

CHAPTER 5

Rivalries as a Testing Methodology

The previous chapter showed that the rivalry approach gives us a new perspective on war and conflict. Yet the rivalry approach does not merely require new ways of theorizing about the traditional foci of studies of international conflict. Such changes must be accompanied by new ways of testing hypotheses. This means alterations in the research design and methodologies employed in the traditional causes-of-war literature. These new tests often have significant advantages vis-à-vis previous methodologies. In this chapter, we illustrate how the rivalry approach transforms standard testing procedures. The rivalry approach provides new means to investigate classic hypotheses, often addressing weaknesses in standard methodologies.

Our critique and reformulation of existing methodologies addresses itself not only to large- N statistical studies, but encompasses comparative case study techniques as well. Though much separates the large- N from the small- N researcher, from our perspective, the commonalties dominate: both are cross-sectional and use the dispute, crisis, or war as the object of study. These two elements characterize most of the causes-of-war literature. The rivalry approach alters these two traditional testing features, as well as transforming the previous, limited attempts to incorporate rivalries in the research design. The rivalry approach does utilize cross-sectional analyses, but it is the rivalry, not the dispute, that is the cross-sectional unit. Furthermore, we shall see that other cross-sectional units can be created. We argue that testing techniques based on rivalries serve to replace or supplement the dispute or dyad-year as the unit of analysis.

Rivalries also fundamentally involve time. In the temporal dimension of a rivalry lies fertile new ground for hypothesis testing. This proves crucial in defining new—and often better—control groups. The causes-of-war literature looks at wars compared to nonwar disputes. Using rivalries, we can compare war with nonwar periods *within* a rivalry; hence by definition we control

for many possibly confounding variables. These and related concerns are addressed in the section below that emphasizes testing implications that flow from longitudinal aspects of rivalries.

Creating new tests also means creating new dependent variables that the theory or model in question implies. For example, the democratic peace literature revolves around the proposition that democracies do not fight wars, but it implies other propositions as well. If a rivalry changes from a nondemocratic to democratic dyad, then the democratic peace hypothesis suggests that the rivalry will end. In short, the rivalry approach increases significantly the means for testing empirically a theory or model.

As we noted in the last chapter, changing the focus of analysis to rivalry results in some fundamental shifts in the way we think about traditional questions about war and peace. Yet it also means that we conduct research differently than in the past. These changes go beyond using rivalries only as a case selection device, replacing disputes or crises as the objects of study. As we discuss in the next section, some of the early work on rivalries has moved beyond the traditional causes-of-war framework, but has not truly adopted the rivalry approach in all its manifestations. Thus, beyond reviewing those efforts, we also address here the testing implications of moving beyond rivalries as a simple device to select a population of cases for study.

We begin by reviewing the history of rivalry use in international conflict research as a prelude to our discussion of how the rivalry approach extends, and we believe improves, these limited efforts. We then consider the changes brought by the rivalry approach to testing methodology, organized under three of the concerns previously mentioned: case selection, cross-sectional analysis, and longitudinal analysis. We again use deterrence analyses as our primary example (although we also occasionally make reference to the democratic peace, the subject of our empirical application in the next chapter), as deterrence hypotheses have proven difficult to test using standard research designs (see Harvey 1998).

The History of Rivalries and Testing

In the early 1980s, the concept of rivalry, specifically enduring rivalry, was introduced as a part of the testing of some hypotheses about international conflict (the first to do this, to the best of our knowledge was Wayman 1982).¹ As we noted in the previous chapter, rivalries were not necessarily an explicit part of the theorizing in those tests; they served as background conditions to select cases for testing. The studies using rivalries shared a number of characteristics in the early period.

¹Rivalries appear relatively early in the international conflict literature, but much of that early work is not interested in rivalries in general, but rather one specific rivalry, for example the U.S.-Soviet or Arab-Israeli rivalries. That is, the concept of rivalry did not play any role in the research design or testing elements of the research of what were essentially descriptive case studies.

First, the focus was almost entirely on *enduring* rivalries, not the broader continuum of rivalries.² Second, despite the use of rivalries to select cases, virtually all these studies employed cross-sectional designs and thereby largely ignored the longitudinal elements of conflict relations that are the centerpieces of the rivalry concept. Typically after choosing a set of rivalry cases, scholars still employed disputes, wars, dyad-years, and other traditional units of analysis in conducting their analyses. Third, most of these studies made the implicit assumption that state behavior in enduring rivalries is different from that in other contexts. We have shown in chapter 3 that this is largely true, but our conclusions are limited to certain specific characteristics and are empirically derived. Studies that used enduring rivalries as a case selection device made this assumption with respect to their phenomenon of interest (e.g., arms races, deterrence) and did so without apparent empirical support (although there may have been some theoretical reasons to suspect so). Fourth, all of the studies saw some things in enduring rivalries and their characteristics that were essential for the phenomenon under study to occur. These characteristics were assumed to be constant and unique to enduring rivalries. They served only as justification in case selection and thereby the focus on enduring rivalries. Finally, these sets of studies were not generally concerned with rivalries per se, be it their origins, dynamics, or termination (subjects that are addressed in the second part of this book). Neither were these studies influenced by new questions or insights of the rivalry approach, as illustrated in the last chapter. Rather, they distinctly were testing classic hypotheses within the causes-of-war framework, but applied to a new set of cases.

Enduring rivalries as a background condition have appeared in a number of theoretical contexts. The most obvious use of enduring rivalries as a background condition is in the study of power transitions (Wayman 1982; Geller 1993; Wayman 1996). Power transition theory (Organski 1958; Organski and Kugler 1980; Kugler and Lemke 1996) posits that war will occur when a challenging state approaches and then surpasses a formerly dominant state. This implies that some kind of competition exists between the states, and, unlike many other international conflict models, there is a dynamic element to this approach. The focus on enduring rivalries performs several functions. First, the existence of an enduring rivalry indicates some long-standing competition between two states and establishes that war is an ongoing threat (rather than a hypothetical, perhaps low, possibility). Studying dyads other than enduring rivals might be open to the criticism that the states involved were not directly attuned to the actions of the other state or that the other state was not its primary rival. Enduring rivalries also present the dynamic and longitudinal feature that is implicit in the power transition model, although in practice, many analyses

²Yet using today's standards of operational measures, what many of these studies were actually studying was lesser as well as longer and more severe rivalries; they often used two- or three-dispute minima to signify the existence of enduring rivalries. In effect, many of these studies were considering, in our language, proto- and enduring rivalries.

of power transition still ignore this element and adopt a static cross-sectional design.

A second, and similar, use of enduring rivalries selects cases in the study of arms races and war (Diehl 1985a; Diehl 1985b; Diehl and Kingston 1987; Diehl and Crescenzi 1998). Arms races are a highly probable manifestation of rivalries because the building up of weapons and military capability is likely to be part of the rivalry competition, as are war and other phenomena (see Goertz and Diehl 1993 for a comparison of the rivalry and arms race concepts). Implicitly then, enduring rivalries indicate the proper context for an arms race to occur, and indeed it may not be possible for a true arms race to occur outside of one; by definition, an arms race is a militarized relationship. As a case selection device, looking at enduring rivalries provides some assurance that new weapons acquisition is directed at a specific opponent and that such increases are viewed as threatening. We know that enduring rivalries indicate the ongoing threat of war between two states, and arming decisions might be a result of that expectation of future conflict. Looking at cases other than enduring rivalries might not reveal true arms races, but mutual military buildups by two states that are not directed exclusively or primarily at each other, but rather against third parties. In those instances, inferences about the alleged effects of those buildups on conflict behavior would be misguided.

Similar to the studies of arms races and war, analyses of the substitutability of alliances and arms races also rely on enduring rivalries to identify the context under which such a phenomenon is likely (Sorokin 1994; Diehl 1994b). The substitution effect requires first an ongoing security threat that presents states with choices on how to respond, of which arming and forming alliances are prominent options. Again, the pull of the future (future threats) is a significant part of the calculation. States are not capable of arming rapidly in order to meet immediate threats, and many alliances cannot be negotiated and executed that quickly. Enduring rivalries present the ongoing security threat and future conflict expectations in which state security choices can be observed. Enduring rivalries have long lives, and thus scholars can observe trade-offs between arms races and alliances over time, especially as the formation of the latter is not common on a year-to-year basis. Enduring rivalries have also been used to study the impact of system-level variables on conflict behavior (Huth, Gelpi, and Bennett 1992) and to test the robustness of the democratic peace (Maoz 1997), the latter of which we ourselves address in the next chapter.

As we illustrate in more detail in the following sections, deterrence studies are also a popular venue for the use of enduring rivalries in case selection. Scholars must identify instances in which deterrence is even being attempted and therefore they must demonstrate that the relevant parties regard each other as serious security threats and that there is some significant probability that one or both sides will use military force against the other. Enduring rivalries are the context under which these conditions are fulfilled (Huth and Russett 1993;

Lieberman 1994, 1995); in another population of cases, one might be unsure of whether any deterrent threats are conceivable or likely. Thus, the past of a rivalry establishes the bases under which threats and reputational elements are developed. The expectation of the future conflict, provided by rivalries, is the central part of deterrence; deterrence, at least of a general variety, is inherently future oriented. We develop more of these applications below.

Table 5.1 summarizes the different foci for which rivalries have been used in the research design. The common ground lies in that they assume some conflict context that is represented by enduring rivalries. The use of rivalries as a background condition and a case selection device is an improvement in many instances over the past testing designs that considered all possible or all contiguous dyads. Indeed, most models or theories that assume preexisting conflict could be tested using rivalries. In that sense, the use of the rivalry concept has improved the validity of the findings on important subjects such as alliances and arms races as well as deterrence and power transition. One need but compare Russett 1963, Huth and Russett 1984, and Huth and Russett 1993 to see the evolution in case selection. We see this as a positive trend. Nevertheless, we regard such efforts as incomplete and not fully reflective of the insights that can be gained from the rivalry approach. In the following sections, we demonstrate how rivalries can be incorporated into research designs beyond their use in case selection. Yet we begin with an analysis of case selection and show that the rivalry approach involves more than a specification of a population of cases.

Case Selection

In moving from theory to empirical testing, one inevitably confronts the problem: what cases are appropriate for an empirical test? On the one hand, the theory or the hypothesis itself should define the limits and the universe to which it applies. In practice, a large region of indeterminacy surrounds most case selection. The deterrence literature illustrates the controversy that case selection can provoke, notably when cases are “selected on the dependent variable.” Many studies only look at deterrence failures, much as many other studies of international conflict select cases based on the occurrence of wars, crises, or militarized disputes. This relates fundamentally to the general problem of what should be the control group in the analysis. The generic situation may be termed the “dog that doesn’t bark” problem, and it occurs frequently in conflict research. The temptation is to select the cases where the dog barks (Achen and Snidal 1989; Most and Starr 1989). As with the democratic peace, we need a control group of nonconflict dyads. Yet it is not obvious which “nonbarking dogs” we should use. In the worst-case scenario, selecting only on the occurrence of the dependent variable, there is no control group at all. We shall return to this general problem of defining the comparison group at various points because it lies at the heart of testing problems, and it shows the value of the rivalry approach in helping resolve those problems.

TABLE 5.1: The History of Rivalry Testing

Author	Question	Population
Wayman 1982	Does a shift in relative power produce war?	Major power rivalries, 1816–1976
Diehl 1985a	Do arms races increase chances for war?	Twenty-two major-power enduring rivalries, 1816–1980
Huth, Bennett, and Gelpi 1992	How does interaction between system structure and leaders' risk propensity affect initiation of militarized interstate disputes?	Eighteen great power rivalries, 1816–1975
Geller 1993	What is the effect of static and dynamic power differentials on the conflict patterns of dispute-prone dyads?	Thirty-two enduring rivalries, 1816–1986
Huth and Russett 1993	Under which circumstances are challengers likely to initiate militarized disputes?	Fourteen enduring rivalries, 1948–82
Diehl 1994b	Under what conditions and to what extent do alliances substitute for or complement military buildups?	Twenty-two major-power enduring rivalries, 1816–1976
Kinsella 1994a	What is the impact of superpower arms transfers on enduring Third World rivalries?	U.S. and Soviet arms transfers into Middle Eastern and Persian Gulf enduring rivalries, 1949–88
Lieberman 1994	Do deterrence successes rarely occur and do leaders challenge deterrence, despite credible threats, due to political vulnerabilities?	Israel–Egypt enduring rivalry, 1948–77
Sorokin 1994	What mix of arms and alliances does a state choose to maximize its security?	Austria-Hungary (1880–1913), France (1880–1913), Israel (1963–88), Syria (1963–88), and their allies and enduring rivals
Kinsella 1995	What is the impact of an arms transfer link between superpower and regional rivalries?	U.S. and Soviet arms transfers into Middle Eastern, Persian Gulf, South Asian, and the Horn of Africa rivalries, 1950–90
Lieberman 1995	Does deterrence work and, if so, what makes it work?	Egypt–Israel enduring rivalry, 1948–73
Maoz 1997	Do realist and cultural critiques undermine the empirical validity of the democratic peace proposition?	Different samples, one of which is 15,805 enduring dyad-years, 1816–1986
Diehl and Crescenzi 1998	Is the arms race–war connection meaningful or spurious?	Sixty-three enduring rivalries, 1816–1992

Achen and Snidal (1989) critique a number of scholars for selecting on the dependent variable, that is, not providing a control group of deterrence success. We propose that the rivalry approach provides a solution for which Achen and Snidal were looking. Given the amount of scholarly attention that has been devoted to this issue, we regard this as a nontrivial claim. Huth and Russett's (1993) use of enduring rivalry is on the right track, but remains rooted in the basic cross-sectional causes-of-war methodological framework.

Disputes and crises provide the immediate context for most discussions of deterrence success.³ Although the distinction between general and immediate deterrence seems widely accepted, most studies choose deterrence in crisis situations as their focus, and thereby focus on the latter form of deterrence. For Huth and Russett (1984), as well as others, immediate deterrence arises *once a crisis or dispute is under way*. Scholars disagree about which crises should be included as deterrence cases (see Lebow and Stein 1990 and the Huth and Russett 1990 response), but they agree that immediate deterrence, by definition, takes place in a crisis context. All sample from the universe of crises and disputes; for example, Huth and Russett depend heavily on Correlates of War Project data sets. Similar problems arise in the much of the arms race–war literature (Wallace 1979; Diehl 1985a; Sample 1997) in which scholars study only the impact of arms on dispute escalation; that is, they study only cases in which militarized conflict has already occurred. Thus, circumstances in which arms races inhibit or deter conflict are ignored, and therefore scholars begin with biased sample of cases (Diehl and Crescenzi 1998).

Because Achen and Snidal have been at the center of this debate, we may ask how they conceive of the universe of deterrence cases. A consistent theme in their discussion is that one tests deterrence in situations in which there already exists a hostile relationship: “Suppose that there are 100 countries in the world, each of them in a deterrence relationship [!] with an average of 5 of the others. Thus, each year there are 250 war-prone dyads . . . Presumably the relevant population [for studying deterrence] for any time period consists of those states that have a ‘serious’ dispute with another state, where ‘serious’ is defined by the potential for war . . . pacific intentions are easily confused with successful deterrence, and vice versa. Rational deterrence theory does contain some minimal psychological content: for example the initiator must realize that the defender exists and threatens to defend” (1989, 161–64). In rivalry terms, general deterrence becomes relevant most clearly *after* the hostile relationship is under-way; deterrence theory not does relate to the birth of rivalries, but is more concerned with their maturity, that is, with youth through old age. Within rivalries, “pacific intentions” are usually assumed by both sides not to exist. Achen and Snidal’s hypothetical cases do not include the U.S.-Canadian or the Dutch-Belgian ones, as they include cases in which there was already a hostile

³We use deterrence success and nonfailure as synonyms. From the absence of a challenge, one cannot necessarily conclude that the deterrent threat is the explanation.

relationship. Within the rivalry literature, the existence of militarized disputes has been taken as the behavioral indication that such hostile relationships exist. The concept of rivalry based on dispute incidence assumes that hostile relations continue between dispute outbreaks (something that may not always be true) and hence justifies considering all these disputes as part of the same rivalry.

Scholars can use the temporal duration of rivalries in addition to this pre-existing conflict requirement of deterrence to choose cases for testing general deterrence. The beginning of a rivalry becomes the case selection device: in concrete terms this could be, perhaps, the first one or two militarized disputes.⁴ The first dispute or two serves to ascertain the existence of “underlying hostility” (to use Lebow’s 1981 expression), the existence of general deterrence relationship. We take this as an indication that states have entered into a potential deterrence relationship. Indeed this is consistent with our notion of a quick “lock-in” to the rivalry relationship, discussed in the second half of the book. States realize that they are in it for the long haul only after a couple of crises or wars. The U.S. developed its deterrence strategy vis-à-vis the USSR after the crises of the mid-1940s. In deterrence terms, general deterrence strategies are put into place after some immediate deterrence failures (or successes). The population for testing is the rest of the rivalry life-cycle until the underlying hostility is gone (i.e., peace). We thus have a set of rivalries and dyad-years in which to study deterrence. Some of these rivalries will have no further disputes, whereas others will go on to become enduring rivalries with multiple crises and wars. We do not claim that such a population covers all deterrence relationships, but it does include hundreds of rivalries and thousands of years.

One can see this proposal as a continuation of the research direction taken by Huth and Russett (1993). They argued that enduring rivalries provide a suitable framework within which one can study general deterrence and solve the control group problem. Failure of immediate deterrence in a crisis results in serious military conflict or war. Failure of general deterrence thus becomes dispute or crisis initiation. The standard problem with studying general deterrence lies in determining the control group of deterrence success. The use of enduring rivalries permits this to be defined as nondispute years in the rivalry. The rivalry is assumed to continue in the interdispute periods, providing a control group of general deterrence nonfailures. Huth and Russett do not move beyond this case selection use of rivalries. Yet it is useful to step back and ask: on which populations should one test general deterrence? The main differences between the approach of Huth and Russett and ours are in that according to the latter (1) one excludes the rivalry formation disputes/years from the analysis, that is, those years during which the deterrence relationship is being set up, and (2) one uses not just enduring rivalries, but all “deterrence relationships” (which would

⁴There are other possible indicators that could indicate a rivalry has begun: alliance treaties, lower-level hostility, military planning data, et cetera. Unfortunately, the rivalry literature has relied on disputes, and thus we continue with this particular approach.

include some rivalries that do not become enduring because deterrence is successful in eliminating or minimizing militarized conflict in the relationship).

Standard methodology only looks backward at events before war, but studying deterrence involves looking at events after crises as well. This is crucial in testing deterrence theory. “He [Huth 1988] does not consider the possibility that a policy of unyielding firmness in one crisis might also lead a potential aggressor never again to challenge the status quo. . . . Statistical work would have to look *forward* from the cases, that is, analyzing future behavior, rather than just *backward*, that is assessing past behavior. Do states that practice conciliatory or bullying behavior get involved in fewer cases with the same adversaries in the future?” (Sagan 1991, 84). Huth and Russett respond to this criticism by including dyad-years until the end of the rivalry: “We begin a rivalry with the first publicly expressed claim of one government to the territory of the other and end it only when a formal agreement resolving competing claims is implemented[,] that is, when both parties accept it” (1993, 63). Rivalries end when they move out of the “war-prone” category, into what we have called “peace.” This usually happens sometime after the last dispute/war. Excluding these postcrisis, prepeace periods constitutes a serious selection bias.

Using rivalries to select cases in the manner outlined above identifies dyads that fit the “war-proneness” criteria that Achen and Snidal (1989) require much better other procedures that they (briefly) suggested. Other typical procedures, such as territorial contiguity (used to define “politically relevant” dyads), have a majority of unsuitable cases (as Lebow and Stein 1990 are quick to point out), because most contiguous dyads have peaceful relations. Selection bias most obviously may enter in through the initial disputes and wars that are used to select rivalries (when the deterrence relationship is being established), but which are then eliminated from actual testing.⁵

To see the grasp that the cross-sectional approach has on the comparative case study methodology, we look to studies that explicitly compare the same dyad over time. Lieberman’s (1994) study of the Arab-Israeli rivalry points in the right direction from a rivalry perspective: one needs longitudinal comparative case studies (Hybel 1990 and Hopf 1994 provide two examples of a comparative, longitudinal case study). The parallel with quantitative studies is striking—in both cases the cross-sectional approach is hegemonic. The point to recall here, underappreciated in the deterrence literature, is the assumption that states are already in a hostile relationship. Deterrence theory does not address why state A and state B get into a situation in which each needs to deter the other. This is not something specific to rational deterrence theory. If one

⁵In our analysis of enduring rivalries, we argue that the beginning of an enduring rivalry is when states lock-in to the rivalry. Testing deterrence during this period thus may lead to misleading conclusions about deterrence in an ongoing rivalry.

examines Lebow's model of crisis (1981, 337), one sees that it starts with "underlying hostility." The concept of a rivalry provides a natural way of testing hypotheses that make preexisting conflict assumptions.

Achen and Snidal (1989) criticize much of the deterrence literature for its failure to choose an appropriate control group—or any control group at all. They focused their attacks on comparative case studies of deterrence. But they do not provide a viable proposal about how to solve the problem, as their solutions focus more on large- N studies. The rivalry approach steps in to give some concrete suggestions. Huth and Russett (1993) illustrate how this works; we have just extended and developed the logic inherent in their research design. Central to Lieberman's (1995) argument is that deterrence was successful in some of the interwar years in the Egyptian-Israeli rivalry. Fundamentally his project involves comparing deterrence failures (i.e., wars) with deterrence success (i.e., nonwar periods). Thus, even within single-case studies, the rivalry approach can provide a control group without adding additional cases for comparison.

As we noted in the previous chapter, using rivalries as a case selection device should not be a substitute for theorizing; rivalry characteristics should be incorporated into the model rather than buried under case selection. We have made this case theoretically; now let us illustrate the methodological implications. These considerations go beyond selecting a suitable population or ensuring a proper control group, worthy goals in and of themselves. Implicitly, the case selection mechanism implies a necessary condition relationship: the rivalry is a necessary condition for the relationship or process. For example, some underlying conflict is necessary for an arms race to lead to war. There are many ways to model necessary conditions, one of which is through interaction terms (Goertz 1994). Crudely, this can be expressed as, for example, $ATTACK = \beta_i(DETERRENCE\ EFFECTIVENESS\ VARIABLES \times RIVALRY)$, where a score of zero on the rivalry variable eliminates the possibility of the process leading to war. For Huth and Russett (1993), this means that cases in which $RIVALRY$ is 0 are excluded from analysis, and when $RIVALRY$ equals 1 they are included. But a more reasonable interpretation of their theoretical framework would suggest that (1) there is no radical break between enduring and lower-level rivalries and (2) characteristics of the rivalry relationship influence the effectiveness of the deterrent threat. Point 1 means that one need not code rivalry dichotomously and that it might be reasonably treated as a continuous variable. Many of the "deterrence effectiveness variables" are in fact characteristics of the rivalry in question (point 2). But this is just our contention: rivalry is really part of the theoretical framework, and its theoretical role should not be obscured by its use as a case selection device. Thus, one may want to model the relationship as $\beta_i = f(RIVALRY\ CHARACTERISTICS)$. This states that the effectiveness of a factor in deterrence (i.e., β_i) depends upon the rivalry in which that deterrent factor is embedded. Although we regard case selection as

important, the rivalry approach also encourages greater theoretical attention to the underlying bases of case selection and therefore better specification of the model underlying the relationship under scrutiny.

Although the rivalry approach makes a significant contribution to case selection, its utility is not confined to this aspect of research design. Rivalries also have an impact on other testing aspects, and we turn to the longitudinal implications of the rivalry approach for testing in the next section.

Testing Hypotheses over Time in Rivalries

Rivalries exist over time, and this quality has some methodological implications for how to design research. In this section, we look at the longitudinal implications of the rivalry approach with particular reference to the use of interrupted time-series and multiple data points for analysis.

The temporal duration of an enduring rivalry proved crucial in Huth and Russett's (1993) testing of general deterrence. The continued existence of a rivalry in between militarized disputes and wars simultaneously provided a control group and defined deterrence success. But the methodological implications of the temporal duration of rivalries extend beyond this provision of a valid control group. In deterrence studies, the focus of analysis remained the dispute and war, more specifically dyad-years with or without disputes. We have taken a modest step away from the causes-of-war methodology, but once the case is chosen, the individual character of rivalries tends to disappear. Nevertheless, another application of the time-series component uses that element more directly. It takes advantage of the fact that we can compare rivalries before and after some critical event. We call this the interrupted time-series aspect of rivalry methodology. The democratic peace suggests just such an analysis. It implies that if a rivalry moves from a nondemocratic to a democratic dyad, it should end shortly thereafter or at very least exhibit considerably less conflict. The democratization of the dyad should fundamentally change the character of the rivalry relationship.

Similar to deterrence, the democratic peace literature has been plagued by the problem of defining an appropriate control group. Instead of using the interdispute periods as control groups (as in deterrence testing), the control period is the rivalry before democratization. Compare this simple solution to the problems encountered with cross-sectional causes-of-war methodology. Much of the debate revolves around the definition of "relevant dyads," pairs of states that could have a dispute (i.e., "dogs that could bark"). The usual solution is to take some combination of contiguity and (major) power status to define a control group (see also Lemke 1995; Reed 1998). Inevitably, this misses some cases and includes other improbable ones. With the interrupted time-series design for rivalries, the relevant dyad problem evaporates because one is using the same dyad in the test and control groups.

The interrupted time-series methodology shows the limitations of using just enduring rivalries for case selection. The democratic peace dramatically illustrates this because we would have only one true case to analyze if we restricted ourselves to enduring rivalries (see chapter 6). Furthermore, there is no reason to assume that the critical event will occur after the 20 years required for a rivalry to become enduring. War hypotheses rarely refer specifically to enduring rivalries. Hence, in terms of case selection, we need not limit ourselves to a particular subset of rivalries. We have seen this already in the deterrence case. None of the Huth and Russett hypotheses is necessarily specific to *enduring* rivalries; isolated and proto-rivalries fit, to varying degrees, within their theoretical framework.

The above point becomes critical in examining conflict management and termination hypotheses. Early democratization of a rivalry may prevent it from becoming an enduring one. Hence proto- and isolated rivalries are just as important, if not more so, than enduring rivalries in the research design. In terms of control group language, the isolated and proto-rivalries are “successes” compared to the enduring rivalries, which represent conflict-ridden failures. Because proto- and isolated rivalries do not become enduring ones (i.e., the cycle of militarized conflict is broken), they might be taken as evidence that general deterrence or conflict management was effective. Even within enduring rivalries, the interrupted time-series design can be useful. For example, Gibler (1997) investigated the impact of alliances that were also territorial settlement treaties on rivalry behavior. He noted that rivals have significantly fewer disputes and less chance for war after such a settlement than in its absence. The comparison of the same rivalry prior to the key event (and which establishes a baseline of behavior) can be compared with rival behavior after the key event, here the removal of territorial dispute issues from the relationship.

Implicitly or explicitly, the interrupted time-series characteristic lies at the heart of tests of conflict termination hypotheses. Sometimes this remains implicit, as with event history analysis (Allison 1984; Yamaguchi 1991). For example, Bennett (1996) defines his dependent variable as the end of the enduring rivalry. Implicit is a comparison of that end, hence future peaceful relations, with the rivalry period. If one were to use event data (Hensel 1997), one might expect to see a standard interrupted time-series effect of reduced conflict after the critical event. Of course, other hypotheses naturally lend themselves to interrupted time-series tests. For example, this methodology provides a natural test of the power transition hypothesis. One would compare the power transition periods to the nonpower transition phases of the rivalry in order to evaluate whether the behavior of the two states differs substantially.

The temporal component of rivalry also can enter more directly into the testing of theory. We have seen that the Huth and Russett research design uses the interdispute period to define a control group of general deterrence success. Focusing on the temporal aspect of a rivalry suggests other dependent variables.

One aspect of general deterrence success is how long it works until it fails the next time (or the rivalry ends).

One of the key aspects of putting deterrence into the context of rivalries is that it permits us to ask what happens following instances of deterrence. The standard approach examines states of affairs before the deterrence success or failure. There is no real afterlife. Except for the endgame of rivalries, there is always an "after." This "after" period can provide new ways to evaluate hypotheses. In most instances, the "after" is also a "before." Thus, one measure of the effectiveness of general deterrence strategies is how long they are successful. Huth and Russett (1993) use disputes/war occurrence as the definition of general deterrence failure. One aspect of general deterrence is when it breaks down, but another is explaining how long it succeeds. Here we have a new dependent variable: "the waiting" time until the next deterrence breakdown.

The same definition of success can be used for conflict management analyses. Bercovitch and Diehl (1997) look at the impact of international mediation attempts on conflict within enduring rivalries. They are not so much concerned with the immediate success of such efforts, but rather with a longer-term impact on rivalry conflict patterns. Thus, they choose as dependent variables the occurrence of future disputes, the severity of those disputes, and the waiting times until those next disputes. We perform similar analyses in chapter 10. The focus on rivalries, and their longitudinal character, permits this kind of analysis, whereas more static approaches would limit the choice of dependent variables to those immediately surrounding the mediation attempt or crisis that precipitated the intervention.

Finally, we note that dynamic hypotheses, almost by definition, require time-series data. Much in deterrence theory refers to credibility and reputation, which are created and modified as states interact. The most obvious way that the historical experience of the rivals comes into deterrence models is through the concept of reputation and threat credibility. Reputation itself makes explicit reference to past behavior. One can say that the credibility of a current deterrent threat depends, in part, on the behavior of the state in previous (deterrence) situations.

In these generic formulations, rivalries do not necessarily play a role. One can search all previous or recent deterrence situations to help evaluate the credibility of the current threat. Rivalries can enter into the evaluation by suggesting that all this information is not equally valuable. In particular, the experience within the rival dyad will outweigh evidence from other interactions. Both Lieberman (1994) and Huth (1988) take this approach. It is previous interactions within the same dyad (i.e., with the same rival) that influence beliefs. This is all the more convincing in enduring rivalries. In many of these cases, the attention of decision makers and bureaucracies is focused primarily on the behavior of the other rival. The emphasis of rivalries on dyadic history parallels the

psychology and organizational behavior of states and decision makers. Investing in a reputation only makes sense—in a cost-benefit sense—if the likelihood of future interactions is high: one invests (pays costs) in a reputation in order to reap future gains. If there are no future gains, then the initial investment is a loss. One desirable conceptual definition of an enduring rivalry states that it exists when decision makers see themselves in a state of militarized conflict for the indefinite future, and make plans assuming that future crises are likely. These rivalry dyads then constitute exactly the sort of situation in which leaders find it worthwhile to invest in reputation.

In our reading, the principal way that the history and context of rivalries enter into rational deterrence theories is through credibility and reputation effects. Part of the theoretical implications of the rivalry approach for deterrence lies in the implication that there are perhaps other ways in which current crisis behavior depends on the rivalry context. For example, Huth (1988) includes two rivalry variables, what he calls “past behavior” of defender and attacker. Given the natural emphasis on deterrence strategy, he investigated if previous strategies influenced outcomes in current cases. Hence, he differentiated between “conciliatory” and “intransigent” strategies. He found no difference between strategies; both decreased the likelihood of extended deterrence success. From the rivalry perspective, the first question we would ask is about the mere existence of previous conflict. Huth’s results support the hypothesis that the more well entrenched the rivalry, the harder it is to deter the other side. This becomes clear when one understands that Huth coded cases with no history as 0 and cases with some history as 1. Combining the past behavior variables permits the quite legitimate interpretation of his analysis as a test of “rivalry effects.”⁶ Yet a broader rivalry perspective would suggest a more refined measure of those effects, including perhaps multiple variables (based on a given theoretical framework) and accompanying interval level measures. Some of those variables might include the outcomes of previous confrontations (e.g., compromise, capitulation, stalemates) or the relative success of the bargaining strategies employed (e.g., coercive, accommodative), to offer a few possibilities.

Rational deterrence theory tells us that deterrence success depends on the cost-benefit analysis for each crisis. Threats become more credible and the level required for an effective threat increases as the rivalry continues and becomes even better entrenched. This suggests that there are other aspects between crises than just credibility and reputation. How does one get at the motivational aspects of deterrence, the costs one side is willing to bear, and hence the magnitude required for a successful deterrent threat? In the “contextless” world of many rational choice models, these are exogenously determined. Yet in the real world of state relations, they are determined, in part, by the past relations of

⁶Fearon (1994b) explicitly confirmed this in his reanalysis of Huth’s 1988 data. In recent work on territorial enduring rivalries, Huth (1996a) also found that a history of militarized disputes was strongly associated with conflict escalation.

the rivals. To say that current credibility and reputation depend on past behavior is to put forth some sort of learning or cognitive theory. Lieberman (1994) proposes that it is through deterrence failure (i.e., actually having to carry out deterrent threats) that a rival learns that a threat is credible. He argues that we should not be surprised to see deterrence failure early in the rivalry followed by deterrence success.

In summary, much in deterrence as well as crisis theory links crises to each other (reputation and credibility). This, according to some, has been overemphasized: “because the second wave [of deterrence theory] overestimates the importance of commitment and the extent to which outcomes are interdependent, it exaggerates the costs of retreats and the advantages of victories” (Jervis 1979, 319). On the other hand, these links have not been much studied in the quantitative literature: “another aspect of deterrence theory requiring further research is the role of reputation, or the impact of past behavior of the adversaries. This variable has received little attention in the quantitative empirical literature on deterrence, not only because of the difficulty of constructing an operational indicator for a large- N study but also because of the limited development of this variable in the theoretical literature” (Levy 1988, 511).

It is only through looking at behavior over time, and ideally in rivalries, that we are able to pick up this behavior. And this requires different methodological approaches, most notably the adoption of dynamic analyses and models. Most suitable would be a model similar to that of Muncaster and Zinnes (1993), which permits an updating of behavior over time as past interactions and their outcomes serve to redefine the current relationships of pairs of states in a given system. In their model, the outcome of a dispute between two states, A and B, influences the relative hostility level in their future relationship. Furthermore, the dispute between A and B may also serve to alter the relationships between other pairs of states in the system. This type of model, with some modification, provides the framework under which the past, and issues of credibility and reputation, can be analyzed.

We can contrast this with the interrupted time-series application of rivalries. In that situation, we divide the rivalry by the occurrence of some critical event. Or, we can break rivalries into beginning, middle, and end, suggesting that these contexts matter in terms of the impact of important variables. In a related theoretical literature, almost all theories of crisis include “context,” “background,” or “past experience” variables.⁷ For example, Brecher (1979) frequently described “past experience” as a core input into various phases of the crises he analyzed. Rivalries provide one way to begin to specify *what* background factors matter and *how* they influence crisis behavior. Vaguely, these authors suggest that it matters exactly when within the rivalry the crisis takes

⁷Not all studies, however, do this. For example, George and Smoke (1974) include no past behavior variables in their list of questions.

place. The hypotheses remain almost always very fuzzy and not explicitly formulated. With the rivalry concept in hand, the possibility of investigating these kinds of relationships now exists. The rivalry approach provides not only new tests for hypotheses that have already been exposed to intensive quantitative testing, but suggests means for quantitative or qualitative testing of hypotheses that have been around for a while with no tests. Here, we can include as another example the frustration-aggression hypothesis in the deterrence literature, which remains untested, to our knowledge.

Instead of using the interdispute period as part of the dependent variable, it appears as the independent variable in the “frustration-aggression” deterrence hypothesis: can early deterrence success lead to failure in later crises? This pattern fits easily into the frustration-aggression framework. Deterrence success means that the rival, “frustrated” by the successful threat, becomes more aggressive in the next dispute. “Deterrence can severely frustrate an adversary who is strongly motivated to change a status quo that he regards as invidious, especially when he feels it legitimate to do so. The consequences of continued frustration are not easily predictable and are not always favorable to the deterring power. Deterrence success in the short run is not always beneficial in the longer run; the adversary may become more desperate to mount a challenge and may proceed to acquire greater resources for doing so” (George and Smoke 1989, 182). “The success of immediate deterrence can weaken general deterrence in the future by increasing the loser’s grievances, convincing it that the state is a grave menace, increasing its incentives to stand firm in the next confrontation or even leading it to fight in order to change a situation which has become intolerable or is expected to deteriorate” (Jervis 1989, 199). This contrasts with the standard hypothesis that deterrence success leads to more success in the future.⁸ By looking at rivalries over time and adopting a dynamic approach, perhaps we can understand why some patterns appear in certain rivalries.

These issues imply another general class of time-series analyses: the tests of the beginning of rivalries versus the middle versus the end of rivalries. All cross-sectional time-series analyses ignore when the dispute takes place within the lifetime of the rivalry. But implicit in many deterrence hypotheses lies the idea that deterrence operates differently in a mature period of an enduring rivalry than in its beginning. One would like to see an evaluation of Huth and Russett’s (1993) model in terms of different rivalry periods: does it work equally well in all of them?

The longitudinal research design may also be essential for detecting empirical patterns and drawing proper conclusions about a given phenomenon. For example, consider a study of deterrence that analyzes 10 deterrence attempts in 10 rivalries, with those deterrence attempts successful 90 percent of the time.

⁸Other explanations can be given for the success-then-failure pattern, such as “designing around deterrence,” complacency, and the like.

TABLE 5.2: Patterns of Deterrence Failure in Rivalries: Pattern 1

Rivalry	Dispute									
	1	2	3	4	5	6	7	8	9	10
1	X	O	O	O	O	O	O	O	O	O
2	O	X	O	O	O	O	O	O	O	O
3	X	O	O	O	O	O	O	O	O	O
4	O	X	O	O	O	O	O	O	O	O
5	X	O	O	O	O	O	O	O	O	O
6	X	O	O	O	O	O	O	O	O	O
7	X	O	O	O	O	O	O	O	O	O
8	O	X	O	O	O	O	O	O	O	O
9	X	O	O	O	O	O	O	O	O	O
10	X	O	O	O	O	O	O	O	O	O

Note: O = success, X = failure.

A typical static and cross-sectional analysis would consider only circumstances surrounding the different deterrence attempts and ignore any interconnection between events. Yet what if the pattern of deterrence failure was evenly spread out among the 10 rivalries, but all the failures occurred near the outset of those rivalries, as illustrated in table 5.2. The longitudinal design would permit a scholar to identify this pattern and likely draw a different set of conclusions than a cross-sectional design would. Several possibilities include that deterrence failure is very likely in the early stages of any relationship (as suggested by Lieberman 1994). More sophisticated would be a model of deterrence based on learning or signaling in which several interactions and at least one failure are needed in order to establish a stable relationship.

Another possible pattern in deterrence failure is that those failures are concentrated in a small number of rivalries, two according to the example given in table 5.3. From this design, one might conduct queries that investigate differences across rivalries as opposed to differences across disputes, the latter of which is suggested by a cross-sectional design. More than just using the interrupted time-series design, the longitudinal design of the rivalry approach in general will assist the analyst in detecting patterns and interpreting results that are not possible, even discouraged, by the traditional causes-of-war approach.

We view the testing possibilities offered by the time-series aspect of rivalry as a framework for addressing the various critiques made about the deterrence literature, but also for a wider range of other conflict concerns. The rivalry approach gives a methodological framework for testing numerous conflict, conflict management, and conflict termination hypotheses. This becomes possible because of two fundamental characteristics of rivalries: they exist over time

TABLE 5.3: Patterns of Deterrence Failure in Rivalries: Pattern 2

Rivalry	Dispute									
	1	2	3	4	5	6	7	8	9	10
1	O	O	O	O	O	O	O	O	O	O
2	O	O	O	O	O	O	O	O	O	O
3	O	O	O	O	O	O	O	O	O	O
4	O	O	O	O	O	O	O	O	O	O
5	X	X	O	O	X	O	O	X	O	X
6	X	O	X	O	X	O	O	X	X	O
7	O	O	O	O	O	O	O	O	O	O
8	O	O	O	O	O	O	O	O	O	O
9	O	O	O	O	O	O	O	O	O	O
10	O	O	O	O	O	O	O	O	O	O

Note: O = success, X = failure.

and they eventually move into nonrivalry phases. These characteristics allow one to define new tests that avoid many of the problems that plague traditional cross-sectional analyses, particularly those that need to define control groups composed of nonconflict dyads.

Cross-Sectional Aspects of Rivalry Methodology

In addition to its time-series components, the rivalry approach provides new cross-sectional analyses. By now the reason for this should be clear. We started our volume with an analysis of our grounding concept: militarized rivalries. We suggested that they could serve as a fundamental object of study, analogous to militarized international disputes or crises. The natural tendency of the causes-of-war methodology is to use rivalries in cross-sectional time-series, with the dyad-year as the unit of analysis. To emphasize the additional testing possibilities that the rivalry approach gives, we generally rely on neither the dispute nor the dyad-year as the statistical or theoretical focus of analysis in the our study of the democratic peace in the next chapter. Our purpose is not to argue that standard methodologies are wrong, but rather that other testing possibilities exist. We see this as an exercise in enlarging the methodological tool-kit of students of international peace and conflict.

The democratic peace is usually defined in terms of the absence of war between two democratic states, but one can reformulate the question in terms of rivalries: what flows from the democratic peace framework and applies to rivalries? The answer to this question can influence our view of democratic peace phenomena. For example, if we have 20 different of disputes between democratic states, they can form rivalries in numerous ways. These range from 20

isolated, one-dispute rivalries to one 20-dispute enduring rivalry. Clearly, the former supports the democratic peace proposition more than the latter does. We explore some of these possibilities and provide actual numbers to these formulations in the next chapter.

The cross-sectional possibilities of the rivalry approach remain tremendously underexploited in the rivalry literature. Being generous, one can find only a handful of studies that test hypotheses with explicit use of the rivalry as the primary focus of analysis. For example, Gibler (1997) looked at the effect of territorial settlement treaties on the termination of rivalries. In those cases, however, the primary subject of those analyses was rivalries. That is, he was attempting to draw generalizations about rivalries and their dynamics and thus the choice of rivalries as the focus of analysis was obvious.

We think it makes sense to start with the rivalry as the phenomenon to be explained before moving to cross-sectional time-series. In chapter 11, we argue that political shocks are almost a necessary condition for enduring rivalry development and termination. A first step is to ascertain the percentage of rivalries that terminated at the time of a political shock. The N here is obviously the number of enduring rivalries. But it also illustrates the conceptual hold of traditional methodologies. We know of very few analyses that report statistics using rivalry as the unit of comparison, even simple ones like the proportions reported by us in that chapter. For example, Bennett in his studies (e.g., 1997a) never provides simple statistics with rivalry as the unit of analysis, it is always event history analysis with dyad-years.

Rivalry as the unit of analysis is an essential tool for the evaluation of cross-sectional time-series results. For example, Huth and Russett examined the predictive power of their model across disputes, but not across rivalries (1993, 69). From the rivalry viewpoint, the cross-sectional question can be stated: "Does their model work better for some rivalries than others?" Knowing which rivalries fit and which do not fit their model provides us with crucial information, both in order to evaluate the theory as well as to suggest improvements, as we have seen above.

All cross-sectional time-series analyses need to be complemented by an examination at the level of the rivalry. This principle is important for theoretical and statistical reasons. Most researchers know about autocorrelation problems and their solutions, but the same potential problem exists over space. If the residuals for one rivalry are all positive, then independence of error term assumptions are violated and parameter estimates are biased (see Raknerud and Hegre 1997). Theoretically, knowing that the model works poorly for some rivalries likely implies that improvements to the model are necessary.

Between the extremes of the rivalry-as-a-whole and the dyad-year lie other cross-sectional possibilities. We have mentioned that the democratic peace has been plagued with problems of defining "relevant dyads" and how to code multiple-year wars. The rivalry approach suggests that rivalry equivalent of

the dyad-year can be peaceful and militarized periods. Instead of dividing a relationship into years and coding them as war and peace, we break relations into rivalry and nonrivalry periods. Each period counts as one observation. Although we still have to define relevant dyads, the dubious ones (e.g., France vs. Bolivia) count only as one case (assuming there are no disputes between the two) instead of the x number of dyad-years under usual procedures. Instead of the huge N s, with which statistical significance is all but assured, as one encounters in studies of the democratic peace (e.g., $N = 185,000+$ for Farber and Gowa 1995), we get a more reasonable universe of cases.

In summary, the cross-sectional unit can take various forms. We have seen four ways that this can work. First, we take the rivalry as the unit of analysis. Second, we divide relationships into rivalry and nonrivalry periods. Third, we divide the rivalry into beginning, middle, and end periods. Finally, of course, we can split rivalries into dyad-years. We have focused our attention on the first three because they are radically underrepresented in the current empirical literature.

We emphasize again that cross-sectional analyses frequently should include all rivalries, not just enduring ones. Otherwise, the hypothesis may not be fully tested. In particular, conflict termination propositions need to consider proto- and isolated rivalries in order to get a clear evaluation of the hypothesis. This relates to the yet unexplored question of why some rivalries become enduring while others do not. Here the proto- and isolated rivalries become the control group. The survival analysis of Cioffi-Revilla (1998) shows that using only enduring rivalries can significantly bias the results. He found that if one looked only at enduring rivalries, one would claim that, counterintuitively, they weaken over time. Once all rivalries are included in the analysis, however, it becomes clear that this is what happens at the end of enduring rivalries, not necessarily what happens before rivalries reach maturity (i.e., approximately 20 years).

Conclusion

The rivalry approach addresses methodology, research design, and testing procedures, with implications that apply to virtually all standard (and probably nonstandard) hypotheses in the international conflict literature. We have just focused primarily on deterrence studies to make our point. A number of studies have utilized rivalries in case selection and testing, but only a small proportion of the possibilities have been exploited in the process. Although rivalries have been used for testing for over a decade, it is rare to find studies that include both time-series and cross-sectional analyses (an exception is Gibler 1997).

International conflict phenomena are multifaceted; by definition, different dependent variables explore different facets. Our argument is that the rivalry approach illuminates aspects of deterrence, the democratic peace, and other classic hypotheses hidden from traditional methodologies. Many of these

new testing procedures investigate the implications of some general proposition. The general democratic peace idea has many wide-ranging ramifications. As scholars investigate them, we understand better what the democratic peace is. In terms of deterrence, using rivalries can help to understand how deterrence success and failure relate to one another. Because of the multidimensional character of these phenomena, there is no single “rivalry test,” as multiple possibilities exist in every circumstance. Given the character of our data, indicators, and theory, this proves to be a real plus. If results of these multiple tests—time-series to cross-sectional and combinations of the two—converge, then our confidence in the results should be strongly reinforced.

We have argued that beyond providing additional tests, the rivalry approach solves some serious testing problems. For example, it provides a framework for case selection in the controversial deterrence literature, with much more validity than current approaches. Similarly, using rivalries to examine the democratic peace helps avoid many problematic features of standard methodologies. We conclude by stressing that many of these techniques have hardly begun to be exploited, not to mention the research design possibilities that remain unknown to us. We see the next chapter as an extended, but preliminary, exploration of some of the proposals we have put forth here. Again, they do not exhaust the methodological vein that constitutes part of the rivalry approach. One must evaluate theoretical approaches on multiple dimensions. On the methodological, research design dimension the rivalry approach gives us new—and often better—means for getting an empirical grasp on many significant hypotheses in the war-and-peace literature. For this reason alone, it merits attention and respect.

