A Unified Theory and Test of Extended Immediate Deterrence

Curtis S. Signorino
Ahmer Tarar
University of Rochester

Work in progress
Comments welcome

December 7, 2001

Abstract

We present a unified theory and test of extended immediate deterrence — unified in the sense that we employ our theoretical deterrence model as our statistical model in the empirical analysis. The theoretical model is a straightforward formalization of the extended immediate deterrence logic in Huth (1988) and Huth and Russett (1984, 1988), coupled with private information concerning utilities. Contrary to Huth (1988), our empirical analysis suggests that alliances, nuclear weapons, military arms transfers, and foreign trade all affect deterrence success. Our model correctly predicts over 93% of the potential attacker’s actions and almost 88% of the crisis outcomes. Finally, we find evidence that the likelihood of deterrence success and of war are not monotonically related to the variables involved in the deterrence calculus, which contradicts a fundamental assumption of most previous studies.
Contents

1 Introduction 3

2 A Strategic Model of Extended Immediate Deterrence 4
  2.1 Uncertainty Concerning Utilities ........................................... 6

3 Empirical Analysis 9
  3.1 Variables and Data ................................................................. 11
  3.2 Strategic Probit Analysis .......................................................... 13
    3.2.1 The Defender’s Utility for War .......................................... 15
    3.2.2 The Attacker’s Utilities ...................................................... 15
    3.2.3 The Probability of Deterrence Success and of War .................. 16
    3.2.4 Model Fit ........................................................................... 24

4 Concluding Remarks 25
1 Introduction

What factors affect deterrence success or failure? The deterrence literature is one of the most exhaustive in international relations, and the logic of deterrence has been extensively studied within both government and academia by scholars from a variety of disciplines. Scholars have investigated the impact of conventional and nuclear balance of forces, interests at stake, reputation from past crises, crisis bargaining strategies, military alliances, geographic contiguity, degree of uncertainty, international system structure, and domestic politics (George & Smoke 1974; Alexandroff & Rosen- crance 1977; Waltz 1981; Mearsheimer 1983; Huth and Russett 1984; Betts 1985, 1987; Huth 1988; Mueller 1989; Huth 1990; Waltz 1990; Weber 1990; Langlois 1991; Huth, Gelpi, Bennett 1993; Huth & Russett 1993; Fearon 1994; Hopf 1994; Paul 1995).1 The logic is continuously put under the microscope of rigorous empirical testing, and subsequently refined. It is no wonder, then, that even the informal rational deterrence literature tends to be transparent in its logic, with much attention paid to the sequencing of moves and to the incentives and expected behavior of other states (see, for example, George and Smoke 1974, 101-3).

Recent research by Signorino (1999) and Signorino and Yilmaz (2000), however, suggests that previous empirical tests of deterrence theories are highly problematic. The heart of the problem is that deterrence is generally considered to be a strategic interaction, but is empirically investigated using non-strategic statistical models such as logit and probit.2 Signorino (1999) demonstrates how failure to incorporate strategic interaction into statistical tests results in faulty inferences. Signorino and Yilmaz (2000) mathematically prove that using logit to analyze data generated by strategic interaction induces the equivalent of omitted variable bias. The upshot of this recent methodological research is that a statistical model needs to be structurally consistent with the theory it is testing. Strategic models imply a particular structural relationship between the regressors and the dependent variable. Typical logit and probit models imply a different structural relationship.

In this study, we present the first unified theory and test of extended immediate deterrence — unified in the sense that we employ our theoretical deterrence model as our statistical model in the empirical analysis. The theoretical model is a straightforward formalization of the extended immediate deterrence logic in Huth (1988) and Huth and Russett (1984, 1988), coupled with private information concerning utilities. We construct our deterrence model in such a way that it guarantees positive probabilities over all actions and outcomes and, therefore, can be used in statistical estimation. That is, our theoretical model is our statistical model.

---

1 For a collection of articles that illustrate the diverse array of approaches that have been used to study deterrence, see also Stern, Axelrod, Jervis, and Radner (eds.) 1989.

2 We use the term “strategic” in the usual game-theoretic sense where players must condition their behavior on the expected behavior of others.
We analyze this model using data from Huth (1988) and Huth and Russett (1988). Contrary to Huth (1988), our empirical analysis suggests that alliances, nuclear weapons, military arms transfers, and foreign trade all affect deterrence success. In contrast to Huth and Russett (1988), we find that the latter three variables, as well as the immediate balance of forces, influence the defender’s decision to defend its protege. In terms of model fit, our model correctly predicts (actually, postdicts) over 93% of the potential attacker’s actions and almost 88% of the crisis outcomes. Finally, we find evidence that the likelihood of deterrence success and of war are not monotonically related to the variables involved in the deterrence calculus. This contradicts a fundamental structural assumption of previous studies using logit and probit.

The paper proceeds as follows. In the next section we present our theoretical model, which is a straightforward formalization of some of the extended immediate deterrence literature, coupled with private information concerning payoffs. Following that, we specify the utilities of the model in terms of regressors. Using data from two previous studies of extended immediate deterrence, we then conduct the empirical analysis. We discuss the factors that influence deterrence success and the decision to go to war, and assess model fit. We conclude by summarizing the results and noting possible avenues for future research.

2 A Strategic Model of Extended Immediate Deterrence

A distinct advantage of the relatively transparent rational deterrence logic is that it allows for straightforward translation into a formal model. An excellent example is the literature on extended immediate deterrence (Huth 1988, 1990; Huth and Russett 1984, 1988). In extended immediate deterrence, a “defender” nation is trying to deter a potential aggressor from attacking one of its allies or “proteges.” Henceforth, we will refer to the defender nation simply as the “defender,” the potential aggressor as the “attacker,” and the defender’s ally or protege that is being threatened simply as the “protege.” The deterrence situation is considered “extended” in that the defender is trying to deter an attack on a third nation rather than on itself, and “immediate” in that the attacker has made threats and the defender counterthreats, so that the deterrence attempt takes place in a crisis atmosphere in which the use of force may be imminent (for the distinction between “immediate” and “general” deterrence, see Morgan 1983; Danilovic 2001). Of primary interest in this literature is the interaction between the attacker and the defender.

Figure 1 displays this interaction in the form of a simple extensive form game. Here, the (potential) attacker can either attack (A) or not attack (Ā). If the attacker chooses not to attack, the deterrence success results in a status quo (SQ) outcome. If, on the other hand, the attacker chooses to attack, deterrence has clearly failed, and the defender must decide whether to come
Here, a (potential) attacker must decide whether to use force against the defender’s protege. If deterrence succeeds and the attacker does not attack ($\overline{A}$), the status quo (SQ) results. If deterrence fails and the attacker attacks the protege, then the defender must decide whether to aid its protege. Defending against the attacker results in war (War). Not defending results in capitulation (Cap).

to the aid of its protege. If the defender chooses to defend ($D$) against the attacker, war (War) results. If the defender does not defend ($\overline{D}$) its protege, then we regard the defender as having capitulated (Cap).

It is certainly true that more complicated formal deterrence models have been developed than that depicted in Figure 1 (see, for instance, Fearon 1994; Kilgour and Zagare 1991; Kugler and Zagare 1987; Powell 1990; Werner 2000; Zagare and Kilgour 1993, 2000). However, we employ this model for a number of reasons. First, we believe that it most closely represents the logic of the extended immediate deterrence literature (Huth 1988, 1990; Huth and Russett 1984, 1988, 1993; Wu 1990). Second, not only has this literature undertaken rigorous empirical testing, but data exists for testing the model presented in Figure 1. This is not a trivial issue, given that most data collection in international relations (and political science more generally) has been undertaken without regard to the structure of formal models. Third, given that this study represents the first example of a unified theory and test of deterrence, we prefer to begin with a simple model rather than a more complex one. Achen and Snidal (1989, 151) verbally describe this model, calling it “the simplest version of rational deterrence theory.” Zagare and Kilgour (2000, Ch 3) formally analyze complete and incomplete information versions of this model. This, therefore, seems to be a good place to start. Finally, although selection effects are always an issue in samples generated by individuals making choices, Signorino (2001) suggests that correctly modeling strategic calculations is generally more important than modeling correlated errors (as in typical selection models). In
Figure 2: The Deterrence Model with Uncertainty Concerning Utilities. The utilities shown in the figure represent each player’s true utilities, broken into their observable $U_i(\cdot)$ and unobservable $\pi_{ij}$ components. For example, let the defender’s true utility for war be written as $U_d(War) = U_d(War) + \pi_{d1}$. The attacker observes $U_d(War)$, but knows only the distribution of the unobserved $\pi_{d1}$. We assume that the analyst also does not observe the $\pi_{ij}$.

In other words, failure to model the correlation between decisions in this model and the choices made prior to it is usually far less deleterious than failure to model the strategic deterrence calculations within Figure 1. As we will demonstrate later, this simple model actually appears to go a long way in explaining extended immediate deterrence outcomes. However, before proceeding to the empirical analysis, we must further specify the model.

### 2.1 Uncertainty Concerning Utilities

It is unlikely that the participants of a deterrence crisis (or almost any situation, for that matter) perfectly observe each other’s utilities. It is also unlikely that the analyst, in conducting the empirical analysis, can perfectly specify the actors’ utilities. Fortunately, relaxing this assumption not only provides a model that is more satisfying theoretically, but also one that can be used as the basis of our statistical estimation.

Figure 2 displays the same strategic situation as in Figure 1, but assumes that the attacker and defender do not perfectly observe each other’s utilities. We also assume that the analyst does not perfectly observe the actors’ utilities. Instead, we assume that the true utility for an outcome can be represented as consisting of an observable component and an unobservable (or private) component.
For example, let the defender’s utility for war be represented as

\[ U^*_d(War) = U_d(War) + \pi_{d4} \]

where \( U^*_d(War) \) is the defender’s true utility for war, \( U_d(War) \) is the component of the true utility that the attacker and the analyst observe, and \( \pi_{d4} \) is the component that is private information to the defender. From the attacker’s and analyst’s perspective, \( \pi_{d4} \) is a random variable. We assume that the attacker and the analyst know only the distribution of \( \pi_{d4} \).

If, as depicted in Figure 2, we make this assumption concerning each of the players’ utilities, we can derive equilibrium choice probabilities for each of the actions and outcomes in the game (see Signorino 2000 for details on deriving the choice probabilities of various strategic probit models). We assume that the payoff perturbations (i.e., the \( \pi_{ij}s \)) are independently and identically distributed \( N(0,\sigma^2) \). Let \( p_d \) denote the probability that the defender defends its protege and \( p_a \) the probability that the attacker attacks the protege. Conversely, let \( p_\pi \) and \( p_\pi^* \) denote the probabilities that the defender does not defend and that the attacker does not attack, respectively. Assuming that the actors maximize their true (expected) utility at their decision nodes, the strategic probit choice probabilities for the deterrence model in Figure 2 are easily derived as

\[ p_d = \Phi \left[ \frac{U_d(War) - U_d(Cap)}{\sqrt{2}\sigma^2} \right] \]

\[ p_a = \Phi \left[ \frac{p_dU_a(War) + p_\pi^*U_a(Cap) - U_a(SQ)}{\sqrt{\sigma^2 \left( p_d^2 + p_\pi^2 + 1 \right)}} \right] \]

where \( \Phi(\cdot) \) is the standard Normal cumulative distribution function, and where \( p_\pi = 1 - p_d \) and \( p_\pi^* = 1 - p_a \).

Notice that the equilibrium choice probabilities reflect the extended immediate deterrence logic of the extensive form game and the uncertainty of the players concerning each other’s payoffs. The numerators of Equations 1 and 2 express the difference in observed expected utility for the options associated with each decision node. For example, the probability \( p_d \) that the defender aids its protege is an increasing function of the difference in the defender’s observed utility for war relative to its observed utility for capitulation: the higher the defender’s observed utility for war relative to capitulation, the higher the probability (from the attacker’s and analyst’s perspective) that the defender will defend its protege.

Similarly, the numerator of Equation 2 is simply the difference between the attacker’s observed expected utility for attacking and its observed utility for not attacking. The attacker’s observed expected utility for attacking, \( EU_a(A) \), is a lottery over the capitulation and war outcomes, based on the attacker’s belief \( p_d \) about whether the defender will defend its protege: \( EU_a(A) = p_dU_a(War) + \)
\( p_a U_a(Cap) \). The higher the attacker’s observed expected utility for attacking relative to its observed utility for the status quo, the higher the probability (from the analyst’s perspective) that the attacker will attack.

The denominator of each probability equation is a variance term, reflecting the amount of uncertainty regarding the unobserved component of the true utilities. A large \( \sigma^2 \) relative to the observable components reflects greater uncertainty on the part of the actors and the analyst, resulting in strategic choice probabilities closer to a coin toss over the options at each decision node. When the players and the analyst have more accurate information about the true utilities — i.e., when \( \sigma^2 \) is small — the choice probabilities approach 0 and 1, and the deterrence model in Figure 2 approaches that of a game of perfect and complete information. Note that when \( \sigma^2 = 0 \), the game is exactly one of perfect and complete information, and our assumption that states maximize their utility at each decision node implies subgame perfection.

It should also be noted that Equations 1 and 2 do not represent mixed strategies. Rather, they are the beliefs of the attacker and the analyst, based on their assumptions of utility maximizing behavior, uncertainty concerning the \( \pi_{ij} \), and the structure of the game. \( p_d \) is the belief of both the attacker and the analyst about whether the defender will fight. \( p_a \) represents the analyst’s belief about whether the attacker will attack, given the attacker’s (and analyst’s) belief about whether the defender will defend. Except in a few knife-edge situations, the underlying behavioral model assumes that the attacker and the defender play pure strategies from their perspective.\(^3\) The twist (relative to conventional game theory) is that the empirical analyst is assumed to know only the distribution of the \( \pi_{ij} \). Therefore, the analyst can only make probabilistic statements about the equilibrium choices.

With that said, the equilibrium outcome probabilities follow directly from the action probabilities. Let \( p_{sq} \), \( p_{cap} \), and \( p_{war} \) be the probabilities of the status quo, capitulation, and war outcomes, respectively. Because of the independence assumption, the probability of any given outcome is simply the product of the action probabilities along its path. Hence,

\[
\begin{align*}
    p_{sq} &= p_a \pi \\
    p_{cap} &= p_a p_d \\
    p_{war} &= p_a p_d
\end{align*}
\]

We now have a strategic deterrence model that is also a statistical (i.e., probabilistic) model. As long as there is some uncertainty concerning the true utilities (on the part of the states and the analyst), we are guaranteed positive probabilities over all actions and all outcomes in the model, and

\(^3\)The defender chooses \( D \) if and only if \( U_d(War) > U_d(Cap) \). The attacker chooses \( A \) if and only if \( p_a U_a(War) + p_d U_a(Cap) > U_a(SQ) \).
we can use this theoretical model directly in our statistical estimation. In doing so, the deterrence theory and its test are unified.

3 Empirical Analysis

The typical empirical analysis, not only in the deterrence literature but in much of the international relations literature, begins with a list of hypotheses drawn from extant theory. In these cases, the hypotheses to be tested almost invariably involve unconditionally monotonic relationships between the dependent variable and the regressors. In the current context, an example of such an hypothesis would be

\[ H: \text{The likelihood of war decreases as the balance of forces increasingly favors the defender.} \]

As Signorino and Yilmaz (2000) show, even the simplest strategic model often implies nonmonotonic — or at least only conditionally monotonic — relationships between the dependent variable and the regressors. That our strategic theories often imply nonmonotonic or conditionally monotonic relationships suggests that typical hypothesis statements are problematic, especially when there is no clear link from the hypothesis to a well-specified model. Because of that and because we have a statistical model that is also our theoretical model, we take a slightly different approach here.

In conducting hypothesis tests, we are usually interested in assessing whether some explanatory variable has an effect on the the phenomenon of interest. In the present context, we might hypothesize that the balance of military forces or the defender’s possession of nuclear weapons affect the attacker’s decision to attack. Alternatively, we might hypothesize that other variables, such as past crisis behavior or current bargaining behavior, contribute to deterrence success (Huth 1988). Finally, in addition to the effects of particular variables, we may want to assess how well a model explains deterrence outcomes where the model includes everything from the game structure to the included regressors. All of this can be accomplished using the statistical strategic deterrence model described above and the relevant regressors.

Deciding which variables enter each of the utilities in the game and how to estimate the parameters associated with those variables are not trivial matters. Ideally, theory should be the guide,

\[ ^4 \text{There are simply too many examples to cite. For some of the better empirical analyses along these lines, see Huth (1988); Huth, Gelpi, and Bennett (1993).} \]

\[ ^5 \text{The relationship in the above hypothesis is monotonic because it implies that as we increase the value of the explanatory variable, the dependent variable always decreases, holding all other variables constant. The hypothesis is also unconditionally monotonic because it is assumed that the monotonicity and its direction hold for every possible set of values at which the other variables could be held constant. Conditional monotonicity implies monotonicity for every set of values at which the other variables are held, but allows the direction of that monotonic relationship to differ, depending on the values at which the other variables are held.} \]

\[ ^6 \text{We are not suggesting that hypothesis testing as a method of inference is problematic, only the (typical) listing of unconditionally monotonic hypotheses with no clear (i.e., derivable) link to a strategic model.} \]
Figure 3: Specification of the Attacker’s and Defender’s Utilities in Terms of Regressors. The figure displays the specification of the observable components of the utilities in terms of regressors. (To simplify the presentation, we have dropped the private $\pi_{ij}$ components.) The attacker’s utility for the status quo is a function $X_{11}\beta_{11}$ of explanatory variables, its utility for capitulation is estimated as a constant $\beta_{130}$, and its utility for war is a function $X_{14}\beta_{14}$ of explanatory variables. The defender’s utility for capitulation is normalized to zero and its utility for war is a function $X_{24}\beta_{24}$ of explanatory variables.

not only for the structure of the interaction, but also for the specification of the utilities. In a perfect world, we would have variables representing the “primitives” of state preferences. Indeed, the functional form of the utility equations should also be theoretically justified. With little else to go on, our approach is to specify the set of utilities for a player as simply as possible and with an eye towards differences in utilities, since it is the size of the utilities relative to each other that determine the equilibrium choice probabilities.

Figure 3 shows the general specification of the utilities employed in the subsequent data analysis. Here, the attacker’s observed utility for the status quo is a linear function $\alpha_{11} + X_{11}\beta_{11}$ of explanatory variables, where $\alpha_{11}$ is a constant and $\beta_{11}$ is a vector of coefficients to be estimated, its observed utility for the defender’s capitulation is estimated as a constant $\alpha_{13}$, and its observed utility for war is a linear function $X_{14}\beta_{14}$ of explanatory variables. In this manner, we are able to differentiate the attacker’s utility for war from his utility for capitulation, and his utility for attacking from his utility for not attacking (i.e., the status quo).\footnote{If we included a constant in all three utilities, the model would be unidentified.} The defender’s utility for capitulation is normalized to zero and we treat her utility for war as a function $\alpha_{24} + X_{24}\beta_{24}$ of explanatory variables.\footnote{The defender’s utility for the status quo does not affect the equilibrium choice probabilities $p_d$ and $p_a$, which is...}
The estimation method we employ is detailed in Signorino (2000). The equilibrium outcome probabilities in Equations 3–5 are used as the basis of maximum likelihood estimation. Let \( y_{sq,i} = 1 \) if the crisis in observation \( i \) results in a status quo outcome, and zero otherwise. Let \( y_{cap,i} = 1 \) if the crisis results in capitulation by the defender, and zero otherwise. Let \( y_{war,i} = 1 \) if the crisis results in war between the attacker and the defender, and zero otherwise. The log-likelihood to be maximized (with respect to the \( \beta \)s) is

\[
\ln L = \sum_{i=1}^{N} \left[ y_{sq,i} \ln p_{sq,i} + y_{cap,i} \ln p_{cap,i} + y_{war,i} \ln p_{war,i} \right]
\]  

(6)

One generally cannot estimate the effects parameters (i.e., the \( \beta \)s) and the variance parameter \( \sigma \) individually. As with most other discrete choice models, they are not all individually identified. As in standard (i.e., nonstrategic) probit estimation, we normalize \( \sigma^2 \) to one. Parameter estimates are therefore actually estimates of the \( \beta \)s and \( \sigma \) to scale.

### 3.1 Variables and Data

We are fortunate in that data is available both on the outcomes of our model and on many substantively interesting variables often hypothesized to affect extended immediate deterrence. The bulk of the data used here is from the previous studies of Huth (1988) and Huth and Russett (1988), which examine fifty-eight extended immediate deterrence crises from 1885 to 1983.

The dependent variable in this study codes which of the outcomes \{SQ, Cap, War\} occurred in each of the fifty-eight crises. In the context of our model, the dependent variable examined in Huth (1988) codes whether the attacker attacked or not, \( A \) vs \( \overline{A} \), respectively. Huth and Russett (1988) followed Huth (1988) with an analysis of the defender’s actions (defend or not defend) in those twenty-four cases in which the attacker attacked. In the context of our model, Huth and Russett (1988) provide data on whether the defender defended or not, \( D \) versus \( \overline{D} \), given that the attacker used force against the protege. The sequence of actions coded in these two studies match the actions in our deterrence model and, therefore, provide all the information we need to code the outcome for each observation.

Most of our explanatory variables are drawn from Huth (1988). Rather that repeat their operationalizations, we refer the reader to Huth (1988) for the complete details. In general, they can be grouped under the following headings:

**Balance of Forces:** Whether the defender possessed nuclear weapons (\( NUCLEAR = 1 \) if the defender possessed nuclear weapons, 0 otherwise). The immediate balance of forces (\( IBF \) as why we do not provide a specification for it.
a ratio of the defender-protege to the attacker — i.e., $IBF > 1$ implies a stronger defender-
protege, and $IBF < 1$ implies a stronger attacker. The short-term balance of forces ($SBF$).
The long-term balance of forces ($LBF$).

**Defender’s Interests at Stake**: Whether the defender and protege had a military alliance ($MI-
LALL = 1$ if yes, 0 otherwise). The percentage of the protege’s arms imports that come
from the defender ($MILARM$), scaled from 1–10. The protege’s share of the defender’s total
merchandise imports and exports ($FORTRADE$), scaled from 0–10.

**Defender’s Reputation from its Last Extended Immediate Deterrence Crisis**: Whether
the defender successfully deterred an opponent in its last crisis ($PASTDET = 1$ if yes, 0 other-
wise). Whether the defender came to its protege’s aid in its last crisis, if the opponent was
not deterred ($ARMED$). Whether the defender capitulated in its last crisis, if the opponent
was not deterred ($CAPITU$). All three of these variables equal zero when the defender has
never been in an extended immediate deterrence crisis before.

**Defender’s Reputation from its Last Crisis, if any, with the Current Attacker**: Whether
the defender adopted a bullying strategy or forced the attacker to make critical concessions
in order to avoid armed conflict, or both ($PUTDOWN$). Whether the defender and attacker
avoided a military confrontation, but failed to resolve the underlying issues of the dispute
($STALEMATE$). Whether the defender retreated under diplomatic and/or military pressure
from the attacker in order to avoid armed conflict ($DIPLO$). All three of these variables equal
zero if the attacker and defender have never been in a crisis before.

**Defender’s Bargaining Behavior in the Current Crisis**: Whether the defender has adopted
a “firm-but-flexible” strategy in diplomatic negotiations until now ($FIRMFLEX$). Whether
the defender has responded proportionally to the military preparations of the attacker ($TFT$).

**Others**: Whether the attacker and defender are territorially contiguous ($CONTIGAD$). Whether
the attacker and protege are contiguous ($CONTIGAP$). Whether the defender and protege

---

9 $IBF$ is measured as the ratio of the defender-protege ground troops to the potential attacker’s ground troops, including only those troops that are at forward positions and that can be deployed to the scene of the battle immediately. $SBF$ includes each side’s standing ground and air forces and first class of trained reserves; it measures each side’s ability to reinforce the troops that are deployed at or near the scene of the battle, as measured by $IBF$. Huth (1988, 61-2) defines $LBF$ as “the capacity of the defender and protege and the potential attacker to build up their existing armed forces (army, air, and naval manpower) and to maintain an increased level of fighting strength by mobilizing the economy and civilian population for war . . . Each state’s existing military capabilities (percentage share of world military personnel and military expenditures) were multiplied by the sum of that state’s industrial and demographic resources (percentage share of world steel production, industrial fuel consumption, urban, and total population). The ratio of defender’s and protege’s capabilities to potential attacker’s capabilities was then calculated.”

10 We use strict land contiguity. Thus, for instance, we don’t consider Turkey and Cyprus to be contiguous.
are contiguous (CONTIGDP). Whether or not the defender was a democracy (DEMDEF). Whether or not the attacker was a democracy (DEMATT).

3.2 Strategic Probit Analysis

Based on the strategic deterrence model and using the preceding regressors, a total of four strategic probit regressions were conducted, representing different theoretical perspectives — e.g., a “realist” balance of forces model, a model based on the defender’s interests, a reputation and bargaining model, and a final model that combined all three. We found that, with the exception of STALEMENT, the variables representing the defender’s reputation from past crises were not significant, nor were the contiguity variables. Finally, DEMATT was also not significant.

Perhaps not surprisingly, the combined model far outperformed the other three, and we report in Table 1 the maximum likelihood estimates for only that model. The four columns in Table 1 are not four different models, but estimates of the four utility functions shown in Figure 3. For example, column 1 displays the estimates associated with the variables entering into the attacker’s utility for the status quo — i.e., they are the and from . Similarly, column 2 shows the estimate for the attacker’s utility for the defender’s capitulation (), column 3 shows the estimates for the attacker’s utility for war (), and column 4 shows the estimates for the defender’s utility for war ( and ). Standard errors are shown below the estimates. Estimates with one asterisk are statistically significant at (two-tailed), estimates with two asterisks are significant at , and estimates with three asterisks are significant at . Finally, the mean log-likelihood, the percentage of outcomes (war, status quo, or capitulation) correctly predicted (actually, postdicted), and the percentage of the attacker’s actions (attack or not attack) correctly predicted are displayed at the bottom to provide a sense of how well the model fares.

To briefly summarize the results, contrary to Huth (1988), the results suggest that alliances, nuclear weapons, military arms transfers, and foreign trade all affect deterrence success. In contrast to Huth and Russett (1988), the results suggest that the latter three variables, as well as the immediate balance of forces, influence the defender’s decision to defend its protege. The model correctly predicts over 93% of the attacker’s actions and almost 88% of the outcomes. We now discuss the results in more detail.

---

11To determine this, we use the POLITY III data set, which contains information on the regime characteristics for all of the states in the international system for the time period 1800-1994 (Jaggers and Gurr 1996). We use the commonly used method (see, for instance, Rousseau, Gelpi, Reiter, and Huth 1996; Schultz 1999) of subtracting the eleven-point autocracy score from the eleven-point democracy score, to create a measure ranging from -10 (entirely autocratic) to 10 (entirely democratic). If the difference exceeds 5, the state is coded as a democracy (the results are identical if we use a threshold of 6 or 7 instead).

12The results of the other three models are available upon request from the authors.
Table 1: Strategic Probit Regression. *The table displays the results of the strategic probit regression based on the model in Figure 3. The four columns report the maximum likelihood estimates of the coefficients associated with the variables entering into the defender’s utility for war (\(\hat{\beta}_{24}\)), the attacker’s utility for war (\(\hat{\beta}_{14}\)), the attacker’s utility for the status quo (\(\hat{\beta}_{11}\)), and the attacker’s utility for the defender’s capitulation (\(\hat{\beta}_{130}\)), respectively.*

<table>
<thead>
<tr>
<th>Variable</th>
<th>(U_d)(War)</th>
<th>(U_a)(War)</th>
<th>(U_a)(SQ)</th>
<th>(U_d)(Cap)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-11.10**</td>
<td>-7.91***</td>
<td>5.23*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.70</td>
<td>3.00</td>
<td>2.87</td>
<td></td>
</tr>
<tr>
<td>Nuclear</td>
<td>6.87**</td>
<td>-0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.77</td>
<td>1.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate Balance</td>
<td>6.51**</td>
<td>-6.24**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.63</td>
<td>2.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short-term Balance</td>
<td>-1.92</td>
<td>-6.77**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.47</td>
<td>3.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long-term Balance</td>
<td>3.68*</td>
<td>2.22**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.98</td>
<td>1.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Military Alliance</td>
<td>19.93**</td>
<td>8.77**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.88</td>
<td>3.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arms Transfers</td>
<td>-2.37***</td>
<td>-0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.91</td>
<td>0.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foreign Trade</td>
<td>5.63**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tit-for-Tat</td>
<td></td>
<td>9.27***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FirmFlex</td>
<td></td>
<td>3.41**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stalemate</td>
<td>10.13***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Democratic Defender</td>
<td>6.27**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean lnL</td>
<td>-0.273</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCP Outcomes</td>
<td>87.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCP Deter</td>
<td>93.1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard errors are shown below parameter estimates. N=58. ***p < .01. **p < .05. *p < .10. (two-tailed)
3.2.1 The Defender’s Utility for War

Let us first turn to the defender’s utility for war. The last column of Table 1 shows that every variable in the defender’s utility except the short-term balance of forces ($SBF$) is statistically significant at at least the $p < .10$ level. Possession of nuclear weapons, immediate and long-term balance of forces favoring the defender-protégé, higher levels of trade between the defender and protégé, an alliance between the defender and protégé, and a stalemate between the defender and attacker in their last crisis all increase the likelihood that the defender will fight to defend the protégé. All of these results fit with a “common sense” intuition of what should affect crisis behavior. In contrast to “selection-effects” arguments made by Fearon (1994a), we find that observable elements of the military balance and the defender’s interests at stake influence the defender’s decision to defend, even as late as the immediate deterrence stage of the crises. In contrast to Huth and Russett (1988), we find that nuclear-armed defenders are more likely to defend their protégés than are non-nuclear defenders.

Interestingly, we find that democratic states are more likely to go to war to defend their protégés than are authoritarian states. A possible explanation for this is Fearon’s (1994b) audience-cost theory of crisis bargaining, which stipulates that leaders who face domestic audience costs (perhaps electoral costs) for publicly escalating a crisis and then backing down are less likely to back down in public crises than are leaders who do not face such costs. The data set analyzed here consists of cases of immediate deterrence, in which threats and counterthreats have already been made, and hence potential audience costs have been generated. If democratic leaders on average face greater audience costs for backing down in public crises than their authoritarian counterparts, they should be more likely to fight to defend their protégés after potential audience costs have been raised, other things equal. Indeed, this is what we find. Interestingly, the regime type of the attacker is not significant. Of course, if potential attackers select on crises in which they do not expect to pay large audience costs, this should not be surprising.\(^\text{13}\)

3.2.2 The Attacker’s Utilities

Now consider the attacker’s utility for the status quo (the first column of Table 1). Tit-for-tat, or proportional responses by the defender to the attacker’s military preparations in the current crisis, increases the attacker’s utility for the status quo. Firm-but-flexible diplomatic bargaining by the defender also increases the attacker’s utility for the status quo. This is consistent with Huth (1988), who argues that tit-for-tat bargaining indicates that the defender is resolved to defend its protégé, but does not provoke the attacker by putting its reputation and credibility on the line, as a more

\(\text{\textsuperscript{13}}\)Note that the defender was a democracy in 31 of the 58 cases, but the attacker was a democracy in only 12 cases.
aggressive/bullying bargaining strategy by the defender might (i.e., tit-for-tat allows the attacker to back down without losing face).

Now consider the attacker’s utility for war. The third column of Table 1 shows that a defender-protege advantage in immediate and short-term balance of forces both decrease the attacker’s utility for war relative to the other outcomes. That is, the greater the defender-protege’s advantage in the immediate and short-term balance of forces relative to the attacker’s, the lower the attacker’s utility for war. These results are intuitive.

However, a higher defender-protege advantage in long-term balance of forces actually increases the attacker’s utility for war. This is an anomalous finding, for which we offer the following explanation. Holding the short-term balance of forces (SBF) constant, increasing the long-term balance (LBF) indicates that the defender-protege’s untapped military potential is increasing (relative to the attacker’s). Therefore, holding SBF constant, higher values of LBF indicate that the defender-protege may be a more formidable foe in the future, and therefore it is better to fight the defender-protege now rather than wait until it has converted its untapped military potential into actual military might. Other things equal, a state’s utility for war against a foe who has substantial untapped military potential may be higher than it is against one who doesn’t, because the former type of foe poses more of a threat in the future than the latter type. That is, a “window of opportunity” explanation may be responsible for this finding. For instance, in the years leading up to World War I, many German officials were concerned that Russia’s potential was such that she would soon be much more powerful, and if war was inevitable anyway, it was better to fight her now rather than later (Taylor 1954, 511, 515, 522, 527-28; Rich 1992, 435-6).

3.2.3 The Probability of Deterrence Success and of War

As in other discrete choice models (such as multinomial or ordered probit), interpreting the relationship between the dependent and independent variables simply by examining the regression results is difficult. A better means for assessing those relationships is by determining how predicted probabilities change as the values of the explanatory variables change. One advantage of our strategic probit analysis is that we can assess the impact of the explanatory variables not only on the probability of deterrence success (to which Huth 1988 is limited), but also to any of the other actions or outcomes of the model. In addition to examining the effects of the explanatory variables on the probability of deterrence success \( p_{sq} \), we will also analyze their impact on the probability of war between the attacker and the defender \( p_{war} \). For both of these, we use the equilibrium probabilities in Equations 1–5 and the estimates reported in Table 1.

\footnote{Our thanks to Robert Walker for suggesting this interpretation.}
Before proceeding, we should note that in the context of our deterrence model, relevant variables can affect the attacker’s behavior in two ways: directly through its utility for the various outcomes and indirectly through its belief \( p_d \) about whether the defender will defend. The attacker attacks if and only if its true expected utility for attacking is greater than its true utility for the status quo [i.e., if and only if \( p_d U^*_a(War) + \beta U^*_a(Cap) > U^*_a(SQ) \)]. \( p_d \) is a function of the explanatory variables in the defender’s utility for war, \( X_{24} \) (Equation 1). Hence, those variables affect not only the defender’s decision to defend, but also the attacker’s decision to attack, albeit indirectly, through \( p_d \). In other words, variables that are statistically significant in the defender’s utility for war but not in the attacker’s utilities, such as NUCLEAR, still affect the attacker’s decision to attack or not, because they affect his estimation of whether or not the defender will defend (they thus affect the attacker’s expected utility for attacking). Variables that are statistically significant in both the attacker’s and the defender’s utilities, such as IBF, have a direct as well as an indirect effect on the attacker’s decision to attack or not.

It is typical in analyses of fitted values or first differences to hold “all other variables” (i.e., other than the one being varied) constant at some value, usually their means. However, six of the explanatory variables in Table 1 are binary, and their means are values we would never observe.
in the data. Moreover, although it is not commonly done, it might be substantively interesting to examine the impact of the explanatory variables in situations other than the one represented by their means. Therefore, to provide a more nuanced picture of the explanatory variables’ effects on deterrence success and on war, we calculate predicted probabilities holding (1) all other variables constant at their minimum values (this is the “Min” column in Table 2), (2) all other continuous variables at their mean values and the binary variables at 0 (the “Mean0” column), (3) all other continuous variables at their mean values and the binary variables at their median values (the “Mean” column), (4) all other continuous variables at their mean values and the binary variables at 1 (the “Mean1” column), and (5) all other variables at their maximum values (the “Max” column).

In general, the “Min” case is one where the defender does not possess nuclear weapons, where the balance of forces greatly favor the potential attacker, where no military alliance exists between the defender and protege, where military arms transfers and foreign trade between the defender and protege are low, and where a firm-but-flexible diplomatic and a tit-for-tat military bargaining strategy were not used by the defender. In other words, the “Min” case is one where we would not usually expect the attacker to be deterred. Cases in which the variables approximate the “Min” case include Germany’s 1914 attack on Belgium (World War I), with Britain as Belgium’s defender, as well as Turkey’s 1974 attack on Cyprus, with Greece as Cyprus’ (potential) defender. In the former case, the balance of forces favored Germany quite strongly ($IBF = .39$, $SBF = .22$, $LBF = .1$), Britain did not of course possess nuclear weapons, there was no formal alliance between Britain and Belgium, and foreign trade and arms transfers between the two were rather low ($FORTRADE = 3$, $MILARM = 2$). In the latter case, the balance of forces strongly favored Turkey ($IBF = .65$, $SBF = .34$, $LBF = .16$), Greece did not possess nuclear weapons, there was no alliance between Greece and Cyprus, and foreign trade and arms transfers between the two were also rather low ($FORTRADE = .9$, $MILARM = 1$).

The “Max” case is exactly the opposite: the defender possesses nuclear weapons, the balance of forces greatly favor the defender-protege, there is a military alliance between the defender and protege, there are high levels of foreign trade and arms transfers between the two, and a firm-but-flexible and a tit-for-tat military bargaining strategy are used by the defender. In other words, the “Max” case is one where we would expect the attacker to be deterred. A case in which the variables approximate the “Max” case is Guatemala’s 1975 threat (successfully deterred) to attack and capture the British colony of Belize, with Britain as Belize’s defender. The immediate balance of forces slightly favored Britain ($IBF = 1.3$), and the short-term and long-term balances strongly favored her ($SBF = 4.04$, $LBF = 4.02$). Britain and Belize are coded as having an alliance, and

---

15Huth and Russett (1984, 511-12) regard a colonial relationship as implying a formal military commitment, and hence code it as an alliance.
Britain possessed nuclear weapons. Although Belize accounted for virtually none of Britain’s foreign trade ($FORTRADE = 0$), almost all of Belize’s arms imports were from Britain ($MILARM = 10$).

For reference, the actual values of the five cases are displayed in Table 2. Of course, no crisis in the data perfectly matches the combination of values expressed by any of the five cases. They are simply references or ideal types for the analysis.

We address the effect of each variable in turn, referring to Figures 4 and 5 for the effects of the continuous explanatory variables on the probability of deterrence success and of war. Tables 3(a) and (b) display the effects of the binary explanatory variables on the probability of deterrence success and of war. In Figures 4 and 5, to avoid cluttering the graphs, we only present the “Min,” “Mean,” and “Max” cases. In Tables 3(a) and (b) we present all five cases. Note that deterrence success is equivalent to a status quo outcome. Hence, the probability of the capitulation outcome can be determined at any time by adding the probabilities of deterrence success and of war, and subtracting that from one.

**Nuclear Weapons**

Scholars have come to very different conclusions concerning the effects of nuclear weapons on deterrence success — and on international stability more broadly. Some conclude that possessing nuclear weapons increases the likelihood of deterrence success (Alexandroff & Rosencrence 1977; Huth 1990; Waltz 1990; Weber 1990; Huth, Gelpi, Bennett 1993; Fearon 1994). Others conclude that nuclear weapons have no effect at all (Mueller 1989; Paul 1995). Finally, Waltz (1981) argues that either no proliferation or a relatively fast proliferation of nuclear weapons increases the chance of war.

We see from Table 1 that, in the context of our extended immediate deterrence model, the defender’s possession of nuclear weapons is not a statistically significant component of the attacker’s utilities, although it is in the defender’s utility (namely, it increases the defender’s utility for war, and hence increases the likelihood that the defender will fight to defend its protege). Through the probabilities $p_d$ and $p_{\bar{d}}$ in the attacker’s expected utility for attacking, the defender’s possession of nuclear weapons indirectly affect the attacker’s decision to attack or not, as discussed earlier. Tables 3(a) and (b) display the predicted probabilities of deterrence success and of war, respectively, when the defender does ($NUCLEAR=1$) and does not ($NUCLEAR=0$) possess nuclear weapons. The general effect of the defender’s possession of nuclear weapons is to increase the probability of deterrence success (because the defender is more likely to defend, and the attacker anticipates this), thereby decreasing the probability of war (although the defender is more likely to defend, resulting in war, in those cases in which the attacker does choose to attack). In the “Mean” case, the defender’s possession of nuclear weapons dramatically increases the probability that the attacker is
deterred, with a corresponding reduction in the probability of war. In the “Min” case, other factors favor the attacker, who is unlikely to be deterred regardless of whether or not the defender possesses nuclear weapons. In the “Mean”, “Mean1” and “Max” cases, nuclear weapons have no discernible effect on deterrence success — regardless of whether or not the defender possesses nuclear weapons, other deterrent factors are sufficient to prevent an attack, and, therefore, war.

Betts (1987) argues that during the 1948 Berlin Blockade, the U.S.’s threats to use nuclear weapons against the Soviet Union were vague and ambiguous, and their impact on the final outcome was unclear. In contrast, our model suggests that nuclear weapons played a very significant role in deterring the Soviet Union. Given the values of the explanatory variables in the Berlin crisis (e.g., that the U.S. possessed nuclear weapons), our model correctly predicts deterrence success with probability $p_{sq} = .99$ — i.e., that the Soviets would not seriously challenge the Allied airlift, and would eventually end the ground blockade without further escalating the crisis. Interestingly, the model predicts that if the U.S. did not have nuclear weapons, the probability of deterrence success would have dropped to virtually zero, the probability of war would have been $p_{war} = .62$, and with probability $p_{cap} = .38$ the U.S. and Britain would have surrendered Berlin.

Regarding the U.S. attempts to deter a Chinese invasion of the (Taiwanese) Nationalist-held islands of Quemoy and Matsu in 1954-55, Betts (1987) notes that Eisenhower and Secretary of State Dulles made a number of public statements suggesting that nuclear weapons would be used to defend the islands (Betts 1987, 59). Again, our model correctly predicts deterrence success with probability $p_{sq} = .99$. Once more, the model makes the provocative prediction that if the U.S. did not have nuclear weapons, the probability of deterring China would have been virtually zero, with probability $p_{war} = .14$ the U.S. would have fought to defend the islands, and with probability $p_{cap} = .86$ would have surrendered them.

Although our model indicates that nuclear weapons have had a major impact on deterrence success in a number of cases, the model predicts that nuclear weapons played no role in the U.S.’s 1961 successful deterrence of a threat by North Vietnam and its local Pathet Lao allies to overrun Laos’s capital. The model (correctly) predicts deterrence success with probability one, both with and without nuclear weapons.

**Military Alliance**

Recall from Table 1 that a military alliance between the defender and protege increases the attacker’s as well as the defender’s utility for war. From Table 3(b) we see that the general effect of a military alliance is to increase the probability of war. From Table 3(a) we see that a defender-protege military alliance has no discernable effect on the probability of deterrence success for any of the five cases considered. For the “Min” and “Mean0” cases, other factors favor the attacker, who
is thus unlikely to be deterred regardless. However, in these two cases, the presence of a defender-protege alliance increases significantly the likelihood that the defender will defend its protege, and hence the probability of war.

_Tit-for-Tat Military Preparations_

Recall from Table 1 that the defender’s use of a tit-for-tat, or proportional, military bargaining strategy increases the attacker’s utility for the status quo. Tables 3(a) and (b) show that tit-for-tat generally increases the probability of deterrence success, and hence decreases the probability of war. In the “Mean” case, the defender’s use of tit-for-tat increases the probability of deterrence success from 0 to .93 and (hence) decreases the probability of war from .35 to .02. Similarly, in the “Mean” case, the defender’s use of tit-for-tat increases the probability of deterrence success from .09 to 1 and (hence) decreases the probability of war from .91 to 0. The defender’s use of a tit-for-tat military bargaining strategy can have a powerful role in deterring the attacker.

_Democratic Defender_

Recall from Table 1 that democratic defenders have a higher utility for war, i.e. they are more likely to go to war to defend their proteges. Because potential attackers can anticipate this (namely, that democratic leaders on average pay high audience costs for backing down in a public crisis), they are more likely to be deterred when faced with a democratic defender than an authoritarian one, as seen in Table 3(a). In the “Mean” case, the probability that the attacker is deterred is 0 when the defender is authoritarian, and .99 when the defender is a democracy. The corresponding probability of war drops from .35 to .01 (although the defender is more likely to defend, resulting in war, in those cases in which the attacker does choose to attack). For the “Mean” case, the probability of deterrence success increases from .93 to 1 in the change from an authoritarian to a democratic defender, and the corresponding probability of war decreases from .02 to 0. In immediate deterrence crises, potential attackers are unlikely to attack the proteges of democratic defenders, knowing that the democratic leader is unlikely to back down and hence incur audience costs.

_Immediate Balance of Forces_

Recall from Table 1 that the defender’s utility for war increases as the immediate balance of forces increasingly favors the defender-protege over the attacker (i.e., as \(IBF\) increases), and the attacker’s utility for war decreases. As Figure 4(a) shows, the monotonicity and direction of the effect of the immediate balance of forces is the same as in Huth (1988): as the immediate balance of forces increases in favor of the defender-protege, the probability that the potential attacker is deterred increases (because the defender is more likely to defend, and the attacker anticipates this). However, the rate of increase differs substantially depending on the values at which the other variables are held constant. In the “Max” case, other deterrent factors are sufficient to ensure that
the attacker does not attack, and IBF has no effect.

In the “Min” case, the probability of deterrence success is zero as IBF ranges from 0 to about 2.1. After that, the probability increases rapidly and reaches 1 by the time IBF has increased to about 2.3, after which the probability levels off and stays at 1. In the “Mean” case, the trend is similar, although the dramatic increase in the probability of deterrence success occur much earlier, at very low levels of IBF. Interestingly, the dramatic increase in the probability of deterrence success in this case occurs well before the defender-protege has achieved parity in immediate balance of forces with the attacker (i.e., well before $IBF = 1$). This suggests that under certain conditions (i.e., for certain combinations of values of the other explanatory variables), the defender’s positioning of forces at or near the scene of the battle can have something of a “tripwire” effect. Full parity of (immediate) forces is not necessary to deter the attacker. If the attacker must overrun forces of the defender to attack the protege, the attacker can be confident that the defender is likely to intervene. Hence, the presence of a sizable defender force at or near the scene of the (potential) battle has a potent deterrent effect, and full parity is not necessary.

Figure 4(b) displays the effect of the immediate balance of forces on the probability of war between the attacker and the defender. In the “Max” case, the attacker is always deterred and so the probability of war is always zero. Of more interest is the “Min” case, which displays a nonmonotonic relationship between the immediate balance of forces and the probability of war. For $IBF < 2.1$, the attacker is never deterred. For $IBF < 1.4$, the defender never defends, and hence the probability of war is 0. As $IBF$ increases beyond 1.4, the defender becomes more likely to defend (as the immediate balance of forces is increasingly in its favor), and hence the probability of war rises (the attacker is still always attacking). However, once $IBF$ reaches 2.1, the attacker becomes increasingly less likely to attack, as the immediate balance of forces is increasingly in the defender’s favor, and hence the probability of war starts declining (although the defender is more likely to defend, resulting in war, in those cases in which the attacker does choose to attack). Therefore, $p_{war}$ is a nonmonotonic function of $IBF$.

For similar reasons, the same occurs in the “Mean” case. For very low values of $IBF$, as $IBF$ increases, the attacker is less likely to attack. However, the defender is more likely to defend, resulting in war, in those cases in which the attacker does choose to attack. For very low values of $IBF$, this second effect dominates the first, and so the probability of war initially increases. However, as $IBF$ increases even more, the probability of deterrence success approaches 1, and hence the probability of war starts decreasing, despite the fact that the defender is more likely to defend, resulting in war, in those cases in which the attacker does choose to attack. Therefore, $p_{war}$ is a nonmonotonic function of $IBF$.

*Long-Term Balance of Forces*
Recall from Table 1 that as the long-term balance of forces increasingly favors the defender (i.e., as $LBF$ increases), the attacker’s as well as the defender’s utility for war increases. As seen in Figures 4(c) and (d), for the “Max” case, other factors favor the defender, and hence the probability of deterrence success is 1, and hence the probability of war is 0, regardless of the value of $LBF$. For the “Min” case, other factors favor the attacker, who is thus never deterred. However, once $LBF$ increases beyond about 2.4, the probability of war starts increasing, because the defender is more likely to defend the protege, as the long-term balance of forces is increasingly in its favor.

For the “Mean” case, for very low values of $LBF$, the probability of deterrence success increases as $LBF$ increases. This is because the defender is more likely to defend the protege, and war is not yet an attractive option for the attacker. Therefore, the probability of war declines. For very high values of $LBF$, the probability of deterrence success starts decreasing as $LBF$ increases even more, because war is now becoming an attractive option for the attacker, since war provides an opportunity for the attacker to destroy a foe with substantial untapped military potential who might be much more powerful in the future. Therefore, the probability of war starts increasing.

**Foreign Trade**

The “foreign trade” variable ($FORTRADE$) measures the protege’s share of the defender’s total merchandise exports and imports and, thus, how economically valuable the protege is to the defender. Recall from Table 1 that as the level of trade between the defender and protege rises, the more likely the defender is to defend the protege. The attacker can anticipate this and hence the probability of deterrence success rises as $FORTRADE$ increases for the “Mean” case, as seen in Figure 5(a). The probability of war is a nonmonotonic function of $FORTRADE$ for this case, for the same reason as it was for $IBF$. For very small values of $FORTRADE$, the probability of deterrence success rises as $FORTRADE$ increases. However, the defender is more likely to defend, resulting in war, in those cases in which the attacker does choose to attack. For very low values of $FORTRADE$, this second effect dominates the first, and so the probability of war initially increases. However, as $FORTRADE$ increases even more, the probability of deterrence success approaches 1, and hence the probability of war starts decreasing, despite the fact that the defender is more likely to defend, resulting in war, in those cases in which the attacker does choose to attack. Therefore, $p_{\text{war}}$ is a nonmonotonic function of $FORTRADE$.

In the “Max” case, other deterrent factors ensure that the attacker is always deterred and hence the probability of war is always 0. In the “Min” case, other factors favor the attacker, who is thus never deterred regardless of the value of $FORTRADE$. However, as $FORTRADE$ increases beyond about 1.4, the defender becomes more likely to defend, and hence the probability of war rises.

**Military Arms Transfers**
Recall from Table 1 that the defender’s utility for war decreases as its share of the protege’s arms imports (MILARM) increases. That is, the defender is less likely to defend the protege the greater the protege’s reliance on the defender for its arms imports. The attacker can anticipate this, and hence the probability of deterrence success decreases as MILARM increases for the “Mean” case, as seen in Figure 5(c). This leads to a nonmonotonic relationship between the probability of war and MILARM, for reasons similar to those above. As MILARM increases from 0 to about 7.6, the attacker is always deterred and hence the probability of war is always 0. However, as MILARM increases beyond 7.6, the attacker is more likely to attack because it knows that the defender is less likely to defend. Initially, the first effect dominates the second and hence the probability of war rises. However, the probability that the attacker attacks soon approaches 1 and as MILARM increases even more, the defender is less likely to defend and hence the probability of war starts decreasing.

3.2.4 Model Fit

On a final note, although the relevance of model fit is a much debated subject, we believe it provides information that, when not stretched beyond its intended use, can be helpful. Signorino (1999) also suggests that poor model fit may be a sign of a misspecified model, which would cast doubt on the parameter estimates.

In discrete choice models such as logit and probit, a commonly used measure of goodness of fit is the percentage of the outcomes (i.e., realizations of the dependent variable) correctly predicted. Table 1 provides the percentage of the the attacker’s actions (i.e., deterrence success versus failure) correctly predicted and the percentage of the model’s outcomes (status quo, capitulation, or war) correctly predicted . The strategic deterrence model correctly predicts (actually, postdicts) almost 88% of the outcomes and over 93% of the attacker’s actions. That level of model fit is impressive for the international relations literature (Huth 1988 correctly predicts 84% of the attacker’s actions, and Huth and Russett 1988 correctly predict 83% of the defender’s actions), where it is not uncommon for discrete choice models to barely predict above the null model of always predicting the most frequent category. The fact that our statistical model is also a theoretical deterrence model, the fact that it is populated by statistically significant and theoretically justified regressors, the fact that most of the regressors have effects that seem reasonable and can be explained in the context of the model — all of this in combination with the extremely high predictive power provides substantial support for the model as an explanation of extended immediate deterrence. Having said that, the percentage of the outcomes correctly predicted (88%) is not as high as the percentage of the attacker’s actions correctly predicted (93%). There is still room for improvement.
4 Concluding Remarks

We have presented the first unified theory and test of extended immediate deterrence. Contrary to Huth (1988), our empirical analysis suggests that alliances, nuclear weapons, military arms transfers, and foreign trade all affect deterrence success. In contrast to Huth and Russett (1988), we find that the latter three variables, as well as the immediate balance of forces, influence the defender’s decision to defend its protege. In terms of model fit, our model correctly predicts over 93% of the potential attacker’s actions and almost 88% of the crisis outcomes. We also find evidence that the likelihood of deterrence success and of war are not consistent with the typical (monotonic) structural assumptions of previous studies using logit and probit.

Substantively, the model predicts that as the immediate and short-term balance of forces increasingly favor the defender-protege, the attacker is more likely to be deterred. Surprisingly, as the long-term balance of forces increasingly favors the defender-protege, the attacker’s utility for war (as well as the defender’s) increases. This suggests that potential attackers have a higher utility for going to war with defenders with substantial untapped military potential than those without, as they see a window of opportunity to attack a foe who might be much more powerful in the future.

We find that the defender’s possession of nuclear weapons can play a significant role in achieving deterrence success. In line with Huth (1988), we find that tit-for-tat military preparations and a firm-but-flexible diplomatic bargaining strategy by the defender indicates to the attacker that the defender will defend its protege, but does not provoke the attacker by putting its reputation and credibility on the line, as a more bullying/aggressive bargaining strategy by the defender might (i.e., tit-for-tat allows the attacker to back down without losing face). Tit-for-tat military preparations and firm-but-flexible diplomatic bargaining by the defender help achieve deterrence success.

We find that defenders are more likely to defend proteges with whom they have a formal military alliance. Surprisingly, attackers also have a higher utility for war when there is a defender-protege military alliance. The greater the protege’s share of the defender’s total merchandise exports and imports, the more likely the defender is to fight to defend its (economically valuable) protege. On the other hand, the more heavily reliant the protege is on the defender for its arms imports, the less likely the defender is to come to the protege’s aid. This suggests that a defender will often send a great deal of arms to the protege precisely when it does not expect to fight itself if the protege is attacked. If there was a previous dispute between the defender and the attacker which was left unresolved in a stalemate, the defender is more likely to fight the attacker to defend its protege. Finally, we find that democratic defenders are more likely to go to war to protect their proteges than are authoritarian defenders, a finding that is consistent with the hypothesis that democratic leaders face greater audience costs (Fearon 1994b) and are hence less likely to back down in public
crises.

Although our model fares quite well, we believe a number of avenues of research would prove fruitful. First, more data needs to be collected. As we mentioned earlier, we are fortunate as it is to have the Huth (1988) and Huth and Russett (1988) data, which allow us to code the outcomes of our model and provide many substantively interesting variables often hypothesized to affect extended immediate deterrence. However, there are only fifty-eight observations, which makes estimation difficult for larger models (both in terms of number of regressors, as well as number of decision nodes, and hence outcomes, in the model). Additional data would not only assist in estimation, but would allow us to assess model fit by testing out-of-sample predictions. Huth and Russett (1988) suggest that these fifty-eight observations appear to comprise the entire universe of extended immediate deterrence crises in the international system over the time period of 1885 to 1984; if that is true, then additional observations can only come from other time periods.\textsuperscript{16}

Second, we raised the issue of selection bias in the text, but did not attempt to extend the model any further than that presented here. It is somewhat surprising that our simple deterrence model goes as far as it does. Nevertheless, there are a few puzzling estimates that may be due to how states selected themselves into the current sample. The next obvious step would be to incorporate the prior round or two of interaction into the current model (see Fearon 1994a for a theoretical model that does this). The fact that that would add more outcomes is somewhat problematic in the current context, since we would need to specify utilities and possibly additional explanatory variables for those outcomes, placing an even larger burden on the fifty-eight observations available. Fortunately, analyzing a model that incorporates prior rounds of interaction (such as Fearon's) would expand the potential population of cases beyond simply those of immediate deterrence (general deterrence could be considered as well). That is, additional observations could be obtained from the same temporal domain as the Huth (1988) and Huth and Russett (1988) data. All this suggests that theoretically-informed statistical analysis of the factors that influence crisis behavior remains a fruitful and exciting avenue for future research.

\textsuperscript{16}Danilovic (2001) has created a new data set of extended immediate deterrence encounters among the major powers from 1895 to 1985, and it only contains 46 observations.
References


Holding all other variables at

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
<th>Min</th>
<th>Mean₀</th>
<th>Mean</th>
<th>Mean₁</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclear</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0</td>
<td>.99</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Military Alliance</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>.99</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Tit-for-Tat</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>.99</td>
<td>.09</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0</td>
<td>.93</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Dem. Defender</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>.93</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0</td>
<td>.99</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

(a) Probability of Deterrence Success

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
<th>Min</th>
<th>Mean₀</th>
<th>Mean</th>
<th>Mean₁</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclear</td>
<td>0</td>
<td>0</td>
<td>.35</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0</td>
<td>.01</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Military Alliance</td>
<td>0</td>
<td>0</td>
<td>.35</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Tit-for-Tat</td>
<td>0</td>
<td>0</td>
<td>.35</td>
<td>.01</td>
<td>.91</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0</td>
<td>.02</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Dem. Defender</td>
<td>0</td>
<td>0</td>
<td>.35</td>
<td>.02</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0</td>
<td>.01</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

(b) Probability of War

Table 3: Effects of Variables on Deterrence Success and on War. *Table (a) shows the probability of deterrence success when each of the binary explanatory variables takes on the values of zero and one, holding the other explanatory variables constant at their minimum, mean/median, and maximum values. Table (b) displays the same, but for the probability of war.*
Figure 4: Effect of Balance of Forces. Figures (a) and (b) display the effect of the immediate balance of forces (IBF) on the probability of deterrence success and war, respectively, holding all other variables constant at their minimum (solid line), mean/median (long dash), or maximum (short dash) values. Figures (c) and (d) show the effect of the long-term balance of forces (LBF).
Figure 5: Effect of Foreign Trade and of Military Arms Transfers. *Figures (a) and (b) display the effect of foreign trade on the probability of deterrence success and war, respectively, holding all other variables constant at their minimum (solid line), mean/median (long dash), or maximum (short dash) values. Figures (c) and (d) show the effect of military arms transfers.*