Political Analysis Advance Access published October 15, 2007

doi:10.1093/pan/mpm029

Statistical Backwards Induction: A Simple Method for Estimating Recursive Strategic Models

Muhammet Ali Bas
Department of Government, Harvard University, CGIS N209,
1737 Cambridge Street, Cambridge, MA 02138
e-mail: mbas@gov.harvard.edu (corresponding author)

Curtis S. Signorino
303 Harkness Hall, Department of Political Science,
University of Rochester, Rochester, NY 14627
e-mail: curt.signorino@rochester.edu

Robert W. Walker
Department of Political Science, Center for Applied Statistics,
Washington University in Saint Louis, Campus Box 1063,
One Brookings Drive, St. Louis, MO 63130
e-mail: rww@wustl.edu

We present a simple method for estimating regressions based on recursive extensive-form games. Our procedure, which can be implemented in most standard statistical packages, involves sequentially estimating standard logits (or probits) in a manner analogous to backwards induction. We demonstrate that the technique produces consistent parameter estimates and show how to calculate consistent standard errors. To illustrate the method, we replicate Leblang's (2003) study of speculative attacks by financial markets and government responses to these attacks.

1 Introduction

Strategic interaction is a fundamental consideration in the study of political choice. The choices of members of Congress to seek reelection or retire is almost certainly related to the decisions of potential challengers (Carson 2003). Executives must decide whether or not to veto legislation in light of the potential for a legislative override (Carson and Marshall 2004). In international politics, Signorino and Tarar (2006) examine the strategic sources of extended deterrence and Leblang (2003) shows that the strategic interaction between governments and markets is a critical element in understanding currency crises. All of the aforementioned theoretical insights require empirical tests that roughly conform to the important structural elements of the decision-making environment.

Authors' note: Our thanks to Kevin Clarke, John Londregan, Jeff Ritter, Ahmer Tarar, and Kuzey Yilmaz for helpful discussions concerning this paper. A previous version was presented at the 2002 Political Methodology Summer Meeting.

Despite the increased use of choice (or game) theoretic explanations in modern political science research, there is often a subtle disconnect between theories and the empirical tests of these theories. "Indirect" statistical tests of formal models generally fail to properly characterize the hypothesized relationships in statistical testing.[1] Techniques for ameliorating this problem have been previously presented (Signorino 1999, 2002, 2003; Lewis and Schultz 2003; Signorino and Tarar 2006). However, these methods generally require researchers to program and then optimize a frequently complex, problem-specific model.

In this paper, we simplify the estimation of statistical strategic models in an effort to bridge the disconnect between theory and empirical analysis. Our motivation is primarily practical with respect to most contemporary research in political science. Signorino (1999, 2002, 2003) provides a general framework for estimating statistical strategic models, along with different ways of conceptualizing the uncertainty that can make strategic models "statistical." Based on this, the program *Strat* was developed, allowing scholars to estimate strategic models relevant to their substantive research.[2] The types of games that can be estimated in *Strat*, however, are limited to six or seven relatively simple—though quite common—decision structures. Scholars desiring to estimate parameters in a strategic model that is not included in *Strat* must therefore construct and estimate their own maximum likelihood models.

In the following sections, we first demonstrate a useful simplification that allows recursive statistical strategic models to be estimated with simple variants of logit and probit—and in a way that is consistent with the underlying theory. The method is very similar to the game-theoretic notion of backwards induction. We also show how to correctly estimate the standard errors (SEs) using this simplified technique. Finally, using the proposed method, we replicate Leblang's (2003) study of currency crises.

2 Referent Example

It will be helpful to employ an example throughout our analysis. Figure 1 displays the structure of our referent model. In this game, player A must choose between playing Right (R) or Left (L). If A chooses Right, player B must then choose whether to play right (r) or left (ℓ). Playing left results in outcome $R\ell$. If B chooses right, (Rr) is the outcome. The players utilities are shown at the terminal nodes—e.g., $U_{\mathrm{A}}(L)$ is state A's utility for the L outcome, and $U_{\mathrm{B}}(R\ell)$ is state B's utility for the $R\ell$ outcome.

Structurally, the model is quite simple. The fact that players do not make their decisions simultaneously (or without knowledge of other players' moves) makes this a fully *recursive* model. Assuming the players and we as analysts have complete information concerning the utilities, then for any specification of the utilities, the model can be solved via backwards induction for the subgame perfect equilibrium.

The logic of backwards induction is relatively straightforward. We start at the end nodes of the game and then ask what choice the player will make at the preceding decision node. With perfect and complete information, we can determine which option will be selected. Given that knowledge, we can step back to the next preceding decision node and determine what that player will do. For example, consider Fig. 1. Let us first suppose that A's preferences are $Rr \succ L \succ R\ell$ and B's are $Rr \succ R\ell$.[3] If we start with B's decision, we can see that, if A chooses R, B will choose r, leading to Rr. Player A knows this, so her decision

[1] Signorino and Yilmaz (2003) focus on this problem.

[2] "*Strat*: A Program for Estimating Statistical Strategic Models" is available at www.rochester.edu/College/PSC/signorino.

[3] $x \succ y \succ z$ means that option x is preferred to y, that y is preferred to z, and that (via transitivity) x is preferred to z. We omit L in B's preferences because it is not relevant for this game.

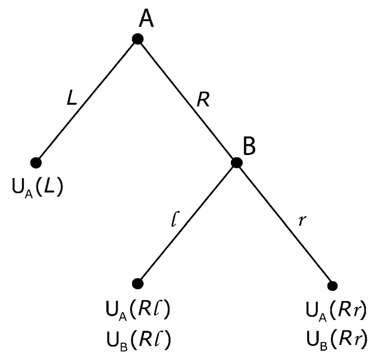


Fig. 1 Two-player game. Player A must choose whether to play Left (L) or Right (R). If Player A plays R, Player B must then choose whether to play Left (ℓ) or Right (r).

between L and R is really a decision between outcomes L and Rr. Given A's preferences above, A will choose R, leading to the subgame perfect equilibrium outcome Rr. Now suppose A's preferences are as stated above, but B's are exactly the opposite: $R\ell \succ Rr$. In this case, if A chooses R, then B will choose ℓ. Because A prefers L to $R\ell$, the equilibrium outcome will be L. It is important to note that it is this ability to work from the end of the game back toward its beginning that allows us to employ the statistical technique we will later present.

In its present form, the model is not statistical. To construct a statistical strategic model, we need to assume some form of uncertainty on the part of the players and/or analyst. There are a number of ways we might do so: agent error, private information by the players, or complete information by the players but imperfect measurement (or specification of regressors) by the analyst (Signorino 2003). Of the three, the one that is consistent with our simplified estimation technique is the agent error specification.[4] We now give Fig. 1 an agent error representation.

2.1 *The Statistical Strategic Model*

To fully specify the statistical model, we assume that there is a shock to the players' expected utilities for their actions. Consider B's decision. B's utilities will be specified as

$$U^*_{\mathrm{B}}(\ell) = U_{\mathrm{B}}(\ell) + \alpha_\ell = U_{\mathrm{B}}(R\ell) + \alpha_\ell, \tag{1}$$

$$U^*_{\mathrm{B}}(r) = U_{\mathrm{B}}(r) + \alpha_r = U_{\mathrm{B}}(Rr) + \alpha_r, \tag{2}$$

where $U^*_B(\cdot)$ is considered the "true" utility, $U_B(\cdot)$ is the component of the utility that is observable to the other players and to the analyst, and α is a random private component, observable only to player B. We assume B maximizes her (true) utility.

Because we as the analysts do not observe the α terms, we can only make probabilistic statements about whether B is likely to play right or left. Once a density is assumed for α, derivation of the choice probabilities is straightforward, following traditional random utility models, as demonstrated in a strategic context in Signorino (1999, 2003).

[4] For those who do not like the Non-Nash interpretation of the agent error specification, Signorino (2003) shows that, for simple models such as that in Fig. 1, it will yield almost identical results as the (strictly Nash) private information version. Researchers should be aware, however, that the more complicated the game structure, the more the agent error model will diverge from the private information model.

We will assume throughout this paper that the α are distributed Type I Extreme Value, leading to logit probabilities.[5] B's probabilities, p_r and p_ℓ, of playing right and left, respectively, are

$$p_r = \frac{e^{U_B(Rr)}}{e^{U_B(R\ell)} + e^{U_B(Rr)}}, \tag{3}$$

$$p_\ell = \frac{e^{U_B(R\ell)}}{e^{U_B(R\ell)} + e^{U_B(Rr)}}. \tag{4}$$

Now consider A's choice between playing Right and Left. If A plays Left, the game ends. However, A's decision to pick Right depends on what she thinks B will do. Because the α's are private information, A is uncertain of B's action and, therefore, must estimate the probability that B will play right or left. Therefore, A's utility for playing Right is an expected utility over the lottery consisting of B's choice, with probabilities as above: p_r, and p_ℓ. Player A's utilities are thus

$$U_A^*(L) = U_A(L) + \alpha_L \tag{5}$$

$$U_A^*(R) = EU_A(R) + \alpha_R \tag{6}$$

$$= p_\ell U_A(R\ell) + p_r U_A(Rr) + \alpha_R. \tag{7}$$

We will assume that the analyst shares the same uncertainty as the players. With the same assumptions as before, the probabilities of A playing Right versus Left are logit probabilities. The twist here is that they are based on an expected utility calculation and do not take the same form as logit probabilities in typical statistical analyses. A's probabilities, p_R and p_L, of picking Right and Left, respectively, are

$$p_L = \frac{e^{U_A(L)}}{e^{U_A(L)} + e^{EU_A(R)}} = \frac{e^{U_A(L)}}{e^{U_A(L)} + e^{p_\ell U_A(R\ell)+p_r U_A(Rr)}}, \tag{8}$$

$$p_R = \frac{e^{EU_A(R)}}{e^{U_A(L)} + e^{EU_A(R)}} = \frac{e^{p_\ell U_A(R\ell)+p_r U_A(Rr)}}{e^{U_A(L)} + e^{p_\ell U_A(R\ell)+p_r U_A(Rr)}}. \tag{9}$$

For any specification of the utilities, (p_ℓ, p_r, p_L, p_R) are the equilibrium probabilities of the statistical strategic model. Since we have assumed that the uncertainty enters as Type I Extreme Value perturbations to the action utilities, (p_ℓ, p_r, p_L, p_R) is a Logit Quantal Response Equilibrium (see McKelvey and Palfrey (1998) and Signorino (1999)). Because the α terms are assumed to be independently distributed, the equilibrium probabilities for the outcomes are just the product of the choice probabilities along each outcome's path:

$$\Pr[L] = p_L \tag{10}$$

[5]The techniques demonstrated here can be implemented in exactly the same fashion if one assumes the α are normally distributed, resulting in probit probabilities.

$$\Pr[R\ell] = p_R \cdot p_\ell \tag{11}$$

$$\Pr[Rr] = p_R \cdot p_r \tag{12}$$

Having specified the statistical strategic model, a natural question concerns how the equilibria of the statistical model compare to the subgame perfect equilibria of the model with complete information—i.e., without the shocks to players' expected utilities. The intuition for this is best explained with reference to the variance of the error terms, assuming the observed utilities are held constant. As the variance of the error terms goes to zero, the model becomes one of players with perfect and complete information, who maximize their utility at each decision point. Therefore, in the limit, the equilibria are subgame perfect. When the variance of the error terms is nonzero but extremely small relative to the observed utilities, the statistical equilibria often resemble a smoothed version of the subgame perfect equilibria. As the variance of the error terms increases, the equilibria may at times resemble smoothed versions of the subgame perfect equilibria, but may at other times look very different from the subgame perfect equilibria because the choice probabilities affect expected utility calculations.

The last step in specifying the model is to assign regressors to the utilities and then estimate the parameters accompanying those regressors. The statistical strategic model is just a strategic, random utility model. Therefore, the equilibrium choice and outcome probabilities provide a probability model that can be used in maximum likelihood estimation. The method of estimation—and, specifically, a new, simpler method for doing so—is the subject to which we now turn.

3 Statistical Backwards Induction

We provide a technique in this section that allows practitioners to estimate recursive strategic models (of any depth or breadth) requiring only variations of logit or probit in standard statistical packages. This technique has a number of nice properties in relation to the system estimator currently used. Before turning to the new procedure, we first specify the system estimator.

3.1 *The System Estimator*

To date, empirical analyses employing statistical strategic models have employed a "system" approach to estimation—where all parameters are estimated simultaneously for the entire model (see, e.g., Signorino 1999, 2002, 2003; Guo 2002; Carson 2003, 2005; Signorino and Yilmaz 2003; Carson and Marshall 2004; Carter 2005; Quackenbush 2005; Signorino and Tarar 2006).[6] To compare the system estimator to Statistical Backwards Induction (SBI), it will be helpful to change our notation slightly.

Consider the game in Fig. 2, which is a more fully specified version of the model in Fig. 1. Let us now assume that our data is coded as

$$y_{\mathrm{A}} = \begin{cases} 1, & \text{if } \; U^*_{\mathrm{A}}(R) \geq U^*_{\mathrm{A}}(L), \\ 0, & \text{if } \; U^*_{\mathrm{A}}(L) > U^*_{\mathrm{A}}(R), \end{cases} \tag{13}$$

[6] *Strat* currently implements a systems approach to maximum likelihood estimation of the parameters. This is also referred to as a "full information maximum likelihood" estimator.

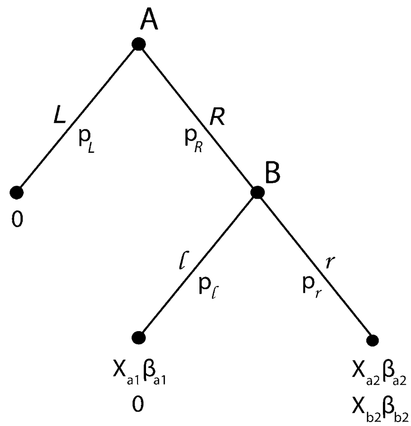


Fig. 2 Model with regressors (X) and parameters (β).

and

$$y_{\text{B}} = \begin{cases} 1, & \text{if } U^*_{\text{B}}(r) \geq U^*_{\text{B}}(\ell), \\ 0, & \text{if } U^*_{\text{B}}(\ell) > U^*_{\text{B}}(r), \end{cases} \tag{14}$$

where $y_{\text{A}} = 1$ and $y_{\text{B}} = 1$ correspond to A and B choosing R and r, respectively. The equilibrium action probabilities are the same as before. According to the new coding, they are $p_R = \Pr(y_{\text{A}} = 1)$ and $p_r = \Pr(y_{\text{B}} = 1)$.

Figure 2 also displays a very simple specification of regressors for the utilities. Here, we have normalized $U_{\text{A}}(y_{\text{A}} = 0) = 0$ and $U_{\text{B}}(y_{\text{A}} = 1, y_{\text{B}} = 0) = 0$. The remaining utilities are specified with a single regressor and parameter:

$$U_{\text{A}}(y_{\text{A}} = 1, y_{\text{B}} = 0) = X_{a1}\beta_{a1},$$

$$U_{\text{A}}(y_{\text{A}} = 1, y_{\text{B}} = 1) = X_{a2}\beta_{a2},$$

$$U_{\text{B}}(y_{\text{A}} = 1, y_{\text{B}} = 1) = X_{b2}\beta_{b2}.$$

Assuming we have data for all outcomes and regressors, then the system estimator would maximize the log of the following likelihood with respect to the regression coefficients β_{a1}, β_{a2}, and β_{b2}:

$$L = \prod^{N} p_L^{(1-y_{\text{A}})} \cdot (p_R p_\ell)^{y_{\text{A}}(1-y_{\text{B}})} \cdot (p_R p_r)^{y_{\text{A}} y_{\text{B}}},$$

where the observation index i, i=1, 2, ..., N, has been dropped. In sum, for the system estimator, we construct a probability model for all of the actions and outcomes, and then use that to form the likelihood to be maximized. All parameters for all of the players are then estimated simultaneously.

3.2 *SBI Estimator*

In order to illustrate estimation by SBI, it is helpful to think of the strategic model as a recursive system of equations (see also Signorino 2002; Signorino and Yilmaz 2003). The main insight is that a recursive system of equations can be estimated equation-by-equation.

Therefore, a recursive statistical strategic model can be estimated by univariate analogs of backwards induction.

3.2.1 The strategic model as a recursive system of equations

Consider the strategic model from the previous section and the data coding in equations (13) and (14). These imply the following system of latent variable equations:

$$y_{\mathrm{A}}^{*} = U_{\mathrm{A}}^{*}(R) - U_{\mathrm{A}}^{*}(L), \tag{15}$$

$$y_{\mathrm{B}}^{*} = U_{\mathrm{B}}^{*}(r) - U_{\mathrm{B}}^{*}(\ell), \tag{16}$$

where the data are coded as $y_j = 1$ if $y_j^* \geq 0$ and $y_j \geq 0$ if $y_j^* < 0$, for $j \in \{A, B\}$. Given the regressor specification in Fig. 2, the system can be written as

$$y_{\mathrm{A}}^{*} = p_\ell X_{a1}\beta_{a1} + p_r X_{a2}\beta_{a2} + \varepsilon_{\mathrm{A}}, \tag{17}$$

$$y_{\mathrm{B}}^{*} = X_{b2}\beta_{b2} + \varepsilon_{\mathrm{B}}. \tag{18}$$

If we assume that the α perturbations to the expected utilities are distributed Type I Extreme Value, then the ε_j terms are distributed logistic (McFadden 1974). The resulting choice probabilities $p_r = \Pr[y_{\mathrm{B}}^* > 0]$ and $p_R = \Pr[y_{\mathrm{A}}^* > 0]$ will be logit probabilities.[7]

It bears reiterating that this system of latent variable equations is completely consistent with the statistical game with agent error—it is just a different way of writing the same model. Moreover, the probabilities are all logit probabilities, but, again, where the expected utility calculations are explicitly modeled.

Because the α shocks are assumed independent, the ε terms are also independent. Therefore, equations (17) and (18) are not "linked" through their error terms. Moreover, the system of equations is recursive, which, in theory, would imply that we can "start from the bottom and work up," estimating first the equation for y_{B}^* and then using that information in the estimation of y_{A}^*. Indeed, it is not coincidental that the game-theoretic model can be solved via backwards induction and that this statistical model can be estimated in an analogous fashion. The only problem with this setup so far is that, although the equations for y_{B}^* and y_{A}^* result in logit probabilities, the systematic component of y_{A}^* does not take the same functional form as in typical logit models. This is, of course, due to the expected utility calculation. If Fig. 2 generated the data and we were to use logit with the typical first-order linear $X\beta$ specification for each equation, then our parameter estimates would be biased and inconsistent (see Signorino and Yilmaz 2003). However, as we will show, this problem can be easily overcome.

3.2.2 The basic SBI procedure

SBI requires three components common to empirical political science: (1) standard logit or probit regression (e.g., in *Stata*, SPSS, etc.), (2) the ability to calculate predicted probabilities from the aforementioned regression, and (3) the ability to generate new variables by multiplying two existing variables. The basic idea for the SBI is to estimate the system equation-by-equation using standard logits, but transforming regressors in expected utility calculations into "new" regressors that can be used in standard logit estimation. Consider

[7] Probit probabilities result from an assumption of normally distributed error terms.

our strategic system of equations in equations (17) and (18) describing Fig. 2. For our referent example, the basic SBI estimation procedure is the following:

1. B's Choice. Since $y_{\mathrm{B}}^{*} = X_{b2}\beta_{b2} + \varepsilon_{\mathrm{B}}$ does not require information concerning y_{A}^{*}, logit may be used to estimate $\hat{\beta}_{b2}$. Once $\hat{\beta}_{b2}$ is obtained, calculate $\hat{p}_r$ and $\hat{p}_\ell = 1 - \hat{p}_r$.
2. A's Choice. Substitute $\hat{p}_\ell$ and $\hat{p}_r$ into A's equation, giving

$$y_{\mathrm{A}}^{*} = \hat{p}_\ell X_{a1}\beta_{a1} + \hat{p}_r X_{a2}\beta_{a2} + \varepsilon_{\mathrm{A}}.$$

To use standard statistical packages, we must first create the transformed regressors $Z_{a1} = \hat{p}_\ell X_{a1}$ and $Z_{a2} = \hat{p}_r X_{a2}$. Logit may then be used to estimate the parameters in the equation:

$$y_{\mathrm{A}}^{*} = Z_{a1}\beta_{a1} + Z_{a2}\beta_{a2} + \varepsilon_{\mathrm{A}}.$$

It is important to note that in each stage, the transformed latent variable regression equation is exactly the same as the original, just written in a different way. Therefore, although the parameters in the transformed equations are associated with transformed regressors, the interpretation of the parameters is the same as in the original equation. For example, $\hat{\beta}_{a2}$ is still the estimated effect of X_{a2} on A's utility for outcome ($y_{\mathrm{A}} = 1, y_{\mathrm{B}} = 1$). The transformation of the regressors is only done to allow for estimation in standard computer packages, and it does not affect the interpretation of the parameter estimates.

SBI implies a number of nice properties above and beyond the ease of implementation. Logit estimation in each stage will yield consistent estimates of the parameters and functions of the parameter estimates, such as the equilibrium choice and outcome probabilities. The consistency of the parameter estimates derives from their status as maximum likelihood estimators. Obviously, system estimators would also yield consistent estimates. Indeed, because the equations are recursive and because the error terms are uncorrelated, we would expect both methods (system and SBI) to yield virtually identical results. However, they may not—and there are additional benefits to using SBI.

First, SBI often provides a feasible estimator, when the system estimator might be much more difficult to estimate. The likelihood function for the system estimator is not guaranteed to be globally concave. That is not to say that there are multiple extrema, but rather that the likelihood function may only be quasi-concave. It is easy to show just from plotting the likelihood that as the number of observations becomes small and as the variance of the error term becomes small, the likelihood function becomes increasingly "step-like," reflecting the fact that the equilibrium probabilities closely approximate the subgame perfect step response. The effect of this is that system estimation may have problems with weak (or fragile) identification. In contrast, logit and probit have globally concave likelihood functions, resulting in much easier and more rapid numerical optimization.[8]

Second, SBI is fast. Optimized algorithms exist for implementing logit (or probit and their variants) in a host of software packages, while maximizing a user-specified likelihood is often a slow and painstaking process that heavily depends on the quality of starting values. Across a host of tests and implementations, the SBI strategic model iterates and converges more rapidly and requires fewer iterations than an identical systems approach.

Third, SBI retains more data for estimation. In practice, the system method requires all of the data appearing anywhere in the game to be sent simultaneously to a procedure that calculates the log likelihood of each observation. Although a number of techniques exist for dealing with missing data, listwise deletion is still often the default method. In this

[8]Monte Carlo simulations suggest a rather rapid decay in the convergence properties of system estimators (in *Stata*) that posed no problem for SBI.

case, if any observation contains a missing data point, the entire row is generally deleted. So for example, if a variable in A's decision contains missing data, then the entire row will be deleted, even if that variable does not appear in B's decision. This produces a needless loss of data in estimating B's parameters. Because the system is recursive, and because the SBI approach estimates each equation individually, missing data outside a given subgame will not affect the sample for the equation associated with that subgame. Therefore, the iterative approach is able to use as much data as possible in estimating the parameters.

Finally, we show in the appendix that the SBI estimator produces consistent estimates of β_{a1}, β_{a2}, and β_{b2}. It also produces consistent estimates of the SE of β_{b2}. However, because $\hat{p}_\ell$ and $\hat{p}_r$ are substituted into A's equation and the resulting $Z_{a1} = \hat{p}_\ell X_{a1}$ and $Z_{a2} = \hat{p}_r X_{a2}$ variables are treated as data (without estimation uncertainty), SBI will yield biased estimates of the SEs of β_{a1} and β_{a2}. Fortunately, this is easily remedied by a nonparametric bootstrap.[9]

3.3 *SBI with Bootstrapped SEs*

To retain the simplicity of SBI, we rely on a simulation method for calculating the SEs—one that adds only a single step to the procedure outlined in the previous section. The approach we use is a form of the nonparametric bootstrap. The bootstrap procedure is quite simple. M iterations of the bootstrap are run. During each iteration $m = 1, 2, \ldots, M$ of the bootstrap, a sample is randomly drawn with replacement from the original data. SBI is then used to estimate parameters governing B's choice, generate new regressors, and estimate A's parameters (which are saved). After the M iterations of the bootstrap procedure, the SEs of A's parameter estimates are calculated by simply taking the standard deviations (SDs) of the saved estimates.

As mentioned previously, the bootstrap correction is only necessary for the estimation of *A*'s parameters. Because *B*'s actions do not depend on any auxiliary parameters, the first-stage maximum likelihood SEs are consistent. It is only when we turn to choices that depend on the expected choices of others that a correction for the presence of a random action probability is necessary. With this need for correction in mind, we compare the system and SBI approaches for estimating statistical strategic models in a Monte Carlo experiment.

3.4 *Monte Carlo Analysis*

For the Monte Carlo analysis, we assumed a similar regressor specification as in Fig. 2. We have normalized $U_A(y_A = 0) = 0$ and $U_B(y_A = 1, y_B = 0) = 0$. The remaining utilities are specified as follows:

$$U_A(y_A = 1, y_B = 0) = X_{a2}\beta_{a2},$$

$$U_A(y_A = 1, y_B = 1) = X_{a3}\beta_{a3} + X_c\beta_{a3c},$$

$$U_B(y_A = 1, y_B = 1) = \beta_{b0} + X_{b3}\beta_{b3} + X_c\beta_{b3c},$$

[9] As with any method, there are limitations to this one as well. The main limitation here is really a modeling issue: it assumes the underlying model is recursive. It is important to note, however, that some common games are not recursive—for example, games with simultaneous moves and those with Bayesian updating. Although these games can be written as a system of equations, they generally do not have the simple structure (and nice properties) just demonstrated. For examples of this, see the simultaneous move games in McKelvey and Palfrey (1998) and the signaling games in Lewis and Schultz (2003), Wand (2005), and Signorino and Whang (2007). In these cases, researchers will need to derive and program the more complicated structural model.

Table 1 Monte Carlo results

	n = 500		n = 5000	
Parameter estimates	$\hat{\beta}_{a3}$	$\hat{\beta}_{a3c}$	$\hat{\beta}_{a3}$	$\hat{\beta}_{a3c}$
System	1.096 (0.224)	1.090 (0.223)	1.037 (0.061)	1.034 (0.058)
SBI	1.047 (0.227)	1.043 (0.226)	1.024 (0.065)	1.021 (0.063)
Standard errors	$\widehat{\mathrm{se}(\hat{\beta}_{a3})}$	$\widehat{\mathrm{se}(\hat{\beta}_{a3c})}$	$\widehat{\mathrm{se}(\hat{\beta}_{a3})}$	$\widehat{\mathrm{se}(\hat{\beta}_{a3c})}$
System	0.215 (0.054)	0.211 (0.055)	0.062 (0.004)	0.061 (0.004)
Bootstrapped SBI	0.255 (0.083)	0.252 (0.084)	0.067 (0.006)	0.067 (0.006)

Note. The first number in each cell is the mean of the Monte Carlo density. The number in parentheses is the SD.

where $\beta_{a2} = \beta_{a3} = \beta_{a3c} = \beta_{b3} = 1$, $\beta_{b3c} = -1$, and $\beta_{b0} = \pi$. X_c is a common regressor that appears in both players' utilities. Data were generated based on the behavioral assumptions of the game, with X uniformly distributed over $[-2\pi, 2\pi]$, ε distributed logistic, with $V(\varepsilon) = \pi^2/3$. Simulations were run for sample sizes $n = 500$ and $n = 5000$. Once data were generated, the parameters were estimated using both the system and SBI methods, and correct SEs for SBI are calculated using the bootstrap with 1000 bootstrap iterations. Simulations were repeated 5000 times for $n = 500$ and 2000 times for $n = 5000$ to form the densities of the estimators.

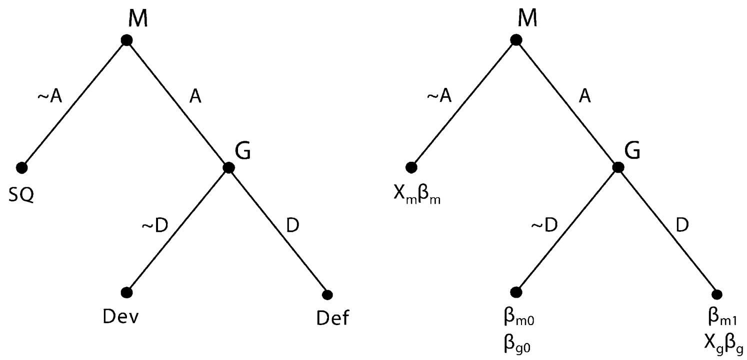
Table 1 summarizes the results of the Monte Carlo analysis for β_{a3} and β_{a3c}. The first number in each cell is the mean of the Monte Carlo density. The number in parentheses is the SD. As we previously noted, the SBI estimator will produce consistent estimates for the parameters (i.e., the β's) for all players in the game, as well as for the SEs for B. The upper section of Table 1 displays the Monte Carlo sampling distributions of $\hat{\beta}_{a3}$ and $\hat{\beta}_{a3c}$ for the system and SBI estimators. First, note that both the SBI and system estimators are consistent; they both recover the true parameter values on average.[10] Second, although the SBI estimator uses less information than the system estimator, it is nearly as efficient for samples of at least 500. Indeed, it is also useful to note that the SD of the system estimator's density is the gold standard for the SE estimators. So, for samples of 500, we would like the average estimated SE to be close to 0.22, whereas for $n = 5000$, it should be close to 0.06.

For the Monte Carlo, A's estimated SEs were calculated in two ways. The system SEs are based on the (estimated) asymptotic variance from systems estimation. The "bootstrapped SBI" SEs are produced by running the bootstrap procedure on the SBI estimator and then taking the SD of the saved parameter estimates.

The lower section of Table 1 displays the average estimated SE based on system and bootstrapped SBI estimation. The system SEs are obviously very close to the SDs of the sampling distributions in the upper section. More importantly for our purposes here, the bootstrapped SBI SEs are also quite close. As is to be expected, the bootstrapped SBI SEs are slightly larger than the system SEs. This is because the SBI method is itself less efficient than the system method. Moreover, the bootstrap method will induce some inefficiency. Nevertheless, the sampling properties of the SBI method and of bootstrapping imply that the bootstrapped SBI SEs will be consistent. In fact, we can see that as the sample increases to 5000, the difference in the SEs is quite small.

Although not reported here, the SBI SEs for B's utilities will be larger than the corresponding system SEs. SBI uses a smaller sample to estimate the coefficients at "lower"

[10]Although not shown here, the results for β_{a2} are similar to those reported in Table 1.



(a) Speculative Attack Game (b) With Regressors

Fig. 3 Leblang's model of speculative attacks on currencies.

nodes, so this again comes as little surprise. Because the system is recursive, SBI SEs for this step are consistent. However, SBI is less efficient than the systems estimator, which simultaneously utilizes all available information.

With a theoretical model and its statistical counterpart in hand, we illustrate the power of this technique with real data in the next section.

4 Data Analysis

A substantial political economy literature has examined speculative attacks on currency pegs—arbitrage opportunities may lead individual investors to take financial positions that combine to "attack" the declared par value of a currency (in terms of some other currency). Given an attack, governments confront a difficult choice: to expend resources in defense of the par value or to allow the currency to devalue, with all the concomitant distributional issues accompanying devaluation. Leblang (2003) analyzes the strategic aspect of such speculative attacks using a system estimator based on the model depicted in Fig. 3.[11] Leblang's (2003) analysis provides a nice opportunity to compare system estimation versus SBI.

The variables that measure market (M) utilities for the status quo, defense, and devaluation and government (G) utility for defense and devaluation are listed in Table 2. In general terms, markets are argued to attack currencies when there is a disconnect between the fixed exchange rate and the equilibrium rate of exchange (proxied by Reserves, Real Exchange Rate Overvaluation, Credit Growth), a history of questionable fixes (proxied by Prior Attacks), and/or greater structural incentives to challenge a government's resolve (proxied by Contagion, U.S. Interest Rates, and Debt Service). Similarly, governments are argued to defend pegged exchange rates when it is politically expedient (proxied by Unified Government, Campaign/Election, Right Government, Export Sector, and Post-election) given the resources at their disposal with which to combat speculative attacks (proxied by Interest Rates, Capital Controls, and Reserves).

The specification of the utilities with regressors is shown in Fig. 3b. Here, regressors and parameters subscripted with m are associated with market utilities. Those subscripted with g are associated with government utilities. The replication results using system and

[11]To be clear, he analyzes a sample of monthly data from states that are nominally democratic—countries with Polity scores greater than 5—with pegged exchange rates.

Table 2 Variables and measures from Leblang (2003)

Variables	*Utility*	*Sign*	*Measures*
Unified Government	G	+	Binary: party that controls the executive also controls the lower house of the legislature[a]
Export Sector$_{t-1}$	G	–	$\text{Log}\left(\frac{\text{Exports}}{\text{GDP}}\right)_{t-1}$: size of the export sector relative to GDP[b]
Campaign/Election	G	+	Binary: 3-month campaigns and the month of elections[a]
Postelection	G	–	Binary: 3 months postdating every election month[a]
Right Government	G	+	Binary: governments classified Left, Center, and Right according to Left-Right positions on state control of the economy[a]
Interest Rates$_{t-1}$	G	+	Deflated discount rates, money market rates, or deposit rates depending on availability[c]
Capital Controls$_{t-1}$	G, M	+, +	Binary: controls on the capital account[d]
Reserves	G, M	+, +	$\text{Log}\left(\frac{\text{Reserves}}{\text{BaseMoney}}\right)_{t-1}$: total reserves minus gold to base money [M0][c]
Real Exchange Rate Overvaluation	M	–	Hodrick-Prescott residuals of the monthly real exchange rate[c]
Credit Growth$_{t-1}$	M	–	Rate of growth in domestic credit[c]
U.S. Interest Rates$_{t-1}$	M	–	Interest rate on 90-day U.S. deposits[c]
Debt Service$_{t-1}$	M	–	Total interest and debt repayment in foreign currencies[c]—International Monetary Fund, short-, and long-term loans—as a percentage of exports[b]
Contagion	M	–	Number of speculative attacks outside country *i* in month *t*
Prior Attacks	M	–	Number of speculative attacks for country *i* in all months preceding *t*

[a]Database of Political Institutions, Beck et al. (2001) updated by Leblang (2003).
[b]World Development Indicators, (World Bank n.d.).
[c]International Financial Statistics, (International Monetary Fund n.d.b).
[d]Annual Report on Exchange Arrangements and Exchange Restrictions.

SBI estimators are presented in Table 3. The columns labeled *System* replicates the system estimator results in Leblang (2003) Table 2. The columns labeled *SBI* replicate Leblang's results, but using SBI. The table is divided into an upper and lower section. In the upper section, the parameter estimates for the constant terms (β_{m0}, β_{g0}, β_{g1}) are presented. In the lower section, the estimates associated with regressors in X_m and X_g are presented. Bootstrapping is used to correct the bias in SEs for markets for reasons described in Section A.2.

The tiny differences in system versus SBI estimates confirm that the simplified SBI method is very similar to the full system estimator. Furthermore, SBI produces inferences that do not at all diverge from the systems estimator. The magnitudes and signs of the estimates do not substantially differ between the system and SBI estimators; and the SEs are quite similar.

Substantively, with respect to government decisions to defend or devalue, *Unified Government*, *Export Sector*, *Postelection* periods, and *Capital Controls* fail to reject the null hypothesis of no effect for both estimation techniques. By contrast, *Campaign/Election*

Table 3 Replication of Leblang's results

	System		*SBI*	
	Market	*Government*	*Market*	*Government*
Constants				
β_{g0}		−0.43 (0.78)		0.20 (0.78)
β_{m0}	−3.66* (0.30)		−4.05* (0.47)	
β_{m1}	−3.14* (0.29)		−3.41* (0.50)	
	β_m	β_g	β_m	β_g
Variables				
Unified Government		−0.35 (0.35)		−0.07 (0.45)
Export Sector$_{t-1}$		−0.20 (0.17)		−0.29 (0.23)
Campaign/Election		1.66* (0.75)		2.23* (0.93)
Postelection		1.06 (0.59)		1.10 (0.74)
Right Government		−0.94* (0.45)		−1.55* (0.65)
Interest Rates$_{t-1}$		1.93* (0.64)		1.33* (0.69)
Capital Controls$_{t-1}$	−0.45 (0.25)	0.07 (0.75)	−0.42 (0.47)	0.67 (0.79)
Reserves$_{t-1}$	0.23* (0.06)	0.31* (0.17)	0.29* (0.07)	0.59* (0.21)
Real Exchange Rate Overvaluation	−0.44* (0.09)		−0.46* (0.18)	
Credit Growth$_{t-1}$	−0.06* (0.03)		−0.07* (0.04)	
U.S. Interest Rates$_{t-1}$	−0.05 (0.06)		−0.05 (0.06)	
Debt Service$_{t-1}$	−0.03 (0.05)		−0.03 (0.05)	
Contagion	−0.12* (0.05)		−0.13* (0.05)	
Prior Attacks	−0.12* (0.05)		−0.12* (0.05)	
n	7240	7240	7240	88
Log likelihood	−482.02		−432.27	−49.79

Note. The first number in each cell is the mean of the Monte Carlo density. The number in parentheses is the SD. $*p \leq 0.05$.

periods, *Right Governments*, the level of *Interest Rates*, and foreign exchange *Reserves* exert statistically significant impacts on the decision to defend or devalue resulting in identical inferences from systems and SBI estimation. In short, government utilities showcase minimal differences between systems estimation and SBI. For market decisions to initiate speculative attacks, *Capital Controls*, *U.S. Interest Rates*, and levels of *Debt Service* have no statistically discernible impact, while *Real Exchange Rate Overvaluation*, the rate of domestic *Credit Growth*, foreign exchange *Reserves*, *Contagion*, and *Prior Attacks* influence the decision without regard to estimation technique. Finally, there is no statistically significant difference between the market utilities for defense and devaluation in either the systems or SBI estimates as represented by the constant. This implies that there is no evidence that markets necessarily prefer devaluations to defenses given a speculative attack.

Assessing model fit, the bottom row of Table 3 shows a very small difference between the log likelihoods of the system and SBI approaches to estimating this statistical strategic model. Comparing the sum of the log likelihoods from the SBI estimator to the system log likelihood, the resulting difference is 0.0466.[12] Thus, this replication exercise demonstrates a simplified technique for the analysis of currency crises that is both useful and closely mirrors its system equivalent.

[12]This result is obtained by subtracting the sum of SBI log likelihoods from the system log likelihood.

5 Conclusion

Our objective in this paper has been to provide a simple method for estimating recursive, statistical strategic models. Researchers can now use common commands (e.g., logit, probit, bootstrap) in their favorite statistical package to estimate these models. Moreover, the technique is flexible, allowing researchers to estimate models corresponding to a wide array of strategic situations.

As we have shown, this user-friendliness does come with a small cost: the SBI estimator is less efficient than the full information system estimator. However, for most data sets, the difference should be negligible. Indeed, we found no substantive differences between the system and SBI replications of the Leblang (2003) analysis.

The above cost, we believe, is well offset by the introduction of a large tool set for political science scholars. Deriving statistical models that are consistent with theoretical models is difficult enough. For many researchers, the programming requirements generally associated with these specialized models are often a roadblock to their implementation. By removing that roadblock, scholars can get on with the business of substantive research concerning strategic political behavior.

Funding

The National Science Foundation (SES-9817947, SES-0213771, and SES-0413381); the Peter D. Watson Center for Conflict and Cooperation.

Appendix A: Properties of the SBI Estimator

The general problem of multistep estimation is considered in a broad econometric literature. Consistency is almost certain, because maximum likelihood estimators are consistent, as are functions of them. That said it is instructive to investigate the equation-by-equation technique and to provide a general analysis of its properties. We first ask, is SBI estimation consistent?

A.1 *Consistency*

The SBI strategic estimator is a special case of the two-step maximum likelihood estimators described in Murphy and Topel (1985). Thus, consistency of the SBI estimator follows from the consistency of two-step maximum likelihood estimation procedures. For simplicity, consider a simple case with two models, in which one model is embedded in the other:

$$y_1 = f_1(x_1, \theta_1), \tag{A1}$$

$$y_2 = f_2(x_1, x_2, \theta_1, \theta_2), \tag{A2}$$

Two-step maximum likelihood first estimates the parameter vector θ_1 by maximum likelihood. θ_2 is then estimated by maximum likelihood with $\hat{\theta}_1$ inserted in place of θ_1 as if it were known (Greene 2000). The consistency of the first step follows from the consistency of maximum likelihood estimators (MLEs). Murphy and Topel (1985) show that estimating θ_2 from $f_2(y_2|x_1, x_2, \hat{\theta}_1, \theta_2)$ is asymptotically equivalent to estimating it from $f_2(y_2|\ x_1, x_2, \theta_1, \theta_2)$; therefore, the second step is also consistent (Murphy and Topel 1985).

In our equation-by-equation strategic method, the first step consists of estimating B's choice probabilities by logit (or probit) and then using these estimates as if they were known in estimating A's choice probabilities with logit (or probit). Thus, it follows that both steps of the SBI estimator provide consistent estimates of players' choice probabilities.

A.2 *Efficiency*

Our analysis of efficiency is based on the considerable econometric literature on multistep maximum likelihood estimators, see Murphy and Topel (1985) and Newey and McFadden (1994). We first examine the properties of multistep estimators.

Newey and McFadden (1994, Theorem 6.2, 2180) characterize the general conditions where the presence of a first-step estimator influences asymptotic inference.[13] Subject to certain regularity conditions, the key to valid asymptotic inference relies on the consequences of consistency in the first-stage estimates. If consistency of the first-stage estimates is critical for the consistency of subsequent estimates, Newey and McFadden (1994) prove that the (estimated asymptotic) SEs for the second stage are inconsistent. With reference to SBI estimation, Newey and McFadden show that the SEs are generally incorrect because consistency of the estimated action probabilities "below" is critical for consistency of the parameters. Upon reflection, this may not be critical. Given the common practice of testing zero null hypotheses, we are more concerned with falsely rejecting true null hypotheses; indeed, the power of tests is seldom if ever discussed in applied research. Unfortunately, a further difficulty for the equation-by-equation technique is suggested by Newey (1984).

Newey (1984) argues that second-stage SEs of equation-by-equation estimators will be incorrect and too small (in a positive semidefinite sense). An intuition arises from the remark of Karaca-Mandic and Train (2003, 401) that "the covariance matrix of the second-stage estimator includes noise introduced by the first-stage estimates." Newey (1984) demonstrates that sequential generalized method of moments estimators will have undersized asymptotic SEs when the covariances between the stages are zero and the (estimated) quantities are assumed to be known. Recall that the SBI estimation technique begins by calculating B's action probabilities and employing predicted values $\hat{p}_r$ as the expected actions of B to form new regressors $Z_{\mathrm{A}} = [\hat{p}_r X_{a2}, \hat{p}_\ell X_{a1}]$. Employing the theory of the partitioned inverse, Newey partitions the estimation problem into first and second stages to demonstrate that a zero covariance among the stages (which we assume to be true because the players' choices are independent) and treatment of the predictions as data requires that the estimated (asymptotic) SEs will be too small because they fail to account for the sampling variance of the predicted probabilities. We first turn to an analytic characterization of the problem.

A.2.1 Efficiency: Analytics

For a two-player model as in Fig. 2, we can derive the SE correction for the second stage of the SBI estimator.[14] As a special case of maximum likelihood estimation, the first

[13] Their analysis is of the general class of method of moments estimators.
[14] The derivation follows closely Karaca-Mandic and Train (2003).

step of the equation-by-equation estimator finds $\hat{\beta}_B$ that solves the moment condition $g_B(\beta_B) = 0$ and then the second stage finds $\hat{\beta}_A$ that satisfies $g_A(, \beta_A, \hat{\beta}_B) = 0$, where

$$g_B(\beta_B) = \frac{1}{N_B}\sum_{i=1}^{N_B}\frac{\delta L_B(y_{Bi}\hat{\beta}_B)}{\delta\beta_B}, \tag{A3}$$

$$g_A(\beta_A, \hat{\beta}_B) = \frac{1}{N}\sum_{i=1}^{N}\frac{\delta L_A(y_{Ai}; \hat{\beta}_A, \hat{\beta}_B)}{\delta\beta_A}. \tag{A4}$$

Taking a first-order Taylor's expansion of first- and second-stage moment conditions around the true parameter values β_B^* and β_A^* we get

$$g_B(\beta_B^*) - R_1(\hat{\beta}_B - \beta_B^*) \stackrel{A}{=} 0, \tag{A5}$$

$$g_A(\beta_A^*, \beta_B^*) - R_2(\hat{\beta}_B - \beta_B^*) - R_3(\hat{\beta}_A - \beta_A^*) \stackrel{A}{=} 0, \tag{A6}$$

where $\stackrel{A}{=}$ denotes asymptotic equality and

$$R_1 = -\operatorname{plim}\frac{\delta g_B(\beta_B^*)}{\delta\beta_B},$$

$$R_2 = -\operatorname{plim}\frac{\delta g_A(\beta_A^*, \beta_B^*)}{\delta\beta_B},$$

$$R_3 = -\operatorname{plim}\frac{\delta g_A(\beta_A^*, \beta_B^*)}{\delta\beta_A}.$$

Let us further assume that R_1, R_2, and R_3 are nonsingular square matrices.

A.2.2 *B*'s Decisions

We can use (A5) to construct

$$\sqrt{N_B}(\hat{\beta}_B - \beta_B^*) \stackrel{A}{=} \frac{1}{\sqrt{N_B}}\sum_{i=1}^{N_B} R_1^{-1}\frac{\delta L_B(y_{Bi}\hat{\beta}_B)}{\delta\beta_B},$$

$$\sqrt{N_B}(\hat{\beta}_B - \beta_B^*) \stackrel{A}{=} \frac{1}{\sqrt{N_B}}\sum_{i=1}^{N_B}\Theta_B,$$

where $\Theta_B = R_1^{-1}\frac{\delta L_B(y_{Bi}\hat{\beta}_B)}{\delta\beta_B}$. Using Θ_B as defined and the information matrix equality, we can write

$$\hat{\beta}_B \stackrel{A}{\sim} N(\beta_B^*, \mathrm{Var}(\Theta_B)/N_B),$$

to characterize the asymptotic distribution of $\hat{\beta}_B$, a standard MLE. Of particular interest, the "bottom" of the tree contains all of the relevant information with which to assess the final node of a recursive decision problem. As a result, estimates obtained from the terminal node of an extensive form set of choices have all of the desirable properties of a univariate maximum likelihood estimator. The SBI estimator is consistent and efficient at the terminal node of a recursive decision problem.

A.2.3 *A*'s Decisions

Unfortunately, the same cannot be said for "upper" decisions precisely because of the sampling distribution of the parameters $\hat{\beta}_{\mathrm{B}}$ and the associated sampling distributions of functions of these parameters—$\hat{p}_\ell, \hat{p}_r$. Turning our attention to (A6) and the parameters governing A's choices, we have

$$
\begin{aligned}
(\hat{\beta}_{\mathrm{A}} - \beta^*_{\mathrm{A}}) &\overset{\mathrm{A}}{=} R_3^{-1}(g_{\mathrm{A}}(\beta^*_{\mathrm{A}}, \beta^*_{\mathrm{B}}) - R_2(\hat{\beta}_{\mathrm{B}} - \beta^*_{\mathrm{B}})) \\
&\overset{\mathrm{A}}{=} R_3^{-1}\left(\frac{1}{N}\sum_{i=1}^{N}\frac{\delta L_{\mathrm{A}}(y_{\mathrm{A}i};\hat{\beta}_{\mathrm{A}},\hat{\beta}_{\mathrm{B}})}{\delta\beta_{\mathrm{A}}} - R_2R_1^{-1}\frac{1}{N_{\mathrm{B}}}\sum_{i=1}^{N_{\mathrm{B}}}\frac{\beta L_{\mathrm{B}}(y_{\mathrm{B}i}\hat{\beta}_{\mathrm{B}})}{\delta\beta_{\mathrm{B}}}\right) \\
&\overset{\mathrm{A}}{=} \frac{1}{N}R_3^{-1}\left(\sum_{i=1}^{N}\frac{\delta L_{\mathrm{A}}(y_{\mathrm{A}i};\hat{\beta}_{\mathrm{A}},\hat{\beta}_{\mathrm{B}})}{\delta\beta_{\mathrm{A}}} - R_2R_1^{-1}\frac{N}{N_{\mathrm{B}}}\sum_{i=1}^{N}y_{\mathrm{A}i}\frac{\delta L_{\mathrm{B}}(y_{\mathrm{B}i}\hat{\beta}_{\mathrm{B}})}{\delta\beta_{\mathrm{B}}}\right)
\end{aligned}
$$

$$
\sqrt{N}(\hat{\beta}_{\mathrm{A}} - \beta^*_{\mathrm{A}}) \overset{\mathrm{A}}{=} \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\Theta_{\mathrm{A}i},
$$

where

$$
\Theta_{\mathrm{A}i} = R_3^{-1}\frac{\delta L_{\mathrm{A}}(y_{\mathrm{A}i};\hat{\beta}_{\mathrm{A}},\hat{\beta}_{\mathrm{B}})}{\delta\beta_{\mathrm{A}}} - R_3^{-1}R_2R_1^{-1}\frac{N}{N_{\mathrm{B}}}y_{\mathrm{A}i}\frac{\delta L_{\mathrm{B}}(y_{\mathrm{B}i}\hat{\beta}_{\mathrm{B}})}{\delta\beta_{\mathrm{B}}}.
$$

This allows us to write the asymptotic distribution of $\hat{\beta}_{\mathrm{A}}$ as

$$
\hat{\beta}_{\mathrm{A}} \overset{\mathrm{A}}{\sim} N(\beta^*_{\mathrm{A}}, \mathrm{Var}(\Theta_{\mathrm{A}})/N).
$$

In implementation, Θ_{A}, R_1, R_2, R_3 are calculated using the estimated parameters, and the plim evaluations are replaced with their sample approximations.[15] If we assume that N/N_{B} is finite, $\mathrm{Var}(\Theta_{\mathrm{A}})$ is given by

$$
\mathrm{Var}(\Theta_{\mathrm{A}}) = \frac{1}{N}\sum_{i=1}^{N}\hat{\Theta}_{Ai}\hat{\Theta}'_{\mathrm{A}i}. \tag{A7}
$$

The SE correction described above applies only for the two-stage SBI estimator or when there are only two decision nodes to estimate. As the number of decision nodes increases, the number of iterations of the SBI estimator also increases and deriving the exact SE correction becomes practically impossible. A more practical approach that goes well with the ease of the SBI estimator is bootstrapping, a nonparametric method for correctly estimating upper stage SEs. This approach uses simulation methods to recover the sampling distribution of the probabilities.

[15]Readers familiar with Maddala (1983) and Murphy and Topel (1985) will note the similarities.

Appendix B: R and *Stata* Programs for Bootstrapping SEs

In this section, we present example R and *Stata* programs for running the SBI estimator for the model in Fig. 2 that we used for our Monte Carlo experiment. As we describe in Section 3 in more detail, $U_A(y_A = 0)$ and $U_B(y_A = 1, y_B = 0)$ are normalized to zero. The remaining utilities are specified as follows:

$$U_A(y_A = 1, y_B = 0) = X_{a1}\beta_{a1}$$

$$U_A(y_A = 1, y_B = 1) = X_{a2_1}\beta_{a2_1} + X_{a2_2}\beta_{a2_2}$$

$$U_B(y_A = 1, y_B = 1) = \beta_{b2_0} + X_{b2_1}\beta_{b2_1} + X_{b2_2}\beta_{b2_2}.$$

In the following R code, we define a simple function `SBI` that runs two consecutive logit regressions to estimate A and B's utilities. We first put the data matrix into the form of DATA $= [y_A, y_B, X_{b2_1}, X_{b2_2}, X_{a1}, X_{a2_1}, X_{a2_2}]$. We then use R's canned bootstrapping function boot from the `boot` library to calculate the bootstrapped SEs.

```
SBI ← function(Data, ind){

    BSdata ← Data[ind,]        # Draw the bootstrap sample
    YA ← BSdata[,1]            # A's choices
    YB ← BSdata[,2]            # B's choices; unrealized=NA
    XB ← BSdata[,3:4]          # B's regressors

# Step 1: Logit to estimate B's utilities and calculate p̂r
    logitB ← glm(YB ~ XB,family=binomial(link=''logit''))
    bB ← logitB$coefficients
    XB ← cbind(1,XB)
    pr ← exp(XB %*% bB)/(1+exp(XB %*% bB))

# Step 2: Transform A's regressors with p̂r
    XA1 ← BSdata[,5]
    XA2 ← BSdata[,6:7]
    ZA1 ← (1−pr)*XA1
    ZA2 ← cbind(pr*XA2[,1],pr*XA2[,2])

# Step 3: Logit for A's utilities
    logitA ← glm(YA ~ ZA1+ZA2-1,family=binomial(link=''logit''))
    logitA$coefficients
}

# Step 4: Calculate standard errors with bootstrapping

boot(DATA, SBI, 500)
```

The above SBI model can also be estimated in *Stata* very easily. In order to do so, we first define a function called sbi that runs two logits and estimates A and B's utilities. Assuming that we have our variables YA, YB, XB21, XB22, XA1, XA21, and XA22 loaded in *Stata*'s memory, we use *Stata*'s canned bootstrap function to calculate the SEs.

```
capture program drop sbi
program sbi, rclass

* Step 1: Run a logit to estimate B's utilities and calculate p̂_r

    logit YB XB21 XB22 if YA
    predict pr

* Step 2: Transform A's regressors with p̂_r

    gen ZA21 = XA21*pr
    gen ZA22 = XA22*pr
    gen ZA1 = XA1*(1-pr)

* Step 3: Logit for A's utilities

    logit YA ZA21 ZA22 ZA1, nocons
    logitA = e(b)
    drop pr ZA1 ZA21 ZA22

    ret scalar XA21 = logitA[1,1]
    ret scalar XA22 = logitA[1,2]
    ret scalar XA1 = logitA[1,3]

end

* Step 4: Calculate standard errors with bootstrapping

bootstrap ``sbi'' r(XA21) r(XA22) r(XA1), reps(500) dots
```

References

Beck, Thorsten, George Clarke, Alberto Groff, Philip Keefer, and Patrick Walsh. 2001. New tools in comparative political economy: The database of political institutions. *World Bank Economic Review* 15:165–76.

Carson, Jamie L. 2003. Strategic interaction and candidate competition in U.S. house elections: Empirical applications of probit and strategic probit models. *Political Analysis* 11:368–80.

Carson, Jamie L. 2005. Strategy, selection, and candidate competition in U.S. house and senate elections. *Journal of Politics* 67:1–28.

Carson, Jamie L., and Bryan W. Marshall. 2004. Checking power with power: A strategic choice analysis of presidential vetoes and congressional overrides. American Political Science Association Conference, September 2004, Chicago, IL.

Carter, Timothy A. 2005. United Nations peacekeeping: Intervention decisions and efficacy. Ph.D. thesis, Department of Political Science, University of Rochester.

Greene, William H. 2000. *Econometric analysis.* 4th ed. New York, NY: Prentice-Hall, Inc.

Guo, Gang. 2002. Party recruitment and political participation in mainland China. Ph.D. thesis, Department of Political Science, University of Rochester.

International Monetary Fund. n.d.a. *Annual report on (exchange arrangements and) exchange restrictions* (vol. various years). CD-ROM. Washington, DC: International Monetary Fund.

International Monetary Fund. n.d.b. *International financial statistics* (vol. various years). CD-ROM. Washington, DC: International Monetary Fund.

Karaca-Mandic, Pinar, and Kenneth Train. 2003. Standard error correction in two-stage estimation with nested samples. *Econometrics Journal* 6:401–7.

Leblang, David A. 2003. To defend or to devalue: The political economy of exchange rate policy. *International Studies Quarterly* 47:533–59.

Lewis, Jeffrey B., and Kenneth A. Schultz. 2003. Revealing preferences: Empirical estimation of a crisis bargaining game with incomplete information. *Political Analysis* 11:345–67.

Maddala, G. S. 1983. *Limited-dependent and qualitative variables in econometrics.* Cambridge, UK: Cambridge University Press.

McFadden, Daniel L. 1974. Conditional logit analysis of qualitative choice behavior. In *Frontiers in econometrics*, ed. Paul Zarembka, 105–42. New York: Academic Press.

McKelvey, Richard D., and Thomas R. Palfrey. 1998. Quantal response equilibria in extensive form games. *Experimental Economics* 1:9–41.

Murphy, Kevin M., and Robert H. Topel. 1985. Estimation and inference in 'two-step' econometric models. *Journal of Business and Economic Statistics* 3:370–80.

Newey, Whitney K. 1984. A method of moments interpretation of sequential estimators. *Economic Letters* 14:201–6.

Newey, Whitney K., and Daniel L. McFadden. 1994. Large sample estimation and hypothesis testing. In *Handbook of econometrics, volume 4*, ed. Robert F. Engle and Daniel L. McFadden, 2111–245. Amsterdam: Elsevier Science.

Quackenbush, Stephen L. 2005. Strategic interaction and general deterrence deterrence. Columbia, MO: Department of Political Science, University of Missouri, Columbia.

Signorino, Curtis S. 1999. Strategic interaction and the statistical analysis of international conflict. *American Political Science Review* 93:279–97.

Signorino, Curtis S. 2002. Strategy and selection in international relations. *International Interactions* 28(1): 93–115.

Signorino, Curtis S. 2003. Structure and uncertainty in discrete choice models. *Political Analysis* 11:316–44.

Signorino, Curtis S., and Ahmer R. Tarar. 2006. A unified theory and test of extended immediate deterrence. *American Journal of Political Science* 50:586–605.

Signorino, Curtis S., and Whang, Taehee. 2007. Uncertainty and learning in statistical strategic models. Working paper.

Signorino, Curtis S., and K. Yilmaz. 2003. Strategic misspecification in regression models. *American Journal of Political Science* 47:551–66.

Wand, Jonathan. 2006. Comparing Models of strategic choice: The role of uncertainty and signaling. *Political Analysis* 14:101–120.

World Bank. n.d. *World development indicators* (vol. various years). Washington, DC: World Bank Development Data Group.