

Statistical Analysis of Strategic Interaction with Unobserved Player Actions: Introducing a Strategic Probit with Partial Observability

Mark David Nieman

Department of Political Science, University of Alabama, Tuscaloosa, AL, USA

e-mail: mdnieman@ua.edu (corresponding author)

Edited by Patrick Brandt

The strategic nature of political interactions has long captured the attention of political scientists. A traditional statistical approach to modeling strategic interactions involves multi-stage estimation, which improves parameter estimates associated with one stage by using the information from other stages. The application of such multi-stage approaches, however, imposes rather strict demands on data availability: data on the dependent variable must be available for each strategic actor at each stage of the interaction. Limited or no data make such approaches difficult or impossible to implement. Political science data, however, especially in the fields of international relations and comparative politics, are not always structured in a manner that is conducive to these approaches. For example, we observe and have plentiful data on the onset of civil wars, but not the preceding stages, in which opposition groups decide to rebel or governments decide to repress them. In this article, I derive an estimator that probabilistically estimates unobserved actor choices related to earlier stages of strategic interactions. I demonstrate the advantages of the estimator over traditional and split-population binary estimators both using Monte Carlo simulations and a substantive example of the strategic rebel–government interaction associated with civil wars.

1 Introduction

Making up a large part of social interactions, strategic interactions are increasingly becoming the focus of scholarly theorizing.¹ This focus on strategic theorizing, however, has not been fully matched with equal attention to devising statistical tools for modeling strategic interactions (Morton 1999). Importantly, empirical analyses that ignore the theorized underlying strategic relationships (e.g., by either applying traditional models of discrete events, such as logit or probit, or selection models, such as bivariate probit) produce biased estimates and incorrect inferences (Signorino 1999, 2002, 2003; Signorino and Yilmaz 2003).²

Author's note: Previous versions of this article were presented at the 2012 Society for Political Methodology Summer Meeting in Chapel Hill, NC; the 2013 St Louis Area Methods Meeting in Iowa City, IA; the 2013 Society for Political Methodology Summer Meeting in Charlottesville, VA; and the 2014 meeting of the Midwest Political Science Association in Chicago, IL. The author would like to thank Fred Boehmke, Olga Chyzh, Curt Signorino, Laron Williams, Sara Mitchell, Cameron Thies, Scott Cook, Walter Mebane, John Freeman, Jay Goodliffe, Jude Hays, Doug Dion, Leah Windsor, Eleonora Mattiacci, Patrick Brandt, and the anonymous reviewers for their helpful comments. Replication data are available on the Dataverse site for this article, <http://dx.doi.org/10.7910/DVN/28662>.

¹The study of strategic interactions is ubiquitous within political science and the social sciences more broadly. For example, congressional legislation is constrained by the threat of presidential vetoes (Matthews 1989; Cameron 2000) and filibusters (Krehbiel 1998; Dion et al. 2011), while state legislators are constrained by the use or threat of state initiatives (Gerber 1996; Boehmke and Patty 2007). Scholars of comparative politics observe that parties strategically change their ideological positions in order to maximize their electoral returns (Adams and Somer-Topcu 2009; Williams and Whitten 2014) and time parliamentary elections to coincide with their perceived governing successes (Smith 2003), while governments adjust their extractive efforts and tax rates on capital and labor in relation to geographical and political neighbors (Franzese and Hays 2008; Hays 2009; Thies, Chyzh, and Nieman 2015). International states act strategically in crisis bargaining (Slantchev 2005; Signorino and Tarar 2006), conflict management (Kydd 2006; Favretto 2009; Gent and Shannon 2011), and even when negotiating and complying with treaties (von Stein 2005; Chyzh 2014).

²Although see Carrubba, Yuen, and Zorn (2007).

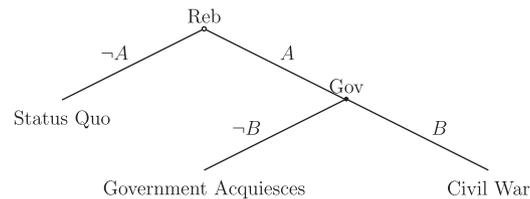


Fig. 1 Dyadic interaction of civil war onset.

Signorino and his co-authors have proposed accounting for strategy by using a *strategic logit* or *probit*, which incorporate players' (or the analyst's) uncertainty directly into statistical estimation via backwards induction (Signorino 1999, 2003; Signorino and Yilmaz 2003; Bas, Signorino, and Walker 2008). The proper implementation of strategic probits and logits, however, is often made impossible by the *outcome*- rather than *actor*-specific structure of available data: while there are data on the aggregated outcomes of an interaction, there is no record of each player's actions at each of the interaction stages.

Civil wars, for example, are often theorized as the outcome of an interaction between an opposition group and the government (see Fig. 1) (Fearon and Laitin 2003; Collier and Hoeffler 2004; Cunningham, Gleditsch, and Salehyan 2009). The opposition group, or political minority (Reb), makes a decision of whether to challenge the existing governing structure (e.g., demand greater autonomy or independence, demand a role in government), while the government (Gov) decides whether to acquiesce to their demands. Within the theorized strategic process, this means that political minorities compare their expected utility from challenging (and potentially fighting) the government to their value of the status quo (i.e., view the government as legitimate). In order to make such a comparison, political minorities must determine the probability that, if challenged, the government would fight rather than acquiesce to their demands. As a result, opposition groups that place greater value on the opportunity to achieve their goals within the existing political structure (e.g., have sufficient opportunity to affect policy and resource distribution) are less likely to challenge the status quo political arrangement (Thyne and Schroeder 2012; though see Cunningham 2011).³ Political minorities that are unable to achieve their goals within the existing political structure, on the other hand, are more likely to challenge the status quo. In the event of such a challenge, the government has a choice to acquiesce and concede to their demands or fight. Note that only the latter action by the government (in response to the minority's decision to challenge) results in a civil war.⁴

Unfortunately, most of the existing large-*n* data sets aggregate these kinds of interactions between the opposition and government, by recording a single outcome-specific event, such as "civil war/no civil war." A proper application of a strategic probit or logit, in the meantime, also requires data on the "government acquiesces" event of the interaction that would capture whether the opposition made a demand to which the government acquiesced. Such data for the player-specific decisions, however, are often unavailable, as most data-collection efforts are either altogether divorced from theoretical modeling or undertaken with a very specific theoretical process in mind.⁵ In the presence of strategic interactions, this disconnect between theoretical models and data collection may result in a type of selection bias within the collected data—bias against non-outcomes or observations in which the outcome of interest was not observed.

To demonstrate, one may also think of the existing versus theoretically required data for the study of strategic interactions between industrial firms (potential polluters) and the Environmental Protection Agency (EPA). While the recorded data contain information on the final *observed*

³This argument is similar to the democratization argument made by Acemoglu and Robinson (2001, 2006), who suggest that autocratic elites extend voting suffrage only when the threat of revolt is credible.

⁴Unprovoked government violence against the opposition, in contrast, would be coded as repression, rather than a civil war, by most political data sets.

⁵The Correlates of War project, for example, was started with a rationale to intentionally avoid being guided by one theoretical model and instead gather data that are suited for testing various theories (Singer 1972, 245).

outcome of this interaction, that is, whether the EPA has issued a fine against a firm, we know that in actuality the interaction has three possible outcomes: (1) a firm does not pollute, (2) a firm pollutes but the EPA does not detect the violation, and (3) a firm pollutes and the EPA detects the violation and issues a fine or other punishments. Note that in either of the first two cases, our data would indicate “no fine.” Since we are unlikely to have accurate data on whether a firm actually violated the law, we are unable to distinguish between the two theoretically different “no fine” events that are grouped together within the data. Observing many “0”s, or “no fines,” therefore, may suggest that no firms are over-polluting or that the EPA has not caught firms that are over-polluting. The analyst’s interpretation of these non-event outcomes, therefore, will produce very different inferences and policy assessments.

A possible solution, of course, is to collect additional data, paying careful attention to theoretical processes. This, however, is not always feasible. Industrial firms are unlikely to volunteer accurate information regarding their compliance with the EPA standards, and the EPA does not have the resources or access necessary to assess the pollution levels of each industrial firm. Proper theoretical modeling of the strategic interactions between political minorities and the government, as described above, does not only require data on civil war occurrence, but also on whether the opposition made any demands (challenged the government) in the first stage. These data are necessary to identify whether: (1) the political minorities and the government were able to negotiate a settlement short of a civil war or (2) the opposition was content with the status quo.⁶ As I demonstrate in what follows, a failure to include information on the two types of non-war events may produce biased statistical estimates.⁷

This conflation of the zero outcomes also has important substantive implications. An existing puzzle within the civil war literature, for example, is the relationship between civil war onset and gross domestic product (GDP) per capita, a nation-level variable. Do increases in GDP per capita reflect deterrence on the part of the government (Fearon and Laitin 2003) or higher opportunity costs for potential insurgents (Collier and Hoeffler 2004)? An estimator that allows for a statistical separation of the two zero outcomes may help shed light on this puzzle.

In this article, I derive a statistical solution to the theoretical puzzles driven by the limitations in the available data: a strategic probit with partial observability (SPPO). Based on Signorino’s strategic probit, SPPO probabilistically estimates unobserved actor choices at unobserved stages of the strategic interaction for outcome-specific data that only contain information on the interaction’s final binary outcome. This estimator corresponds to the strategic logic underlying many political interactions and outperforms both traditional and split-sample binary choice models in a set of Monte Carlo simulations. I also show that SPPO is robust to common model misspecification problems, such as inclusion of irrelevant and confounding covariates. Finally, to demonstrate its advantages within a more substantive framework, I apply SPPO to re-estimate the causes of civil war onset using data from Fearon and Laitin (2003).

2 Statistical Models of Strategic Interactions

The use of strategic models has greatly increased since Signorino (1999, 2003) first introduced them to political science from experimental economics (McKelvey and Palfrey 1995, 1996, 1998). Following Signorino (2003), I use the simple model displayed in Fig. 2 to illustrate the logic of the strategic approach with two rational, utility-maximizing players.

⁶Note that deterrence and acceptability of the existing political arrangement can be separated by selection and placement of covariates, as I demonstrate in the substantive application.

⁷Note that the recent upsurge of civil war scholarship using dyadic data does not alleviate this problem. Dyadic studies of civil war only include cases from observed civil wars—where there are data on rebels—and not cases where a civil war was avoided because of government concessions or deterrence. This introduces the possibility of selection bias in these studies, as they only include cases where the opposition groups have strategically selected into the civil war. Cunningham (2013), for example, examines civil war onset using data on opposition movements seeking self-determination in effort to account for the dyadic nature of civil war onset. This is an important step in accounting for biases caused by treating civil war onset as a unitary state-level phenomenon. The data, however, still suffer from selection problems regarding which groups seek self-determination and which are placated by government concessions.

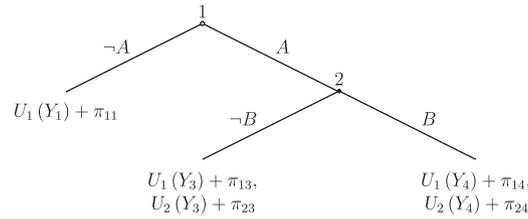


Fig. 2 A strategic model with private information.

Each player's payoff is comprised of an observable utility, denoted $U_i(Y_j)$, and a private information component, π_{ij} , where i is the player and j is the payoff. Player 1 can choose A or $\neg A$. If Player 1 chooses $\neg A$, the game ends, Y_1 is observed, and Player 1 receives the payoff $U_1(Y_1) + \pi_{11}$. If Player 1 chooses A , then Player 2 chooses either B or $\neg B$. If Player 2 selects $\neg B$, Y_3 is observed, Player 1 receives $U_1(Y_3) + \pi_{13}$, and Player 2 receives a payoff of $U_2(Y_3) + \pi_{23}$. If Player 2 chooses B , Y_4 is observed, Player 1 receives $U_1(Y_4) + \pi_{14}$, and Player 2 receives $U_2(Y_4) + \pi_{24}$. The model is strategic in that, in order to maximize her own payoff, Player 1 must take into account the expected action by Player 2.

Following Signorino's approach, π_{ij} is known only to player i , while the other player (and the analyst) know only its distribution. Assuming that π_{ij} follow independent and identical normal distributions with mean 0 and variance σ^2 , and that players are rational utility-maximizers, the strategic choice probabilities are

$$p_B = \Phi \left[\frac{U_2(Y_4) - U_2(Y_3)}{\sqrt{2\sigma^2}} \right], \quad (1)$$

$$p_A = \Phi \left[\frac{p_B(U_1(Y_4)) + (1 - p_B)(U_1(Y_3)) - U_1(Y_1)}{\sqrt{p_B^2\sigma^2 + (1 - p_B)^2\sigma^2 + \sigma^2}} \right], \quad (2)$$

where Φ is the normal cumulative density function.

If we assume that Player 2 plays B if and only if $U_2(Y_4) + \pi_{24} > U_2(Y_3) + \pi_{23}$ and that Player 1 selects A if and only if $p_B(U_1(Y_4) + \pi_{14}) + (1 - p_B)(U_1(Y_3) + \pi_{13}) > U_1(Y_1) + \pi_{11}$, then equation (1) represents both Player 1 and the analyst's belief that Player 2 will choose B , while equation (2) reflects the analyst's belief that Player 1 will select A . Solving the game yields three possible outcomes, with each being the product of the choice probabilities:

$$p_{Y_1} = 1 - p_A \quad (3)$$

$$p_{Y_3} = p_A(1 - p_B) \quad (4)$$

$$p_{Y_4} = p_A p_B. \quad (5)$$

One advantage of Signorino's approach is that it permits us to derive an estimator that directly reflects the structure of the formal model (Signorino 1999, 2003, 2007). Empirically, one can represent $U_i(Y_j)$ with a set of regressors, such that $U_i(Y_j) = X_{ij}\beta_{ij}$, while the private information is assumed to take a specified distribution, such as normal or logistic. Signorino (2007, 487) notes that "assuming one has data for the players' decisions and regressors for the utilities, then one can estimate parameters via maximum likelihood estimation (MLE)." Making the standard assumption that $\sigma^2 = 1$, the parameters are recovered by maximizing the following equation:

$$L = \prod_{i=1}^n \text{P}(Y_{1,i} = 1)^{y_{1,i}} \text{P}(Y_{3,i} = 1)^{y_{3,i}} \text{P}(Y_{4,i} = 1)^{y_{4,i}}. \quad (6)$$

Unfortunately, one does not always have data for the players' decisions. Signorino and Tarar (2006, 588) describe one reason for this, noting that most data-collection efforts are undertaken without regard for the underlying structure of the theoretical model. The solution for this, of course, is to collect additional data, while paying careful attention to theoretical processes. This solution, however, is not always feasible, given financial and time constraints, coordination problems, language limitations, and other factors. Moreover, in some cases, the data may no longer exist for collection. In order to collect data for cases of civil war as described above, for example, one would have to collect data not only on opposition groups in cases where civil wars have occurred, but also in states with no civil war outcomes. Including data on the latter cases is necessary to determine why groups within these states did not turn to violence (e.g., deterrence by the government or content with the existing arrangements). Finally, in order to fit the framework of large- n analysis, such data-collection efforts must reach a sufficient spatial and temporal scope.

In practice, many types of currently available data are outcome- rather than actor-specific. Rather than providing information for each of the two decisions observed in Fig. 2, many existing data sets only provide information associated with the outcome where Player 1 “challenges” and Player 2 “fights” (outcome denoted as Y_4 in Fig. 2). Within such data sets, the outcome variable Y_4 is only “partially observed,” since it is the result of unobserved *joint* decisions of two players, rather than decisions by a single decision-maker (Poirier 1980).⁸ Consequently, such data sets also conflate the other two terminal nodes (Y_1 and Y_3) into a single “non-event” outcome. The two non-events, outcomes Y_1 and Y_3 , however, also result from two distinct interactive processes. Grouping these outcomes together effectively treats the observed outcome as an additive function of actor utilities and ignores the conditional nature of Player 2's choices, producing biased estimates of Player 2's utility. Signorino and Yilmaz (2003, 556–57) show that ignoring this conditionality results in omitted-variable bias, where the omitted variables are effectively higher-order nonlinear terms of a Taylor series expansion.

In short, most of the existing estimation techniques are unable to address this problem. *Traditional probit/logit models* mistreat the strategic model as an additive function (see Signorino and Yilmaz 2003). This ignores the conditional nature of Player 1's choices, as depicted in Fig. 2 and equation (2). Even models that account for inflation of zeros, such as *split-sample probit (SSP)/logit models*, are insufficient. Split-sample models assume two distinct “types” of Player 1—one who never engages with Player 2 (the zero-inflated or relevancy equation) and one who does interact with Player 2 (the traditional probit/logit equation) (Xiang 2010).⁹ By assuming two “types,” these estimators treat the behavior of Player 1 as independent of Player 2; hence, they ignore strategic behavior and are inconsistent with theoretical models of strategic interactions, such as that in Fig. 2.

3 Strategic Probit with Partial Observability

I address this problem by deriving an estimator—SPPO—that probabilistically estimates player actions for the types of outcome-aggregated data described above. In contrast to a split-population binary choice estimator, SPPO explicitly accounts for strategic behavior on the part of the actors. The estimator relaxes the data availability assumption required by traditional multi-stage strategic models (e.g., Signorino 2003), and instead uses information from the regressors representing observed utilities to predict unobserved player actions. If we assume π_{ij} is independent and identically normally distributed with mean 0 and variance 1, the likelihood takes the form

$$L = \prod_{i=1}^n P(Y_i = 1)^{y_i} P(Y_i = 0)^{1-y_i}, \quad (7)$$

⁸Like the present study, Poirier (1980) examines outcome variables resulting from the joint actions of utility-maximizing actors. Poirier's study examines only bivariate probit models, however, and does not consider strategic behavior.

⁹See also Harris and Zhao (2007) and Bagozzi et al. (2014) for models using a split-sample ordered probit.

where

$$P(Y_i = 1) = p_A p_B \quad (8)$$

$$P(Y_i = 0) = (1 - p_A) + p_A(1 - p_B) = 1 - p_A p_B. \quad (9)$$

The estimator corresponds to the choice probabilities described earlier. Equation (8) is the probability that we observe a “1” in the data, that is, where both A and B occur. The part where the estimator extends the existing strategic probits is equation (9), which pools the two “non-event” or “0” outcomes in the hypothesized data (Y_1 and Y_3). Thus, SPPO uses the observed portion of the utility (the regressors) to separate and identify the types of “0,” just like split-population models. Unlike split-population models, however, SPPO explicitly takes into account the strategic nature of the interaction. It does this in two ways: (1) Player 2’s expected behavior is included in Player 1’s utility calculation (i.e., p_B is a constituent term in equation (8)) and (2) by incorporating $p_A(1 - p_B)$, which represents the probability that Player 2 acquiesces when Player 1 selects A .

Maximizing the likelihood produces parameter estimates of the strategic relationship with unobserved player actions, where the non-events are pooled (or cannot be separated). The estimator corresponds to the structure depicted in Fig. 2 and adheres to the same backwards induction logic commonly theorized in strategic games.¹⁰

4 Monte Carlo Analyses

In order to demonstrate the advantages of SPPO over traditional and split-sample binary estimators, I conduct two sets of analyses using Monte Carlo simulations.¹¹ The first set of analyses compares three estimators—SPPO, SSP, and traditional probit—assuming a data-generating process (DGP) consistent with the strategic model outlined earlier. This analysis is useful because it demonstrates the extent of bias in parameter estimates that results from ignoring the strategic relationship. In the second set of analyses, I assess the robustness of the SPPO estimator when the model is misspecified. That is, I examine whether the estimator recovers the correct parameter estimates when irrelevant variables—variables not included in the DGP—are included. This is important, because strategic models are especially sensitive to bias: given the conditional relationship of the first-stage estimates on the predicted probabilities of the second stage, any bias in the estimates of the second stage will carry over to the estimates of the first stage (Leeman 2014).

4.1 Strategic DGP

Let us assume the following DGP:

$$y^* = \begin{cases} Y_1 & \text{if } U_1^*(Y_1) \geq U_1^*(Y_3) \text{ and } U_2^*(Y_3) \geq U_2^*(Y_4) \\ & \text{or } U_1^*(Y_1) \geq U_1^*(Y_4) \text{ and } U_2^*(Y_4) > U_2^*(Y_3) \\ Y_3 & \text{if } U_1^*(Y_3) > U_1^*(Y_1) \text{ and } U_2^*(Y_3) \geq U_2^*(Y_4) \\ Y_4 & \text{if } U_1^*(Y_4) > U_1^*(Y_1) \text{ and } U_2^*(Y_4) > U_2^*(Y_3) \end{cases}$$

¹⁰The same identification and exclusion restrictions that affect other strategic models are true for SPPO as well. Namely, the same variable must be excluded from at least one equation per player in order for the model to be identified (Lewis and Schultz 2003).

¹¹All materials necessary to replicate simulations, figures, and tables from this article are available in Nieman (2015).

where, from the perspective of the analyst,

$$U_1^*(Y_1) = X_{11}\beta_{11} + \pi_{11}, \tag{10}$$

$$U_1^*(Y_3) = \pi_{13}, \tag{11}$$

$$U_1^*(Y_4) = X_{14}\beta_{14} + X_c\beta_{14c} + \pi_{14}, \tag{12}$$

$$U_2^*(Y_3) = \pi_{23}, \tag{13}$$

$$U_2^*(Y_4) = X_{24}\beta_{24} + X_c\beta_{24c} + \pi_{24}, \tag{14}$$

and $y = 1$ if $y^* = Y_4$, and $y = 0$ otherwise. I set the parameter values as $\beta_{11} = \beta_{14} = \beta_{24} = \beta_{14c} = \beta_{24c} = 1$. The explanatory variables X_{ij} represent single regressors and are normally distributed with the mean of 0 and variance of 1, $N(0, 1)$. In addition, both players' utilities may include the same or common regressors.¹² To reflect this, both players' utilities from outcome Y_4 , ($U_1^*(Y_4)$ and $U_2^*(Y_4)$), include a common regressor X_c . Finally, π_{ij} is drawn from a normal distribution with mean 0 and variance 1 and is independently and identically distributed across observations. I run 2000 simulations with 5000 observations each. The DGP corresponds to the theoretical model displayed in Fig. 2.

The specification of the SPPO estimator is the same as the DGP and adheres to the likelihood depicted in equation (7). The SSP model is specified as

$$y_s^* = X'_{11}\beta_{11} + X_c\beta_{14c} + \epsilon_1, \tag{15}$$

$$y^* = X_{14}\beta_{14} + X_{24}\beta_{24} + X_c\beta_{24c} + \epsilon_2, \tag{16}$$

where $y_s = 1$ if $y_s^* > 0$ and 0 otherwise, and $y = 1$ if both $y^* > 0$ and $y_s = 1$ and 0 otherwise. Equation (15) represents the “selection” equation and equation (16) represents the “interaction” equation (i.e., traditional probit equation).

Finally, the traditional probit model is specified as

$$y^* = X'_{11}\beta_{11} + X_{14}\beta_{14} + X_{24}\beta_{24} + X_c\beta_{24c} + \epsilon, \tag{17}$$

where $y = 1$ if $y^* > 0$ and 0 otherwise.

Figure 3 displays the results of the simulations comparing the kernel density of the estimates from the SPPO (solid line), SSP (long dashed line), and traditional probit (short dashed line). SPPO always captures the true value while the SSP and traditional probit results are biased, consistently underestimating the true coefficients.

Table 1 compares the recovered coefficients and root mean squared errors (RMSEs) associated with each estimator. As was the case with Fig. 3, the results in Table 1 show that SPPO is able to recover estimates approximating the true value, while neither SSP nor the traditional probit model do. While SSP outperforms traditional probit, due to the former's ability to model outcomes as results of multiple processes, it is unable to account for the strategic nature of multi-player interactions. The RMSEs tell much the same story, with SPPO producing the lowest combined levels of bias and variance among models, SSP being the second best, and probit consistently performing worst.

Finally, Table 2 reports in-sample fit statistics for the Monte Carlo simulations. Because SPPO, SSP, and traditional probit are non-nested in their functional forms, ordinary fit statistics, such as AIC, BIC, and comparing log-likelihoods, cannot discriminate between them (Clarke 2001; Clarke and Signorino 2010). Instead, I employ two tests that are designed to evaluate the model fit of

¹²In the civil war example above, both the government and the rebels' utilities from war may be affected by the states' natural resource endowment; that is, territories rich in natural resources, such as oil or diamonds, make for a more desirable prize.

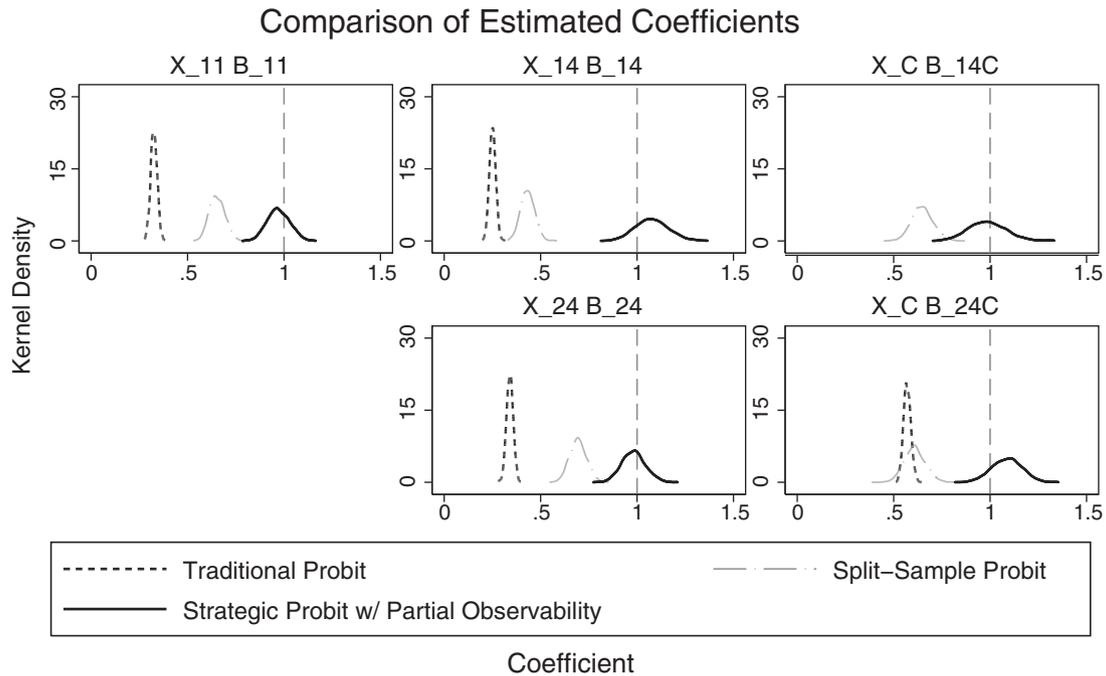


Fig. 3 Comparison of recovered coefficients across models with a strategic DGP.

Table 1 Comparison of estimated coefficient, standard error, and RMSE across models with a strategic DGP

Regressor	True Value	Recovered coefficient (standard error)			RMSE		
		Probit	SSP	SPPO	Probit	SSP	SPPO
Player 1: Equation Y_1							
X_{11}	1	0.326 (0.020)	0.653 (0.041)	0.972 (0.058)	0.689	0.402	0.248
Player 1: Equation Y_4							
X_{14}	1	0.251 (0.017)	0.433 (0.037)	1.078 (0.083)	0.762	0.599	0.308
X_c	1	—	0.649 (0.056)	0.987 (0.090)	—	0.425	0.315
Player 2: Equation Y_4							
X_{24}	1	0.341 (0.021)	0.694 (0.043)	0.982 (0.059)	0.675	0.371	0.250
X_c	1	0.570 (0.022)	0.610 (0.055)	1.087 (0.069)	0.454	0.457	0.286

Notes: $RMSE = \sqrt{Bias^2 + Variance}$. Because the traditional probit is a single equation model, it estimates only one parameter for X_c , which is displayed in Player 2: Equation Y_4 .

non-nested models: the Vuong test (Vuong 1989) and Clarke's (2003) distribution-free test. The Vuong test compares the average log-likelihoods of two models. The null hypothesis of the Vuong test is that the average log-likelihood ratio of two models is zero. If the first model is closer to the true specification, then the average log-likelihood ratio is significantly greater than zero. If the second model is closer to the true specification, then the average log-likelihood ratio is significantly less than zero. Clarke's distribution-free test, on the other hand, tests whether the median log-likelihood ratio is significantly different than zero. The null hypothesis is that half of the individual log-likelihoods are above zero, and half are below. If the first model is closer to the true specification, then the ratio is positive. If the second model is closer to the true specification, then the ratio is negative. Both the Vuong and Clarke tests can be modified to correct for the degrees of

Table 2 Comparison of average model fit with a strategic DGP

	<i>Probit</i>	<i>SSP</i>	<i>SPPO</i>
Clarke test			
$\sum_i^n ll_{SPPO,i} - ll_{M_A,i} > 0$	3891.2	2524.7	—
Positive, one-side test (<i>p</i> -value)	<0.001	0.380	—
Negative, one-side test (<i>p</i> -value)	>0.999	0.625	—
Equal, two-side test (<i>p</i> -value)	<0.001	0.290	—
Vuong test			
Vuong	710.055	48.016	—
SE	11.223	3.756	—
<i>t</i> -Statistic	171.337	14.507	—
<i>p</i> -Value	<0.001	<0.001	—
Prediction			
% of observation correct	63.8	80.9	80.2
% of observation correct if <i>Y</i> = 1	75.0	56.2	64.8
% False positive	40.8	9.2	13.6
% Correctly predicted unobserved <i>Y</i> ₃	—	—	41.3

Notes: *M_A* is the alternative model listed in the column, while *i* is the individual observation.

freedom.¹³ I also look at the percent of cases that are correctly classified as an “event” or “non-event,” the percent of “events” that are correctly identified, and the number of false positives.¹⁴

The in-sample fit statistics from Table 2 show that SPPO outperforms the traditional probit model. Both the Clarke and Vuong tests reject the null that the models are the same and provide support for the SPPO. The traditional probit is only able to correctly classify approximately 64% of observations, while both SPPO and SSP are able to correctly classify approximately 80%. The probit model is better able to correctly predict an “event” when *Y* = 1; however, it does so by over-predicting the number of “events,” as evident by the high rate of false positives. On the other hand, SSP and SPPO are more difficult to distinguish. The Clarke test is unable to reject the null that the models are the same, while the Vuong test is able to reject the null and offers support for SPPO. While SPPO is slightly better able to correctly identify an “event,” it also has a slightly higher rate of “false positives.”¹⁵ Finally, unlike either probit or SSP, SPPO is also able to probabilistically identify the third (unobserved) type of outcome, associated with the second type of zeros. In the above set of analyses, the rate of correctly classified unobserved events is approximately 41.3%. While not very high, this rate is substantially greater than either the random rate of 33.3% or, which is more common, altogether ignoring this qualitatively distinct type of event. In short, while providing some support for SPPO, the results also highlight the difficulty of distinguishing between non-nested models, especially when the outcome variable of interest is only partially observed.

SPPO is the only available binary choice estimator that corresponds to the strategic logic common to many political science theories, and social interactions more generally. In contrast to biological or medical experiments where actors are truly selected into one population or another, most social interactions imply that one actor’s behavior affects that of another actor; hence, social actors rarely act without any consideration of one another’s actions. Ignoring this strategic

¹³I employ the Schwarz (1978) correction to each of the non-nested tests. The corrected Vuong test is $LR_n(\hat{\theta}_n, \hat{\gamma}_n) - [(\frac{p}{q}) \ln n - (\frac{q}{p}) \ln n]$, where LR is the log-likelihood ratio, $\hat{\theta}$ and $\hat{\gamma}$ are the model estimates, and *p* and *q* are the number of estimated coefficients for model *f* and *g*, which are the two models being compared (Vuong 1989; Clarke 2001). The adjusted Clarke test multiplies the individual likelihoods by the average Schwartz correction (Clarke 2003, 78). That is, model *f* is multiplied by $(\frac{p}{2n}) \ln n$ and model *g* is corrected by $(\frac{q}{2n}) \ln n$.

¹⁴For both probit and SSP, I classify an “event” *Y_i* as correctly identified as 1 if *Y_i* = 1 and $P(Y_i|X) > 0.5$. Analogously, an event is classified as correctly identified as 1 if *Y_i* = 0 and $P(Y_i|X) \leq 0.5$. For SPPO, an “event” *Y_i* is classified as correctly identified as 1 if the observed outcome *Y_i* = 1 and $P(Y_4) > P(Y_3)$ and $P(Y_4) > P(Y_1)$, and correctly identified as 0 otherwise, given that the observed outcome *Y_i* = 0.

¹⁵Despite its impressive performance on fit statistics, recall that SSP does not fare well in terms of unbiasedness of recovered coefficients, as shown in Table 1.

behavior during estimation has been shown to induce omitted-variable bias, resulting in biased parameter estimates and potentially incorrect inferences (Signorino and Yilmaz 2003).

4.2 Robustness to Misspecification and Confounding Regressors

The Monte Carlo analysis conducted in the previous section assumed the researcher's knowledge of the correct model specification: that is, the estimated model included all of the relevant regressors and no irrelevant regressors. In practice, however, researchers seldom know the true underlying DGP and are likely to misspecify the equations associated with one or more players' utilities (i.e., include theoretically irrelevant and possibly confounding regressors). In the application of strategic models, such possible misspecification is thought to be especially problematic: given that Player 1's action is conditioned by the expected behavior of Player 2, the bias and inefficiency from Player 2's utilities (i.e., equation (1)) also affect parameter estimates for Player 1 (i.e., equation (2)). In order to assess the robustness of the SPPO estimator to model misspecification, I conduct the following additional analyses.

The true specification remains as depicted in equations (10)–(14). The estimated model, however, is misspecified to approximate several common specification errors. First, to approximate the most common specification error—the inclusion of irrelevant or spurious covariates—the model includes a spurious covariate X_s , drawn from the normal distribution with the mean 0 and variance 1 and uncorrelated with any other regressors or the outcome variables. Second, I approximate two possible model specifications associated with strategic estimation: (1) incorrectly including Player 1's regressor in an equation associated with Player 2's utility and (2) including a player's regressor in the wrong utility. As a result, the estimated model is different from the true model in three ways: (1) Player 1's utility for outcome Y_4 includes a spurious regressor X_s , (2) Player 2's utility from outcome Y_4 includes regressor X_{14} , which belongs in Player 1's utility for outcome Y_4 , and (3) Player 1's utility from outcome Y_1 incorrectly includes the regressor X_{14} , which belongs into Player 1's utility for outcome Y_4 . Notably, (2) and (3) induce a spurious correlation between utilities. As a result, the estimated model looks the following way:

$$U_1^*(Y_1) = X_{11}\beta_{11} + X_{14}\beta_{1m} + \pi_{11}, \quad (18)$$

$$U_1^*(Y_3) = \pi_{13}, \quad (19)$$

$$U_1^*(Y_4) = X_{14}\beta_{14} + X_c\beta_{14c} + X_s\beta_s + \pi_{14}, \quad (20)$$

$$U_2^*(Y_3) = \pi_{23}, \quad (21)$$

$$U_2^*(Y_4) = X_{24}\beta_{24} + X_c\beta_{24c} + X_{14}\beta_{2m} + \pi_{24}, \quad (22)$$

where β_{1m} represents the estimation parameter on the regressor X_{14} placed incorrectly in Player 1's utility from outcome Y_1 ; β_s represents the parameter on the spurious regressor X_s ; and β_{2m} represents the parameter on the regressor X_{14} , incorrectly placed in Player 2's utility from outcome Y_4 . Note that, as a result of this misspecification, regressor X_{14} appears in three of the estimated equations.

Table 3 displays both the true and recovered coefficients from this model. First, we can see that the model is able to correctly identify the spurious regressor X_s as irrelevant, as the estimated coefficient on this regressor is close to 0 in absolute value and is not statistically significant. Second, the estimator is also able to discern the misplacement of Player 1's regressor into Player 2's equation: the estimated coefficient on X_{14} in Player 2's equation is -0.043 and not statistically significant. The estimator runs into more trouble, however, discerning the correct placement of regressor X_{14} between the two equations associated with Player 1's utilities. Unlike in the true model, in both of the relevant Player 1's utilities, the coefficients on X_{14} have incorrect values; moreover, the recovered coefficient is statistically insignificant in the equation the variable belongs to, while statistically significant in the wrong equation. This result, however, is not as discouraging

Table 3 Estimates recovered by SPPO for the misspecified model

<i>Regressor</i>	<i>True value</i>	<i>Recovered coefficient (standard error)</i>
Player 1: Equation Y_1		
X_{11}	1	0.972 (0.057)
X_{14}	0	-0.672 (0.195)
Player 1: Equation Y_4		
X_{14}	1	0.198 (0.247)
X_c	1	0.916 (0.084)
X_s	0	0.003 (0.064)
Player 2: Equation Y_4		
X_{24}	1	0.998 (0.059)
X_c	1	1.146 (0.070)
X_{14}	0	-0.043 (0.063)
% Correct		80.2
% Correct if $Y = 1$		64.7
% False positive if $Y = 0$		13.6
% Correctly predicted unobserved Y_3		42.7

as it may seem, as it leads to the correct theoretical inference: the regressor decreases the utility from outcome Y_1 , which is theoretically equivalent to increasing the utility from outcome Y_4 —the intended inference in the true model. The recovered coefficient, however, is of smaller magnitude than in the true model, possibly due to the conditional relationship between players' utilities and the difficulty of separating effects in the presence of partial observability.

In other words, the proposed estimator performs well under the first two types of misspecification (spurious variables and inclusion of regressors in the wrong player's equation). Misplacing regressors among the utilities of the same player, however, raises some concerns. Fortunately, the third type of misspecification—placing the same regressors in several of the same player's equations—is rather uncommon in the published research that uses strategic estimators.¹⁶ Such misspecification also seems unlikely from the theoretical point of view: conventionally, regressors that affect a player's utility from Action A are going to decrease its utility from Action $\neg A$, which means that for capturing the effect of these regressors on the player's utilities, it should suffice to include them in one of the utilities, but not in both.

Finally, the inclusion of spurious variables does not have much affect on the estimator's ability to recover correct coefficients on the remaining regressors X_{11} , X_c , and X_{24} , as well as on the model fit statistics. The numbers of correctly predicted events, false positives, and correctly classifying the unobserved outcome are virtually unchanged, despite the misspecification.

5 Re-Examining Civil War Onset

To highlight the benefits of SPPO in more applied terms, I re-estimate Fearon and Laitin's (2003) seminal study on civil war onset. Fearon and Laitin (2003) evaluate several hypotheses regarding the onset of civil wars. Their study aims at identifying the factors that increase the likelihood of civil war onset, specifically (1) factors that drive local populations to rebel against the government and (2) factors that decrease the governments' ability to effectively police its territory and engage in counterinsurgency practices (Fearon and Laitin 2003, 75–76). Notably for the purposes of the current project, the structure of the Fearon and Laitin's data, like most civil war data, does not distinguish cases where potential rebels preferred the status quo (either because they chose not to

¹⁶I am unaware of any research in which the same regressor is placed in multiple equations of the same player. It is more common to place the same regressor in both players' utilities (see Leblang 2003; Carson 2003, 2005; Signorino and Tarar 2006; Holyoke 2009; Carter 2010; Helmke 2010; McLean and Whang 2010; Bas 2012a; König and Mäder 2014). Even if misspecified, the Monte Carlo simulations suggest that the latter is unproblematic.

take up arms or were deterred) from the cases where the state acquiesced to challenges rather than responding with violence. I re-estimate Model 1 of Table 1 from Fearon and Laitin's study, which examines the onset of civil war for all countries from 1945 to 1999 (Fearon and Laitin 2003, 84).

Fearon and Laitin (2003) articulate the theorized mechanisms associated with each variable within their empirical model, and separate the covariates into (1) the factors associated with the actions of the insurgents, (2) the factors associated with the actions of the state, or (3) both of the above. Many of the mechanisms are theorized to adhere to a strategic process; for example, stronger central governments are expected to *deter* potential insurgents from taking up arms (Fearon and Laitin 2003, 80). In the original analysis, however, these theoretical expectations are tested using logistic regression, which conflates the causal mechanisms associated with each actor's explanatory variables, and ignores the strategic nature of behavior.

For example, while *GDP per capita* is one of many variables included in the analysis, its effect—and the underlying cause of its effect—have been the subject of significant scholarly debate (cf. Collier and Hoeffler 2004; Fearon 2005). Fearon and Laitin (2003, 80) discuss three mechanisms associated with GDP per capita: (a) proxy for state's police and military capacity; (b) the power projection of the central administration to reach into rural society; and (c) opportunity costs for potential rebel recruits. They lament that, “[t]hough we try below, it is difficult to find measures to distinguish among these three mechanisms associating a low per capita income with civil war onset.” The advantage of SPPO estimator is in its ability to do exactly that—separate the deterrence mechanisms based on state capacity (a and b) from the opportunity cost mechanism (c).

Given the theorized strategic relationship at the heart of government—insurgent interaction—this interaction can be modeled as a two-player strategic game with incomplete information. The two players are the *Rebels* and the *Government*. Corresponding to the logic of the game depicted in Fig. 2, the *Rebels* choose to either accept the legitimacy of the government or pick up arms and challenge the existing governing structure. The *Government* decides whether to accept their demands or respond with force. As described above, *Rebels* must compare their expected utility from challenging to that of accepting the status quo. In doing so, of course, the *Rebels* must take into account the expected behavior of the *Government*. This produces three possible outcomes: one, in which the status quo is maintained, another, in which the *Government* Acquiesces, and finally a third type of outcome, in which the interaction results in a civil war.

5.1 Placement of Regressors

I assign each variable to players' outcome utilities according to the causal mechanisms posited by Fearon and Laitin (2003, 78–82). Importantly, some variables are theorized to be a part of more than one player's outcome utility. *Oil exporter*, for example, is expected to decrease the *Government's* utility for war as it is associated with less effective state apparatuses and bureaucracies and, hence, less effective policing that might capture or deter potential insurgents ($X_{24}\beta_{24}$). At the same time, *oil exporter* also represents a “prize” held by whoever is controlling the state ($X_{14}\beta_{14}$) (Fearon and Laitin 2003, 81). Likewise, *GDP per capita* both represents an opportunity cost for joining an insurgency ($X_{11}\beta_{11}$), and provides a deterring effect by increasing state capacity, representing the government's military strength ($X_{24}\beta_{24}$).

Noteworthy, regressors in the *Government* war equation, such as *GDP per capita*, capture the idea of deterrence, as these regressors' effects condition the *Rebels'* expected return on challenging the status quo (recall equation (2)). In other words, *Government's* probability of fighting the insurgents decreases the *Rebels'* expected utility from challenging the existing governing structure, all else equal.

In practice, choosing the placement of individual regressors for each player, and especially determining the appropriate player utility equation, can be quite difficult (Leeman 2014; Signorino and Tarar 2006, fn 12). These difficulties are as true for SPPO as they are for any other multi-stage model. When it comes to strategic models, like SPPO, however, the silver lining is that this particular estimator highlights the importance of relying on theory to not only

determine which variables matter, but more importantly, *how* they matter.¹⁷ The emphasis on theory and causal mechanisms discourages “garbage can” models, which include a number of covariates with little thought or justification regarding the functional form or the relationship among the covariates, that is, direct, indirect, or conditional effects, monotonic or non-monotonic relationships, etc. (Achen 2005; Ray 2005). On account of this, strategic models can provide leverage when evaluating competing underlying causal mechanisms prescribed to some variables, as is the case with *GDP per capita*. As noted previously, Fearon and Laitin facilitate the current study’s task of placing regressors by providing very clear descriptions of the causal process(es) through which each variable affects civil war.¹⁸

5.1.1 Rebels’ status quo regressors (X_{11})

The opposition’s decision to operate within the existing political structure (utility from the *status quo*) is theoretically affected by three factors: *GDP per capita*, *democracy*, and *political instability*.¹⁹ *GDP per capita* has a positive affect on the opposition’s value for the status quo, as it proxies the opportunity cost that a potential recruit must pay to join an insurgency. According to Fearon and Laitin (2003, 80), “[r]ecruiting young men to the life of a guerrilla is easier when the economic alternatives are worse.” The costs of joining an armed rebellion are also higher in *democracies*, as democratic institutions provide alternative means of voicing and acting on political grievances. Democratic political structures give all groups access to political power and are usually associated with less discrimination and repression of political minorities (Fearon and Laitin 2003, 79). Finally, *political instability* has a negative effect on the value of the status quo, as this variable captures weakness or ineffectiveness on the part of the government (Fearon and Laitin 2003, 81). Such instability could manifest itself in a lack of rule of law, reducing the opportunity costs of joining insurgents and perhaps even increasing the benefits of doing so, as insurgents may serve as an alternative to police (e.g., Kalyvas and Kocher 2007).

5.1.2 Rebels’ civil war regressors (X_{14})

Fearon and Laitin identify a number of variables affecting the opposition’s choice to violently challenge the central government: *population size*, *mountains*, a state’s *oil export*, *ethnic* and *religious fractionalization*, and *prior war*. *Population size* has a positive effect on the utility for war, as it proxies the number of the potential recruits available to rebel organizations (Fearon and Laitin 2003, 81). Insurgents are also likely to derive benefits from rough terrain, operationalized as *mountains*, because “the numerical weakness of the insurgents implies that, to survive, the rebels must be able to hide from government forces” (Fearon and Laitin 2003, 80). Oil production and sales create an economic prize for controlling state power, or at least specific territory, enhancing the utility of war for *Rebels* in *oil exporting* countries (Fearon and Laitin 2003, 81).

Political and social groups’ value for war increases in the presence of discrimination and grievances based on cultural and religious differences (Fearon and Laitin 2003, 78–79). While not especially swayed by the cultural argument, Fearon and Laitin note two specific theoretical mechanisms that lead to cultural grievances. According to the first mechanism, often highlighted by journalists and nationalist politicians, contrasts in cultural practices give rise to deep-seated animosities between different ethnic and religious groups (e.g., Huntington 1996). Second, differences in cultural practices may be associated with institutional impediments for particular groups’ upward economic mobility, which increases these groups’ utility for seeking greater autonomy or independence (e.g., Anderson 1983). In the empirical model, these cultural factors are

¹⁷See Carter (2010) for an example where the author pays special attention to the causal logic of the included regressors in a strategic model.

¹⁸When justifying the placement of each regressor in what follows, I include page number references to Fearon and Laitin’s variable-specific theoretical justifications for how each variable is expected to affect *Rebels* and/or *Government’s* expected utilities.

¹⁹See Fearon and Laitin (2003) for a complete description of variables.

operationalized as *ethnic* and *religious fractionalization*. Finally, insurgents may be less likely to initiate a conflict in the immediate aftermath of a civil war because (a) they have obtained their goal (independence, increased autonomy, a change in policy, etc.) or (b) they lack resources, recruits, and leadership after fighting and either losing or reaching a stalemate with the government.²⁰

5.1.3 Government's civil war regressors (X_{24})

Government's decision to fight back rather than acquiesce (utility for civil war) is affected by *GDP per capita*, whether the insurgent territory is *contiguous* to the rest of the country, state reliance on *oil exports*, whether the state itself is *new* or recently independent, *political instability*, and *prior war*. *GDP per capita* operates as a proxy for a state's police and military capacities, where greater values indicate a stronger state. Fearon and Laitin (2003, 80) argue that "[i]nsurgents are better able to survive and prosper if the government and military they oppose are relatively weak." Stronger states are better equipped to respond to rebel groups' challenges and, therefore, derive greater utility from war than weaker states. Governments are less effective at responding to challenges from *non-contiguous* territories, as physical separation from the state's center increases the costs of power projection (Fearon and Laitin 2003, 81).

States with heavy reliance on export of natural resources, such as *oil*, tend to have weaker bureaucratic apparatuses, including police and security resources, than states with similar economic development but less reliance on natural resource exports (Fearon and Laitin 2003, 81). Newly independent states lack the support of their former imperial power and may also lack experienced militaries; hence, *new state* is expected to decrease *Government's* utility from civil war (Fearon and Laitin 2003, 81). *Political instability* is expected to reduce the state's ability to fight insurgents. In the words of Fearon and Laitin (2003, 81), "political instability at the center, which may indicate disorganization and weakness," creates "an opportunity for a separatist or center-seeking rebellion." Finally, *prior war* may operate in one of two ways. On the one hand, in the immediate aftermath of a civil war, the government may be weakened and less able to effectively combat new challenges. On the other hand, recent experience of fighting insurgents may strengthen the government's preparedness and ability to combat new challenges in the near future.

Finally, it is important to keep in mind that all regressors included in the *Government's* civil war utility will also indirectly affect the calculation for *Rebels'* utility for challenging (Acquiescence + Civil War) (recall equation (2)).²¹

5.2 Empirical Analysis

Table 4 compares the results of replicating Fearon and Laitin (2003) using a traditional probit and SPPO.²² The coefficients from the SPPO model are interpreted in the same manner as other strategic models: they are associated with the value an actor places on each *utility*.²³ This means that positive values indicate that the actor places greater value (utility) on a specific outcome, yet the values do not directly reflect the overall probability of an outcome. In Table 4, for example, positive parameters on the variables under the status quo equation suggest that increases in these variables are associated with an increase in *Rebels'* utility from the status quo, making the probability of a challenge less likely. Moreover, coefficients for *Rebels* represent their values for each utility *after accounting for the expected action of Government*. That is, the factors contributing to *Government's*

²⁰*Prior war* is the only variable for which (Fearon and Laitin 2003, 82–83) provide no theoretical justification, though they do note a methodological justification. Hence, I determine the placement of this particular regressor in players' utilities by drawing from the bargaining literature on war (e.g., Werner 1999; Filson and Werner 2002; Chiba 2011).

²¹The *Rebels'* acquiescence (X_{13}) is captured by a constant.

²²Although the two estimated models include the same variables, SPPO estimates a larger number of parameters, since several of the variables are included in more than one equation, as described above.

²³As in the case with probit/logit regression model, the reported coefficients are parameter estimates of the variable effects, assuming that the variance is held constant (Signorino 1999, 2003, 2007). Bas (2012b) develops a strategic model where the variance is specified as a function of regressors and is also estimated.

Table 4 Replication of Fearon and Laitin (2003): Comparing traditional probit and SPPO

<i>Traditional probit</i>		<i>SPPO</i>	
		<i>Rebels' (Status Quo):</i>	
Prior war	-0.391** (0.130)	GDP per capita	0.208** (0.067)
GDP per capita	-0.135** (0.028)	Democracy	-0.012 (0.010)
Population	0.108** (0.031)	Instability	-0.220 (0.273)
Mountainous	0.091** (0.034)	Constant	6.247** (1.866)
Non-contiguous state	0.179 ⁺ (0.122)	<i>Rebels' (Acquiescence):</i>	
Oil exporter	0.352** (0.123)	Constant	4.474** (2.001)
New state	0.757** (0.163)	<i>Rebels' (War):</i>	
Instability	0.259** (0.101)	Population	0.413* (0.225)
Democracy	0.008 (0.007)	Mountainous	0.313* (0.175)
Ethnic fractionalization	0.087 (0.157)	Oil exporter	1.467 ⁺ (1.023)
Relig fractionalization	0.128 (0.209)	Ethnic fractionalization	0.308 (0.616)
Constant	-3.224** (0.303)	Relig fractionalization	0.272 (0.751)
		Prior war	-1.682* (0.991)
		<i>Government (War):</i>	
		GDP per capita	0.022 (0.119)
		Non-contiguous state	0.500 (0.393)
		Oil exporter	-0.098 (0.468)
		New state	2.517 (3.187)
		Instability	0.342 (0.460)
		Prior war	0.274 (0.620)
		Constant	-0.768 ⁺ (0.525)
Observations	6327		6327
Log-likelihood	-481.419		-477.849
% of observation correct	98.3		98.3
% of observation correct if Y=1	0.0		0.9
% False positive	0.0		0.0
Clarke test:			
$\sum_i^n ll_{Probit,i} - ll_{SPPO,i} > 0$		5090	
Positive, one-side test (<i>p</i> -value)		<0.001	
Negative, one-side test (<i>p</i> -value)		>0.999	
Equal, two-side test (<i>p</i> -value)		<0.001	
Vuong test:			
Vuong		-29.828	
SE		12.536	
<i>t</i> -Statistic		-2.379	
<i>p</i> -Value		0.009	

Notes: ***p* <0.05, **p* <0.1, two-tailed; ⁺*p* < 0.1, one-tailed. Standard errors in parentheses.

value for civil war affect *Rebels'* expected utility for both civil war and acquiescence indirectly through *p_B* (recall equation (2)).

Although both probit and SPPO recover coefficients with the same signs, there are several differences in the statistical significance of these coefficients (*non-contiguous state*, *new state*, *instability*).²⁴ What is of greater consequence, however, is that, unlike probit, SPPO allows for a more nuanced separation of alternative theoretical mechanisms. For instance, while *GDP per capita*, *oil exporter*, *instability*, and *prior war* are all statistically significant in the probit model, each of these

²⁴SPPO necessarily introduces greater uncertainty around parameter estimates: the probabilistic estimation of the unobserved player actions from the first stage of the interaction leads to more conservative parameter estimates than those of the traditional probit estimators. In contrast, in the presence of strategic interaction, the traditional probit produces fairly tight confidence intervals around *biased* parameter estimates, as illustrated in Fig. 3.

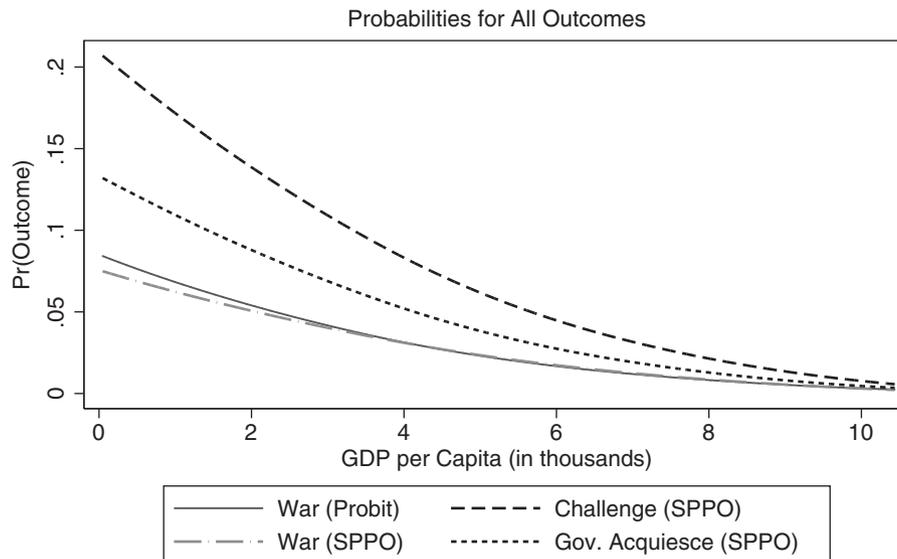


Fig. 4 Comparison of the predicted probabilities for challenge, acquiescence, and civil war for traditional probit and SPPO.

variables may affect the onset of civil war through several alternative theoretical mechanisms, often associated with different players. SPPO permits for evaluating each of the player-specific and outcome-specific theoretical mechanisms, while probit estimates the “combined” effect of treating these (often nonlinear) effects in an additive manner.

To further demonstrate this advantage of SPPO, let us focus on how the two models separate the competing theoretical mechanisms linking civil war and *GDP per capita*, identified by Fearon and Laitin (2003) and Collier and Hoeffler (2004), among others. The traditional probit model (Model 1 of Table 4) shows that *GDP per capita* has a negative and statistically significant effect. While this suggests that *GDP per capita* reduces the likelihood of civil war onset, we cannot determine which of the two alternative theoretical mechanisms are supported by this result. In other words, a negative coefficient on *GDP per capita* in the probit model may indicate that (1) states with greater capacities are more efficient at deterring insurgents or (2) prospective rebels are less likely to challenge the state in the presence of higher opportunity costs or (3) both.

In contrast, the model estimated using SPPO provides some insights in regard to these processes. *GDP per capita* is positive and statistically significant in *Rebels’* status quo equation. This indicates that as *GDP per capita* increases, potential rebel groups place more value on the status quo and are less likely to challenge the government. In contrast, *GDP per capita* is statistically insignificant in *Government’s* civil war equation. This suggests that government strength does not systematically deter insurgencies.

Figure 4 presents the predicted probabilities of several outcomes across varying values in *GDP per capita*. At first glance, it appears that the traditional probit and SPPO estimators provide similar results, as evidenced by the predicted probabilities of war for probit (solid line) and SPPO (long dashed line). Both models show a declining probability of war as *GDP per capita* increases. Upon closer examination, however, the SPPO model reveals additional information about the effect of *GDP per capita* on the *Rebel–Government* interaction, by also allowing to construct the probability of *Rebels’* challenging (dashed line) and *Government’s* acquiescing (dotted line). We see, for example, that governments are almost twice as likely to acquiesce to a challenge than to fight it.

The results are theoretically meaningful, as they distinguish between two distinct types of non-conflict events: (a) those associated with support for the legitimacy of the state, and (b) those associated with challenges of this legitimacy, yet do not turn violent due to appeasement from the government. The results are methodologically meaningful because they highlight and provide a

possible solution to a selection problem for civil wars and other outcomes that result from strategic processes, yet lack data structured in a manner that permits the use of the existing strategic models (e.g., Signorino 2003).²⁵ The results are substantively meaningful because they demonstrate that civil wars are just one outcome generated from the strategic interaction between opposition groups and governments; ignoring this underestimates the degree of domestic conflict in world politics. In addition, scholars and practitioners may want to know why some challenges are resolved through non-violent means, while others escalate; separation of the two event types, made possible by the method proposed here, is a first step in conducting research in this direction.

It is worth noting a few other differences between the two sets of results. First, Table 4 reveals several differences in the in-sample fit statistics. The distribution-free Clarke test and the Vuong tests disagree on which model has a better specification: the former identifies the probit, while the latter provides support for SPPO. Both the probit and SPPO models correctly identify 98.3% of observations, though both have difficulty correctly classifying the observations when $Y = 1$.²⁶ The models' difficulty of predicting war becomes less surprising when put in the context of rareness of civil war observations in the data—106 civil wars among the 6237 country-year observations. Lowering the threshold for classifying observations as war to $\Pr(\text{war}) > 0.1$ rather than 0.5 increases the number of correctly predicted wars for both models, though the probit model identifies more false positives.²⁷

Second, there is a high correlation ($r = 0.84$) between the two models' prediction of *war* outcome for the individual observations, and so is the correlation between the individual observations' predicted probabilities of *status quo* from SPPO and $\neg\text{war}$ from the probit model ($r = 0.81$). More interestingly, the predicted outcomes from the two models also diverge: the correlation between the *acquiescence* outcomes from SPPO and *war* from probit is $r = 0.57$. This result again highlights the theoretical value-added of SPPO over probit: while probit treats all $\neg\text{war}$ outcomes as peaceful “non-events,” SPPO identifies two types of “non-events,” one of which, *acquiescence*, is generally identified as *war* by the probit model.

6 Conclusion

Strategic interactions are at the heart of many political processes. This observation has drawn scholarly attention, more recently in the context of selecting the most appropriate estimation technique and identifying the repercussions of ignoring the strategic nature of interactions (Signorino 2003; Signorino and Yilmaz 2003). This article builds on this line of scholarship by developing an estimator—SPPO—that accounts for strategic interactions when the structure of existing data does not permit the implementation of the traditional strategic models. The results highlight the importance of explicitly matching the underlying theoretical process to an appropriate estimation technique. Monte Carlo simulations demonstrate that neither traditional nor split-sample binary choice estimators capture strategic processes.

SPPO has many potential applications for the cases of theorized strategic interactions but limited data on actor choices. In addition to civil war, SPPO could be applied to crisis escalation in interstate conflict, intra-party discipline in the American and comparative contexts, coup d'états and coup-proofing, or regulatory/policing scenarios, such as the EPA example provided earlier. SPPO offers a statistical technique that allows for modeling rather than ignoring the strategic processes in the presence of limited data on actor choices.

²⁵Signorino and Yilmaz (2003) and Kenkel and Signorino (2011) provide an alternative method of estimation based on a Taylor series expansion.

²⁶Observations where $\Pr(\text{war}) > 0.5$ are classified as war predictions for the probit model. Observations where the $\Pr(\text{war}) > \Pr(\text{status quo})$ and $\Pr(\text{war}) > \Pr(\text{acquiesce})$ are classified as war predictions for the SPPO model.

²⁷Using the lower threshold, probit correctly identifies 14.2% of wars with a false-positive rate of 1.4%, while SPPO correctly identifies 13.2% of war with a false-positive rate of 1.0%.

Appendix: Stata Code

Given a theory that expects a strategic process such as that depicted in Fig. 2, SPPO can be estimated using the procedure outlined below. Consistent with Fig. 2, I use the argument style i_j to represent player i and outcome j . Once the program is defined, DV represents the binary outcome variable, IVA1 represents the regressors making up the utility for Player A's Y1 outcome, IVA3 represents the regressors making up the utility for Player A's Y3 outcome, and IVA4 represents the regressors making up the utility for Player A's Y4 outcome. IVB3 and IVB4 represent the regressors making up the utility for Player B's Y3 and Y4 outcomes, respectively.

```
#delimit;
program define sppo_lf, rclass;
  args lnf A_Y1 A_Y3 A_Y4 B_Y3 B_Y4;
  tempvar pun chal;
  quietly gen double `pun' = normal((`B_Y4'-`B_Y3)/sqrt(2));
  quietly gen double `chal' = normal((`pun'*A_Y4' + (1-`pun')*A_Y3'-A_Y1')/
    (sqrt(`pun'^ 2+ (1-`pun')^ 2+1)));
  quietly replace `lnf' = ln((1-`chal')+(`chal')*(1-`pun')) if $ML_y1==0;
  quietly replace `lnf' = ln((`chal')*(`pun')) if $ML_y1==1;
end;
ml model lf sppo_lf (DV = IVA1) (= IVA3) (= IVA4) (= IVB3) (= IVB4);
ml maximize, diff;
```

References

- Acemoglu, D., and J. A. Robinson. 2001. A theory of political transitions. *American Economic Review* 91(4):938–63.
- . 2006. *Economic origins of dictatorship and democracy*. Cambridge, UK: Cambridge University Press.
- Achen, C. H. 2005. Let's put garbage-can regressions and garbage-can probits where they belong. *Conflict Management and Peace Science* 22(4):327–39.
- Adams, J., and Z. Somer-Topcu. 2009. Policy adjustment by parties in response to rival parties' policy shifts: Spatial theory and the dynamics of party competition in twenty-five post-war democracies. *British Journal of Political Science* 39(4):825–46.
- Anderson, B. 1983. *Imagined communities*. London: Verso.
- Bagozzi, B. E., D. W. Hill, W. H. Moore, and B. Mukherjee. 2014. Modeling two types of peace: The zero-inflated ordered probit model in conflict research. *Journal of Conflict Resolution*. Forthcoming.
- Bas, M. A. 2012a. Democratic inefficiency? Regime type and suboptimal choices in international politics. *Journal of Conflict Resolution* 56(5):799–824.
- . 2012b. Measuring uncertainty in international relations: Heteroskedastic strategic models. *Conflict Management and Peace Science* 29(5):490–520.
- Bas, M. A., C. S. Signorino, and R. W. Walker. 2008. Statistical backwards induction: A simple method for estimating recursive strategic models. *Political Analysis* 16(1):21–40.
- Boehmke, F. J., and J. W. Patty. 2007. The selection of policies for ballot initiatives: What voters can learn from legislative inaction. *Economics and Politics* 19(1):97–121.
- Cameron, C. M. 2000. *Veto bargaining: Presidents and the politics of negative power*. Cambridge, UK: Cambridge University Press.
- Carrubba, C. J., A. Yuen, and C. Zorn. 2007. In defense of comparative statics: Specifying empirical tests of models of strategic interaction. *Political Analysis* 15(4):465–82.
- Carson, J. L. 2003. Strategic interaction and candidate competition in U.S. House elections: Empirical applications of probit and strategic probit models. *Political Analysis* 11(4):368–80.
- . 2005. Strategy, selection, and candidate competition in U.S. House and Senate elections. *Journal of Politics* 67(1):1–28.
- Carter, D. 2010. The strategy of territorial conflict. *American Journal of Political Science* 54(4):969–87.
- Chiba, D. 2011. The strength of cease-fire agreements and the duration of postwar peace. Paper presented at the Annual Meeting of the Peace Science Society (International), Los Angeles, CA.
- Chyzh, O. 2014. Can you trust a dictator? A strategic model of authoritarian regimes' signing and compliance with international treaties. *Conflict Management and Peace Science* 31(1):3–27.
- Clarke, K. A. 2001. Testing nonnested models of international relations: Reevaluating realism. *American Journal of Political Science* 45(3):724–44.
- . 2003. Nonparametric model discrimination in international relations. *Journal of Conflict Resolution* 47(1):72–93.
- Clarke, K. A., and C. S. Signorino. 2010. Discriminating methods: Tests for non-nested discrete choice models. *Political Studies* 58:368–88.
- Collier, P., and A. Hoeffler. 2004. Greed and grievance in civil war. *Oxford Economic Papers* 56(4):663–95.

- Cunningham, D. E., K. S. Gleditsch, and I. Salehyan. 2009. It takes two: A dyadic analysis of civil war duration and outcome. *Journal of Conflict Resolution* 53(4):570–97.
- Cunningham, K. G. 2011. Divide and conquer or divide and concede: How do states respond to internally divided separatists? *American Political Science Review* 105(2):275–97.
- . 2013. Actor fragmentation and civil war bargaining: How internal divisions generate civil conflict. *American Journal of Political Science* 57(3):659–72.
- Dion, D., F. J. Boehmke, W. MacMillan, and C. R. Shipan. 2011. *War of words: The filibuster as a war of attrition*. Working paper.
- Favretto, K. 2009. Should peacemakers take sides? Major power mediation, coercion, and bias. *American Political Science Review* 103(2):248–63.
- Fearon, J. D. 2005. Primary commodity exports and civil war. *Journal of Conflict Resolution* 49(4):483–507.
- Fearon, J. D., and D. D. Laitin. 2003. Ethnicity, insurgency, and civil war. *American Political Science Review* 97(1):75–90.
- Filson, D., and S. Werner. 2002. A bargaining model of war and peace: Anticipating the onset, duration, and outcome of war. *American Journal of Political Science* 46(4):819–37.
- Franzese, R. J., Jr., and J. C. Hays. 2008. Interdependence in comparative politics: Substance, theory, empirics, substance. *Comparative Political Studies* 41(4/5):742–80.
- Gent, S. E., and M. Shannon. 2011. Decision control and the pursuit of binding conflict management: Choosing the ties that bind. *Journal of Conflict Resolution* 55(2):710–34.
- Gerber, E. R. 1996. Legislative response to the threat of popular initiatives. *American Journal of Political Science* 40(1):99–128.
- Harris, M. N., and X. Zhao. 2007. A zero-inflated ordered probit model, with an application to modelling tobacco consumption. *Journal of Econometrics* 141(2):1073–99.
- Hays, J. C. 2009. *Globalization and the new politics of embedded liberalism*. Oxford: Oxford University Press.
- Helmke, G. 2010. The origins of institutional crises in Latin America. *American Journal of Political Science* 54(3):737–50.
- Holyoke, T. T. 2009. Interest group competition and coalition formation. *American Journal of Political Science* 53(2):360–75.
- Huntington, S. 1996. *The clash of civilizations and the remaking of the world order*. New York: Simon and Shuster.
- Kalyvas, S., and M. Kocher. 2007. How “free” is free-riding in civil war? Violence, insurgency, and the collective action problem. *World Politics* 59(2):177–216.
- Kenkel, B., and C. S. Signorino. 2011. Data mining for theorists. Paper presented at the Annual Meeting of the Society of Political Methodology.
- König, T., and L. Mäder. 2014. The strategic nature of compliance: An empirical evaluation of law implementation in the central monitoring system of the European union. *American Journal of Political Science* 58(1):246–63.
- Krehbiel, K. 1998. *Pivotal politics*. Chicago: University of Chicago Press.
- Kydd, A. H. 2006. When can mediators build trust? *American Political Science Review* 100(3):449–62.
- Leblang, D. 2003. To devalue or to defend? The political economy of exchange rate policy. *International Studies Quarterly* 47(3):533–60.
- Leeman, L. 2014. Strategy and sample selection: A strategic selection estimator. *Political Analysis* 22(3):374–97.
- Lewis, J. B., and K. A. Schultz. 2003. Revealing preferences: Empirical estimation of a crisis bargaining game with incomplete information. *Political Analysis* 11(4):345–67.
- Matthews, S. A. 1989. Veto threats: Rhetoric in a bargaining game. *Quarterly Journal of Economics* 104(2):347–69.
- McKelvey, R. D., and T. R. Palfrey. 1995. A statistical theory of equilibrium in games. *Games and Economic Behavior* 10:6–38.
- . 1996. A statistical theory of equilibrium in games. *Japanese Economic Review* 47(2):186–209.
- . 1998. Quantal response equilibria for extensive form games. *Experimental Economics* 1(1):9–41.
- McLean, E. V., and T. Whang. 2010. Friends or foes? Major trading partners and the success of economic sanctions. *International Studies Quarterly* 54(2):427–47.
- Morton, R. 1999. *Methods and models*. Cambridge, UK: Cambridge University Press.
- Nieman, M. N. 2015. *Replication data for: Statistical analysis of strategic interaction with unobserved player actions: Introducing a strategic probit with partial observability*. [http://dx.doi.org/10.7910/dvn/28662unf:5:8m7tqrotzpb3khba9vwx==dataverse\[distributor\]v1\[version\]](http://dx.doi.org/10.7910/dvn/28662unf:5:8m7tqrotzpb3khba9vwx==dataverse[distributor]v1[version]).
- Poirier, D. J. 1980. Partial observability in bivariate probit models. *Journal of Econometrics* 12(2):209–17.
- Ray, J. L. 2005. Explaining interstate conflict and war: What should be controlled for? *Conflict Management and Peace Science* 20(2):1–31.
- Schwarz, G. 1978. Estimating the dimension of a model. *Annals of Statistics* 6(2):461–64.
- Signorino, C. S. 1999. Strategic interaction and the statistical analysis of international conflict. *American Political Science Review* 93(2):279–97.
- . 2002. Strategy and selection in international relations. *International Interactions* 28(1):93–115.
- . 2003. Structure and uncertainty in discrete choice models. *Political Analysis* 11(4):316–44.
- . 2007. On formal theory and statistical methods: A response to Carrubba, Yuen, and Zorn. *Political Analysis* 15(4):483–501.
- Signorino, C. S., and A. Tarar. 2006. A unified theory and test of extended immediate deterrence. *American Journal of Political Science* 50(3):585–605.

- Signorino, C. S., and K. Yilmaz. 2003. Strategic misspecification in regression models. *American Journal of Political Science* 47(3):551–66.
- Singer, J. D. 1972. The “Correlates of War” project: Interim report and rationale. *World Politics* 24(2):243–70.
- Slantchev, B. L. 2005. Military coercion in interstate crises. *American Political Science Review* 99(4):533–47.
- Smith, A. 2003. Election timing in majoritarian parliaments. *British Journal of Political Science* 33(3):397–418.
- Thies, C. G., O. Chyzh, and M. D. Nieman. 2015. The spatial dimensions of state fiscal capacity: The mechanisms of international influence on domestic extractive efforts. *Political Science Research and Methods*. Forthcoming.
- Thyne, C. L., and R. D. Schroeder. 2012. Social constraints and civil war: Bridging the gap with criminological theory. *Journal of Politics* 1(1):1–13.
- von Stein, J. 2005. Treaties constrain or screen? Selection bias and treaty compliance. *American Political Science Review* 99(4):611–22.
- Vuong, Q. H. 1989. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* 57(2):307–33.
- Werner, S. 1999. The precarious nature of peace: Resolving the issues, enforcing the settlement, and renegotiating the terms. *American Journal of Political Science* 43(3):912–34.
- Williams, L. K., and G. D. Whitten. 2014. Don’t stand so close to me: Spatial contagion effects and party competition. *American Journal of Political Science*. Forthcoming.
- Xiang, J. 2010. Relevance as a latent variable in dyadic analysis of conflict. *Journal of Politics* 72(2):484–98.