Advance Access publication May 18, 2005 *Political Analysis* (2006) 14:121–129
doi:10.1093/pan/mpi020

Learning about Learning: A Response to Wand

Kenneth A. Schultz
Department of Political Science, Stanford University,
Stanford, CA 94305-6044
e-mail: kschultz@stanford.edu

Jeffrey B. Lewis
Department of Political Science, UCLA,
Los Angeles, CA 90095-1472
e-mail: jblewis@ucla.edu

1 Introduction

We welcome the opportunity to respond to Wand's careful and detailed analysis of our paper (Lewis and Schultz 2003).[1] With the discipline's increasing inclination to move toward fully structural strategic choice models (e.g., Signorino 1999; Morton 1999), the issues that Wand raises are important to consider, as they bear on crucial questions of model construction and interpretation. Moreover, his work has allowed us to consider more carefully the properties of an estimator that we are in the process of applying to actual data.

We should start by noting that we are all in total agreement on the main implication of Wand's discussion: decisions about the specification of the theoretical model, including the nature and location of stochastic elements, have profound implications for empirical analyses built on such models. It was precisely this realization that motivated our original effort (see particularly Lewis and Schultz 2001). Signorino (1999) had forcefully made the case for capturing the strategic structure in empirical estimators, but his interest was primarily in internalizing the extensive form of the game. His use of quantal response equilibrium (QRE), and the information structure it implied, responded to the need to induce non-degenerate probability distributions over the outcome nodes. For the game discussed in that paper—Bueno de Mesquita and Lalman's (1992) "international interaction game"—this move made sense, since the model, like the QRE used by Signorino, assumed symmetric information. By contrast, we came to this project motivated by the burgeoning literature on the role of asymmetric information in crisis bargaining (e.g., Morrow 1989; Fearon 1995, 1997). Our goal was to build an estimator from a model that captured the informational assumptions highlighted by the existing theoretical literature. To do this, we reasoned, the empirical estimator had to internalize not only the extensive form of the game but also the appropriate information structure. The model and estimator proposed in Lewis and Schultz (2003) were designed to match as closely as possible the simplest models used in the existing literature.

We then sought to establish that the informational assumptions mattered—that applying a QRE to same extensive form, without capturing the information structure stressed in the

[1]All references to Wand throughout are to Wand 2005.

© The Author 2005. Published by Oxford University Press on behalf of the Society for Political Methodology.
All rights reserved. For Permissions, please email: journals.permissions@oupjournals.org

literature, could lead to different inferences from the same observational data. To be clear: It was not our intention to propose a method for distinguishing between our model and one with symmetric information solved using QRE. Nor was it our intention to show that there would always exist differences between the two. Rather our aim was to make a more limited "existence proof" that there could be differences in the comparative statics of the models and in the inferences drawn from estimators built from them. Given our belief that the model with asymmetric information was the proper starting point for analysis, establishing *any* substantive difference between that model and the one solved with QRE was sufficient to prove the value of internalizing the asymmetric information structure.

Wand's criticism of our article has two related elements. The main claim is that the differences we identified were a product of arbitrary choices that served to restrict the amount of stochastic variability in our model. Wand argues that the differences were due not to the fact that our model has asymmetric information and the QRE has symmetric information, but rather to the arbitrary manner in which we introduced asymmetric information. In other words, had we chosen to introduce asymmetric information in a different and perhaps more general way, the contrasts we uncovered would have largely gone away. The second argument that Wand makes is that, in interpreting the figures that served as our demonstration proof, we erroneously attributed differences between the models to the role of learning, which is permitted in our model but not in the QRE version.

We have two main responses two these points:

1. Had our main purpose been to generate comparative statics that could empirically discriminate between the information symmetry and asymmetry, then our assumptions would have been both arbitrary and unreasonable. As Wand has ably demonstrated, differences in the distribution of types between our model and the QRE variant did not allow for a clean comparison that isolated the effects of private information and learning. As just noted, however, this was not our main purpose. Nor did we intend to suggest that one could never specify a game such that its PBE has similar comparative statics to the QRE variant. Rather, we sought to develop a model that accurately captured the assumptions found in the theoretical literature on crisis bargaining. In this respect, Wand's analysis actually underscores our subsequent contention that the QRE model was not appropriate given our interests. As we will show, to generate a model with asymmetric information whose PBE has the same properties as the QRE, one needs to make questionable assumptions about the preference orderings actors can possess.
2. Wand is substantially correct in showing that our discussion overemphasized the role of learning in the examples presented, and in two specific instances, the interpretations we gave were unwarranted. One of our motivations for using a game theoretic model to inform our empirical analysis is that strategic settings can generate complex and counter-intuitive relationships (Lewis and Schultz 2003, p. 346). In trying to supply simple and intuitive interpretations of some of the relationships we uncovered in our article, this motivation should also have served as a warning. Wand has appropriately reminded us of how muddy these waters sometimes are.

Nonetheless, this does not mean that learning and strategic manipulation of information have no impact on the observed behavior in our model. Moreover, even though learning may appear to be "negligible" over much of the range of parameters used in our figures, that does not mean that the model will never permit substantial amounts of learning. We show that the maximum amount of learning that is possible in the model depends crucially on the rate of challenges that are observed in equilibrium. At the relatively low challenge rates that are observed empirically, updating can be quite pronounced.

2 Assumptions and Arbitrariness

There are two ways in which Wand sees our assumptions as arbitrary. The first has to do with the number and location of the stochastic shocks. While we put stochastic terms in three of the eight utilities, Wand suggests that a more general model would allow for stochastic variability in seven of the eight of utilities (though two of these disturbance terms—those that appear in *B*'s payoff from war and its payoff from *A*'s backing down—are constrained to be equal).[2] The second arbitrary assumption on our part was to normalize the variance of the disturbance terms to one. Since we had already normalized the value of the disputed good to one, this decision entailed loss of generality, one that Wand shows drives the nonmonotonicity in one of our examples.

The second point is the easiest to address. The loss of generality caused by this normalization is something that we were completely open about and one that we did not even attempt to justify. Wand (2005) quotes from this section verbatim, but it is worth noting what we wrote both before and after the passage he excerpted: "This normalization is not innocuous. . . . Whether or not this assumption is plausible—it most likely is not—it is not one that can be tested in the data" (Lewis and Schultz 2003, p. 356). We could be so cavalier about this because, for the purposes of making an existence proof, it was sufficient to show that there exists a set of parameters such that the models generate different comparative statics. In retrospect, we should have reiterated this caveat in subsequent sections of the article to ensure that our results were appropriately qualified. We should also note that, in the empirical work that we are currently conducting, we have been able to relax this restriction by allowing to vary the constants that we had previously fixed to one (Lewis and Schultz 2005). As we showed in our original paper (p. 356), adding covariates to the analysis makes the relaxation of this restriction possible. Hence, whether or not there is nonmonotonicity in the estimated relationships is something that we can now allow the data to tell us.

Wand's argument about the placement of stochastic terms in different payoffs requires more thorough consideration. Although Wand's alternative specification may be intuitively more appealing on the grounds that it imposes fewer restrictions, generality is not necessarily a virtue in this context.[3] Adding unbounded stochastic shocks to every payoff eliminates the analyst's ability to impose sensible restrictions on the payoff orderings that are possible. Indeed, while Wand emphasizes the amount of uncertainty in the distribution of player types, the main differences between our model and his alternative PBE arise from the fact that the latter puts positive probability on preference orderings that we sought to rule out. For example, since *A*'s payoff from the status quo and its payoff from concessions by *B* are both subject to unbounded shocks in Wand's model, there is nonzero probability that *A* strictly prefers the status quo to getting the good without a fight. In order words, this assumption creates a type of *A* that would choose the status quo *even if B were to concede with probability one.* In the simple numerical examples we used, in which the difference between the status quo and the value of concessions was normalized

[2]*B*'s payoff from the status quo has no impact on the game, so there is no point in adding a stochastic term to this quantity.

[3]Although Wand's alternative PBE model is more general in the sense of having fewer restrictions, our model cannot be said to be nested within his because of the equality constraint that he imposes on two of *B*'s shocks.

to one, the probability associated with such a preference ordering is $\Phi(\frac{0-1}{\sqrt{2}}) = 0.24$. A somewhat smaller fraction of types sees the status quo as the most preferred outcome, in which case they have a dominant strategy not to challenge. Similarly, because *B*'s payoffs are all subject to shocks, there are types of *B* that would prefer to concede even if *A* were certain to back down.

The existence of such types runs counter to the initial assumption of the game that there is a good in dispute—that is, there is something that *B* has that *A* would like to have (Lewis and Schultz 2001, p. 348). Since in theory there are always goods that one state could transfer to another, allowing such preferences orderings in the population is not unproblematic and, at a minimum, requires some additional assumptions about the game. One way to justify a shock to *A*'s status quo payoff is to assume that there are ex ante costs (or benefits) to issuing a challenge. It is plausible that some of the actions required to make a challenge, such as mobilizing and deploying troops, generate such sunk costs (Fearon 1997); however, it is not clear that the appropriate way to introduce these costs is through unbounded and privately observed shocks. Doing so forces us to assume that the unobservable component of these costs is potentially quite large—large enough to swamp the value of getting the good without a fight. Whatever one thinks about this assumption, it is certainly not trivial given the large proportion of *SQ* outcomes that are observed in actual data (Lewis and Schultz 2005). Analysts will have to decide whether they want their estimator to account for these outcomes by inferring that a sizable share of the population prefers the status quo to getting concessions.

We fully appreciate that Wand's purpose in designing the PBE2 was not to propose a more realistic model, but rather to show that one could construct a game with asymmetric information whose PBE has similar properties as the game with symmetric information solved using QRE. The exercise nicely demonstrates what one must do to create such apparent equivalence: In order to make fully rational actors of the PBE behave in observationally similar ways to the error-prone agents envisioned by the QRE, it is necessary to assume that former include types who want to do what the latter only do by mistake. The perceptual error that leads a quarter of all agents in the QRE to choose the status quo even if *B* is certain to concede becomes, in the PBE2, a rational choice.

To be fair, our model is not entirely immune from this problem, since the stochastic shock on *A*'s "audience cost" for backing down creates types for whom backing down is the most preferred outcome. The existence of this shock can also lead to counterintuitive effects, since *B* might infer from a challenge not that *A* is willing to fight but that *A* really wants to back down. The fact that this inference is possible explains why, in some of the examples Wand gives, the posterior probability that *A* will fight is lower than the prior probability.[4] Why then introduce such a shock? The somewhat unsatisfying answer is that at least two of *A*'s payoffs must have stochastic elements in order to ensure that every outcome happens with nonzero probability for every configuration of parameters.[5] Hence, our model introduced just enough stochastic variability to avoid the "zero likelihood" problem. Putting in additional shocks may add generality to the model, but it is possible to

[4]It is relatively easy to see that the direction of updating will depend upon the initial bias in the prior. If $\bar{a} < \bar{W}_A$ then the prior probability that *A* will fight, P_F, is greater than 0.5 and updating will lead to $P_{F|C} > P_F$. If $\bar{a} > \bar{W}_A$ then $P_F < 0.5$ and updating will lead to $P_{F|C} < P_F$.

[5]Of course, it is not necessary that such shocks represent asymmetric information, and in an earlier version of this article, we considered a case in which the shock on *A* payoff from backing down was known to *B* but not to us (Lewis and Schultz 2001). While this assumption rules out situations in which *B*'s posterior is less than its prior, it still admits the possibility that some types of *A* most prefer the *BD* outcome.

Table 1 Estimated parameters for two sets of outcome probabilities

	Outcome 1	*Outcome 2*
Probability:		
P_{CD}	0.15	0.30
P_{SQ}	0.50	0.41
P_{BD}	0.25	0.21
P_{SF}	0.10	0.08
Parameter:		
$\bar{a}$	−0.72	−1.13
$\bar{W}_A$	−1.33	−1.77
$\bar{W}_B$	−1.98	−2.50

Note. The first column shows the game payoffs implied by the given set of outcome probabilities. In the second column, the probability of "concede" is doubled from its value in the first column, holding the relative probabilities of the other outcomes fixed.

be overly general. If the goal is to rule out unreasonable payoff orderings, then an argument could be made to include fewer stochastic terms.[6]

To see how this difference in permissible types matters, it is useful to revisit one of the results that we presented and Wand reinterpreted: the observation in Fig. 5, column 3, of his paper that, as the proportion of *CD* outcomes increases, our model and his can make opposite inferences about the direction of change in $\bar{a}$ and $\bar{W}_A$. What accounts for this? As we all note, an increase in the proportion of *CD* outcomes must derive in part from a decrease in *B*'s value for war, $\bar{W}_B$, which makes *B* less likely to resist. In our model, this change creates a very strong temptation for more types of *A* to make a challenge: after all, all types in our model would make a challenge if the probability of resistance were essentially zero. Hence, in order to keep the probability of *SQ* from going to zero, challenges have to become more costly at the same time that *B*'s probability of resistance goes down. This is accomplished by lowering the estimates of $\bar{a}$ and $\bar{W}_A$. In the two variants considered by Wand, by contrast, there is a sizeable fraction of types who, either through their preference ordering (PBE2) or by mistake (QRE), will choose the status quo even as the *B* becomes less likely to resist. The estimators built on those models, therefore, have to encourage more challenges in order to get the fraction of *CD* outcomes to increase beyond a certain point. This requires increasing $\bar{a}$ and $\bar{W}_A$.

To see a numerical example of what is going on in our model, consider what happens when the frequency of the *CD* outcome changes from 0.15 to 0.30, holding the relative frequency of all other outcomes constant. Table 1 shows the estimates generated by our PBE model for each scenario. As already noted, the estimates for all three parameters decrease as the frequency of *CD* outcome increases. To see precisely why, consider what the outcome distribution would be if only $\bar{W}_B$ decreased, without changing the other two parameters. The answer is striking: decreasing only $\bar{W}_B$ from the first value to the second causes the frequency of the status quo to drop from 0.50 to 0.014. That is, *SQ* outcomes virtually disappear as almost every type of *A* now has an incentive to make a challenge. In order to move the status quo back up to its new level of 0.41, this temptation to challenge

[6]Alternatively, it might be worth exploring the introduction of shocks with bounded support.

has to be restrained by decreasing $\bar{a}$ and $\bar{W}_A$. By contrast, in the QRE and PBE2 models, the same change in the outcome distribution leads to a larger decrease in the estimate of $\bar{W}_B$ but virtually no change in the other parameters.

Put another way, our model and Wand's two variants give different answers to the question of how the status quo outcome can remain relatively frequent even as *B*'s probability of resistance decreases. For our model, the only possible answer is that war and backing down must become relatively less attractive in order to restrain the temptation to challenge. In the other two models, at least some of the answer comes from the fact that a sizable fraction of potential challengers simply do not want *B*'s concessions or, in the QRE formulation, they do want those concessions, but they have made a mistake in evaluating their options.

3 The Amount of Learning in the Model

Wand demonstrates that the differences between our model and the QRE variant in the comparative statics that we presented are not due to the fact that there is a significant amount of updating in the former. He also argues that, in both PBE models, the amount of learning is "negligible." Though it is true that we overstated the degree to which learning accounted for the differences evident in our figures, looking for the effects of learning only through the difference between the posterior and prior understates the role that learning and signaling play in these models. Moreover, we can show that substantial amounts of learning are possible in our model, particularly when the probability of a challenge is low.

At the most fundamental level, of course, all equilibrium behavior in a PBE takes place in the shadow of learning. One cannot even identify the equilibrium without looking at how different strategies would influence beliefs and how those beliefs would subsequently influence strategies. Many strategy-belief combinations are ruled out as equilibria because the strategies are not sequentially rational given the learning that they would induce. For example, in the game considered here, a maximally informative equilibrium in which only types that would fight make a challenge typically does not exist. The inference that *B* would make upon seeing a challenge, and the action it would subsequently take, would create an incentive for some types of *A* to bluff, thereby undermining the equilibrium. Since Bayesian updating is built into the derivation of the equilibrium strategies, the possibility of updating affects everything, even if little or no updating takes place in an actual equilibrium.

Indeed, under many conditions, a PBE can exhibit "pooling" behavior, which happens when all types behave in the same way. In such cases, no updating can take place in equilibrium, since learning requires that there be meaningful variation in the strategies played by different types. This is particularly likely when the actors have some conflicting interests, as they do here. In such settings, optimal behavior by an informed player may involve denying its opponent the opportunity to learn. For example, if the prior probability that *A* will fight is sufficiently large that *B* is very unlikely to resist based on that prior, unresolved types of *A* have a strong temptation to mimic resolved types and thereby "free ride" on the state's strong reputation. Under these conditions, it is not rational for weaker types of *A* to reveal themselves.

With this in mind, it is instructive to revisit Wand's Fig. 2. The top row of graphs in that figure permit a comparison of the prior and posterior probabilities that *A* will fight. The dotted line corresponding to the QRE shows the prior probability (which is the same for all three models), while the lines for the PBE models show the posterior probability that holds in equilibrium, given the parameters. Comparing this row with the bottom row depicting the probability of a challenge, it becomes apparent that the prior exactly equals the posterior in our PBE1 whenever the probability of a challenge is effectively one.[7] It is precisely under these conditions that no learning can take place: there is simply no variation in the signals that different types send in equilibrium.[8]

Notice, however, that the PBE2 can in some cases exhibit no updating even when the probability of a challenge is less than one. The reason lies in the existence of types that strictly prefer the status quo. Because the frequency of such types is common knowledge, their decision not to challenge conveys no information about what will happen at the final decision node. Hence, even though there is variation in the signals sent in equilibrium, this variation is not informative. This observation suggests that comparing the difference between the posterior and prior is not the only way to compare the amount of learning in the two models. In some cases, differences in the amount of learning are evident in the different challenge rates that support the same posterior.

More generally, we know that the amount of updating in the model must be tied to the rate of a challenges. Figure 1 shows the maximum amount of learning that can take place as a function of the rate of challenge, where learning is defined as the difference between the prior and posterior probabilities that A will fight. The solid line on each panel gives the logical upper bound on the degree of learning given only the structure of the game and the assumptions that $P_F > 1/2$ and that the posterior probability must be at least as large as the prior ($P_{F|C} \geq P_F$). Note that for a given P_F and P_C, the maximum $P_{F|C}$ is $\min(P_F/P_C, 1)$.[9] When $P_C > P_F$, not all challengers can be of the type that would fight and, therefore, $P_{F|C} < 1$. As P_C goes to 1, any differences between the population of all A types and the population of A types that challenge vanish, and $P_{F|C}$ approaches P_F. Thus, in the examples with very high challenge rates presented in our paper and in Wand's response, the lack of learning is not so much a matter of parametric specification as it is a matter of basic probability.

The dotted lines show the maximum amount of learning at each challenge rate given our parameterization of the game. At low rates of challenge, the maximum possible learning approaches the logical maximum. However, as P_C increases the degree of learning falls off quickly. That the maximal amount of learning in our model is (sometimes far) less than the logical upper bound is not all that surprising. Because of the conflictual nature of the bargaining situation considered, maximal revelation of private information would not be expected (Fearon 1995). Further, because in our model some A types prefer backing down to the status quo, the posterior probability of fighting will generally be less than one even as the probability of resistance approaches one. Indeed, the higher degree of learning that occurs in our model when $P_F = 0.8$ as opposed to $P_F = 0.6$, is due to there being relatively fewer As that prefer backing down to the status quo when $P_F = 0.8$.

What this exercise also shows is that it is straightforward to find values of the parameters which, in equilibrium, generate sizable differences between the prior and posterior beliefs. Such learning generally requires that the probability of a challenge be small, which is achieved when $\bar{a}$ and $\bar{W}_A$ are low. For example, consider the case in which $(\bar{a}, \bar{W}_A, \bar{W}_B) = (-3.5, -2.9, -0.5)$. With these parameters, the prior probability that A will fight is 0.66, and the posterior probability that A will fight, given that it has made

[7]We say "effectively" because strictly speaking, the probability of a challenge cannot be equal to one for finite values of the parameters.

[8]The other condition that would prevent learning from taking place is if $\bar{a} = \bar{W}_A$, in which case both the prior and posterior will equal 0.5 for any rate of challenge.

[9]To see this, write $P_F = P_{F|C}P_C + P_{F|\sim C}(1 - P_C)$. Rearranging, $P_{F|C} = P_F/P_C - P_{F|\sim C}\frac{1-P_C}{P_C}$. Maximizing over $P_{F|\sim C} \in [0, 1]$ subject to $P_{F|C} \leq 1$, the maximum $P_{F|C} = \min(\frac{P_F}{P_C}, 1)$.

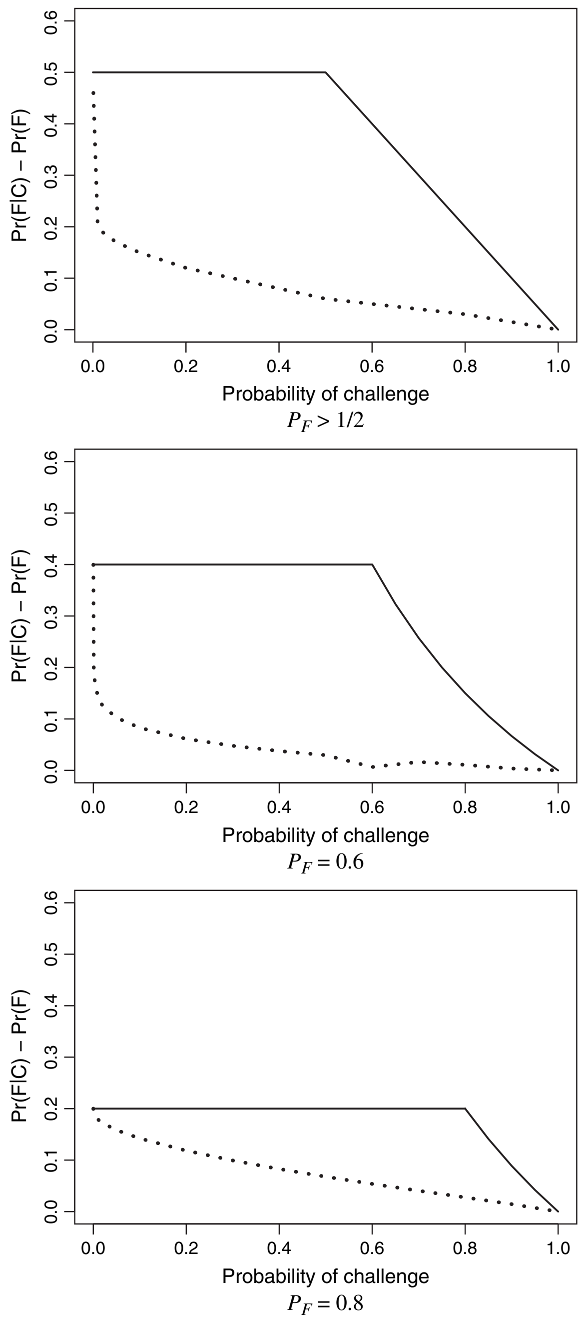

Fig. 1 Learning in the simple crisis bargaining game. Each panel shows the maximum possible "learning" (difference between the posterior and prior probabilities of fighting) in our simple crisis bargaining model as a function of the rate of challenge. The solid lines show the upper bound on the possible learning for any parameterization of the game's payoffs. The dashed lines give the maximum possible learning in our parameterization as described in the text. In the left panel, the prior probability of "fight" (P_F) is constrained only to be greater than 1/2. In the center and right panels, P_F is fixed at 0.6 and 0.8, respectively.

a challenge, is 0.79. The difference between the posterior and the prior is 0.13, but since the maximum possible change given the prior is 0.44, the "proportionate reduction in error" is 30%. Another way to think about this is that, based on its prior beliefs, B would have resisted with probability 0.51, but based on its posterior, it will resist with probability 0.41. Hence, the learning in this case leads to a 20% decrease in the probability of resistance. It should be added that such low values of $\bar{a}$ and $\bar{W}_A$, and the low challenge rates associated with them, are not wholly unrealistic. In the preliminary data analysis we have done so far, the observed frequency of challenges is less than 5% and estimates of $\bar{a}$ and $\bar{W}_A$ tend to be in the neighborhood of -3 (see Lewis and Schultz 2005). Hence, the model does allow for considerable amounts of learning, even if dramatic instances of learning were not evident in the figures we presented.

In the final analysis, this exchange reinforces our initial views and conclusions, even if some of our specific interpretations along the way have been shown to have been mistaken. The assumptions that analysts build into their models matter. Given that many of these assumptions are untestable, the only solid foundation we have to rest on is our theories. Hence, model construction should be governed to as large an extent as possible by those theories. While this exchange has forced us to look more carefully at aspects of our estimator, our view remains that the model we are using is a sensible one to apply on real data—the best place to start given the theoretical literature. Moreover, even if the differences between our model and the QRE variant are not as profound or as inevitable as we originally suggested, nothing in the exchange should be taken as suggesting that an estimator based on the QRE would be appropriate for our needs.

References

Bueno de Mesquita, Bruce, and David Lalman. 1992. *War and Reason: Domestic and International Imperatives.* New Haven: Yale University Press.

Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49(Summer):379–414.

Fearon, James D. 1997. "Sunk Costs versus Tying Hands." *Journal of Conflict Resolution* 41:68–90.

Lewis, Jeffrey B., and Kenneth A. Schultz. 2001. "Limitations to the Direct Testing of Extensive Form Crisis Bargaining Games." Paper presented at the annual meeting of the Midwest Political Science Association.

Lewis, Jeffrey B., and Kenneth A. Schultz. 2003. "Revealing Preferences: Empirical Estimation of a Crisis Bargaining Game with Incomplete Information." *Political Analysis* 11(Fall):345–367.

Lewis, Jeffrey B., and Kenneth A. Schultz. 2005. "Estimating State Preferences in International Crises." Unpublished manuscript, Stanford University.

Morton, Rebecca B. 1999. *Methods and Models.* Cambridge: Cambridge University Press.

Morrow, James D. 1989. "Capabilities, Uncertainty, and Resolve: A Limited Information Model of Crisis Bargaining." *American Journal of Political Science* 33(November):941–972.

Signorino, Curtis S. 1999. "Strategic Interaction and the Statistical Analysis of International Conflict." *American Political Science Review* 93(June):279–298.

Wand, Jonathan. 2005. "Comparing Models of Strategic Choice: The role of uncertainty and signaling." *Political Analysis* doi:10.1093/pan/mpi017.