Constitution was written and ratified. The quotation from Thomas Paine at the opening of this chapter reveals that he thought of the Constitution as a “compact,” and this was probably true for many of those involved in its writing and ratification. Many Americans today probably also

<sup>18</sup> The problem of constraining the agents in government is central in the contractarian approach to constitutions of Brennan and Buchanan (1980, 1985). Merville and Osborne (1990) also stress the principal-agent nature of the constitutional contract, and emphasize that the contract must be self-enforcing. Thus, the need for constitutions to be self-enforcing is *not* what divides the contractarians and the conventionarians.

<sup>19</sup> See my proposals in Mueller (1996a, ch. 21). One of the great weaknesses of the U.S. Constitution has proved to be the difficulty of changing it. See Ackerman (1998).

<sup>20</sup> Even the signing at Runnymede involved only the king and some barons.



feel that they are willing parties to this compact owing to the illustrious status of the “Founding Fathers.” And this sense of belonging may help explain the reverence with which so many Americans hold their Constitution and thus its durability.

### 26.9 Conclusions regarding two-stage theories of social choice

In this chapter we have described the basic elements of a two-stage theory of social choice, where decisions in the first stage are made behind a veil of ignorance, while decisions in the second stage are made with each individual in full knowledge of her personal preferences. We have seen that the implications of this theory depend crucially on the nature of the uncertainty assumed in the first stage. If participants at the constitutional convention can envisage the utilities of every future individual for every possible future action, and are uncertain only with regard to which of these future individuals they will be, they can write all rules governing future actions into the constitution. Postconstitutional politics disappears, and the constitution maximizes a Harsanyi social welfare function. Government will not disappear entirely, because individuals in the postconstitutional stage may have incentives to disobey the stipulations of the constitution, and such cheating must be punished. However, no additional collective choices and thus political institutions need be defined.

On the other extreme, if the constitution framers lack all information for calculating future probabilities and utility payoffs, they are incapable of writing rules into the constitution that will maximize their expected utilities. They are then thrust into Rawls’s world, and some additional normative principles – like Rawls’s two principles of justice – must be invoked to select political institutions for making future collective choices and resolving future conflicts.

If the middle degree of uncertainty is assumed – individuals can judge the likely utilities associated with different future actions, but cannot determine the numbers of individuals who will benefit or be harmed by each action – it may be possible to select voting rules to reveal this information in the postconstitutional stage. It is in this middle area of uncertainty that the principles of public choice come into play. Implicitly it was this middle degree of uncertainty that Buchanan and Tullock (1962) assumed in their analysis of the choice of voting rules and other political institutions at the constitutional stage.

Although we have been able to derive some very precise conditions for the selection of a particular voting rule or the specification of a right to act, we have done so under rather restrictive assumptions – there are only two groups of individuals, they are able to make interpersonal cardinal utility comparisons, and so on. If we were to expand the number of groups with different preferences our ability to define voting rules that would maximize the expected utility of someone at the constitutional stage would decline rapidly (Mueller, 2001). Thus, the real lesson to be learned from this exercise is not that it is possible in a utilitarian theory of constitutions to derive conditions under which a voting rule like the simple majority rule is optimal, but rather how restrictive the assumptions are that one must make to accomplish this task.

On the other hand, we have also limited consideration to the family of qualified majority rules running from dictatorship up to unanimity. The potential for

specifying voting rules that maximize the expected utility of someone at the constitutional stage is greatly enhanced once one allows the constitution framers to consider some of the procedures for revealing individual preferences that have been invented. For example, the point voting procedure discussed in Chapter 8 can be designed to maximize a Benthamite SWF and thus would be an attractive option for individuals at a constitutional convention who wished to select a voting rule that would reveal their preferences for public goods issues in the postconstitutional stage. Under the assumptions that support the probabilistic voting model discussed in Chapter 12, a set of electoral rules that would produce a two-party system would also maximize an SWF. In the forty years since Buchanan and Tullock wrote *The Calculus of Consent* public choice has produced a wide range of candidates for institutions that could be chosen by a group of individuals who sought to write a constitution to maximize their expected utilities when they were uncertain of their future positions under the constitution.<sup>21</sup>

### **26.10 From the normative, two-stage theory of constitutions to hypothesis testing**

As noted at the beginning of this chapter, Buchanan and Tullock's *The Calculus of Consent* can be regarded as both a normative and a positive theory of constitutions. Most of the analysis of two-stage theories of constitutions has tended to be normative, and that has been the approach taken in this chapter. Before leaving these two-stage theories we shall briefly discuss the extent to which they can or have been tested empirically.

There are two ways of thinking about testing the implications of constitutional theories. The first is to think of them as theories about how political rules or institutions translate into outcomes.

Rules → Outcomes.

Much public choice can be regarded as developing and testing theories about this aspect of constitutional political economy. For example, the theories of how different electoral rules determine the number of parties represented in the legislature discussed in Chapter 13 fall into this category. Whether or not a country has a two-party or a multiparty system, or a presidential or a parliamentary system in turn will affect the size and composition of its governmental outlays.<sup>22</sup>

The second way to think about constitutional theories is as theories about how individual preferences translate into political rules, where the relevant preferences in this case pertain to the individuals who write the constitution.

Preferences → Rules.

It is this way of viewing the theory of constitutions that is most closely related to the two-stage theory outlined above. For example, the two-stage theory of constitutions predicts that individuals place rights to undertake specific actions into the

<sup>21</sup> See Mueller (1996a).

<sup>22</sup> See Persson and Tabellini (2000a) and the discussion in ch. 21.

constitution if they envisage significant losses to those prevented from undertaking the action relative to any externality it might cause, and *if* they are uncertain over whether it will be they or someone else whom the community might try to prevent from undertaking the action. The United States was settled in part by people who had escaped from religious persecution in Europe, and at the time its constitution was written, many might still have feared that some future majority might try and prevent them from practicing their religion and thus that the freedom to do so required protection in the constitution. Similarly, many had been subject to arbitrary arrest under British rule and thus might also have felt uncertain about this sort of danger. The existence of several of the rights to act protected in the Bill of Rights of the U.S. Constitution is easily accounted for by the two-stage theory of constitutions.

Slavery also has the characteristics that would from the theory lead us to anticipate a constitutional ban against it – large expected utility losses for those prevented from acting, and relatively small gains from those benefiting from slavery. Why then did the U.S. Constitution originally fail to ban slavery? The obvious answer is that there was no uncertainty among those who wrote and ratified the Constitution about their ever becoming slaves in the future. Uncertainty about future position – real or self-imposed – is an essential element in the two-stage theory of constitutions.

McGuire and Ohsfeldt (1986, 1989) and McGuire (1988) have had some success in explaining voting at the Philadelphia Convention and at the ratifying conventions in terms of the self-interest of the participants. Although some of their interpretations of the data can be questioned (Mueller, 1996a, pp. 62–3), they provide convincing evidence that the Founding Fathers of the United States did not suppress all aspects of narrow self-interest when writing and ratifying the Constitution. Unfortunately, constitutional conventions are such rare events that empirical testing of hypotheses about voting on constitutional provisions is likely to remain an “infant industry” in the public choice field.<sup>23</sup>

#### *Bibliographical notes*

The number of papers that have implicitly adopted the constitutional stage decision as a point of reference is large. See, in particular, Rae (1969); Mueller (1971, 1973, 1996a); Mueller, Tollison, and Willett (1974a,b, 1976); and Abrams and Settle (1976). The field of constitutional political economy now has its own journal, *Constitutional Political Economy*, and the entire 90th volume of the journal *Public Choice* (March 1997) was devoted to the topic.

James Buchanan has expounded upon and defended the two-stage constitutional approach to public choice in numerous essays down through the years. A good sampling of these is contained in the two volumes published in 1986 and 1991. The Liberty Fund is in the process of republishing all of his writings. Riley (2001) presents an excellent analysis of the two-stage approach to constitutions.

Coleman (1988) critically discusses constitutional political economy from a legal perspective. Cooter (2000) applies concepts from game theory to the analysis

<sup>23</sup> For an innovative attempt to test propositions about constitutional design using data on the democratic rules used in condominiums, see Sass (1992).

of constitutional institutions. Ferejohn, Rakove, and Riley (2001) contains several interesting essays on constitutional issues.

Beard's (1913) *An Economic Interpretation of the Constitution* of the United States might well be regarded as a, if not *the*, pioneering contribution to both public choice and constitutional political economy. Certainly, this book can be characterized as "constitutional political economy without romance" to paraphrase a characterization of public choice that James Buchanan has often made. Beard clearly anticipated some of the hypotheses and results of McGuire and Ohsfeldt (1986, 1989) and McGuire (1988), and his book shares with all "economic theories of politics" a certain cynicism about individual motivations and their effects on political outcomes. But Beard's work has been essentially ignored by students of both public choice and constitutional political economy.

Voigt (1997, 1999) has been one of the leading proponents of the development of the positive dimension to constitutional political economy.