ABSTRACT

This paper explores the causes of domestic conflict over foreign policy in the context of a dynamic learning model. One state is composed of two political parties that are uncertain about how another state will respond to cooperative and confrontational policies. To resolve this uncertainty, they must implement policies and update on the basis of the observed outcome. While both parties have in interest in learning whether or not mutual cooperation is possible, their constituents have different costs from conflict and hence different willingness to risk the cooperative gestures that are needed to obtain this information. The model shows that parties whose constituents have low costs for conflict are less likely to change their policy preferences in response to external events, while dovish parties have more fluid preferences. It also shows that under conditions of uncertainty, interstate conflict can lead to a polarization of party preferences, rather than consensus. I then explore these implications using data on party positions in Western democracies during the Cold War. The analysis shows that parties of the left were more likely to change their policy preferences over time than were parties of the right, and that periods of high conflict, such as the resumption of Cold War in the 1980s, generated polarization due to the different responses of left and right parties.
When and why does foreign policy become a subject of domestic political conflict? When do parties articulate and promote distinctive foreign policies and when is there consensus? This paper takes theoretical and empirical cuts at these questions. On the theoretical side, it presents a dynamic learning model in which the actors within a state attempt to learn the type of a rival in order to discover whether or not mutual cooperation is possible. The model yields some insight into how the preferences of domestic actors can evolve over time in response to signals from the international environment—in particular, the conditions that lead to domestic consensus and those that foster polarization. The model then inspires an empirical analysis of the structure of domestic conflict in Western democracies during the Cold War. I use estimates of parties’ position on a hawk-dove continuum to explore how different actors responded to changes in Soviet behavior.

The use of a learning model to explore these questions stems from a sense that uncertainty plays a large role in domestic contestation of foreign policy under conditions of high potential threat. This contrasts with a traditional view that it is the level of threat to a state’s interests that determines the level of agreement within a country. For realists, a serious threat to the state’s sovereignty and territorial integrity creates imperatives that overwhelm differences of interest or opinion at the domestic level. Thus, when the external environment is particularly dangerous, the common interest in acquiring security is particularly salient, and may overcome differences on other, secondary dimensions. It is when such threats lessen that disagreement’s over the national interest begin to grow. Thus, Arnold Wolfers asks us to consider the plight of individuals in a house that is on fire. These individuals may be very different from one another in their tastes, desires,
abilities, and so forth. And yet, for all their differences, they are likely to all share the same goal: to get out of the burning house. Indeed, we need know nothing more about each individual than that he or she wants to avoid death to come to the conclusion that all will adopt the same response their predicament. Now, contrast this situation to one in which the house is merely warm. In this case, we can no longer assume that all individuals will respond in the same way. Some may like the warmth and therefore do nothing; others may turn on a fan or take off a layer of clothing; still others may be so uncomfortable that they prefer to leave and find an air-conditioned spot. In this second scenario, individuals do not all converge on the same strategy because the core interest which they all share—survival—is not threatened. As a result, their behaviors are determined by other factors which vary from person to person.

Of course, as Jervis noted long ago, it is not always clear whether or not the house is on fire, and even if it clearly is, there is likely to be more than one exit (1976, 20). Even Wolfers acknowledges that it is often the case that “the ‘exits’ are not clearly marked” (1962, 15). Thus, actors within a state constantly face questions that do not have easy or obvious answers:

- Does Hitler have modest ambitions or does he seek world conquest?
- Does cooperating with the Soviet Union create the possibilities for an end to Cold War or does it open us up to exploitation?
- Will engagement of China lead to democratization and perpetual peace or is it simply strengthening a future adversary?
- Does Yasser Arafat want peace or is he committed to the destruction of Israel?
Under conditions when domestic actors share the same core interests in the preservation of the state’s sovereignty and territorial integrity, it is uncertainty about the nature of the threats faced and the means of dealing with them that opens up a space for domestic conflict over foreign policy. When such questions do not have clear answers, parties that represent different sets of constituents or hold different beliefs can come to promote different answers.

If this is true, then we can capture the link between uncertainty and domestic conflict in the context of a dynamic learning model in which uncertainty is changing (though not necessarily decreasing) over time in response to feedback. In the model, a generic state A has a choice between a confrontational or cooperative policy and then state B has the same choice. B’s response depends upon its type, which is unknown to actors in A. Based on this response, beliefs are updated and A then chooses again. State A is composed of two parties who represent constituents with different costs from international conflict. The two parties rotate in office on the basis of a retrospective evaluation of their performance in each period.

An important aspect of the model is that learning is active: what actors in A learn depends on the policy they implement. So, for example, in order to learn whether or not the other state will reciprocate cooperative gestures, the government in A must be willing to risk a cooperative action first. By confronting the other state, little is learned about how it would respond to cooperation. Hence, the actions the state takes influence not only its payoffs but also what it can learn in any given period. This aspect of the model is going to predict circumstances in which learning causes polarization of policy preferences because parties might agree that a confrontational policy is leading to bad
outcomes but disagree over whether the alternative would be better. Depending upon actors’ beliefs and preferences, a history of mutual conflict can lead some to advocate a switch to cooperation and others to hold steadfastly to confrontation. The former argue that the other state’s threatening behavior is a response to the confrontational policy of their own state and that cooperation would likewise be reciprocated; the latter argue that the other’s behavior stems from its inherent aggressiveness and that cooperation would therefore be exploited. Unless and until cooperation is actually tried, this kind of disagreement cannot be resolved.

The model has a number of implications, two of which can be probed empirically. The first is that parties that have high costs for international conflict are more likely to revise their policy preferences in response to external events than are parties with lower costs for conflict. While the latter tend to have a stable preference for confrontation under a wide set of conditions, the former are more likely to switch their preferred policies from confrontation to cooperation and, if the outcome of cooperation is bad, back again. The second prediction, as just described, is that under conditions of uncertainty, an increase in international tensions can cause polarization rather than consensus.

After laying out these results, I then explore the insights from the model using data on party positions in Western democracies during the Cold War. Using estimates of parties’ positions on a hawk-dove dimension using their election platforms, I look at how these positions changed over time in response to changes in Soviet behavior and the parties’ left-right ideology. Two results emerge from this analysis, corresponding to the two implications of the model:
1. Parties of the left were more sensitive to changes in Soviet behavior than were parties of the right. While the latter were consistently hawkish throughout the Cold War, the former changed directions on several occasions, especially in response to détente and the resumption of Cold War in the late 1970s.

2. After mid 1960s, parties of the left tended to become more dovish in response to increases in the level of Soviet conflict, while parties of the right did not change or became slightly more hawkish. As a result, détente was actually associated with a general consensus, while the resumption of Cold War in the late 1970s and early 80s was polarizing. This is the opposite of what Wolfers’ story suggests, but it is anticipated by the model.

This paper proceeds as follows. I first present a simple version of the learning model in which state A is a unitary actor. This permits a relatively straightforward introduction to some key intuitions of the model. Section 2 then introduces domestic politics by decomposing state A into two parties that represent constituents with different costs of conflict. This section discusses the additional implications that arise and focuses on the conditions that foster consensus and polarization of policy preferences. The third section then presents the empirical analysis.

1. The Basic Model with a Unitary Actor

The model presented here is an example of class of models known as “multi-armed bandits” (e.g., Rothschild 1974; Berry and Fridest 1985; Easley and Kiefer 1988; Gittins 1989; Banks and Sundaram 1992). The name derives from the image of a gambler trying to decide which of several arms to pull on a slot machine. The gambler is
uncertain about the payoff probabilities associated with each arm. The problem is to
determine an optimal search strategy that maximizes the gambler’s total returns. One
common application of these models is in the design of drug tests: How should a doctor
go about assigning different experimental drugs to patients given that he is unsure about
their effectiveness? In economics, the bandit model and others involving active learning
have been used to analyze the problem of a monopolist trying to learn the demand curve
for his product (e.g., Rothschild 1974; McClennan 1984). To my knowledge, the only
application of the bandit framework in political science is a working paper that considers
experimentation by government agencies (Ting 1999).

Assumptions

Consider an infinitely repeated decision problem. In each period, state A can
either cooperate with or defect against state B. After viewing A’s strategy, B can
likewise either cooperate or defect. Following standard convention, we define $R$ as the
reward for mutual cooperation, $P$ as the punishment for mutual defection, $T$ as the
temptation to defect unilaterally, and $S$ as the payoff for being “suckered” into unilateral
cooperation. I assume that state A’s preference ordering resembles that in a Stag Hunt
game, so that $R > T > P > S$ (e.g., Jervis 1978). With this preference ordering, state A
prefers to cooperate if it believes cooperation will be reciprocated, and it prefers to defect
if it believes that state B will do likewise. Such a state might be referred to as a
“conditional cooperator.” This assumption motivates state A’s desire to learn state B’s
type, since A’s optimal action in any period depends upon what it expects B to do.
Without loss of generality, we can normalize the payoff from mutual defection, \( P \), to zero.

State B is a strategic actor in only a very limited sense. We assume that its response to A’s strategy is governed by an unknown but fixed probability distribution. In particular, if A cooperates, B cooperates with probability \( s_C \) and defects with probability \( 1 - s_C \); if A defects, B cooperates with probability \( s_D \) and defects with probability \( 1 - s_D \).

The probabilities of cooperation are unknown and depend upon the type of state that B is. We assume that \( s_C \) can take either a high value, \( s_C^H \), or a low value, \( s_C^L \), with \( s_C^H > s_C^L \).

We similarly define \( s_D^H \) and \( s_D^L \) as the possible values of \( s_D \). The actual probabilities governing B’s behavior depend upon its type. We assume that there are three possible types:

1. A \textit{conditional cooperator} generally cooperates in response to cooperation and defects in response to defection, or \( s_C = s_C^H \) and \( s_D = s_D^L \).

2. A \textit{hard-liner} generally defects regardless of what A does, or \( s_C = s_C^L \) and \( s_D = s_D^L \).

3. A \textit{paper tiger} generally defects if A cooperates but cooperates if A defects, or \( s_C = s_C^L \) and \( s_D = s_D^H \).

The assumption that B’s strategy has a fixed probability distribution can be made rational by assuming that B is a myopic actor (or an infinite succession of short-lived actors) who can take on various types in any given period, that these types have pure strategy best responses to A’s strategy, and that \( s_C \) and \( s_D \) describe the “type generating function.”
While it would be nice to make B a fully strategic, non-myopic player, this enormously complicates the solution of the problem.

The distinction between these types is rather simple. If state B is a conditional cooperator, the optimal policy is to cooperate. If state B is a hard-liner or paper tiger, the optimal policy is to defect. The difference between these latter two types hinges on the expected costs of a confrontational policy. These costs are lower if B is a paper tiger than if it is a hard-liner, since the former is more likely to be compliant when confronted. A hard-liner, by contrast, is more likely to meet conflict with conflict.

State A knows the possible values of \( s_C \) and \( s_D \) but has inexact beliefs about B’s type. In particular, let \( p_0 \) denote A’s prior probability that B is a conditional cooperator, and let \( q_0 \) denote A’s prior belief that B is a hard-liner; clearly, \( 1 - p_0 - q_0 \) is the prior probability that B is a paper tiger. Although B’s type is fixed, its strategy in any given period is governed by a probability distribution; as a result, B’s strategy is a noisy signal of its true type. At the end of each period, A can update about B’s type using Bayes’ rule. Table 1 shows the formulas for \( p_{t+1} \) and \( q_{t+1} \) as a function of A’s strategy and beliefs and B’s response in period \( t \). An important feature of the updating process is that what state A learns in any given period depends upon the strategy it selected. When state A cooperates, hard-liners and paper tigers respond in the same way. As a result, a strategy of cooperation can help A distinguish conditional cooperators from the other two types but not hard-liners from paper tigers. Indeed, it is easy to show that, if A cooperates in period \( t \), then the probabilities it assigns to these two types do not change relative to one another, or
\[
\frac{q_{t+1}}{1 - p_{t+1} - q_{t+1}} = \frac{q_t}{1 - p_t - q_t}.
\]  

(1)

This is true regardless of how B responds. If B cooperates (defects), then \(p_{t+1}\) increases (decreases) and \(q_{t+1}\) shrinks (grows) by an amount that leaves the ratio in (1) unchanged.

Similarly, defection by A can help it distinguish paper tigers from conditional cooperators and hard-liners but leaves unchanged the relative probabilities of the latter.

Indeed, it is readily apparent from Table 1 that, if A defects in period \(t\),

\[
\frac{q_{t+1}}{p_{t+1}} = \frac{q_t}{p_t},
\]

(2)

again regardless of the outcome. Thus, in choosing strategies, A determines not only its payoff for that period but also what kind of information it will receive.

**Solution**

State A’s decision problem is to find a policy rule that maximizes the current discounted value of its payoff stream. Let \(\pi\) denote a policy function that determines, in each period \(t\), the policy that state A will select as a function of the history of the game to that point. Letting \(\delta \in [0,1]\) represent the policymaker’s per-period discount factor, the expected value of the game given the policy function \(\pi\) and initial beliefs \(p_0\) and \(q_0\) can be written as

\[
V_\pi(p_0, q_0) = \mathbb{E} \sum_{t=0}^{\infty} \delta^t r_t(\pi_t; p_0, q_0),
\]

(3)

were \(r_t\) is the payoff in period \(t\). The optimal policy function, \(\pi^*\), generates higher value than all alternative policy functions for all possible prior beliefs; that is

\[^1\text{To ensure that both of these statements are true, we must assume that } s^{**}_L > \frac{T_L - S}{R - S} > s^*_L.\]
$V_{\pi_0}(p_0, q_0) \geq V_{\pi_0}(p_0, q_0)$ for all $p_0$ and $q_0$. Let $V(p_0, q_0)$ denote the value associated with the priors $p_0$ and $q_0$ given that the state follows the optimal policy.

Because the game is infinitely repeated, the decision problem is identical in every period; only the beliefs change from one period to the next. Thus, $V(p_t, q_t)$ represents the current expected value of the game from any period $t$ onward given beliefs $p_t$ and $q_t$ and assuming that the state enacts the optimal policy in every period. Moreover, if a given policy rule is optimal at period zero, then it is also optimal in any generic period $t$. Thus, we can restrict our attention to a stationary policy function, $\pi(p_t, q_t)$, which determines the policy as a function of current beliefs. The stationarity of the policy function implies that a given set of beliefs will always lead to the same policy regardless of the period $t$; formally, $p_s = p_t$ and $q_s = q_t$ imply that $\pi(p_s, q_s) = \pi(p_t, q_t)$ for all $s$ and $t$. Standard results in stochastic dynamic programming ensure that an optimal stationary policy exists for this problem (e.g., Blackwell 1965, esp. Theorem 7b).

To understand what this policy function looks like, we need to decompose the value function. First, consider the expected payoff to cooperating in period $t$. This payoff has two components: the expected reward in period $t$, and the expected value of continuing the game with updated beliefs. The first part is easy to write down. Let $s^H_c(t) = p_t s^H_c + (1 - p_t) s^L_c$ denote the expected probability that B will cooperate in response to cooperation by A, given beliefs $p_t$. Then, the expected current reward of cooperating is given by

$$E[r_t(C; p_t, q_t)] = s^H_c(t)R + [1 - s^H_c(t)]S.$$ 

(4)
After cooperating, state A knows that it will update its beliefs. With probability \( s_c(t) \), B will cooperate, and A will update as in the upper left-hand box in Table 1; with probability \( 1 - s_c(t) \), B will defect, and A will update as in the upper right-hand box. Let \( p_{t+1}^{CC} \) and \( q_{t+1}^{CC} \) denote the update values of \( p \) and \( q \), respectively, given that both states cooperated in period \( t \). Similarly, let \( p_{t+1}^{CD} \) and \( q_{t+1}^{CD} \) denote the update values of \( p \) and \( q \), respectively, given that A cooperated and B defected in period \( t \). Assuming that it will continue to play optimally in the future, state A knows that the expected value of periods \( t+1 \) on is given by

\[
V(p_{t+1}, q_{t+1}) = s_c(t)V(p_{t+1}^{CC}, q_{t+1}^{CC}) + [1 - s_c(t)]V(p_{t+1}^{CD}, q_{t+1}^{CD}).
\]  

(5)

We can now combine (4) and (5) to determine the expected value of cooperating in period \( t \):

\[
V(C; p_t, q_t) = s_c(t)[R + \delta V(p_{t+1}^{CC}, q_{t+1}^{CC})] + [1 - s_c(t)][S + \delta V(p_{t+1}^{CD}, q_{t+1}^{CD})].
\]  

(6)

Identical reasoning (and analogous definitions) yields the expected value of defecting in period \( t \):

\[
V(D; p_t, q_t) = s_d(t)[T + \delta V(p_{t+1}^{DC}, q_{t+1}^{DC})] + [1 - s_d(t)][D + \delta V(p_{t+1}^{DD}, q_{t+1}^{DD})].
\]  

(7)

We can now re-write the total expected value as

\[
V(p_t, q_t) = \max\{V(C; p_t, q_t), V(D; p_t, q_t)\}.
\]  

(8)

The optimal policy function picks the strategy that maximizes (8) in each period \( t \) given the beliefs in that period.

The recursive form of (6), (7), and (8) makes an explicit solution of the optimal policy impractical for \( \delta > 0 \). We can, however, say a few things about what this policy looks like. First, consider the much simpler problem in which state A is purely myopic,
or $\delta = 0$. In this case, the expected value in each period is just the expected current reward. Thus,

$$V(p_t, q_t) = \max \{s_C(t)R + [1 - s_C(t)]S, s_D(t)T\}. \quad (9)$$

It is easy to show that the optimal policy in each period is to cooperate if and only if

$$p > \alpha - \beta q,$$

where

$$\alpha = [Ts_D^H - (R - S)s_C^L - S] / \gamma,$$ \hfill (10)

$$\beta = T(s_D^H - s_D^L) / \gamma,$$ and

$$\gamma = (R - S)(s_C^H - s_C^L) + T(s_D^H - s_D^L).$$

The myopic policy rule is depicted as a dotted line in Figure 1, which shows the possible combinations of $p_t$ and $q_t$ and the cutoff line implied by (10). For belief pairs that fall under the cutoff line, the optimal policy is to defect; for belief pairs that fall above the line, the optimal policy is to cooperate.

The optimal policy rule for $\delta > 0$ takes a similar form, but we cannot generate an explicit analytical solution because of the recursive form of the value function. Nevertheless, we can get a sense for what the optimal policy looks like through computational methods. We can derive the optimal policy rule computationally by exploiting the following result. Let $V^n(p_0, q_0)$ denote the value of the finite horizon decision problem consisting of only $n$ periods. A standard result is that $V^n$ converges uniformly to $V$, the value of the infinite horizon problem, as $n$ goes to infinity (Blackwell 1965; for a very accessible treatment, see Ross 1983, ch. 2). As a result, the optimal policy rule in a game with a large but finite number of periods is a good approximation of

\footnote{Note that the assumptions laid out in footnote 1 ensures that $\alpha$ falls between 0 and 1 and that the cutoff line hits the frontier $p = 1 - q$ at a positive value of $p$.}
the optimal policy rule in the infinite horizon problem. This is fortuitous, because we can compute the former relatively easily (see, e.g., Easley 1988). For each belief pair \((p, q)\), we determine the optimal policy as if it were the last period of the decision problem. Clearly, this is the policy that maximizes the current reward and is identical to the myopic policy described in (10). Let \(V^1(p, q)\) denote the expected value associated with each belief pair; this is given in (9). We can then determine the expected payoffs of cooperation and defection under the assumption that there are two periods left. From (6) and (7) we know that

\[
V^2(C; p, q) = s_c(t)[R + \delta V^1(p^{CC}, q^{CC})] + [1 - s_c(t)][S + \delta V^1(p^{CD}, q^{CD})],
\]

(11)

\[
V^2(D; p, q) = s_d(t)[T + \delta V^1(p^{DC}, q^{DC})] + [1 - s_d(t)]\delta V^1(p^{DD}, q^{DD}).
\]

(12)

The optimal policy in the second-to-last period is the one that delivers the higher expected value given the beliefs. Again, we let \(V^2(p, q)\) denote the value of the associated with the optimal policy—i.e., the maximum of (11) and (12) for each belief pair. Using the same technique, we can take another step back and consider the optimal policy if there were three periods left in the game, and so on. Because of the geometric discounting, the estimated value function changes less and less with each additional iteration. As a result, a modest number of iterations yields a policy function that is differs from the infinite horizon policy function by an infinitesimal amount.

One cost of this technique is that we cannot estimate the value or policy functions for all possible beliefs; instead, we can only do so for a grid of beliefs pairs. Thus, while the optimal policy function is smooth, we can only estimate a step function. Still, the approximation technique permits us to get a sense of how the optimal policy departs from the myopic one. The solid line in Figure 1 depicts the general form of the optimal cutoff.
While I have drawn the cutoff as a straight line, it need not be exactly linear, though it is monotonically decreasing. Recall that, for belief pairs above this cutoff, the optimal policy is cooperation, while for belief pairs below this cutoff, the optimal policy is defection.

Notice that the optimal cutoff falls beneath the myopic cutoff. This means that there are some beliefs for which it is optimal to cooperate even though the current reward from cooperation is lower than the current reward for defection. Moreover, the range of such beliefs becomes larger as $\delta$ increases—that is, the more state A values future payoffs. This result is due to the fact that mutual cooperation is the best possible outcome of the game. As a result, there is value in trying to learn whether or not state B is a conditional cooperator. When state A has beliefs that fall between the optimal and myopic cutoffs, it expects to receive more in the current period if it defects than if it cooperates. However, a strategy of cooperation is the most efficient way to reveal whether or not state B is a conditional cooperator—information that will be useful in making future decisions. A strategy of defection, by contrast, does not help state A differentiate between a conditional cooperator and a hard-liner, since these types respond to defection in the same manner. Thus, even though cooperation is associated with a lower current reward, the expected value of the information that this strategy will yield compensates for this loss. A state with such “experimental” beliefs expects that defection is the better policy in the present, but it nevertheless cooperates in the hope that the next period will bring good news, that it will revise his beliefs upward, and that it will eventually come to learn that state B is partner for mutual cooperation. A second feature
of this figure is that the gap between the optimal and myopic cutoffs increases with $q$, the belief that state B is a hard-liner. I will return to this observation in a moment.

*The Evolution of Beliefs and Policy*

The optimal policy rule tells us how the state should act for a given set of beliefs. The next thing to consider is how beliefs may evolve over the course of the interaction. The actual path that beliefs follow in any given history depends upon B’s type and the realization of the random draws that determine its response in a particular period. We can, however, make a few general observations about the learning process.

First consider periods in which state A cooperates—that is, periods in which the belief pair $(q, p)$ is above the cutoff line. Such a generic pair is depicted in the top half of Figure 2. As noted earlier, while A is cooperating, it updates about the probability that B is a conditional cooperator relative to the probability that B is a hard-liner or paper tiger. The updated probabilities of the latter two types do not move relative to one another, as shown in (1). Thus, if A’s beliefs start at $(q, p)$ and as long as A continues to cooperate, the belief pair is constrained to stay on the line depicted. This line is such that the original ratio $\frac{q}{1 - p - q}$ is always maintained. If B cooperates, A becomes increasingly sure that it is facing a conditional cooperator; $p$ increases and $q$ decreases by an amount that keeps the ratio constant, moving the belief pair up and to the left. If B defects, A becomes increasingly sure that it is not facing a conditional cooperator; $p$ decreases and $q$ increases by the necessary amount, moving the belief pair down and to the right. After a sufficient number of defections, updating in the manner can cause the belief pair to fall below the cutoff line, and state A will start defecting.
Now consider how beliefs evolve when A defects. The points \((q', p')\) and \((q'', p'')\) in Figure 2 depict two generic belief pairs that cause A to defect. As noted earlier, while A is defecting, it updates about the probability that B is a paper tiger relative to the probability that B is a hard-liner or conditional cooperator. The updated probabilities of the latter two types do not move relative to one another, as shown in (2). Thus, if A’s beliefs start at \((q', p')\) and as long as A continues to defect, the belief pair is constrained to stay on the line depicted. This line is such that the original ratio \(q'/p'\) is always maintained. If B cooperates, A becomes increasingly convinced that it is facing a paper tiger; \(p\) and \(q\) both decrease, and the belief pair moves down and to the left. If B defects, A is increasingly sure that it is not facing a paper tiger; \(p\) and \(q\) both increase, and the belief pair moves up and to the right.

If B continues to defect in response to defection by A, whether or not A changes policies depends on the original ratio of beliefs. If A’s beliefs are like the pair \((q', p')\), then continued defections will eventually cause beliefs to rise above the cutoff, and A will switch to cooperation. On the other hand, if the original ratio of beliefs is like that of \((q'', p'')\), then the path of beliefs can never cross the cutoff line. In this case, the initial belief that B is much more likely to be a hard-liner than a conditional cooperator can never be updated. As defection by A continues to generate defection from B, this only confirms A’s prior belief that its rival is a hard-liner, and A never allows its pessimism to be tested. Any belief pair that falls below the dotted line in the figure has this quality. In what follows, I will refer to these as “pessimistic” beliefs.

We thus have two general findings. First, if a given policy is successful—that is, if it leads to cooperation by the other side—then that policy will be continued for at least
another period. This is evident from the way beliefs move in Figure 2. A successful policy always leads beliefs to move in such a way that the same policy remains optimal. Only failures—that is, defections by the other side—can propel beliefs over the cutoff and cause a change in policy. Failure is thus a necessary condition for policy change. It is not, however, a sufficient condition. Depending upon where beliefs started, it may take a number of failures before beliefs move across the cutoff line. Moreover, as we just saw, there may be some pessimistic beliefs that never permit policy change, even in the face of continued failure. The finding that policy failure is necessary but not sufficient for policy change is well established in the learning literature (e.g., Levy 1994).

The second finding is that learning is not always complete. Even if B is a conditional cooperator, A need not discover this fact, leading to inefficient conflict. If state A’s beliefs ever fall into the range of pessimistic beliefs, then it will continue to defect forever. It does so because it is pessimistic that B is a conditional cooperator, not because it knows this to be true. Indeed, a pessimistic A continues to defect even though it believes that there is some nonzero probability that mutual cooperation is possible. Under these conditions, however, it is not optimal to probe this possibility. The fear of incurring the “sucker” payoff and the belief that this is the most likely outcome mitigate against testing the waters of cooperation—and may prevent state A from learning B’s true type. The logic here is similar to that in Jervis’ (1976) “spiral model” in which conflict can arise when hostile actions tend to reinforce states’ pessimistic initial beliefs about each other’s intentions. Notice, however, that the state cannot similarly be trapped into cooperating if this is not the optimal policy. Because of the way updating takes

---

3 This “stay-with-a-winner” result is common to bandit models (see, e.g., Berry and Fristedt 1985, 72-81).
place during periods in which A cooperates, any belief above the cutoff line can end up below that line after a sufficient number of defections by state B.

The potential inefficiency associated with pessimistic beliefs helps to explain why, as noted above and depicted in Figure 1, the gap between the optimal and myopic cutoff lines increases with $q$. One effect of this is that the range of pessimistic beliefs shrinks as $\delta$ increases. Thus, the more weight state A places on the future, the smaller is the range of beliefs that will “trap” the state in a potentially inefficient outcome. Indeed, as $\delta$ approaches 1, the point at which the optimal cutoff hits the frontier ($q$, 1-$q$) approaches (1,0); in this case, there are no beliefs sufficiently pessimistic that they will support defection indefinitely in the fact of defection by B.

2. Adding Domestic Politics: Consensus and Polarization

I now decompose state A into two political parties which regularly compete for control of office and decision-making authority. In each period, one party is in power, the other in opposition. The governing party monopolizes control over the state’s foreign policy and thus its choice between cooperation and defection. After the government implements its policy and state B responds, the two parties face each other in an election. Voters are assumed to be myopically retrospective: the incumbent’s probability of being retained is increasing in the outcome of that period’s interaction. Thus, the probabilities of reelection have the same ordering as the preferences over outcomes. Let $r^{CC}$ denote the probability that the incumbent is reelected given that A and B both cooperated in that period; define $r^{CD}$, $r^{DC}$, and $r^{DD}$ similarly. Given the assumption of retrospective voting $r^{CC} > r^{DC} > r^{DD} > r^{CD}$. 

18
The parties’ payoff in each period is a function of two things: whether or not they are in office, and the outcome of the international game in that period. Parties value holding office, due both the perks of office and the control it gives over other policy areas; thus, in any period in which the party is in power, it receives some fixed benefit \( B > 0 \). Parties also value the outcome of the game directly because of its effect on their constituents.

I assume that the parties represent constituents who differ in the costs they incur from mutual defection. This heterogeneity can arise from a number of sources depending upon the actual substantive issues at stake. For example, if the interaction describes a potential arms race, in which mutual defection leads to unbridled military spending, heterogeneity would arise if the parties’ constituents differ in their willingness to trade social spending for defense spending. Alternatively, if the substantive issue is cooperation on trade, those groups that can reap the largest benefit from trade also have the most to lose from mutual defection. To capture this kind of heterogeneity, assume that each individual \( i \) has a value for mutual defection \( w_i \). I will say that higher values of \( w_i \) correspond to more “hawkish” preferences, while lower values correspond to more “dovish” preferences. I assume that the two parties care about constituents with different preferences. In particular, the Doves represent constituents whose value for mutual conflict is \( w_D \), and the Hawks represent constituents whose value for mutual conflict is \( w_H \), with \( w_H > w_D \) (see Roemer 1994 for a similar assumption in a model of economic policy). For the purposes of this analysis, I assume that the basic ordering of outcomes is the same for both hawks and doves, or \( R > T > w_H > w_D > S \). Altering this assumption does not substantively change the results.
Solution

As before, parties try to maximize the current discounted value of their payoff stream given their beliefs. The stationary optimal policy for each party tells that party what policy to select given its current beliefs. Of course, only the party in government actually gets to enact its preferred policy; the party in opposition does not take any action in that period. Let \( \pi_H \) and \( \pi_D \) represent the optimal policies of the Hawks and Doves, respectively. For any starting beliefs \( p_0 \) and \( q_0 \), the value of the game for each party given that each plays optimally is:

\[
V_D(p_0, q_0, D_0) = \mathbb{E} \sum_{i=0}^{\infty} \delta^i \left\{ D_i [r_i(\pi_D; p_0, q_0) + B] + (1 - D_i) r_i(\pi_H; p_0, q_0) \right\}, \quad \text{and} \quad (11)
\]

\[
V_H(p_0, q_0, D_0) = \mathbb{E} \sum_{i=0}^{\infty} \delta^i \left\{ (1 - D_i) [r_i(\pi_H; p_0, q_0) + B] + D_i r_i(\pi_D; p_0, q_0) \right\}, \quad (12)
\]

where \( D_i \) is an indicator that equals one in periods in which the Doves are in power and zero otherwise. Notice that \( D_0 \), which indicates the party in power at the beginning of the game, is a new argument in the value function.

As before, we can decompose the value function for each party into current and future rewards. Assume that the Doves are in office in period \( t \) and consider the expected value of cooperating. Given the current beliefs, the probability that B will cooperate in response is \( s_C(t) \), as defined above, and the Doves’ expected current reward from cooperating is the same as given in (4). The value of the rest of the game depends upon the updated beliefs that the parties will take into period \( t+1 \) and whether or not the Doves remain in office. With probability \( s_C(t) \), B will cooperate, and three things will happen: \( p_t \) will be updated to \( p_{t+1}^{CC} \), \( q_t \) will be updated to \( q_{t+1}^{CC} \), and the Doves will be reelected with
probability $r^{CC}$. With the probability $1 - s_c(t)$, B will defect, and $p_t$ will be updated to $p_{t+1}^{CD}$, $q_t$ will be updated to $q_{t+1}^{CD}$, and the Doves will be reelected with probability $r^{CD}$.

Thus, the expect value of the future given that the Doves cooperate is

$$EV_D(p_{t+1}, q_{t+1}, D_{t+1}) = s_c(t)[r^{CC}V_D(p_{t+1}^{CC}, q_{t+1}^{CC}, 1) + (1 - r^{CC})V_D(p_{t+1}^{CC}, q_{t+1}^{CC}, 0)] + [1 - s_c(t)][r^{CD}V_D(p_{t+1}^{CD}, q_{t+1}^{CD}, 1) + (1 - r^{CD})V_D(p_{t+1}^{CD}, q_{t+1}^{CD}, 0)]$$

We can again combine the current and future rewards, as in (6), to give the total expected value to the Doves of cooperating in period $t$. Notice that, if the Doves are not in power in period $t$, a similar calculation yields their expected value in the event that the Hawks cooperate. The expected current reward is the same. The value of the future is similar to that in (13) except that the $r$ terms now refer to the probability of reelection for the Hawks; thus, we need to switch the zeros and ones in the value functions on the right-hand side.

We can perform similar calculations for the value of defection to the Doves and then replace the $D$ subscripts with $H$ subscripts to generate the value function for the Hawks. Since we will again have to solve for the optimal policy using computational methods, there is little value to writing out all of the equations.

The main result is depicted in Figure 3, which shows the optimal cutoffs for the Hawks and the Doves. The interpretation is the same as before: for belief pairs that fall above the cutoff, the party prefers to cooperate, and for belief pairs below the cutoff, the party prefers to defect. Not surprisingly, the cutoff for the Doves is always below that of the Hawks, meaning that Doves are willing to cooperate under a larger range of beliefs than are the Hawks. Because mutual defection is worse for the Doves, they have a greater incentive to risk cooperation and thereby learn whether or not the mutually
cooperative outcome is possible. Not only is cooperation more appealing to Doves in the present, but the information it can reveal is also more valuable to them, since they have greater stake in learning whether or not B is a conditional cooperator.

Notice that the range of pessimistic beliefs is larger for the Hawks, meaning that the Hawks are more likely to be content with mutual defection as an outcome and less willing to risk cooperation. In this respect, however, there is an important difference between the two-party model and the unitary actor model presented above. In the latter, if the state’s beliefs fall in the pessimistic range, it will never switch to cooperation, since it beliefs cannot be updated above the cutoff line. When there are two parties that rotate in office, the fact that one party has pessimistic beliefs does not foreclose cooperation, nor does it prevent that party from revising its beliefs out of the pessimistic range. If beliefs are such that the Hawks are pessimistic but the Doves are not, the Doves can come into office and start cooperating. If this cooperation is successful—i.e., it is reciprocated by B—then the beliefs can be updated in such a way that the Hawks are no longer pessimistic, and indeed they may come to support cooperation. Only if beliefs are such that the Doves are also pessimistic is there no chance of a switch to cooperation.

There are two interesting comparative static results that we can obtain. The first deals with the value of holding office, $B$. Holding everything else equal, an increase in $B$ causes the cutoffs of the two parties to converge toward each other. This should not be surprising, since the larger is the value of holding office, the less important are the differences between the parties’ constituents. Hence, as parties place more weight on their probabilities of reelection and less on the welfare of their particular constituents, the
differences in their policy preferences becomes muted. Put another way, office-seeking concerns tend to make the parties converge to a more moderate position.

Second, we can explore the effect of the voter heterogeneity. For example, fix the payoff from mutual defection to the Hawks’ constituents, $w_H$, and vary the difference between this and the corresponding payoff to the Doves’ constituents, $w_D$. The results of increasing this difference—and hence making the Doves more dovish relative to the Hawks—is depicted in Figure 4. Not surprisingly, as the Doves become more dovish relative to the Hawks, the former’s cutoff drops lower and the distance between the two parties’ cutoffs increases. What may be more surprising is that, as the Doves become more dovish, the Hawks’ cutoff line moves up. The reason is that, while both parties would be better off if they could learn that the other state is a conditional cooperator, cooperation is risky: it is associated with both the best and the worst payoffs, as well as the best and worst probabilities of reelection. As a result, the Hawks have incentive to “free ride” on the Doves’ greater willingness to risk cooperation. After all, if B turns out to be a conditional cooperator, the Hawks can reap the rewards of this knowledge when they come into office in the future; in the present, it makes sense to let the Doves incur the risk of experimentation. Thus, as the Doves become more willing to experiment, the Hawks become less so. This is true even though we have only changed the preferences of the Doves, while leaving those of the Hawks unchanged.

**Empirical Implications and a Preliminary Test**

Some of the results that come out of the model, like the last one, do not have clear observable implications. Here, I consider a few implications that might be measurable
and present a preliminary test using data on party positions in Western democracies during the Cold War. I focus on two implications in particular:

1. Given a history of mutual conflict, in which A is defecting and B generally responds in kind, parties with dovish constituencies are more likely to change their policy positions in response to news—both good and bad—than are parties with hawkish constituencies. This observation stems from the fact that Hawks have a larger range of pessimistic beliefs, meaning that they are less likely to switch to a policy cooperation. If they start with pessimistic beliefs, Hawks will only switch positions if (1) the Doves come into power and initiate cooperation and (2) the cooperation is generally reciprocated rather than exploited. If these conditions are met, then the Hawks can shift policy positions. Otherwise, the Hawks will have a rather stable preference for defection. The Doves, by contrast, are more likely to switch to a policy of cooperation in the response to a history of mutual conflict. They may also switch back to defection if their cooperative gestures are not reciprocated. Thus, there are many conditions under which Doves change their policy preferences in both directions while the Hawks remain consistently in favor of confrontation.

2. A related result is that interstate conflict can polarize rather than unite. Recall that, in response to mutual defection, both $p$, the probability that B is a conditional cooperator, and $q$, the probability that B is a hard-liner, increase in tandem. Even though they share the same beliefs, this updating process can lead to different conclusions about the correct policy. For the Doves, the increase in $p$ draws their beliefs above the cutoff line, and they come to argue that the probability that B is a conditional cooperator is sufficiently high that it is worth risking cooperation. For the Hawks, the coincident
increase in $q$ leads them to argue that the danger that B is hard-liner makes such a gamble unwise. Thus, mutual conflict can turn a situation of consensus—in which both parties support defection—to one of disagreement—in which the Doves advocate cooperation and Hawks do not. Note that this polarization effect is even more pronounced if we were to assume that Hawks and Doves are differentiated, not by their costs of conflict, but by their prior beliefs. In this case, two beliefs pairs that were initially close to one another—such as $(q', p')$ and $(q'', p'')$ in Figure 2—would be driven apart by mutual defection.

I now explore these predictions using data on party positions in Western democracies during the Cold War. Using a method I developed to estimate positions on military policy using data from election platforms (Schultz 2001), I examine (1) how parties’ positions changed in response to Soviet conflict behavior and (2) how reactions of parties depended upon their the costs of conflict to their constituents. For the purposes of these tests, I assume that parties’ general left-right position is a good proxy for the costs of conflict to their constituents. In particular, parties of the left are assumed to have more dovish constituents—in the sense that conflict with the Soviet Union imposed higher costs on them—than did parties of the right. This assumption that left-right ideology correlates with costs of conflict stems primarily from the sense that parties of the left were more reluctant to trade off social spending for military spending and hence had higher opportunity costs for the latter. Thus, periods of conflict with the Soviet Union, to the extent that they necessitated large investments in the military, imposed greater costs on parties of the left and their constituents. It might also be noted that, because the Soviet Union was a “left-wing” enemy, parties of the right also had an ideological hatred for it that was not shared by parties of the left. To the extent that
hawkishness and dovishness reflect such ideological considerations—rather than simply the opportunity costs of military spending—the assumption that parties of the left were more dovish may not be valid in other contexts (e.g., when the enemy is right-wing, like Nazi Germany). Nevertheless, in the context of the Cold War, replacing Doves with Left and Hawks with Right seems justified. Moreover, we will see empirically that parties of the left were consistently more dovish in their policy preferences—that is, more likely to advocate cuts in military spending, disarmament, and disenchantment with NATO—than were parties of the right.

**Dependent Variable: Policy Preferences**

The dependent variable in this analysis is a measure of parties’ preferences on military and foreign policy derived from statements in election manifestos. The data come from the Comparative Manifestoes Project (Volkens et al. 1995). This project has compiled information on election platforms in 15 Western democracies over the period 1945-1988. The data were coded with the explicit goal of permitting comparability across countries and years. As a result, this data set gives better spatial and temporal coverage than could be obtained using alternative sources, such as opinion polls.

The unit of analysis in the data set is the election manifesto. For each election in each country covered by the project, there is one observation for each political party which contested that election. The platforms are coded along 56 issue categories which reflect seven major policy “domains”: external relations, freedom and democracy, political system, economy, welfare and quality of life, fabric of society, and social
groups. Numerical codings are generated by breaking each manifesto down into statements which express single political ideas or issues and then assigning each statement to one of the 56 issue categories. For each of these categories, the data set indicates what percentage of statements in the manifesto express the corresponding viewpoint.

The estimated positions are derived from two issue categories: “Military: Positive” and “Military: Negative.” The first indicates the percentage of statements in the manifesto which express the following kinds of sentiments or positions:

Need to maintain or increase military expenditure; modernizing armed forces and improvement in military strength; rearmament and self-defence; need to keep military treaty obligations; need to secure adequate manpower in the military. (Volkens et al. 1995, 123)

The second measures the frequency of statements which express the opposite sentiment:

Favourable mentions of decreasing military expenditures; disarmament; “evils of war”; promises to reduce conscription; otherwise as [Military: Positive], but negative. (Volkens et al. 1995, 123)

Elsewhere, I discuss a method for estimating parties’ position on a 0 to 10 military policy scale using these two variables (Schultz 2001). Let \( MIL_i \) denote the estimated position of party \( i \) in election year \( t \).

**Independent Variables: Threat and Ideology**

---

4 The countries included in the sample are: Australia, Austria, Belgium, Canada, Denmark, France, West Germany, Italy, Luxembourg, Netherlands, New Zealand, Norway, Sweden, United Kingdom, United States.

5 For more on the coding procedures, see Volkens et al. (1995). The codebook’s appendix contains an excellent description of the coding rules, as well as an example of how they were applied to an actual manifesto.
The main independent variable captures variations in Soviet conflict behavior. I measure this in two different ways. The first indicator comes from two data sets that measure the level and frequency of conflictual and cooperative actions which countries direct at one another: the Conflict and Peace Data Bank (COPDAB) and the World Events/Interaction Survey (WEIS). Both data sets record foreign policy events on a daily basis, indicating the sender and target of the action and the level of conflict or cooperation the action entailed. COPDAB provides a weighting system which makes it possible to aggregate the daily events into larger time periods and to assign each period a score on a single conflict-cooperation dimension (Azar 1993). WEIS was not designed in the same manner, but Vincent (1979) and Goldstein (1992) have devised scaling techniques to translate WEIS event categories into a conflict-cooperation score that is reasonably compatible with those generated by COPDAB (Goldstein and Freeman 1990; Reuveny and Kang 1996). This is good news because the COPDAB data only cover the years 1948-78, while the WEIS data cover 1966-88. Hence, generating a time series long enough to match the period covered by the manifesto data requires splicing the two data sets together.

As a first cut, I develop conflict activity emanating from the Soviet Union against the United States. The main reason is practical: because WEIS data were only available for this dyad, this is the only data series which covers the entire 1948-88 time period. While we can look at the effect of Soviet actions directed at other states in the sample, any such indicator would only be available up to 1978. Theoretically, we know that the interaction between the Soviet Union and the United States was the most prominent

---

6 COPDAB data and description come from Azar (1993). The coding rules for WEIS can be found in McClelland (1978); the actual WEIS data used in this study are from Goldstein and Freeman (1990, 165-7).
feature of the international system in this period. Not surprisingly, the United States is
target of the vast majority of Soviet actions directed at the countries in the sample.
Indeed, if we use the COPDAB data to compare Soviet activity directed against the
United States with Soviet activity against all states in the sample, the correlation between
the two series is 0.83. Thus, to the extent that we can see the states in our sample as a
single bloc facing a common threat environment, it is not unreasonable to use our
measure of Soviet activity against the United States as a proxy for the level of threat
experienced by all of them.7

The two data sets were spliced using a method outlined by Goldstein and Freeman
(1990). We use the COPDAB scores up to 1978 and then multiply the WEIS scores by
2.3 in order to have a comparable measure for the remainder of the period. Let $SUUS_m$
denote the level of conflict directed against the United States by the Soviet Union in
month $m$. Then, for each party $i$ in election year $t$, let $SUUS_{it}$ equal the average level of
conflict over the 12 months preceding the election.

The second indicator of Soviet behavior comes from the Militarized Interstate
Dispute (MID) data set (Jones, Bremer, and Singer 1996). MIDs are events in which at
least one state directed took militarized action against at least one other state, where
militarized action consists of a threat, display, or use of military force. Using this data
set, I determined for each month during the period 1946-1988 how many MIDs the
Soviet Union was currently involved in.8 For each party $i$ in election year $t$, let $SUMID_{it}$

---
7 Having said that, it is important to recognize that Austria and Sweden were technically unaligned
countries during this period. However, it is not unreasonable to assume that their perceptions of the threat
posed by the Soviet Union was highly correlated with those of other countries in the sample. Still, it is
worth being sensitive to their unusual position.
8 This count did not include MIDs in which the Soviet Union was coded as a participant but in which it
took no militarized action.
equal the average Soviet MID involvement in the 12 months preceding the election. Notice that this coding rule treats all MIDs equally, regardless of the level of militarized action reached by the Soviet Union. Figure 5 plots annual averages of $SUUS_m$ and $SUMID_m$ for the period 1945-1990. As is apparent from the figure, the indicators are related but not identical. Indeed, the correlation between the two is only 0.38.

The second independent variables is a measure of general party ideology on the traditional left-right scale. The measure I use is based on surveys conducted by Castles and Mair (1984) and Inglehart and Klingemann (1976). The Castles/Mair study used an expert survey to estimate party positions on an 11-point scale. It is the most comprehensive study undertaken during the period covered by the manifesto data; however, there are a number of parties in the latter that were no longer around by the mid-1980s. The older, Inglehart/Klingemann study estimated party positions by asking voters to place themselves on the left-right scale. They then report the average self-placement of voters that identified with each party. In order to maximize coverage, I combined the two in the following way. Using the parties covered by both studies, I regressed the Castles/Mair score on the Inglehart/Klingemann score and a constant. The results show that the correlation between the two is quite high, with an $R^2$ of 0.81. I then used the coefficients from this regression to impute Castles/Mair scores for parties only covered by the Inglehart/Klingemann study. The variable $RIGHT$ was set equal to the Castles/Mair score when possible and to the predicted score otherwise. It takes values from 0 to 10, with higher values corresponding to more right-wing parties.

---

9 I also generated alternative measures which only counted MIDs in which the Soviet Union used force and which weighted MIDs depending upon the severity of Soviet actions. In general, these indicators performed the same or worse than the unweighted measure.
To test whether reactions to Soviet behavior depended upon ideology, the regressions include an interaction term which measures the product of the Soviet threat measure and the ideology score.

*Results*

The first result, which is easy to show graphically, is that parties of the left exhibited much greater variation over time in their positions, while parties of the right were consistently hawkish throughout the period. Figures 6 and 7 graph the estimated positions of the major right- and left-wing parties, respectively. From Figure 6, it is clear that parties of the right exhibited a great deal of stability. There is a slight dovish turn in the mid-1970s, corresponding to the period of détente, and this quickly reversed as détente broke down in the late 1970s and early 1980s. Figure 7 shows that parties of the left showed much more variability, both across countries and across time. This point is reinforced in Figure 8, which plots the difference between each party’s score and its average score over the entire period as a function of $\text{RIGHT}_i$. It is clear from this figure that left-wing parties exhibited greater deviations from their means than did right-wing parties. This point is not only important analytically, but it also suggests that any regression using these scores as a dependent variable is likely to suffer from heteroskedasticity due to the greater variance in the error terms for more left-wing parties.

I now turn to regression analysis. The data set pools together time series from each of the parties in the sample. Unlike typical pooled time series designs, however, the observations are taken at irregularly spaced elections. As a result, some of the typical
treatments used in this context are not practical (e.g., Beck and Katz 1995).

Nevertheless, we can anticipate and correct for two problems common to such data.

First, as we just saw, there is cross-sectional heteroskedasticity due to the higher variation in the positions of left-wing parties. To correct for this, I employ a technique developed by Harvey (1976; see also Greene 1997, 565-7). In brief, this technique uses maximum likelihood estimation based on the assumption that the errors are distributed normally with mean zero and variance

$$\sigma^2_i = \sigma^2 \exp(\gamma_{RIGHT_i}) .$$

This regression model with multiplicative heteroskedasticity can be estimated by a routine available for Stata 6.0 (STB 42, sg77).

Second, there is good reason to believe that there is autocorrelation within the time series. Parties’ positions are bound to have certain amount of inertia from one election to the next due to the fact that it is not costless to change positions willy-nilly. To deal with this, I include a lagged dependent variable. Not only does this variable eliminate the autocorrelation, but it permits us to interpret the model as one with partial adjustment, in which the parties’ positions change slowly in response to events. Using the coefficient on the lagged dependent variable permits us to estimate both the short- and long-run effects of changes in Soviet behavior on the parties’ positions.

The regressions were run with each of the alternative Soviet threat indicators. The two generated very similar results, but they washed each other out when used together. A preliminary analysis of the data revealed a clear structural break in the series,

---

10 An alternative specification of the variances is to assume that $\sigma^2_i = \sigma^2 \cdot RIGHT_i^{\gamma}$, which is equivalent to taking the natural logarithm of $RIGHT_i$ in expression (14). The results reported below do not depend upon which of these specifications is used.
which most likely took place after 1963. This is confirmed using a Quandt’s method for testing for likely break points (see Beck 1991). Hence, I split the regression into pre- and post-1963 samples. The results are presented in Table 2.

It is clear from the table that there is no systematic relationship between Soviet behavior and party positions in the earlier period. Post-1963, however, a robust relationship emerges, regardless of which indicator of Soviet behavior is used. Because the ideology measure runs from 0 to 10, the coefficients on the Soviet threat variable and the interaction term imply that left parties tended to become more dovish in response to increases in Soviet conflict behavior, while parties of the right became more hawkish or did not change. The magnitude and robustness of this effect are shown in Table 3. The first and third columns show the conditional coefficient on the two Soviet threat indicators for various ideology scores. It is clear that there is a strong negative relationship for parties of the left and no statistically measurable relationship for parties of the center and right. The second and fourth columns take advantage of the partial adjustment interpretation of the model to calculate the long-run relationships implied by the coefficients. One minus the coefficient on the lagged dependent variable reflects the speed with which party positions change from one election to the next in response to changes in international conditions. If we then divide the conditional coefficients by this term, we can estimate how much a change in Soviet behavior would affect the parties’ positions in the long-run or, alternatively, if it were costless for parties to change their positions as much as they wanted from one election to the next. The mean and standard deviation of the Soviet conflict scores are also reported. From these, we learn that if SUUS shifts from one standard deviation above to one standard deviation below its mean,
a party with an ideology score of 3—which roughly corresponds to most European
socialist parties—would change its position by –0.5 in the short-run and by –1.2 in the
long-run.

How can we interpret these results? The absence of a systematic relationship in
the earlier period seems to reflect the fact that many party movements in this period were
unrelated to changes in the Soviet threat. In some cases, shifts occurred as a result of
primarily domestic considerations. For example the German SDP made a dramatic shift
toward the center in 1957 as part of the party’s effort to shed its Marxist roots and
become more mainstream and electable. There may also be international factors not
captured by the model, such as the French war in Algeria and Belgian involvement in the
Congo. It would be nice to include improved measures that control for influences on
military policy that were not central to the Cold War. This was also a period of some
stability in party preferences, probably reflecting the fact that there was a good deal of
consistency in the relationship between the superpowers. The United States was mainly
defecting and the Soviet Union was sometimes reciprocating and sometimes backing
down: there was no clear signal that the policy of containment and deterrence was failing
or succeeding.

After the mid-1960s, things started to unsettle. The Cuban Missile Crisis drove
home the danger of mutual confrontation and planted the first seeds of détente. Then the
Vietnam War reinforced the sense that global containment could be quite costly and
paved the way for much deeper efforts at cooperation. In many respects, the heyday of
détente marked the strongest consensus among left and right, especially in Europe. (In
the United States, the lingering effects of Vietnam exert a polarizing effect, but the early
efforts at cooperation, such as SALT I, were met with bipartisan approval.) By this point, most of the left-wing parties had come to accept NATO and nuclear deterrence, as long as these were coupled with genuine efforts at disarmament and reducing tensions (Eichenberg 1989). Thus, most of the mainstream parties of the left (i.e., excepting some communist and left socialist parties) had moderate positions in the 1970s and were very close to their counterparts on the right.

What split the consensus was the resumption of conflict in the late 1970s and early 1980s. For the right, Soviet actions in the Third World and particularly Afghanistan were proof positive that cooperative moves by the West would be exploited rather than reciprocated: as suspected, the Soviet Union was a hard-line state or, at best, a paper tiger that would bow to superior strength. As early as the mid-1970s, the right began to question the wisdom of détente and call for a return to a more confrontational policy. The elections of Thatcher in Britain and then Reagan in the United States would bring these plans to fruition. To the left, the setbacks of the late 1970s were unfortunate, but not sufficient evidence to prove that cooperation with the Soviet Union was impossible. Many on the left felt that the United States had never really given détente a chance, and that the resumption of Cold War in the early 80s was premature: the experiment with cooperation should continue. Thus, with the return to mutual conflict, party positions tended to polarize along left-right lines. This effect comes through in the regression results: the increase in Soviet conflict behavior in the late 1970s and early 1980s caused left parties to become more dovish and right parties to remain, or become more, hawkish. This polarization was particularly marked in Britain, Germany, and the Netherlands, which witnessed massive protests against the planned deployment of intermediate range
missiles on their soil. As Joffe (1988) shows, these protest movements arose largely in response to dovish turns by socialist parties in those countries. The effect is not confined to Europe, however. In New Zealand, the same period sees a dovish turn by the Labor party that culminates in the dissolution ANZUS. In the United States, the Democrats briefly became more hawkish in response to the crises in Afghanistan and Iran, but moved back in a dovish direction in response to the Reagan build-up and the coincident resumption of Cold War. Thus, an increase in international tensions was polarizing, rather than uniting.

Conclusion

This last result, of course, runs exactly contrary to Wolfers’ claim that a burning house tends to induce similar preferences among its occupants. As I have shown, under conditions of uncertainty, an increase in threatening behavior by a rival state can generate disagreement among parties with different interests. While parties may agree that a confrontational policy is leading to costly conflict, they may disagree about the desirability of the alternate course. For some, a history of mutual conflict generates an incentive to experiment with cooperation; for others, it reinforces the sense that cooperation would be dangerous. Thus, during the Cold War, it was the relaxation of tensions during détente that elicited consensus, and the resumption of conflict in the late 1970s that broke that consensus apart. The fact that the house was on fire led to strong disagreements about which exit the state should take.

Of course, one limitation of this analysis is that the Cold War might not provide sufficient variation in threat conditions to permit an accurate test of this story. While
there are periods of relative tension and calm, the entire period is one in which the West faced a strong and dangerous rival. A more complete test should extend the analysis to the post-Cold War period, in which severity of the threats faced changed quite dramatically—as did the uncertainty surrounding them. Such as test would be a natural extension of this work.


Table 1. Bayesian Updating as a Function of the Period $t$ Strategies and Outcomes

<table>
<thead>
<tr>
<th>State A’s Strategy</th>
<th>State B’s Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperate</td>
<td>Defect</td>
</tr>
<tr>
<td></td>
<td>$p_{t+1} = \frac{p_t s_C^H}{p_t s_C^H + (1 - p_t) s_C^L}$</td>
</tr>
<tr>
<td></td>
<td>$q_{t+1} = \frac{q_t s_C^L}{p_t s_C^H + (1 - p_t) s_C^L}$</td>
</tr>
<tr>
<td>Defect</td>
<td>$p_{t+1} = \frac{p_t s_D^L}{(p_t + q_t) s_D^L + (1 - p_t - q_t) s_D^H}$</td>
</tr>
<tr>
<td></td>
<td>$q_{t+1} = \frac{q_t s_D^L}{(p_t + q_t) s_D^L + (1 - p_t - q_t) s_D^H}$</td>
</tr>
</tbody>
</table>
### Table 2. Party Positions and the Soviet Threat, 1948-88

Multiplicative Heteroskedastic Regression (MLE)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Threat Indicator: <strong>SUUS</strong></th>
<th>Threat Indicator: <strong>SUMID</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-1963</td>
<td>Post-1963</td>
</tr>
<tr>
<td></td>
<td>Constant</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.72)</td>
</tr>
<tr>
<td></td>
<td>Position at last election</td>
<td>0.52***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.077)</td>
</tr>
<tr>
<td></td>
<td>Ideology</td>
<td>0.38***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.13)</td>
</tr>
<tr>
<td></td>
<td>Soviet threat</td>
<td>-0.00012</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0087)</td>
</tr>
<tr>
<td></td>
<td>Soviet threat*</td>
<td>-0.000028</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0013)</td>
</tr>
<tr>
<td>No. obs.</td>
<td>134</td>
<td>332</td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>156.4***</td>
<td>381.0***</td>
</tr>
</tbody>
</table>

*** p<0.01  ** 0.05>p>0.01  * 0.10>p>0.05
Huber-White robust standard errors, with clustering on elections, reported in parentheses.
Table 3. Effect of the Soviet Threat Conditional on Ideology

<table>
<thead>
<tr>
<th>Ideology</th>
<th>Threat Indicator: SUUS</th>
<th>Threat Indicator: SUMID</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient w/ Partial Adj.</td>
<td>Coefficient w/ Partial Adj.</td>
</tr>
<tr>
<td>0</td>
<td>-0.013** -0.032</td>
<td>-0.56** -1.38</td>
</tr>
<tr>
<td>1</td>
<td>-0.011** -0.027</td>
<td>-0.47** -1.16</td>
</tr>
<tr>
<td>2</td>
<td>-0.009** -0.022</td>
<td>-0.38* -0.94</td>
</tr>
<tr>
<td>3</td>
<td>-0.007** -0.017</td>
<td>-0.29* -0.72</td>
</tr>
<tr>
<td>4</td>
<td>-0.005 -0.012</td>
<td>-0.2 -0.49</td>
</tr>
<tr>
<td>5</td>
<td>-0.003 -0.007</td>
<td>-0.11 -0.27</td>
</tr>
<tr>
<td>6</td>
<td>-0.001 -0.002</td>
<td>-0.02 -0.049</td>
</tr>
<tr>
<td>7</td>
<td>0.001 0.002</td>
<td>0.07 0.17</td>
</tr>
<tr>
<td>8</td>
<td>0.003 0.007</td>
<td>0.16 0.40</td>
</tr>
<tr>
<td>9</td>
<td>0.005 0.012</td>
<td>0.25 0.62</td>
</tr>
<tr>
<td>10</td>
<td>0.007 0.017</td>
<td>0.34 0.84</td>
</tr>
<tr>
<td>Mean</td>
<td>30.7</td>
<td>Mean 0.89</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>36.4</td>
<td>Std. Dev. 0.70</td>
</tr>
</tbody>
</table>

*** p<0.01    ** 0.05>p>0.01    * 0.10>p>0.05
Figure 1. The Optimal and the Myopic Policy
Figure 2. The Evolution of Beliefs

\[ \text{Pr(Hard-liner)} = q_t \]
\[ \text{Pr(Conditional Cooperator)} = p_t \]

\((q', p')\), \((q, p)\), \((q'', p'')\)

- B Cooperates
- B Defects

\(0 \leq q_t \leq 1\)
\(0 \leq p_t \leq 1\)
Figure 3. The Optimal Policy for Hawks and Doves
Figure 4. The Effects of Making the Doves more Dovish

\[
\begin{align*}
\Pr(\text{Hard-liner}) &= q_t \\
\Pr(\text{Conditional Cooperator}) &= p_t
\end{align*}
\]
Figure 5. Soviet Conflict Behavior, 1946-1990

The figure shows the Soviet conflict directed at the United States (SUUS) and Soviet military involvement (SUMID) over the years from 1946 to 1990. The data is represented with two lines: one for SUUS and another for SUMID, illustrating the trends and fluctuations in conflict behavior during this period.
Figure 6. Major Parties of the Right
Figure 7. Major Parties of the Left

- German SPD
- Italian PSI
- Italian Communist
- French PS
- UK Labor
Figure 8. Variation in Party Positions as a Function of Left-Right Ideology

Left-Right Score (RIGHT)