In Defense of Comparative Statics Analysis: A Comment on Signorino's Treatment of Empirical Tests of Theoretical Models*

Abstract: Starting in 1999, Curtis Signorino challenged the use of traditional logits and probits analysis for testing discrete choice, strategic models. Signorino argues that the complex parametric relationships generated by even the simplest strategic models can lead to wildly inaccurate inferences if one applies these traditional approaches. In their stead, Signorino proposes generating stochastic formal models, from which one can directly derive a maximum likelihood estimator. We propose a simpler, alternative "methodology" for theoretically and empirically accounting for strategic behavior. In particular, we propose carefully and correctly deriving ones comparative statics from ones formal model, whether it is stochastic or deterministic does not particularly matter, and using standard logit or probit estimation techniques to test the predictions. We demonstrate that this approach performs almost identically to Signorino's more complex suggestion.

Clifford J. Carrubba
Associate Professor of Political Science
Emory University
Department of Political Science
Atlanta, GA 30322
(404) 727-7915
ccarrub@emory.edu

Chris Stanton
Undergraduate
Emory University

Amy Yuen
PhD Student of Political Science
Emory University
ayuen@emory.edu

* We would like to thank Randy Calvert, Mark Hallerberg, Andrew Martin, Eric Reinhardt, and Craig Volden for invaluable feedback on this project.

Political scientists are concerned with the strategic nature of political decision-making. That is, political actors do not simply make choices independent of the anticipated actions and reactions of other political actors. Rather, they make their decisions specifically contingent on those anticipated choices. For example, when one country is considering whether to attack another country, the potential aggressor does so anticipating how the other country is likely to react. Will it respond militarily, or will it back down? Under many conditions, the potential aggressor's decision will be contingent on that expected response.

Starting in 1999, Curtis Signorino used this notion of strategic behavior to challenge how political scientists empirically study political behavior (see also Smith 1999; Lewis and Schultz 2003). In this challenge, Signorino argues that strategic behavior generates complex parametric relationships that confound traditional logit and probit analyses.[1] As a result, he claims, these traditional approaches can lead to wildly inaccurate inferences and should not be used. Signorino proposes an alterative approach of writing stochastic formal models from which one can directly derive maximum likelihood estimators to test strategic theories (Signorino 1999, Signorino 2003, Signorino and Yilmaz 2003).[2]

---

[1] As discussed in more detail below, it is unclear whether Signorino uses the term "traditional" to refer to the use of logits and probits in general, or whether he is referring to just standard functional forms of these estimation techniques (i.e. the linear link). Thus, we leave the term intentionally undefined here as well.

[2] In a minimalist reading of this critique, Signorino simply is critiquing the existing applications of logit and probit. Scholars generally operationalize linear logits and probits, and these "typical" logits and probits fail because relationships between the independent variables, or parameters of the model, and the dependent variables, either observed behavior or outcomes, are often nonlinear. With this reading, Signorino is simply using the technique of writing a stochastic formal model and deriving maximum likelihood estimators directly from the theoretical model as a way of illustrating the problems with typical logit and probit specifications. In a maximalist reading of this critique, Signorino is critiquing the entire

The importance of this challenge cannot be oversold. If Signorino is right, existing discrete model quantitative tests of strategic behavior in political science are potentially deeply flawed. Further, Signorino's identified solution is no quick fix. It requires scholars to re-derive existing deterministic theories as stochastic formal models, and then test the new stochastic model by deriving a maximum likelihood estimator as well. Thus, there is no "off-the-shelf" solution on either the theoretical or empirical front.

This challenge has quickly gained wide notoriety and acceptance. His work is published in the leading journals, including the American Political Science Review (APSR), the American Journal of Political Science (AJPS), and Political Analysis (PA); the work is already widely cited with the 1999 APSR article alone having over 50 citations; and National Science Foundation funded Empirical Implications of Theoretical Models (EITM) summer workshops make a point of teaching graduate students and young faculty his approach. Thus, it appears that Signorino has found a serious problem in the study of strategic behavior, and that his solution is rapidly penetrating the discipline.

---

enterprise of using logits and probits to test strategic models period, and is suggesting that the discipline needs new statistical techniques, i.e. his approach, if one is going to accurately test models of strategic behavior.

While nowhere in his published work does Signorino unambiguously state that he is making the maximalist critique, it is not an implausible conclusion. First, throughout his work, Signorino makes statements that read as critiques of using logits or probits in general. For example, in motivating the logit quantal response model in his 1999 article, Signorino states, "In sum, logit models of international conflict are unlikely to capture the real or theorized structure of strategic interaction" (Signorino 1999:280). Similarly, Signorino concludes his Monte Carlo analysis by stating: "The question posed at the beginning of this section was: How well does traditional logit model strategic interaction?" (Signorino 1999:287). Thus, Signorino often does not qualify his statements by indicating that he is critiquing a particular empirical specification. Second, Signorino motivates using his LQRE approach as if it is necessary for incorporating strategic interdependence into a statistical model: "I analyze… the effects of using a logit model when two states behave strategically. To do this, however, we need a method for incorporating the structure of strategic interdependence into statistical models of conflict." (Signorino 1999:281). Third, and perhaps most importantly, at no point in any of his work does Signorino propose the obvious alternative to "typical" logit and probit specifications or to the LQRE approach, appropriately deriving standard comparative statics approaches to generate predictions and test them in a logit or probit model.

We very much agree with the emphasis Signorino places on careful theoretical and empirical modeling of strategic behavior. However, we believe that Signorino has yet to demonstrate either the inability to use well-established empirical estimation techniques to test strategic theories, or the value-added of his proposed solution to testing formal models of strategic behavior. Rather, we believe that an alternative, simpler "methodology" is available that both allows one to continue to rely upon standard estimation techniques and performs just as well as Signorino's proposed alternative. Simply put, we recommend that scholars directly derive their comparative statics from their strategic models, and use standard estimation techniques to test the predictions generated through the comparative statics analysis.

To demonstrate the validity of our approach, we proceed in three parts. First, we demonstrate that the complex parametric relationships derived by Signorino in his crisis bargaining models to illustrate the failure of traditional estimation techniques can be derived using standard comparative statics. Second, we demonstrate that these complex relationships do not rely upon the stochastic formal modeling assumption. That is, one can get exactly those complex relationships using a simple deterministic formal model. And third, we demonstrate that if one directly derives one's estimators in Signorino's examples, they are traditional variants of standard logits and probits. In sum, we demonstrate the appropriateness of standard estimation techniques in which the functional form of the estimator is based upon comparative statics derived from simple deterministic models.

Nonlinearities, Stochastic Modeling and Comparative Statics Analysis

This section proceeds in three parts. First, we summarize in more detail Signorino's critique of, and solution to, existing empirical work as argued in Signorino (1999) and Signorino and Yilmaz (2003). Second, we reanalyze the parametric relationships from these papers and demonstrate that correctly derived comparative statics capture the central nonlinearities used by Signorino to demonstrate the failure of traditional estimation techniques.[3] Finally, we also demonstrate that making the model stochastic is not central to deriving these nonlinear relationships.

Signorino's Argument

In Signorino (1999), Signorino first challenges the use of traditional logit analysis in the face of strategic interdependence. His challenge starts by characterizing "the problem with traditional methods of estimation" in empirical studies of international conflict (Signorino 1999, pg. 280). The core of this critique is "the typical use of logit", by which he means including an array of plausible covariates in a logit predicting the occurrence of conflict at the nation-year, dyad-year, or monad-dispute level. Signorino states that there are a number of reasons to be wary of this form of analysis, including "if observed actions are the results of (perhaps complex) strategic interaction, then it is unlikely that a simple logit functional form will capture the structure of that strategic interdependence…".

[3] This functional form critique is also central to part of Signorino's (2003) analysis. In the first part of Signorino (2003), he demonstrates that the maximum likelihood estimator one derives from a strategic and nonstrategic model are fundamentally different. This is certainly true. However, this difference arises predictably from the fact that predicted behavior is going to differ across the two models, since incentives and choices differ. If one derived a series of comparative statics based predictions from the two models, one would also arrive at different functional forms to ones logit or probit as well. Thus, in the end, this analysis is also a demonstration of the importance of correctly deriving ones predictions from ones theory.

Focusing on this problem, Signorio argues that one needs "a method for incorporating the structure of strategic interdependence into statistical models of conflict" to demonstrate the failure of the traditional logit. To do so, Signorino derives the logit quantal response equilibrium (LQRE). This is a statistical model in which one writes down a discrete choice formal model, derives a quantal response equilibrium (QRE), and then directly derives the maximum likelihood estimator (MLE) from the statistical formal model.[4] By deriving such an estimator, one can be assured of capturing the behavioral interdependence that arises in strategic situations.

To demonstrate the superiority of the LQRE model over "traditional logits", Signorino characterizes a "typical bilateral crisis game". This game is slightly more complex than the crisis bargaining game used in subsequent work (Signorino and Yilmaz 2003; Signorino 2003). Figure 1 characterizes the game. First, player 1 chooses whether to fight (*F*) or not (~*F*). Once player 1 makes her decision, player 2 then chooses whether to fight (*f*) or not (~*f*). If player 1 choose not to fight and player 2 chooses to fight, player one then must choose whether to fight or not. If neither player chooses to fight, the outcome is the status quo (*SQ*), if player i chooses to fight and player j chooses not to fight, the outcome is capitulation by player *j* (*Cj*), and if both players choose to fight, the outcome is war (*W*).

Assuming that the world operates according to the crisis game, and assuming the QRE solution concept, Signorino generates a Monte Carlo simulation and compares the performance of his LQRE estimation technique to that of the typical logits. In particular, he operationalizes a naïve logit in which both player's military capabilities and assets are

[4] See Mckelvey and Palfrey (1998) for derivation of the QRE solution concept. Also, note that QRE is another name for the agency error stochastic model mentioned in Signorino (2003).

included linearly, a more sophisticated "balance of power" logit in which "military concentration" variables are created from the capabilities and assets variables, and finally another more sophisticated "joint utility" logit in which the joint utility of war, $u_1(W)u_2(W)$, is included as a regressor. Signorino not only demonstrates that only the LQRE model correctly retrieves the underlying parameters that generated the dataset, but that one would actually draw incorrect inferences from the other three logits.

To illustrate why these logit analyses fail, Signorino demonstrates that country 1's military power is nonlinearly related to the probability of war. While the LQRE successfully recaptures this nonlinear relationship, the other logit models do not.

From this analysis, Signorino concludes,

> The question posed at the beginning of this section was: How well does traditional logit model strategic interaction? At least for the simple crisis interaction model here, the answer appears to be: Not very well at all. Perhaps more troubling are the highly significant results in each case, which would be interpreted by the typical researcher as supporting one model or another. Hence, out of a single data set, support could be "found" for a number of different theories of international conflict—all of which are wrong.

Signorino and Yilmaz (2003) extend this illustration of the failure of traditional logit by demonstrating that strategic misspecification is the equivalent of omitted variable bias. Once again, they rely upon a crisis bargaining model to illustrate the effects of strategic misspecification. However, here they use an even simpler model. Figure 2 characterizes the model.[5] The first player chooses whether to attack (*A*) or not (~*A*). If she does not attack the game ends and both players receive the payoff associated with the status quo. If she does attack the second player gets to choose to retaliate (*R*) or not (~*R*). If player 2 retaliates, both players receive their payoffs associated with war. If player 2 does not retaliate, they both receive their payoffs associated with the second player capitulating.

To demonstrate the omitted variable bias, Signorino and Yilmaz show that the linear logit (i.e. a model with parameters included only linearly) omits higher order terms included in a Taylor expansion of the strategic model. Once again, this omission arises from the fact that strategic behavior generates nonlinear relationships among the parameters. In fact, Signorino and Yilmaz demonstrate that the linear logit is adequate only when relationships are *unconditionally* monotonic. Once again, they prove this point through simple illustration.

Signorino's proposed solution to the failure of the traditional logit is to directly derive ones maximum likelihood estimator from a stochastic formal model. In his 1999 paper, this is the LQRE. However, the actual estimator will depend upon the precise error structure one assumes in the formal model (Signorino 2003). While nowhere does Signorino explicitly rule out careful derivation of comparative statics, Signorino also never proposes it as an alterative solution for dealing with complex strategic interdependence.

[5] Drawn from Signorino and Yilmaz (2003) and Signorino (2003).

Comparative Statics Analysis and Deterministic Modeling

Signorino's observations are clearly important. However, they are a critique of poor hypothesis generation and testing, not of estimation failure related to the use of the logit or probit. Careful comparative statics generation is more than adequate for resolving the complex relationships between parameters and outcomes that arise in the presence of strategic interdependence. To demonstrate this point, we reanalyze Signorino's two crisis bargaining models.

First consider the simpler variant characterized in Signorino's later work (Signorino 2003; Signorino and Yilmaz 2003). For simplicity, we will discuss the solution in terms of a random utility version of the model, though in this example the agency error version actually yields the same solution if we define $e_{1w}=e_{1c}$.

The solution to this game is straightforward. Player 2 retaliates when $u_2(war)+e_{2w}>u_2(cap) = 0+e_{2c}$. Thus, if we define G to be the CDF of $e_{2c}-e_{2w}$, we can define the probability of player 2 retaliating as $Q=G(u_2(war))$. Player 1 attacks when the expected utility of going to war is greater than the payoff for the status quo.[6] That is, $EU_1(A) = Q*(u_1(war)+e_{1w}) + (1-Q)*(u_1(cap)+e_{1c}) > 0+e_{1sq}$. Thus, Player 1 attacks when $e_{1sq}-Qe_{1w}+(1-Q)e_{1c}<Q(u_1(war))+(1-Q)u_1(cap)$. Defining F as the CDF of $e_{1sq}-Qe_{1w}+(1-Q)e_{1c}$, the probability that player 1 attacks is $P=F[u_1(cap)-Q(u_2(war))(u_1(cap)-u_1(war))]$.

Having derived equilibrium behavior, we can now derive comparative statics. Equations 1 through 4 are the first derivatives of each player's probabilistic moves with respect to the parameters of interest.

---

[6] The payoff is an expected value because player 1 does not know with certainty what player 2 is going to do if attacked.

$$\frac{\partial Q}{\partial u_2(war)} = g(u_2(war)) \geq 0 \quad (1)$$

$$\frac{\partial P}{\partial u_1(war)} = G(u_2(war)) * f[u_1(cap) - G(u_2(war))(u_1(cap) - u_1(war))] \geq 0 \quad (2)$$

$$\frac{\partial P}{\partial u_1(cap)} = [1 - G(u_2(war))] * f[u_1(cap) - G(u_2(war))(u_1(cap) - u_1(war))] \geq 0 \quad (3)$$

$$\frac{\partial P}{\partial u_2(war)} = -[u_1(cap) - u_1(war)] * g(u_2(war)) * f[u_1(cap) - G(u_2(war))(u_1(cap) - u_1(war))] \quad (4)$$

Equation one is obviously positive, equation two is positive since it consists of a cumulative and density function, and equation three is similarly positive since it consists of one minus a cumulative function and a density function. Equation four is unsigned, because it can be either positive or negative depending on the relative size of the capitulation and war payoffs.

All results are intuitive. Most obviously, player 2 should be more likely to retaliate, the more player 2 values war. Also fairly clearly, player 1 should be more likely to attack the more player 2 values both war and capitulation. These are the two possible outcomes if he attacks, and each has a strictly positive probability of occurring. Finally, and perhaps least obviously, whether player 1's probability of attacking is increasing or decreasing in player 2's value of war depends on whether player 1 values war or capitulation more. If player 1 values capitulation more, the expected value of attacking decreases as player 2's value of war increases, because the war outcome is more likely the more player 2 values war. The opposite relationship holds if player 1 values war more.

These relationships yield exactly the relationships that Signorino and Yilmaz derive in their MLE, which, of course, they have to. In particular, notice that we have derived, through standard comparative statics, both the nonlinear and the "conditionally monotonic" relationships illustrated in figure 5 of Signorino and Yilmaz (2003:562), and reproduced below as figure 3 (with a re-labeling of the axes for ease of exposition). To see why, let us consider the two parameters of interest separately.

First consider the relationship between player 2's value of war and the probability of player 1 attacking. When player 1's value of capitulation is less than player 1's value for war, which is fixed at zero in the figure, the probability of attack is increasing in player 2's value of war, and when the opposite is true the probability of attack is decreasing in player 2's value of war. This conditional relationship is identical to the comparative static derived in equation 4.

Next consider the relationship between player 1's value of capitulation and player 1's probability of attacking. One might suppose that the comparative static yields an unconditionally monotonic relationship, because its sign is unconditionally positive. However, in fact it does not. To see why, notice that equation 3 consists of two parts, the density function, $f(*)$ and $[1-G(u_2(war))]$. As player 2's value of war increases, the value of the first derivative, $\delta P/\delta u_1(cap)$, decreases. Interpreting this as a slope, we see that the probability of player 1 attacking is much less sensitive to player 1's value of capitulation as player 2's value of war increases. Thus, we derive exactly the same relationships through comparative statics that we would if we wrote Signorino's version of a strategic empirical model. The only difference between traditional comparative statics and what we have just done is that we have taken advantage of all of the information provided in the comparative static, not just its sign.

All of this analysis has been performed assuming a stochastic formal model. However, none of these derived relationships rely upon the random utility assumption. That is, we would derive exactly the same relationships between the parameters and the values of interest from a deterministic version of the same model. Using subgame perfection and solving backwards, we know that player 2 will retaliate when $u_2(cap)=0<u_2(war)$. Further, we know that player 1's decision to attack will depend upon what player 1 anticipates player 2 will do. If $u_2(war) > 0$, then player 1 will attack if $u_1(war) > u_1(SQ) = 0$, while if $u_2(war) < 0$, then player 1 will attack if $u_1(cap)>u_1(SQ)=0$.[7]

Now compare the probability of player 1 attacking across the two models. In the deterministic version, player 1's decision to attack depends conditionally on the three parameters. Player 1 attacks when $u_1(war)$ is sufficiently large if $u_2(war)$ is sufficiently large, and attacks when $u_1(cap)$ is sufficiently large if $u_2(war)$ is not sufficiently large. Figure 4 compares this solution to the stochastic version's solution, focusing on player 1's utility for war and player 2's utility for war. As can be seen, the solutions are not as different as one might have initially suspected. When $u_2(war)$ is small, the probability player 2 plays R is small, and both models predict that $p_1(A)$ is basically independent of $u_1(war)$. When $u_2(war)$ is large, the probability player 2 plays R is large, and both models predict that $p_1(A)$ is almost perfectly predicted by $u_1(war)$. The only place that the two model's predictions substantially differ are when $u_2(war)$ is moderate. In this case, the

[7] Note that this solution is identical to the solution for Signorino's regressor error version of a stochastic game theoretic model (Signorino 2003). In Signorino's regressor error model, as above, behavior is deterministic, only the regessor's ability to measure actor utility is stochastic.

deterministic model either predicts no attack with certainty (when $u_2(war)<0$) or attack with certainty (when $u_2(war)>0$), while the stochastic model predicts that $p_1(A)$ probabilistically increases as $u_1(war)$ increases. Thus, aside from some smoothing of the cutpoints, the predicted relationships across the two versions of the model are basically identical.[8]

Next consider the more complex crisis bargaining model from Signorino's 1999 article. For simplicity and brevity, we derive solely the subgame perfect, deterministic conditions that lead to war. Referring back to figure 1, Signorino defines the payoffs for each outcome as follows:

$$u_i(SQ)=D_{ij},$$

$$u_i(C_i)=-A_i,$$

$$u_i(Cj)=A_j, \text{ and}$$

$$u_i(W)=p_iA_j+(1-p_i)(-A_i-M_i),$$

$$\text{where } p_i=M_i/(M_i+M_j).$$

[8] The same findings hold with regards to the probability of war as well. We omit discussion here for the sake of brevity.

On a related note, Signorino (2003) makes a seemingly contradictory point of demonstrating that the type of error matters. That is, it matters whether one assumes regressor error, error on the part of the scholar in measuring parameters of the model, or agent or utility error, error built into the solution concept of the game itself. He demonstrates this point in two ways. First, he assumes that the random utility model is the correct error specification, and then demonstrates that the mean-squared error of the random utility and agent error models are smaller than the regressor error model. Second, still assuming a data generation process based on the random utility model, he demonstrates that the regressor error model will have higher average and maximal deviations from the true probabilities than the random utility or agent error versions. He conjectures, but does not demonstrate, why this would be the case. We believe that all of these differences are driven by the "smoothing" process of the random utility and agent error models. Differences in point predictions along these curves should generate increased mean-squared errors, increased average differences in probabilities, and increased maximal differences in probabilities. While plausible and interesting, this issue is not critical. If one assumes, as we stated above, that deterministic models are approximations off realities and that we generate implicitly probabilistic hypotheses from these deterministic models, then we get the same smoothed relationships whether we rely upon a stochastic or deterministic model. That is, the differences that Signorino identifies are a construct of the assumption that one would operationalize a deterministic model with hard cutpoints if one could.

Substantively, M_i are state *i*'s military assets, Ai are state *i*'s other assets (e.g. land and natural resources), and p_i is the probability that state *i* wins a military conflict.

Assuming state 1 did not initiate conflict, state 1 fights back when $u_1(W_2)>u_1(C_1)$. Substituting and simplifying yields $M_2<A_1+A_2$. If this constraint holds, state 2 chooses to fight when $u_2(W_2)>u_2(SQ)$. Substituting and simplifying yields $M_1<M_2A_1/(A_2+M_2)$. If this constraint does not hold, state 2 chooses to fight when $u_2(C_1)>u_2(SQ)$, which yields $A_1>D_{ij}$. Assuming state 1 did initiate conflict, state 2 fights back when $u_2(W_1)>u_2(C_2)$, which yields $M_1<A_1+A_2$. Finally, state 1 will initiate a fight when $u_1(argmax_2\{W1,C2\}) > u_1(argmax_2\{SQ, argmax_1\{C1,W2\}\})$, where $argmax_i$ identifies the outcome that maximizes the utility for player *i*. For brevity, we do not list out all of the constraints generated by this condition here.

With this solution in hand, let us reconsider Signorino's analysis. Using Monte Carlo analysis, Signorino demonstrates that his strategic model recaptures parameters correctly while the naïve, balance of power, and joint utility models do not. He demonstrates that the failure arises from the inability of the other models to correctly capture the nonlinear relationships between parameters and outcomes that are generated by strategic interdependence.

Figure 5 reproduces Signorino's figure 3 and 4, in which he illustrates this point with regards to the relationship between country 1's military power and the probability of war. The relationship in the strategic model is curvilinear; as country 1's military capability increases, the probability of war increases quickly, drops close to zero, increases again, and then tails off. In contrast, the naïve and balance of power logits estimate the relationship as linear and convex, respectively.

We certainly agree that Signorino has demonstrated the failure of these "typical" logits. However, none of the three logits estimated by Signorino rely upon correctly derived comparative statics. If one derives a comparative static over country 1's military capability from the deterministic solution characterized above, we observe exactly the curvilinear relationship derived in the QRE version of the game.[9] The probability of war is large when country 1's military capabilities are either moderately small or moderately large, while the probability of war is small when country 1's military capabilities are moderate, quite small, or quite large (see appendix for proof). This solution is graphed in figure 6. As with the comparison between the deterministic and stochastic versions of the crisis bargaining game discussed in the previous section, the only difference is that the QRE version of the game "smooths" the cutpoints in the deterministic solution and explicitly makes all of the derived relationships probabilitistic.

In sum, much of Signorino's critique is a functional form critique that has more to do with correctly deriving ones hypotheses from an underlying model, than it has to do with the actual estimation technique used to test those hypotheses. Thus, much of the failure of traditional logit analysis can be resolved with more careful generation of comparative statics. Nothing in the functional form critique requires one to move to generating estimators from stochastic formal models. Next we consider appropriate estimation techniques for testing these models.

[9] We fix $M_2=20$, $A_1=40$, $A_2=40$, and $D_{12}=0$ to generate the graph, as Signorino does.

Estimation Techniques

Can we use standard logits and probits to test strategic models? Signorino seems to imply that one should not use these approaches. For example, in motivating his use of the LQRE to evaluate the failure of traditional logit analysis, Signorino states, "… I analyze solely the effects of using a logit model when two states behave strategically. To do this, however, we *need* a method for incorporating the structure of strategic interdependence into statistical models of conflict {emphasis added}." The implication seems to be that Signorino does not believe that *any* logit can correctly incorporate the structure of strategic interdependence. In fact, standard logits and probits can be perfectly appropriate estimation techniques. Logits and probits are merely techniques that predict the likelihood that the outcome variable takes the value of 1. The strategic interaction can be captured in the specification of the covariates of the logit or probit model. To illustrate this point, let us return to the simple crisis bargaining model used in most of Signorino's work (Signorino 2003; Signorino and Yilmaz 2003; Signorino 2002).[10] We demonstrate our point by showing that probits and logits can be used for predicting both behavior (i.e. moves by players) and outcomes (i.e. whether war occurs or not).

Equations 5 and 6 are restatements of equilibrium behavior derived in the previous section. They characterize player 2's probability of retaliating and player 1's probability of attacking, respectively. As can be seen, if we assume the errors have type 1 extreme distributions, these equations yield simple logits, and if we assume they have normal distributions, these equations yield simple probits. The error distribution in equation five is a simple difference of errors, and the error distribution in equation six is a

[10] Identical results hold for Signorino's LQRE version of the more complex crisis bargaining model. Because deriving through these results would be redundant with those for this simpler model, we omit the additional proofs.

weighted combination of errors. Thus, there is nothing inappropriate in using standard comparative statics in a logit or probit to test predicted behavior in Signorino's simple crisis bargaining model.

$$p_2(R) = \Pr[e_{2c} - e_{2w} \leq u_2(war)] \text{ (5)}$$

$$p_1(A) = \Pr\left[e_{1sq} - Qe_{1w} + (1-Q)e_{1c} \leq u_1(cap) - Q(u_1(cap) - u_1(war))\right] \text{ (6)}$$

Equation 7 characterizes the probability of war in Signorino's model. Assume for the moment that the errors are independent. Again, familiar estimation techniques can be used with this function. For example, if we assume a normal distribution to the errors, the probability of war is a function of two independent, normal distributions. In this case, a perfectly appropriate estimation technique would be sequential probits in which the first probit predicts player one's move and is estimated on the entire dataset, and the second probit predicts player two's move and is estimated only on the observations in which player one attacked. If the errors are not independent, then we have a standard Heckman probit (Heckman 1979). Thus, comparative statics and existing estimation techniques are fine if one wants to test predicted outcomes in Signorino's model.[11] And, more generally, we see that one can use standard estimation techniques in the face of strategic interdependence. This is not to say that one cannot use Signorino's strategic modeling approach. It simply is unnecessary in these cases.[12]

---

[11] The same findings hold if we assume independent errors in Signorino's LQRE version of the more complex crisis bargaining model, as Signorino does in his analysis.

[12] Of course, Signorino's approach does have the benefit of allowing you to test the whole model simultaneously, rather than just particular predictions of the model. If one actually has the necessary data to estimate all of the parameters of the model simultaneously, one might wish to pursue this approach. Our

$$p(war) = p_1(A) * p_2(R)$$

$$= \Pr[\varepsilon_a \le u_1(cap) - Q(u_2(war))(u_1(cap) - u_1(war))] * \Pr[\varepsilon_r \le u_2(war)],$$

$$\text{where } \varepsilon_a = e_{1sq} - Qe_{1w} + (1-Q)e_{1c} \text{ and } \varepsilon_r = e_{2c} - e_{2w} \quad (7)$$

If standard techniques can be used in this example, can we say something about when they will and will not fail more generally? In fact, we can.

First, we can always use a standard logit or probit to predict behavior as long as the errors are type one extreme or normal, respectively. At any given node, a player is making an expected utility calculation over the various options from which he can choose. This structure generates an error term that is a linearly additive combination of weighted errors. Since a linearly additive combination of weighted type one extreme or normal errors is also type one extreme or normal, we can use a probit or logit to predict behavior, no matter how complex the strategic game.

Predicting outcomes is somewhat harder to generalize. However, the findings with regards to the probabilities of attack and war provide a clear roadmap. Think of attack and war as the two outcomes of interest. In the case of attack, there is one decision node, followed by the outcome. This structure yields a probit. In the case of war, there are two decision nodes, followed by the outcome. This structure yields sequential probits if the errors are independent, and a Heckman probit if they are not. Thus, one can think about this structure as a selection model, in which the first stage is the decision to attack, and the second stage is the decision to retaliate. And, more generally, what we observe is that

point is simply that one does not need to pursue this more complex estimation approach in order to test the model. Standard comparative statics and standard estimation techniques are perfectly appropriate.

adding a decision node prior to the outcome of interest yields a selection stage. For example, if we added a third decision node before the war outcome we were trying to predict—e.g. the defending player gets to decide whether to mobilize the military in anticipation of an attack—we would have a two stage selection model. Thus, simply enough, each additional decision node before the outcome of interest adds an additional selection stage to our econometric specification. If the errors are independent, we can always use sequences of logits or probits, no matter how many moves prior to the outcome of interest. If the errors are not independent, we would have to construct a multi-stage Heckman probit. Unfortunately, if errors are not independent, this estimation approach quickly becomes untenable because one cannot evaluate, and thus optimize, the MLE due to the high-dimensional integrals in the likelihood. However, this is a problem with maximum likelihood estimation in general, and will be a binding constraint whether one uses a Heckman approach or Signorino's LQRE approach. Upon hitting this constraint, the researcher would have to turn to other tools, such as Markov chain Monte Carlo estimation methods, which can be used with likelihood functions with high-dimensional Normal integrals (see, for example, Quinn, Martin, and Whitford 1999).

Conclusion

What do we conclude from this analysis? Nontrivially, Signorino (1999) is exactly correct in that one must carefully operationalize predictions generated by strategic behavior. Operationalizing naïve models that atheoretically include a host of covariates and loosely operationalizing predictions predicated on strategic models can both lead to inaccurate inferences. Thus, we completely agree that to correctly test a theory, one must correctly operationalize those tests. However, we also believe that correctly operationalizing comparative statics, even from simple deterministic models, and testing those predictions in a logit or probit is not prima facie inadequate. Correctly derived comparative statics do successfully capture all of the nonlinear and conditionally monotonic relationships that Signorino uses to highlight the failure of traditional logit analyses, and the error structures generated in simple strategic games are appropriately modeled in standard logits and probits. Thus, we do not believe that one must write a stochastic formal model, or that one must derive ones estimator directly from that stochastic model to adequately test a strategic theory. Where traditional estimation techniques become inadequate are the more complex models in which Signorino's recommended estimation technique will also become inadequate.

Figure 1: A Typical Bilateral Crisis Game

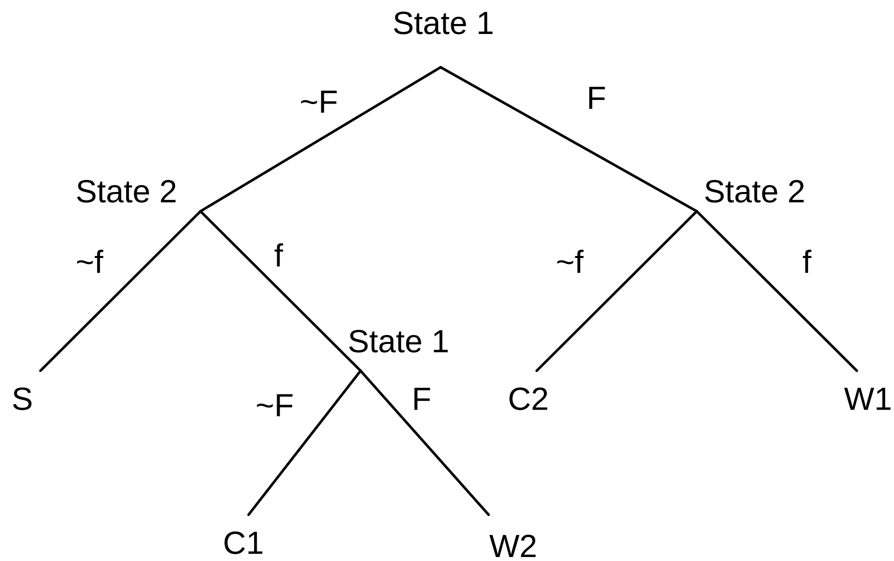


Figure 2: Strategic Deterrence Model

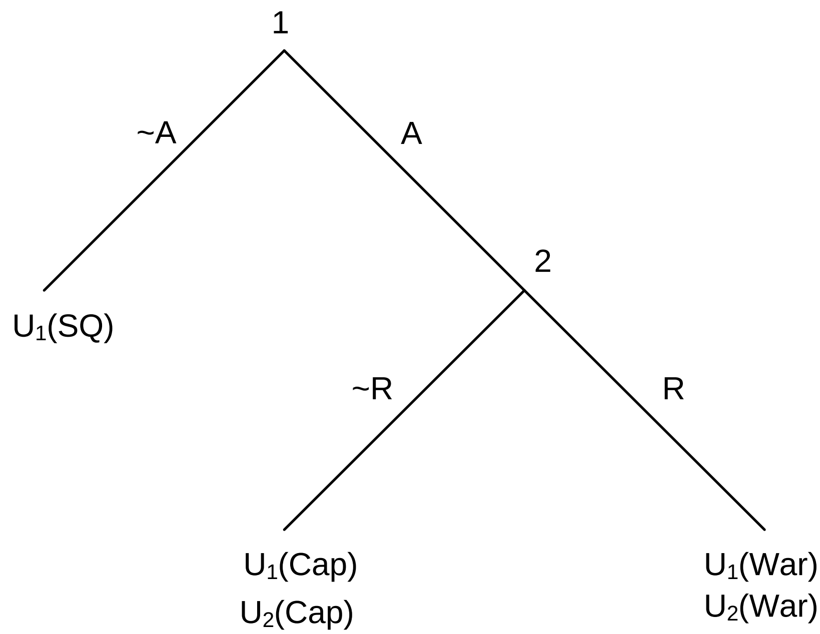
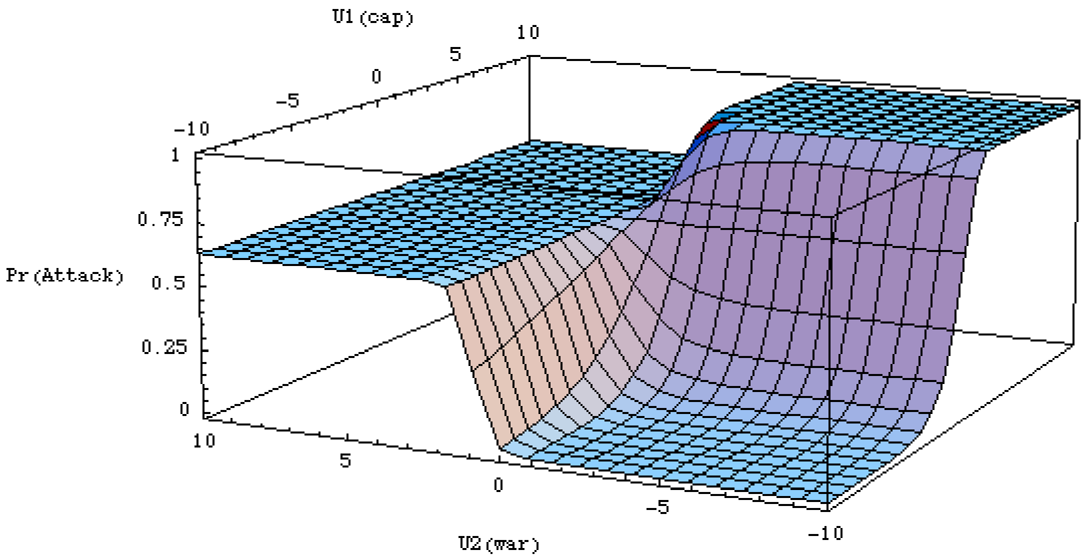


Figure 3[13]: Estimated Probabilities of Attacking: Strategic



[13] This graph was generated using a Type I Extreme Value distribution like Signorino's original graph. The mean and variance, however, are different. In this graph, the mean is zero and the variance is one, creating a graph that is shifted and stretched a little differently, but still reflecting the same dynamics for the game.

Figure 4: Deterministic Predictions of Attacking: Strategic

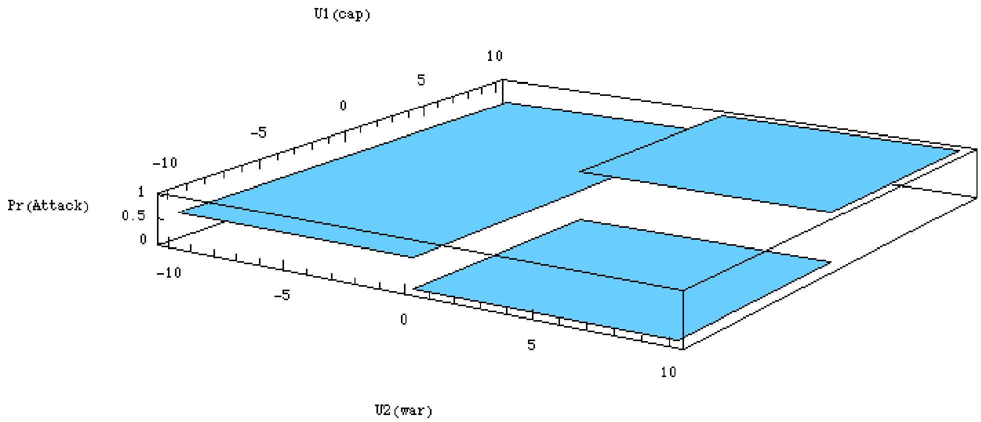


Figure 5: Logit QRE, Naïve Logit and Balance of Power Predicted Probabilities of War

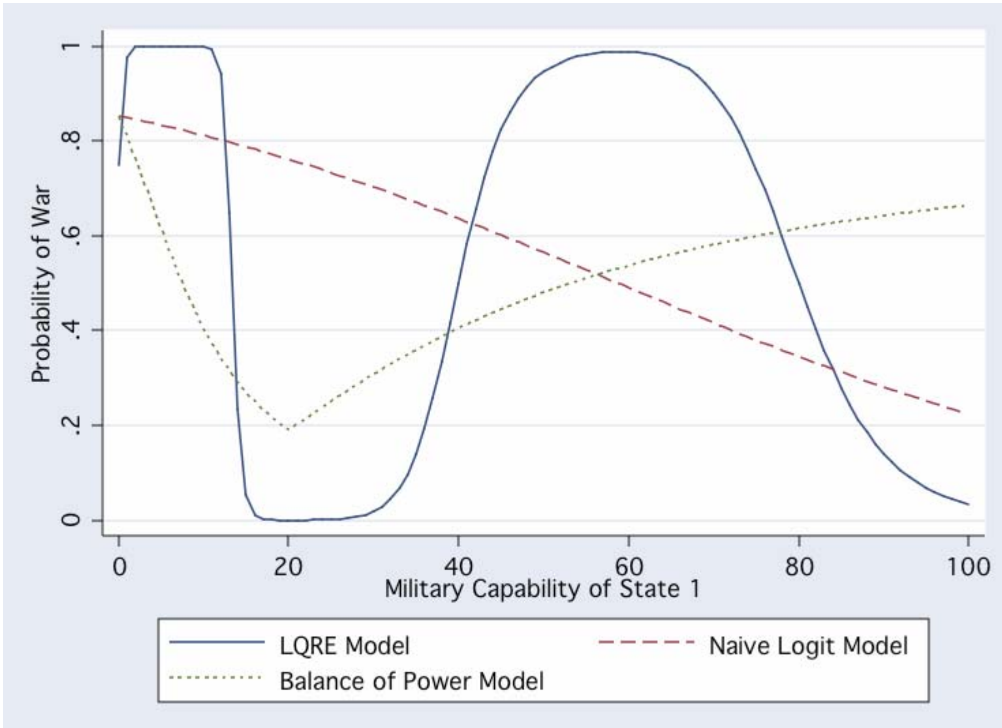
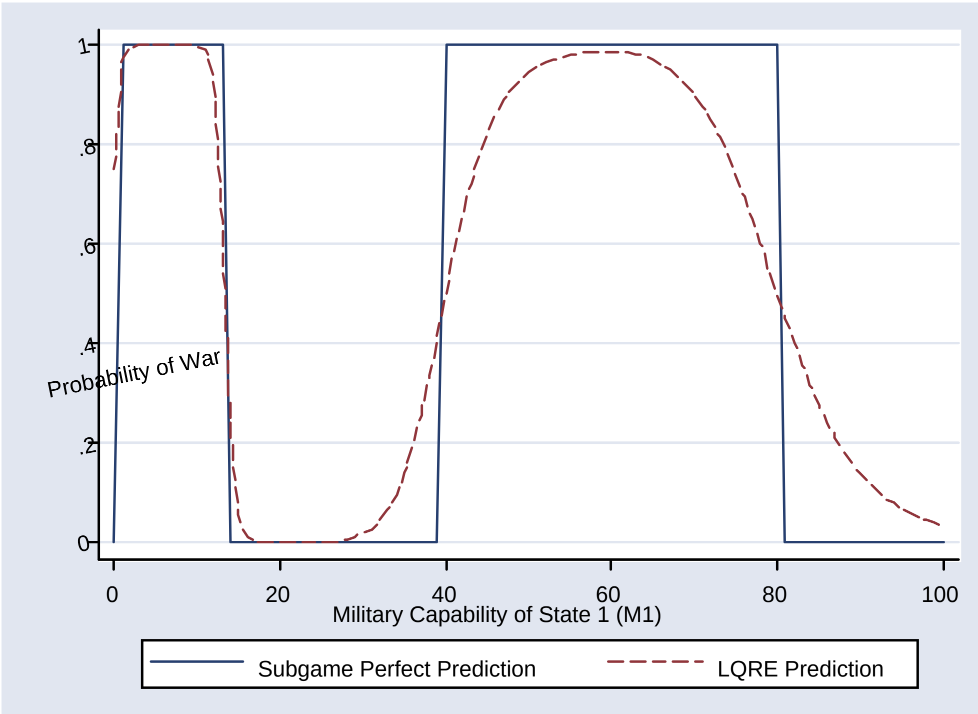


Figure 6: Subgame Perfect vs. LQRE Predictions



Appendix: Comparative Static of Simple Crisis Game (from Signorino 1999):

Using backwards induction, the cutpoints in the deterministic model can be derived at each node of the game, fully characterizing the relationship between the probability of war and State 1's military capability.

The first cutpoint (M_1^l) corresponds to the last decision node in which State 1 decides to fight, given it chose not to fight at the first node but then was challenged by State 2 at the second node.

$$U_1(war) > U_1(cap_1)$$

$$\frac{M_1}{M_1 + M_2} A_2 + \frac{M_2}{M_1 + M_2}(-A_1 - M_1) > -A_1$$

$$M_1(A_1 + A_2 - M_2) > 0$$

Substituting $A_1, A_2 = 40$ and $M_2 = 20$,

$$M_1^l > 0$$

The second cutpoint (M_1^c) marks the upper bound on M_1 for when State 2 will issue a challenge, given State 1 did not challenge at the first node.

$$U_2(war) > U_2(SQ)$$

$$\frac{M_2}{M_1 + M_2} A_1 + \frac{M_1}{M_1 + M_2}(-A_2 - M_2) > D = 0$$

$$M_1^c < \frac{M_2 A_1}{A_2 + M_2}$$

Substituting

$$M_1^c < \frac{800}{60} = 13.33$$

The third cutpoint (M_1^r) marks where State 2 would decide to fight, given that State 1 challenged at the first node. This sets the upper bound on war for M_1.

$$U_2(war) > U_2(cap_2)$$

$$\frac{M_2}{M_1 + M_2} A_1 + \frac{M_1}{M_1 + M_2}(-A_2 - M_2) > -A_2$$

$$M_1^r < A_1 + A_2$$

Substituting

$$M_1^r < 80$$

The final cutpoint, (M_1^m), denotes the amount of military capability State 1 would have to have to challenge at the first node.

$$U_1(war) > U_1(SQ)$$

$$\frac{M_1}{M_1 + M_2} A_2 + \frac{M_2}{M_1 + M_2}(-A_1 - M_1) > D = 0$$

$$M_1^m > \frac{M_2 A_1}{A_2 - M_2}$$

Substituting

$$M_1^m > 40$$

Heckman, James J. 1979. “Sample Selection Bias as a Specification Error.” *Econometrica*. 47: 153-161.

Lewis, Jeffery B., and Kenneth A. Schultz. 2003. “Revealing Preferences: Empirical Estimation of a Crisis Bargaining Game with Incomplete Information.” *Political Analysis*. 11:345-367.

McKelvey, Richard D., and Thomas R. Palfrey. 1998. “Quantal Response Equilibria for Extensive Form Games.” *Experimental Economics*. 1(1):9-41.

Quinn, Kevin M., Andrew D. Martin and Andrew B. Whitford. 1999. “Voter Choice in Multi-Party Democracies: A Test of Competing Theories and Models.” *American Journal of Political Science*. 43: 1231-1247.

Signorino, Curtis S. 1999. “Strategic Interaction and the Statistical Analysis of International Conflict.” *American Political Science Review*. 93(2):279-297.

Signorino, Curtis S. 2002. “Strategy and Selection in International Relations.” *International Interactions*. 28:93-115.

Signorino, Curtis S. 2003. “Structure and Uncertainty in Discrete Choice Models.” *Political Analysis*. 11:316-344.

Signorino, Curtis S., and Kuzey Yilmaz. 2003. “Strategic Misspecification in Regression Models.” *American Journal of Political Science*. 47(3):551-566.

Smith, Alastair. 1999. “Testing Theories of Strategic Choice: The Example of Crisis Escalation.” *American Journal of Political Science*. 42:1254-1283.