Valuing Exit Options
Author(s): Jenna Bednar
Source: *Publius*, Vol. 37, No. 2, Federalism and Formal Theory (Spring, 2007), pp. 190-208
Published by: Oxford University Press
Stable URL: http://www.jstor.org/stable/4624790
Accessed: 21/10/2014 10:10

Valuing Exit Options

Jenna Bednar*

This article examines an important aspect of federalism: the effect of a secession threat on the union's productivity. Productivity requires a compliance maintenance regime with credible punishment. An exit option gives a government the alternative of opting out of the union rather than suffer the disutility of a punishment. Equilibria are characterized over a continuous range of exit option values. The results indicate that only exit options that are superior to union membership improve utility; those of moderate value *decrease* net and individual government utility due to their harmful effect on compliance maintenance. A prescription that emerges from these results is that if the exit option is inferior to the benefit from a thriving union, member governments should voluntarily submit to measures that make exit as costly as possible.

In federalism, we observe no perfectly harmonious unions; instead, even the most stable—the United States since 1865, Switzerland—are characterized by near-constant quibbling and, periodically, more serious disputes. Others, such as Canada, seem to be perennially at the brink of rupture. An emerging body of work studies the institutional design supporting effective federations (Filippov, Ordeshook, and Shvetsova 2004; Volden 2005; de Figueiredo and Weingast 2005; Bednar 2007). In this article, I take up a special aspect of federalism's problem: the effect of an exit option on the productivity of the union and the utility of the governmental members.

A significant literature describes the opportunity provided by the exit option, generally coming to the conclusion that exit options are beneficial for those holding them. They substitute for voice (Hirschman 1970) by being an option to use instead of within-system protest; without contradiction, they also increase (complement) voice (Hirschman 1993, Gelbach 2005, Clark, Golder, and Golder 2006) by improving the threat point or bargaining position. In analyses of decentralized systems, exit options lead to subnational gains because the subnational government is able to extract a greater distributional allocation from the center (Triesman 1999; de Figueiredo and Weingast 2005). In general,

*University of Michigan

Publius: The Journal of Federalism volume 37 number 2, pp. 190–208
doi: 10.1093/publius/pjm001
Advance Access publication 26 March 2007

the better the outside option, the better off the government. That is, exit options improve utility.

A contrasting view builds on Schelling's (1960) insights about the value of commitment to make threats credible. Many solutions to collective action problems require a joint commitment to punish, which becomes a collective action dilemma of its own (Axelrod 1986; Greif, Milgrom, and Weingast 1994). Despite vows to react punitively, if called upon to do so, an agent prefers not to punish if the punishment is costly. If an agent can exit the system rather than punish, the commitment to punish is broken. Unless an agent can precommit to follow through, a threat has no deterrent effect on others capable of foresight. Relatedly, Elster (1979/1984) invokes the image of Ulysses binding himself to his ship's mast to resist the sirens' temptations; by eliminating future options, precommitment mechanisms can overcome the problems that are introduced by changing preferences (for example, due to addiction or an altered state of mind). This second literature focuses on the importance of punishment for joint production. Exit, while not explicitly considered, disrupts the punishment commitment.

In this article, I bring the cautions of the second literature to bear on the study of exit options. If federalism is portrayed as a distributional challenge—the allocation of divisible goods between member governments—then the application of Hirschman and related insights is fitting. But if we consider federalism to be a problem of compliance because its productivity (including the generation of divisible goods) is dependent upon the willingness of governments to punish one another, then exit options have a second effect that may trump the first, imperiling even a mutually beneficial union. This article will show that it is not sufficient for the value of the union to exceed the value of exiting for a state to prefer federal membership to independence. Or, more precisely, we need to specify the value of the union more particularly than expected return when all member governments comply. At the moment of punishment, a Member State may choose to exit rather than incur costly punishment. Therefore, the extent of punishment possible must take exit options into account. The result is that punishments are minimized. But with less severe punishment, high compliance cannot be maintained. The union is less productive and utility from membership drops. A mediocre exit option—one less beneficial than a productive union—makes the optimal productive union untenable. Possession of a mediocre exit option results in a loss of utility.

The article is structured in four sections. First, I define federalism as a compliance problem. Second, I develop a base model of compliance, generating results about the optimal punishment and compliance levels, and making inferences about the effect of an exit option. Third, with a more general model we see the full effect of exit options on compliance and utility. In the fourth section, I conclude.

Federalism's Compliance Problem

A political community may support federalism for a variety of reasons; broad categories include military security, economic efficiency, or improved representation (Bednar 2007). The assignment of authority between federal and state governments is a design problem that depends upon the community's specific weighting of its objectives, since some goals require trade-offs (some demand greater centralization, some greater decentralization). While the optimal structure of every federation will be different—each according to its unique priorities—in each federation, the constitutional division of authority is set to achieve these objectives. The productivity of the union—its ability to meet its goals, and therefore the return to each member government—is a function of the degree of compliance with the constitutional rules allocating authority.[1]

While legislation itself is observable, many factors confound the ability of others to determine compliance, including the multi-dimensionality of policy space, the imperfect mapping of specific policies to constitutional requirements, the multiplicity of legislation that contributes to outcomes, and the more invisible execution of policy. We will focus on actions where overcompliance is not possible. This assumption excludes domains such as troop requisitions, where it is possible that a state could send more troops than federally requested, but includes most domains of intergovernmental activity, such as trade barriers (where overcompliance would imply privileging out-of-state firms over one's own) or activities that generate negative externalities, such as pollution controls, where it is not physically possible to eliminate pollution created by other states.

Imperfect information and costly compliance motivate opportunism. To prevent the federation from falling apart, a variety of safeguards are available: *structural*, fragmenting authority; *citizen/electoral*, where transgressing governments are punished by the voters; *judicial*, through constitutional review; and *political*, where an integrated party system polices politicians within the structure of an interdependent party organization. Although these safeguards have an advantage of targeting a particular government and so are virtually costless for those not punished, none of these safeguards are singly sufficient to sustain the union, as the punishment that each can inflict is too mild to deter major transgressions (Bednar 2007). While the above list of four safeguards may or may not be present in any particular federation, and quite often if present do not function optimally, a fifth safeguard is present in every federation: the capacity of member governments (federal and state) to punish one another's transgressions through retaliatory noncompliance, with realizations ranging from trade barriers to civil wars. While in principle, intergovernmental retaliation might be targeted and its punishment effect exactly calibrated, in practice, its effects are both scattered and may trigger escalation. There is a reason why unions that depend exclusively on intergovernmental retaliation for maintenance (such as the United States under the Articles of Confederation) tend to underperform. Adding the mild safeguards allows a fine-tuning of the compliance maintenance mechanisms. Nonetheless, the intergovernmental retaliation remains necessary to ward off major constitutional breaches. Any compromise to its efficacy will injure the union.

In this article, I show how the utility the members get from the union depends not just on the internal efficiency of the union but also upon the participants' outside options. Stability and the union's productive capacity are both affected by exit options. I show how mediocre exit options—ones that the participants would not choose over membership in the union under normal circumstances—detrimentally affect the productivity of the union, and as the exit option improves, but remains below the offerings of a productive union, either no union is possible, or it will break apart at the first bit of bad luck.

A Base Model of Intergovernmental Compliance

The model will treat the federal union's performance as a function of the extent that each member government complies with the distribution of authorities. As an example, consider the value of a common market. Free trade between Member States is mutually beneficial, and the benefits increase with the number of states that participate. But compliance is costly: each state would like to protect its own industries, even as it enjoys free access to the markets developed below.

We will assume that there is a cost for the government to comply, and therefore it prefers to shirk if it believes it will not be punished or if others will shirk as well. The benefit to each government—its utility—is a function of the actions of other governments, its own noncompliance, and any sanctions. Observations of the actions of other governments are imperfect, reflecting the informational challenges described above; judgments of actions are based upon an imperfect signal received by others, which may be a measure of productivity or harm.[2] To induce compliance, we consider trigger mechanisms. Not meeting the target triggers punishment. Because our interest is to explore the effect of exit on the federation's most severe safeguard, intergovernmental retaliation, we do not add any other constitutionally introduced trigger mechanisms. In general, a trigger mechanism has three components: a threshold, which it compares the signal against, a punishment, and a duration of the punishment. In the base model, we collapse the last two elements into a single parameter, a fine, so that the punishment is realized as a utility loss. I eliminate this simplification in the extended model developed below.

To simplify the presentation, we consider a linear utility function but include curvature in the probability of being punished.[3] Actions are expressed as non-compliance, a continuous variable $x \in [0, 1]$, so that as x increases, opportunism is greater, and it is bounded on the lower end at full compliance, keeping with our assumption that overcompliance is not possible. With costly compliance (it involves sacrifice), each government prefers not to comply, but benefits from the compliance of others. Specifically, here we assume that each government's utility is a fraction α of the sum of the contributions to the federation plus whatever that member holds back. We assume $\alpha < 1$; otherwise no collective action problem exists. Given these assumptions, the single period utility to member 1 from its contributions and those of others (governments 2 through n) can be written as follows:

$$u_1(x_1, x_2, \ldots, x_n) = \alpha \sum_{i=1}^{N} (1 - x_i) + x_1. \tag{1}$$

Although the governments cannot monitor the levels of compliance of one another, they can infer it, subject to uncertainty, by evaluating the signal received about one another's activity. To model uncertainty, we will assume that the members of the federation see a common signal ω as an indicator of the amount of noncompliance. The value of ω is a function of the sum of the noncompliance of the members plus a noise term ϵ, $\omega(\sum x_j, \epsilon) = \theta(\sum x_j) + \epsilon$, where θ is an increasing function. Given this formulation, the noise term could increase or decrease the governments' perceived noncompliance. We assume here that ϵ takes some value between some minimum $\min_\epsilon$ and maximum $\max_\epsilon$ according to a probability distribution F, which has an associated density f.

If the signal exceeds a threshold, T, governments punish by engaging in retaliatory opportunism. It is important to keep in mind that the punishment regime is not the same as withdrawal, or secession from the union. Players still "participate," but they comply less in order to punish one another with the aim of deterring non-compliance.[4] To avoid confusion, we write $Pr(\omega > T)$ to denote the probability that ω exceeds a threshold, T. That is, punishment occurs if $\omega(\sum x_j, \epsilon) > T$. Therefore, the probability of punishment equals one minus the probability that ϵ is less than $T - \theta(\sum x_j)$, expressed formally as $1 - F(T - \theta(\sum x_j))$. In this base model, we will assume that the utility during the punishment phase is Q, strictly less than the utility during normal play.

The threshold T and the bounds on the error term (its minimum and maximum) determine the range of the governments' aggregate behavior, $\sum x_j$ (at least in all interesting cases). Specifically, $\theta(\sum x_j)$ will lie between $T - \max_\epsilon$ and $T - \min_\epsilon$. To see why, let us consider the alternatives. Suppose that $\theta(\sum x_j) < T - \max_\epsilon$. In this case, there is no possibility of punishment: even the worst possible error term will not put the signal ω above the threshold. Therefore, the governments have an incentive to be more opportunistic. Alternatively, if $\theta(\sum x_j) > T - \min_\epsilon$, then no matter what the error term is, punishment is certain. Here too governments might as well be more opportunistic because they maximize immediate gain knowing that punishment is on its way anyway. Thus, in equilibrium, we can assume that $\theta(\sum x_j) \in [T - \max_\epsilon, T - \min_\epsilon]$.

To complete the model, we need only include a discount rate δ. We restrict attention to stationary, symmetric equilibria in which each member makes the same contribution in each period. This allows us to write a recursive equation that gives the value for member i (the present discounted sum of utilities, denoted V_i):

$$V_i(x_1, x_2, \ldots, x_n) = \alpha \sum_{j=1}^{N} (1 - x_j) + x_i + \delta[Pr(\omega > T)(-Q) + V_i(x_1, x_2, \ldots, x_n)] \quad (2)$$

Solving for V_i produces:

$$V_i(x_1, x_2, \ldots, x_n) = \frac{\alpha \sum_{j=1}^{N} (1 - x_j) + x_i - \delta Pr(\omega > T)(Q)}{1 - \delta} \quad (3)$$

Government i chooses x_i to maximize this value. To solve for the first order condition, we can exploit the fact that $Pr(\omega > T) = 1 - F(T - \theta(\sum x_j))$. Taking the derivative of this new expression gives the marginal value of opportunism.

$$\frac{1 - \alpha - \delta f(T - \theta(\sum x_j))\theta'(\sum x_j)Q}{1 - \delta} \quad (4)$$

Recall that f is the density function associated with the probability distribution F. If full compliance were an equilibrium, then this expression would equal zero when $\sum x_j = 0$. We may assume that small deviations are very difficult to notice, implying that $\theta'(0) \sim 0$. In other words, at full compliance the marginal increase in the signal from a slight deviation is approximately zero. Given that assumption, the above expression cannot equal zero at full compliance [because $\alpha \in (0, 1)$, a basic assumption of collective action problems]. Therefore, in equilibrium some noncompliance occurs.[5]

The intuition behind the slippage result is straightforward. A constitutional division of powers, and therefore federalism, must be self-sustaining; it cannot rely upon outside forces to maintain it. Any self-enforcing relationship depends upon players doing what is in their best interest. That is, results focus on utility calculations (payoffs to players) rather than compliance maximization. And, as we will see, full compliance, in general, does not provide players with as much utility as partial compliance. The cost of maintaining full compliance, in terms of the frequency of punishment, makes member governments prefer to tolerate a little indiscretion rather than seek perfection.

While this is sufficient to show inherent opportunism, by making further assumptions about the distribution of the error term and the functional form of θ which enters into the signal, we can derive comparative statics. Our modification is to assume that the error term is uniformly distributed—that it is equally likely to assume any value between $-m$ and m. This implies that $f = 1/2m$. Recall the realistic assumption that small deviations are less likely to be caught than larger ones. One way to capture this mathematically is to assume that $\theta(x_1, x_2, \ldots, x_n) = (x_1 + x_2 + \cdots + x_n)^2$. Recall that $\omega(\sum x_j, \epsilon) = \theta(\sum x_j) + \epsilon$. What the governments observe about one anothers' behavior is a combination of signal and noise: with this assumption that squares the sum of the noncompliance, small degrees of opportunism produce a very small signal to noise ratio, and therefore the deviation cannot be detected. The marginal value of opportunism can now be written as follows:

$$\frac{1 - \alpha - \delta(1/2m)2\left(\sum_{j=1}^{N} x_j\right)Q}{1 - \delta}$$

Notice that the noise term, ϵ, drops out of the derivative because its effect is independent of the action taken by the governments. In this formulation small amounts of opportunism pay, but the marginal value of opportunism falls as the degree of opportunism increases. In equilibrium, the marginal value of opportunism will be exactly zero. Setting the previous expression equal to zero gives the following symmetric equilibrium level of noncompliance:

$$x_i^* = \frac{m(1 - \alpha)}{N\delta Q} \tag{5}$$

Note that noncompliance does not go to zero, no matter how high the discount factor (as long as it is not greater than one!). This contrasts with linear folk theorem results where full compliance is a possibility. Likewise, even very high punishments (high Q) do not eliminate all opportunism. Again, as long as $\alpha < 1$, opportunism is inherent. Slippage—noncompliance—is unavoidable.

We are also now able to compute comparative statics, or the effect of altering parameters on governmental behavior. Notice that the higher α, δ, N, and Q, the less slippage: opportunism decreases as the return on compliance increases, as patience increases, as the number of governments increases,[6] and as the penalty increases. On the other hand, opportunism increases as the range of the noise term increases (m). These are all intuitive results.

A natural question to ask is whether this slippage is meaningful. As constructed, the federal problem has two sources of disutility: (i) punishment periods to maintain compliance incentives, even though no one deviated from the equilibrium, and (ii) the loss due to the equilibrium noncompliance. In the example following, the second source is a greater source of disutility than the first.

A Numerical Example

We can plug in specific numerical values for each of the parameters to help us see the relative effects of slippage and incentive preserving punishments. Suppose that we have a federation of two states plus a central government (so $N=3$) and that the function θ squares the non-compliance, so that $\omega(x_1, x_2, x_3, \epsilon) = (x_1 + x_2 + x_3)^2 + \epsilon$. We will assume that the random error term is uniformly distributed between -1 and 1 so that $m=2$. We further assume that the discount rate, δ, equals 9/10, the return on collective action α equals 7/10, the punishment for triggering the threshold Q equals 2, and that the threshold T equals 1. To solve for the equilibrium, we need only plug these values into equation 5 to find the equilibrium level of opportunism:

$$x_i = \frac{(2)(3/10)}{(3)(9/10)(2)}.$$

Thus, $x_i = 1/9$. Each member contributes 8/9. This creates a utility loss. If each member contributed fully (and if we ignore punishment), each member would get utility of 2.1 $(7/10 \times 3)$ each period. However, owing to opportunism, each member only gets $7/10(8/9)(3) = 56/30$, a per period loss of utility equal to 7/30, or $\sim$0.233.

We can compare this loss from opportunism to the cost of the punishment regimes. To determine the probability of a punishment regime, we first calculate the likelihood that the signal ω exceeds the threshold T (set at 1). From above, we have that $\omega = (1/9 + 1/9 + 1/9)^2 + \epsilon$. This exceeds 1 if and only if $\epsilon > 8/9$. Given our assumptions, this occurs with probability 1/18.[7] If we add in our assumptions that $\delta = 9/10$ and $Q=2$, we get that the expected per period loss due to incentive preserving punishment equals (1/18)(9/10)(2), which equals 1/10.

Thus, in this numerical example, the utility loss due to opportunism (0.233) exceeds the loss due to incentive preserving punishment (0.1). This need not always be the case; if the probability of punishing or the cost of the punishment were to increase, the losses due to opportunism would decrease, and the losses dues to incentive-preserving punishment would increase. At some point, the losses due to opportunism would become smaller than the losses due to incentive-preserving punishment.

To evaluate the effect of an exit option it will be useful to understand the importance of the punishment mechanism on the level of compliance, and therefore utility.

Claim 1 *Given the assumptions of the model, the utility during normal play from being in the federation increases in Q, the amount of the punishment.*

Proof: See the appendix.

Kreps (1990, 517–21), who considers a numerical example with a different functional form, finds a similar result: larger punishments lead to higher utility.

With this baseline intuition established (inherent opportunism, compliance increases in Q, and utility increases in Q), the federalism context invites us to consider an extension missing from more generic analyses of public goods games. In a federation, member governments have the option of exiting the union. Assume that if a participant chooses to exit (the equivalent of not playing the game at all), its expected utility is W, a parameter that is exogenously determined and commonly known. In this symmetric game, $W_i = W_j \ \forall\ i, j$. The decision to exit is permanent. If a player exits, assume the payoff to the remaining participants is $\beta \geq 0$, forever. The withdrawal payoff, W, defines a participation constraint. We can now discuss when that constraint will bind, and the effect of having an exit option on the overall productivity of the union. Let Q_p be the amount of punishment that maximizes utility at the start of the punishment period.

Claim 2 *Given the assumptions of the model, the utility from being in the federation at the start of the punishment regime increases up to* Q_p *and decreases thereafter, where*

$$Q_p = \sqrt{\frac{\alpha N\psi - \psi + \delta\psi^2}{1-\delta}}$$

Proof: See the appendix.

Figure 1 helps to illustrate the results to this point. The severity of punishment varies along the x axis and the y axis is utility. Two curves are graphed: V_n represents expected utility from the union during normal play and V_p is expected utility at the start of a punishment. Claim 1 shows that utility increases asymptotically in Q, the severity of punishment, as represented by V_n. Existing treatments of exit options consider them to be a participation constraint, binding if the value of the exit option exceeds the value of the union. However, this claim concerns only the value of the union in normal play—when no punishment has been triggered. When no exit is available, it is the only value that is calculated. In equilibrium, all governments deviate (transgress) slightly, a minor transgression that is tolerated as unfixable given information constraints. But from time to time, stochastic error causes the observed behavior to cross a safeguard's threshold, triggering the punishment regime. When no exit is available (or when it is so low that it is not worth considering), the punishment can be optimally severe to deter transgressions. But with the potential for exit, a government may decide to leave

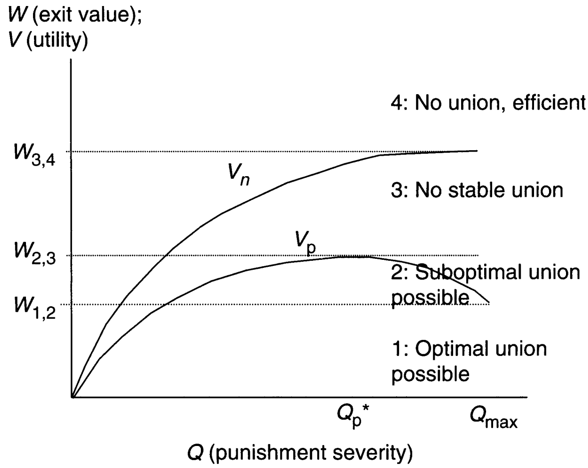

Figure 1 Punishment's effect on utility.

the union rather than suffer the punishment. Claim 2 established that there is an optimal *finite* Q_p* that maximizes utility at the start of the punishment regime. If the punishment is lessened to Q_p*, the equilibrium compliance level during *normal* play is less, as seen by the plotting of V_p.

We can now compare the implications of differently valued exit options. Lines graphed at $W_{1,2}$, $W_{2,3}$, and $W_{3,4}$ represent cut points in exit option utility. (In the next section, I solve for these cut points in terms of V.) These cut points define areas, labeled 1, 2, 3, and 4. Looking at figure 1, we see that in Areas 1 and 4 the exit option has the intuitive effect: if it is especially low, it does not affect the union or utility; if it is especially high, no union is possible, as one would expect. But in Areas 2 and 3, its effect as a participation constraint is more subtle, to the detriment of the union (in Area 3) and utility. We can see the effect duplicated in a more general model.

A More General Model

The mechanics of this extended model are identical to the base model except that equilibrium strategies now define two levels of compliance: that played during *normal* periods, when compliance is maximized, and that played punitively, in the *punishment regime.* Rather than triggering a fine, observations that cross the threshold trigger reactionary non-compliance, where all governments reduce their level of compliance (increase x) for a finite, known number of periods. This modification more closely resembles intergovernmental retaliation.

Let $\mathbf{S}$ denote the set of equilibrium strategies. Given $s \in \mathbf{S}$, we can define $V_n(s)$ to be the expected utility in normal play and $V_p(s)$ as the expected utility at the start of the punishment regime. Let s^* be the strategy that maximizes $V_n(s)$ and $\hat{s}$ be the strategy that maximizes $V_p(s)$. We focus on finite period punishment strategies. We can write a strategy as $s = (x, z, k)$ where x is the extent of noncompliance in normal play, z is the extent of selfish play[8] during a punishment regime, and k is the number of periods of punishment, where $x < z$. In the following claims it will be helpful to define the function $z(x, k)$ which equals the shirking in each of k punishment periods that would sustain a contribution level of x in normal play. We will assume that $z(x, k)$ is continuous and differentiable where defined.[9] We further assume that z is defined in an open neighborhood around x^*, which is just a technical way of saying that you could support higher levels of cooperation in normal play, but it would not be efficient to do so. That is, slippage exists, as stated earlier.

Claim 3 $V_n(s^*) > V_n(\hat{s})$: *The contribution that maximizes expected utility at the start of the punishment regime generates does not maximize normal play expected utility.*

Proof: Let $s^* = (x^*, z(x^*, k))$. For any fixed k, once a sustainable normal period noncompliance x is selected, it determines a unique $z(x, k)$. It follows that z is strictly decreasing in x: as the targeted behavior admits more noncompliance, so x grows, it is easier to sustain, so the punishment regime play, z can be relaxed, with less severe consequences. Therefore, with a fixed k, we can define the payoff at the beginning of normal play solely as a function of x: $V_n^k(x)$. Since x^* maximizes $V_n^k(x)$ and since $z(x, k)$ is defined in the neighborhood of x^*, it follows that $V_n^{k'}(x^*) = 0$.

Again keeping k fixed, we can also define V_p as a function of x.

$$V_p^k(x) = (1 + \delta \cdots + \delta^{k-1})U(\mathbf{z}(x, k)) + \delta^k V_n^k(x)$$

Taking the derivative with respect to x,

$$V_p^{k'}(x) = (1 + \delta \cdots + \delta^{k-1})U'(\mathbf{z}(x, k))z'(x, k) + \delta^k V_n^{k'}(x)$$

at $x = x^*$, $U' > 0$, $z' < 0$, and $V_n^{k'} = 0$ which implies that the derivative is negative. Therefore, x^* cannot maximize V_p^k. If V_n is concave, then $\hat{x} < x^*$. □

Corollary 1 If V_n *is concave,* $\hat{x} < x^*$.

Proof: See proof of Claim 3.

Claim 4 $V_p(s^*) < V_p(\hat{s})$: *The strategy that maximizes normal play expected utility does not maximize expected utility at the start of the punishment regime.*

Proof: The proof follows the same logic as that for Claim 3.

Claim 5 *We can fully characterize the equilibria as a function of W. (1) For* $W < V_p(x^*)$, *contribute* x^* *in normal play. Expected utility is not affected. (2) For* $V_p(x^*) < W < V_p(\hat{x})$, *contribute* $\bar{x}$, where $\hat{x} < \bar{x} < x^*$, *for a utility loss.*

Table 1 Union viability and utility change for varying exit option utilities

Case	Range	Contribution	E.U.	Utility Change	Union?
1	$W < V_p(x^*)$	x^*	$V_n(x^*)$	None	Yes
2	$V_p(x^*) < W < V_p(\hat{x})$	$\bar{x}$	$V_n(\bar{x})$	Decreases	Yes
3	$V_p(\hat{x}) < W < V_n(x^*)$	n/a	W	Decreases	No
4	$V_n(x^*) < W$	n/a	W	Increases	No

(3) For $V_p(\hat{x}) < W < V_n(x^*)$ players do not participate in the union, for a utility loss. (4) For $V_n(x^*) < W$ players do not participate in the union, for a utility gain. In particular, the introduction of an exit option translates into a net utility loss for moderate exit option utilities where $V_p(x^*) < W < V_n(x^*)$.

Proof: Table 1 and figure 2 summarize the results, indicating any change to utility from having an exit option over the case where they do not exist.

1. $W < V_p(x^*)$: The participation constraint does not bind. The union is not affected by the withdrawal option. Governments have a strict preference to participate in the union. Any threat to use the withdrawal option to induce a higher payoff within the union is not credible. Furthermore, it cannot be sustained, as players could reduce the payoff from the punishment regime below the payoff from exit. For example, to induce compliance, participants may wage war on one another, a mechanism more costly—and therefore more effective—than simple noncompliance. Utility is unchanged.
2. $V_p(x^*) < W < V_p(\hat{x})$: The constraint binds. Governments choose the equilibrium contribution, $\bar{x}$ such that $V_p(\bar{x}) = W$. While the union can be sustained, it is not as productive as the union was without the exit option; the introduction of the exit option decreases equilibrium contribution levels. In the symmetric game, all players are strictly worse off by having the exit option.
3. $V_n(\hat{x}) < W < V_n(x^*)$: The constraint binds and makes the union impossible:[10] participants hit the logical limit of how much they can increase $V_p(\hat{x})$ by increasing the contribution during punishment periods. As in Case 2, the players' utility is strictly worse off by having the option to exit.
4. $W > V_n(x^*)$: No union is possible, as the highest expected utility from participation in the union is less than the players' utility from exiting. Overall utility increases.

Corollary 2: *If $V_n(\hat{x}) > V_p(\hat{x})$, then there is a discontinuous drop in utility at $W = V_p(\hat{x})$.*

Proof: See figure 2. As exit option utility increases from $V_p(x^*)$, to compete with the exit options requires an increase in the expected utility at the start of

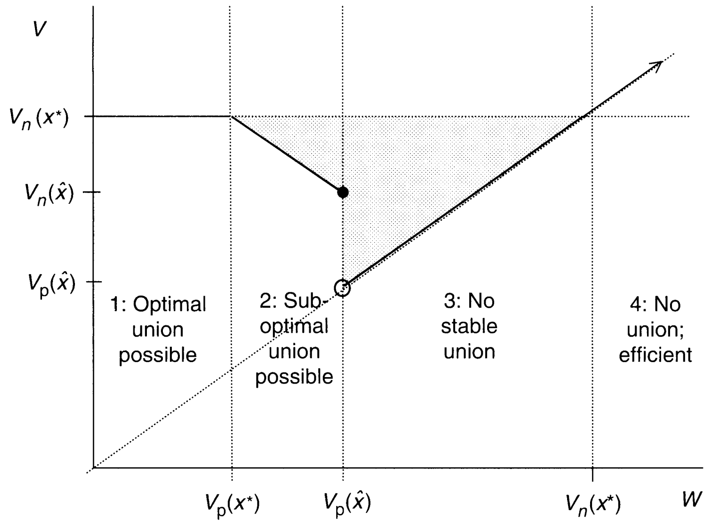

Figure 2 The effect of an exit option (W) on utility (V).

the punishment regime, so that m is increased. Since $\delta z/\delta x < 0$, increasing z implies a decreased normal-play level of compliance. This reduced compliance lowers the union's productivity, lowering utility. At $W = V_p(\hat{x})$, players cannot increase their punishment play expected utility any more, and use their exit option. In so doing, their utility drops discontinuously from $V_n(\hat{x})$ to $V_p(\hat{x})$. The shaded region represents utility loss from the exit option. □

As long as the value of the exit option is greater than the expected utility at the start of the punishment regime, the addition of the exit option affects utility and often, the sustainability of the union. Without any option to quit the game, the optimal contribution in equilibrium, while not at full compliance, is as close as possible. Contribution is sustained by making the punishment regime severe enough that the players want to avoid it. Armed with an option to exit the union, players first compare W to $V_n(s^*)$, the maximal expected utility of participation in normal periods. If the exit option exceeds this amount, naturally they will quit the union, as Case 4 describes. However, we see that even if the value of the exit option is less than the expected value of the union, the union might not be sustainable. As the exit option utility just surpasses the expected utility at the start of the punishment regime (Case 2), players might have some room to adjust the punishment regime utility upward by lessening the severity of the punishment. Players have many options for adjusting the expected utility at the start of the punishment period: they can reduce the severity of the single-period punishment (by complying more), they can reduce the duration of the punishment regime, or they can lower the threshold, thereby increasing the tolerance for noncompliance.

While any one of these modifications raises the expected utility at the start of the punishment regime, they all will lower the maximal utility obtainable during normal periods.[11] While these adjustments may rescue the union, they come at a cost: a less severe punishment induces less cooperative behavior from all members, reducing the utility from the union, and therefore reducing player utility. As Case 3 describes, players are limited in how much they can adjust the punishment utility. The presence of the exit option causes utility to decrease and the union to dissolve. Cases 2 and 3 are worrisome: governments lose utility as a result of having a moderately attractive exit option. We will return to these cases in the subsequent discussion.

Discussion

In every federation exit options exist. Absence of a legal right to secession does not preclude exit, but it may make it more costly. At times they may be so undesirable, or so costly to pursue, that they hardly register as "options". At other times, they are so fantastic that they foreclose any hope for union. But what effect does the exit option have when it is in the grey area, a mediocre to pretty good option? This article extends our understanding of the influence of exit options by showing that unless the option is so desireable that it is a "no-brainer" to leave, the presence of the exit option can only lower the utility of member governments.

The backbone of federalism's compliance mechanism is the credible potential for intergovernmental retaliation. Should one member government abuse its constitutional authority, and mild safeguards fail to deter it, other governments may retaliate with their own noncompliance. This study demonstrates the susceptability of intergovernmental retaliation to exit options. Retaliatory periods are costly for all governments; metanorms are necessary to ensure the credibility of mutual punishment. Exit options are a third alternative to punishment or nonpunishment, a way to avoid both the cost of punishment and the penalties of nonpunishment. The temptation to exit affects compliance in every contingency, although governments would only quit the federation in downturns. Therefore exit options are not universally beneficial, and may actually reduce the utility of those who possess them.

Exit options serve as a participation constraint in two ways. Existing analyses argue that the utility from being in the union must exceed utility derived from independence. As such, it is a participation constraint on *joining* the union. But when the union is self-enforcing, sustenance (and therefore utility) depends upon the compliance enforced by credible threats of mutual retaliation. This form of punishment is costly, and although milder mechanisms minimize its impact, the severe punishment inflicted by intergovernmental retaliation must exist for the union to be optimally productive. When no possibility of exit is present, optimal punishments may be sustained. But when exit is an option, a second utility calculation is made: the anticipated value of the union once punishment is triggered. At that moment—at the start of a punishment regime—expected utility is lower than in normal periods, and may dip below the value of exiting. Therefore, we can define a second participation constraint: that of *sustaining* the union. The punishment severity may be adjusted to boost expected utility at the start of punishment above the exit value, but this compromises compliance levels, reducing utility in normal periods. Therefore, a mediocre exit option—one that is strictly worse than the value of the union performing optimally—may still reduce utility.

At this point, institutional tinkering (beyond self-enforcing strategies) may help to sustain the union and prevent utility loss. For example, see Chen and Ordeshook (1994) for complementary analysis of constitutional secession clauses as a means to preserve federal unions by selecting between equilibria; the existence of a constitutional prohibition of secession may affect beliefs sufficiently to induce punishment of any subunit that attempts to secede. If collective commitment to punish secession attempts is credible, (overcoming the same metanorm challenge that made exit attractive!), then the cost of exit is raised, reducing the utility of it. See figure 2: As long as $W < V_n(x^*)$, utility is either increased or unaffected. Elster's insights are confirmed: Constitutional prohibitions on secession are at the least innocuous, and under conditions specified by Chen and Ordeshook, become a method of tying one's hands to improve utility.

Finally, we can return to the earlier debate about exit as either substitute or complement to voice. When voice is used as a mechanism to induce compliance, as in the federalism case, the relationship between exit and voice is conditional. In particular, in all cases where it is interesting to compare exit and voice—when the utility from the two are comparable—the presence of exit reduces the efficacy of voice, or the capacity to induce cooperation. Therefore, they are substitutes. We therefore confirm Hirschman's (1970) original intuition: exit reduces voice.

Notes

I am grateful to audience members at the 2006 Midwest Political Science Association Meetings and the 2006 American Political Science Association Meetings. I also thank, without implicating, my discussants at these conferences: Ken Kollman, Burt Monroe, and especially John Londregan. Finally, two terrific referees and Craig Volden provided useful comments to improve the article.

1. A unitary government does not suffer from the same compliance problem as a decentralized state, but also cannot capture the benefits of decentralization. Any prescription of federalism must weigh the costs of compliance maintenance, and its feasibility, against anticipated advantages. For some states (Afghanistan is one) the attractions of federalism are dwarfed by the high cost of maintaining intergovernmental cooperation and compliance.

2. In other contexts, examples of imperfect signals include the price of oil as an aggregate signal of the oil production of each member of a cartel (Green and Porter 1984), while in Gilligan and Krehbiel (1987, 1989) the policy realization or outcome is a signal of the intent of the committee that authorized the policy, and would be indexed to indicate the policy source.
3. See Bednar 2006 for a full treatment of both cases and general proofs of the claims that follow.
4. Abreu, Pierce, and Stacchetti (1990) show that efficient equilibria rely on grim trigger strategies. While their result is proven for a finite action space, and so does not translate directly to this article's infinite action space, the intuition is equivalent.
5. Models that sustain full compliance assume linear signals and uniform distributions of the error term. Given these strong assumptions, it is easy to see that full compliance is possible. The first order necessary condition becomes:

$$\frac{-\alpha + 1 - \delta Q}{1 - \delta}.$$

If Q and δ are sufficiently large then this expression is negative, which implies that at the margin opportunism reduces the member's value. This aligns with the intuition from the repeated prisoners' dilemma. If punishment is sufficiently severe, no deviation need occur.

6. Increasing the number of participants increases the benefits of cooperation. It is possible that it would also increase monitoring costs, an effect that would encourage opportunism. The model could be amended to make uncertainty a function of the number of government members.
7. Given that $m = 2$, the distribution is uniform with a value of 1/2. Therefore, the probability that ϵ lies in an interval of length x equals $x/2$.
8. Since z is invoked as part of an equilibrium to induce compliance, it is normatively not accurate to call it transgression or shirking.
9. If no k-period punishment regime would sustain x, then z is not defined.
10. The game-theoretic prediction is for no union. However, if we allow for cognitive constraints, this region may explain cases of unions that break apart at the first sign of trouble.
11. Only an increase in the level of patience could make the wait worthwhile, and in general, we assume that the discount parameter is exogenously determined.

References

Abreu, Dilip, David Pearce, and Ennio Stacchetti. 1990. Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica* 58 (5): 1041–1063.

Axelrod, Robert. 1986. An evolutionary approach to norms. *American Political Science Review* 80 (4): 1095–1111.

Bednar, Jenna. 2006. Is full compliance possible? Conditions for shirking with imperfect monitoring and continuous action spaces. *Journal of Theoretical Politics* 18 (3): 345–373.

Bednar, Jenna. Forthcoming 2008. *The Robust Federation.* New York: Cambridge University Press.

Chen, Yan, and Peter C. Ordeshook. 1994. Constitutional secession clauses. *Constitutional Political Economy* 5 (1): 45–60.

Clark, William Roberts, Matt Golder, and Sona Nadenichek Golder. No date (2006). "Power and Politics: Insights from an Exit Voice and Loyalty Game." University of Michigan manuscript.

de Figuieredo, Rui, and Barry R. Weingast. 2005. Self-Enforcing Federalism. *Journal of Law, Economics, & Organization* 21 (1): 103–135.

Elster, Jon. 1979 (1984). *Ulysses and the Sirens: Studies in Rationality and Irrationality,* revised edition. Cambridge, UK: Cambridge University Press.

Filippov, Mikhail, Peter C. Ordeshook, and Olga Shvetsova. 2004. *Designing Federalism: A Theory of Self-Sustainable Federal Institutions.* Cambridge, UK: Cambridge University Press.

Gelbach, Scott. 2005. A Formal Model of Exit and Voice. University of Wisconsin manuscript.

Gilligan, Thomas W., and Keith Krehbiel. 1987. Collective Decisionmaking and Standing Committees: An Informational Rationale for Restrictive Amendments Procedures. *Journal of Law, Economics, and Organization* 3(2): 287–335.

Gilligan, Thomas W., and Keith Krehbiel. 1989. Asymmetric Information and Legislative Rules with a Heterogeneous Committee. *American Journal of Political Science* 33(2): 459–490.

Green, Edward J., and Robert H. Porter. 1984. Noncooperative Collusion under Imperfect Information. *Econometrica* 52 (1): 87–100.

Greif, Avner, Paul Milgrom, and Barry R. Weingast. 1994. Coordination, Commitment, and Enforcement: The Case of the Merchant Guild. *Journal of Political Economy* 102 (4): 745–776.

Hirschman, Albert O. 1970. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States.* Cambridge, MA: Harvard University Press.

Hirschman, Albert O. 1993. Exit, Voice, and the Fate of the German Democratic Republic: An Essay in Conceptual History. *World Politics* 45 (2): 173–202.

Kreps, David M. 1990. *A Course in Microeconomic Theory.* Princeton, NJ: Princeton University Press.

Schelling, Thomas C. 1960 (1980). *The Strategy of Conflict.* Cambridge, MA: Harvard University Press.

Treisman, Daniel. 1999. Political decentralization and economic reform: A game-theoretic analysis. *American Journal of Political Science* 43 (2): 488–517.

Volden, Craig. 2005. Intergovernmental Political Competition in American Federalism. *American Journal of Political Science* 49 (2): 327–342.

Appendix

Proof of Claim 1:

Recalling equation 5, the equilibrium amount of opportunism is given by the following expression.

$$x_i{}^* = \frac{m(1-\alpha)}{N\delta Q}$$

We will assume that $T = 1$ and that ϵ is uniformly distributed in [0, 1]. Setting $T = 1$ is optimal given this error structure because if $T > 1$, participants would have an incentive to be more opportunistic, thus lowering the benefits of begin part of the federation. And if $T < 1$, the probability of punishment would increase, but there would be no effect on the marginal probability of punishment. Therefore, we are solving for the optimal mechanism.

Recall that the value function equals

$$V_i(x_1, x_2, \dots, x_n) = \frac{\alpha \sum_{j=1}^{N}(1-x_j) + x_i - \delta Pr(\omega > T)Q}{1-\delta}$$

If we plug in our value for x_i we get the following expression for the value function:

$$V_i(x_1, x_2, \dots, x_n) = \frac{\alpha N(1-(m(1-\alpha)/N\delta Q)) + (m(1-\alpha)/N\delta Q) - \delta(m^2(1-\alpha)^2/\delta^2 Q^2)Q}{1-\delta}$$

To simplify the notation, let $\psi = (m(1-\alpha)/N\delta)$. We can rewrite the previous expression as follows:

$$V_i(x_1, x_2, \dots, x_n) = \frac{\alpha N(1-(\psi/Q)) + (\psi/Q) - \delta(\psi^2/Q)}{1-\delta}$$

We want to choose Q to maximize this expression. Taking the derivative with respect to Q gives the following expression.

$$\frac{(\alpha N\psi/Q^2) - (\psi/Q^2) + \delta(\psi^2/Q^2)}{1-\delta}$$

Since $\alpha N > 1$ by assumption, this expression is always positive. This implies that the higher is Q, the higher is utility from the union in expectation. □

Proof of Claim 2:

If we look at the utility to a participant at the beginning of a punishment regime, what we call $V_{\mathrm{p}}(i)$, we get the following expression.

$$V_{\mathrm{p}}(i)(x_1, x_2, \dots, x_n) = \frac{\alpha N(1-(\psi/Q)) + (\psi/Q) - \delta(\psi^2/Q)}{1-\delta} - Q$$

If we take the derivative of this expression we get

$$\frac{(\alpha N\psi/Q^2) - (\psi/Q^2) + \delta(\psi^2/Q^2)}{1-\delta} - 1$$

Setting this expression equal to zero gives

$$\frac{\alpha N\psi - \psi + \delta\psi^2}{Q^2} = (1-\delta)$$

Which if we solve for the Q_p that maximizes the pre-punishment regime payoff, we get

$$Q_p = \sqrt{\frac{\alpha N\psi - \psi + \delta\psi^2}{1-\delta}}$$

Note that the derivative increases up to Q_p and decreases thereafter. Thus Q_p maximizes V_p. □