Estimating Signaling Games in IR
Problems and Solutions

Casey Crisman-Cox\textsuperscript{1} \hspace{1cm} Michael Gibilisco\textsuperscript{2}

\textsuperscript{1}Washington University in St. Louis
\textsuperscript{2}Caltech

March 9, 2018
Crisis signaling in International Relations

- Crisis bargaining with incomplete information is *the* workhorse formal model in international relations
- Applied modelers and EITM scholars frequently use this model since Lewis and Schultz (2003)
- Used to study audience costs, sanction threats, dispute escalation, and more
- We aid their undertaking
Struggles faced

How to handle multiplicity in signaling games?

- Bas et al. 2014; Kurizaki et al. 2015; McLean et al. 2015; Wand 2006; Whang 2010, 2016.
- Apply Signorino (1999)—a game with a unique equilibrium.

Is it worth the computational expense?

Our Contribution

We solve both these questions and aid the application of these models to data by

1. Adapting three estimators from the dynamic games literature and study the crisis-signaling games in IR.
2. Analyzing their performance in a variety of experimental settings.
4. Offering recommendations (and R code) for future work.
Results

1. In all settings, our proposed solutions perform better (and faster) than the current best-practices.
   - Even if there is a unique equilibrium in the signaling game from the DGP.

2. Pseudo-likelihood methods perform the best.

3. In application, current best practices return misleading estimates and underestimate the role of audience costs in economic sanction disputes.
   - Our solutions remedy both problems.
The Model: Lewis and Schultz 2003

\[
\begin{align*}
\text{Challenge} & \quad \text{Not challenge} \\
B & \quad \text{CD} \quad (V_A, C_B) \\
& \quad \text{Resist} \\
& \quad \text{Not resist} \\
& \quad \text{SF} \quad (\tilde{W}_A + \varepsilon_A, \tilde{W}_B + \varepsilon_B) \\
& \quad \text{Not fight} \\
& \quad \text{BD} \quad (\bar{a} + \varepsilon_a, V_B)
\end{align*}
\]
Each side has some private information:

- $A$ has information related to the costs/benefits of war ($\varepsilon_A$) and backing down ($\varepsilon_a$)
- $B$ has information about the costs/benefits of war ($\varepsilon_B$)
- For exposition, all information is i.i.d. standard normal
Perfect Bayesian Equilibria

Strategies: $p = (p_C, p_R, p_F)$

Parameters: $\theta = \left( \bar{a}, C_B, (S_i, V_i, \bar{W}_i)_{i=A,B} \right)$

Result 1 (Jo, 2011)

An equilibrium $\tilde{p}$ exists. There exist smooth functions such that $\tilde{p}$ is an equilibrium if and only if the following hold.

$$
\tilde{p}_C = g(\tilde{p}_R; \theta) \quad \tilde{p}_F = h(\tilde{p}_R; \theta) \quad \tilde{p}_R = f(\tilde{p}_F; \theta).
$$

The functions are best responses for $A (g,h)$ and $B (f)$. 
Overview

Data consists of $D$ games, or dyads. From each dyad $d$, we observe $T$ observations from equilibrium $p^*_d$ and covariates $x_d$.

That is, $p^*_d$ solves equations in Result 1 parameterized by:

\[
\theta(x_d, \beta) = \begin{bmatrix}
S_dA \\
S dB \\
V_dA \\
C dB \\
\bar{W}_dA \\
\bar{W} dB \\
\bar{a}d \\
V dB
\end{bmatrix}
= \begin{bmatrix}
x_dS_A \cdot \beta S_A \\
0 \\
x_dV_A \cdot \beta V_A \\
x_dC_B \cdot \beta C_B \\
x_d\bar{W}_A \cdot \beta \bar{W}_A \\
x_d\bar{W}_B \cdot \beta \bar{W}_B \\
x_d\bar{a} \cdot \beta \bar{a} \\
x_dV_B \cdot \beta V_B
\end{bmatrix}.
\]

Our goal is to estimate $\beta$. 
Traditional ML

Best practices since Signorino (1999), following Rust (1987).

Standard multinomial log-likelihood:

\[
L(\beta | Y) = \sum_{d=1}^{D} \sum_{t=1}^{T} \log \Pr(y_{dt} | p_d^*(\beta))
\]

where

\[
\Pr[y_{dt} | p_d] = \begin{cases} 
(1 - p_{dC}) & \text{if } y_{dt} = SQ \\
p_{dC}(1 - p_{dR}) & \text{if } y_{dt} = CD \\
p_{dC}p_{dR}(1 - p_{dF}) & \text{if } y_{dt} = BD \\
p_{dC}p_{dR}p_{dF} & \text{if } y_{dt} = SF.
\end{cases}
\]

Given \( \beta \), compute an equilibrium \( p_d^*(\beta) \) (may not be unique).
What do equilibria look like?

Under some conditions we can produce a unique equilibria

\[ \theta(x_d, \beta) = \]
\[ S_{di}, C_B = 0 \]
\[ V_{dA} = 1 \]
\[ \bar{W}_{dA} = -1.8 \]
\[ V_{dB} = 1 \]
\[ \bar{W}_{dB} = -2.45 + 0.1x_d \]
\[ \bar{a}_d = -1.2 \]

Note: \( x_d \sim U[0, 1] \).
What do equilibria look like?

But, with only slight tweaks, we find a multiplicity

\[ \theta(x_d, \beta) = \]
\[ S_{di}, C_B = 0 \]
\[ V_{dA} = 1 \]
\[ \tilde{W}_{dA} = -1.9 \]
\[ V_{dB} = 1 \]
\[ \tilde{W}_{dB} = -2.9 + 0.1x_d \]
\[ \bar{a}_d = -1.2 \]

Note: \( x_d \sim U[0, 1] \).
The Problem

Discrete choice MLE from our methods courses:
The Problem

Discrete choice MLE from our methods courses:

\[
\text{Input } \beta
\]
The Problem

Discrete choice MLE from our methods courses:

\[ \Phi(x_d \cdot \beta) \quad \text{Input} \quad \beta \quad \rightarrow \quad \text{Compute} \quad \Pr[y_d | x_d, \beta] \]
The Problem

Discrete choice MLE from our methods courses:

\[
\text{Input } \beta \xrightarrow{\Phi(x_d \cdot \beta)} \text{Compute } \Pr[y_d \mid x_d, \beta] \xrightarrow{} \text{Output } L_d
\]
The Problem

Discrete choice MLE from our methods courses:

\[
\text{Input } \beta \xrightarrow{\Phi(x_d \cdot \beta)} \text{Compute } \Pr[y_d | x_d, \beta] \xrightarrow{} \text{Output } L_d
\]

MLE with multiple equilibria (choice probabilities):
The Problem

Discrete choice MLE from our methods courses:

Input $\beta$ $\Phi(x_d \cdot \beta)$ Compute $Pr[y_d | x_d, \beta]$ Output $L_d$

MLE with multiple equilibria (choice probabilities):

Input $\beta$
The Problem

Discrete choice MLE from our methods courses:

\[
\text{Input } \beta \xrightarrow{\Phi(x_d \cdot \beta)} \text{Compute } \Pr[y_d | x_d, \beta] \xrightarrow{} \text{Output } L_d
\]

MLE with multiple equilibria (choice probabilities):

\[
\text{Input } \beta \xrightarrow{\text{Eq. 1}} \text{Compute } \Pr[y_d | x_d, \beta]
\]
\[
\text{Input } \beta \xrightarrow{\text{Eq. 2}} \text{Compute } \Pr[y_d | x_d, \beta]
\]
\[
\text{Input } \beta \xrightarrow{\text{Eq. 3}} \text{Compute } \Pr[y_d | x_d, \beta]
\]
The Problem

Discrete choice MLE from our methods courses:

\[
\text{Input } \beta \xrightarrow{\Phi(x_d \cdot \beta)} \text{Compute } \Pr[y_d | x_d, \beta] \xrightarrow{\text{Output }} L_d
\]

MLE with multiple equilibria (choice probabilities):

\[
\text{Input } \beta \xrightarrow{\text{Eq. 1}} \text{Compute } \Pr[y_d | x_d, \beta] \xrightarrow{\text{Output }} L_{d,1}
\]
\[
\text{Input } \beta \xrightarrow{\text{Eq. 2}} \text{Compute } \Pr[y_d | x_d, \beta] \xrightarrow{\text{Output }} L_{d,2}
\]
\[
\text{Input } \beta \xrightarrow{\text{Eq. 3}} \text{Compute } \Pr[y_d | x_d, \beta] \xrightarrow{\text{Output }} L_{d,3}
\]
Questions

1. How to select equilibria to evaluate the likelihood?
2. How to do (1) without computing **all** equilibria?
3. Can we eliminate equilibria?
Proposed solutions

Proposal 1: Hope that there is a unique eq. in game from DGP.
  - This only works with almost exact starting values.
  - If not, then likelihood will be evaluated at $\beta$ with multiple eq.

Proposal 2: Try to use refinements.
  - We show almost all equilibria are regular—a strong refinement.
  - Ad hoc criteria would need to be used, while still computing all equilibria.
  - Computationally exhausting on two fronts: Finding the refined eq. and working with discontinuous likelihood.
Proposed solutions

Proposal 1: Hope that there is a unique eq. in game from DGP.

- This only works with almost exact starting values.
- If not, then likelihood will be evaluated at $\beta$ with multiple eq.

Proposal 2: Try to use refinements.

- We show almost all equilibria are regular—a strong refinement.
- Ad hoc criteria would need to be used, while still computing all equilibria.
- Computationally exhausting on two fronts: Finding the refined eq. and working with discontinuous likelihood.

Solution: We adapt approaches from industrial organization and marketing literature.

- Have not been applied to games with signaling incentives.
- Fast, good, and computationally cheap
Pseudo-Likelihood (PL)

A two-step estimator based on Hotz and Miller (1993):

1. Consistently estimate $p_{dR}^*$ and $p_{dF}^*$.
2. Given $\beta$, compute players’ best response to $\hat{p}_{dR}$ and $\hat{p}_{dF}$, then evaluate likelihood with best responses.
3. Estimator maximizes this pseudo-likelihood.

First stage estimates use random forests with predictors $x_d$.
- If $T$ is large, we could use frequency estimator.

This provides a feasible approximation of the true ML.
Nested Pseudo-Likelihood (NPL)

Analyzed in Aguirregabiria and Mira (2007).

Begin with the PL estimates $\hat{\beta}_0$ and first-stage probabilities:

$$\hat{p}_{R,0} = (\hat{p}_{1R}, \ldots, \hat{p}_{DR})$$

$$\hat{p}_{F,0} = (\hat{p}_{1F}, \ldots, \hat{p}_{DF})$$

For the $k$th iteration, define

$$\hat{p}_{dF,k} = h(\hat{p}_{dR,k-1}; x_d, \beta_{k-1}), \quad d = 1, \ldots, D$$

$$\hat{p}_{dR,k} = f(\hat{p}_{dF,k-1}; x_d, \beta_{k-1}), \quad d = 1, \ldots, D$$

$$\hat{\beta}_{NPL}^k = \arg\max_{\beta} PL(\beta \mid \hat{p}_{R,k}, \hat{p}_{F,k}, Y).$$

Repeated until convergence in $(\beta_k, \hat{p}_{R,k}, \hat{p}_{F,k})$. 
Constrained MLE (CMLE)

Proposed by Su and Judd (2012).

Treat the equilibrium quantity $\mathbf{p}^*_R$ as a parameter to be estimated in a full-information setup.

$$\arg\max_{\beta, \mathbf{p}_R} L(\beta, \mathbf{p}_R | \mathbf{Y})$$

s.t. $f \circ h(p_{dR}; x_d, \beta) = p_{dR}$

Maximize likelihood as before, but enforce equilibrium play.

Equivalent to the true ML (but avoids eq. computation and global optimization).

- Incidental parameters $\hat{\mathbf{p}}_R$ – consistent in $T$
Monte Carlo Results: Multiple Equilibria I

RMSE in Signaling Estimators

<table>
<thead>
<tr>
<th>Games</th>
<th>Estimators</th>
<th>CMLE</th>
<th>PL</th>
<th>tMLE</th>
<th>NPL</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>Logged RMSE</td>
<td>2.5</td>
<td>5.0</td>
<td>7.5</td>
<td>10.0</td>
</tr>
<tr>
<td>50</td>
<td>Logged RMSE</td>
<td>2.5</td>
<td>5.0</td>
<td>7.5</td>
<td>10.0</td>
</tr>
<tr>
<td>100</td>
<td>Logged RMSE</td>
<td>2.5</td>
<td>5.0</td>
<td>7.5</td>
<td>10.0</td>
</tr>
<tr>
<td>200</td>
<td>Logged RMSE</td>
<td>2.5</td>
<td>5.0</td>
<td>7.5</td>
<td>10.0</td>
</tr>
</tbody>
</table>
Monte Carlo Results: Unique Equilibrium I

RMSE in Signaling Estimators

<table>
<thead>
<tr>
<th>Estimator</th>
<th>CMLE</th>
<th>PL</th>
<th>tMLE</th>
<th>NPL</th>
</tr>
</thead>
</table>

Within-game observations
Logged RMSE

- 25 Games
- 50 Games
- 100 Games
- 200 Games

Estimator: CMLE, PL, tMLE, NPL
Monte Carlo: recap

1. tML performs the worst in almost all settings.
2. PL and NPL generally perform the best.
Economic Sanctions

- In the spirit of Whang et al. 2015.
- Threat and Imposition of Sanctions data.
- A game $d$ is a **politically relevant** dyad decade, $D = 418$.
- We observe $T = 120$ observations in each decade.
## Specification

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Utilities</th>
<th>Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SQ</strong></td>
<td>$S_A$</td>
<td>Econ. Dep$_A$, Dem$_A$, Contig.</td>
</tr>
<tr>
<td></td>
<td>$S_B$</td>
<td>Fixed to 0 for identification</td>
</tr>
<tr>
<td><strong>CD</strong></td>
<td>$V_A$</td>
<td>Const., Expected Costs$_A$, Ally</td>
</tr>
<tr>
<td></td>
<td>$C_B$</td>
<td>Const., Econ. Dep$_B$, Expected Costs$_B$, Contig, Ally</td>
</tr>
<tr>
<td><strong>SF</strong></td>
<td>$\bar{W}_A$</td>
<td>Const., Econ. Dep$_A$, Dem$_A$, Cap. Ratio</td>
</tr>
<tr>
<td></td>
<td>$\bar{W}_B$</td>
<td>Const., Dem$_B$, Cap. Ratio</td>
</tr>
<tr>
<td><strong>BD</strong></td>
<td>$\bar{a}$</td>
<td>Const., Dem$_A$</td>
</tr>
<tr>
<td></td>
<td>$V_B$</td>
<td>Fixed to 0 for identification</td>
</tr>
</tbody>
</table>
## Results: Point Estimates (abridged)

<table>
<thead>
<tr>
<th></th>
<th>tML Model 1</th>
<th>Pseudo-Likelihood Model 2</th>
<th>Nested-PL Model 3</th>
<th>CMLE Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_B$: Const.</td>
<td>0.81</td>
<td>−0.89*</td>
<td>−2.14*</td>
<td>−4.53*</td>
</tr>
<tr>
<td></td>
<td>(0.74)</td>
<td>(0.31)</td>
<td>(0.80)</td>
<td>(2.24)</td>
</tr>
<tr>
<td>$C_B$: Econ. Dep$_B$</td>
<td>−0.23</td>
<td>1.34*</td>
<td>2.34*</td>
<td>2.83*</td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.59)</td>
<td>(1.11)</td>
<td>(0.63)</td>
</tr>
<tr>
<td>$C_B$: Costs$_B$</td>
<td>−0.08*</td>
<td>0.11</td>
<td>0.12</td>
<td>0.19*</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>$C_B$: Contiguity</td>
<td>−0.24*</td>
<td>0.09*</td>
<td>0.12*</td>
<td>0.10*</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.04)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>$C_B$: Alliance</td>
<td>0.10</td>
<td>0.03</td>
<td>−0.03</td>
<td>−0.02</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.13)</td>
<td>(0.12)</td>
<td>(0.11)</td>
</tr>
<tr>
<td>$\bar{\alpha}$: Const.</td>
<td>−0.57</td>
<td>−2.63*</td>
<td>−2.64*</td>
<td>−2.71*</td>
</tr>
<tr>
<td></td>
<td>(0.76)</td>
<td>(0.09)</td>
<td>(0.13)</td>
<td>(0.10)</td>
</tr>
<tr>
<td>$\bar{\alpha}$: Dem$_A$</td>
<td>0.00</td>
<td>−0.02</td>
<td>0.02</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Log $L$</td>
<td>−4095.30</td>
<td>−3964.53</td>
<td>−3932.45</td>
<td>−3927.91</td>
</tr>
<tr>
<td>$D \times T$</td>
<td>418 × 120</td>
<td>418 × 120</td>
<td>418 × 120</td>
<td>418 × 120</td>
</tr>
</tbody>
</table>

**Notes:** * $p < 0.05$

Standard Errors in Parenthesis
Results: Substantive Implications

1. Models accommodating multiple equilibria generally agree. MLE returns misleading estimates.

2. Models 2–4 all have observations that are “competitive” from Lewis and Schultz (2005). tML has no observation that is competitive.

3. Audience costs: $\bar{a}_d = \beta_{\bar{a},1} + \beta_{\bar{a},2} \cdot \text{Dem}_A$
   - In Models 2–4, we reject the null that $\bar{a}_d \geq 0$ at $p < 0.05/D$ for all dyads $d$.
   - With the tMLE, we cannot reject the null that $\bar{a}_d \geq 0$ at $p < 0.1$ in any dyad $d$. 
Effects of Audience Costs on Threat Initiation

Comparative Statics: USA—CHN, 1990

Prob. of Onset, $p_C$

USA Audience Costs, $\bar{a}$
Effects of Audience Costs on Outcomes

Comparative Statics: USA–CHN, 1990

Probability of Observing Sanctions

USA Audience Costs, $\bar{a}$
More General Results

In 97% of dyads, larger (more negative) audience costs increase likelihood of threat initiation.

- Median effect: 0.03
- 90%-CI: [0.003, 0.122]
- 95%-CI: [0.0004, 0.17]

Linear interpolation suggests large effects:

- If we double $\bar{a}_d$, then the likelihood of threat initiation increases by 8 percentage points.
Suggestions for Scholars

1. Estimate game with PL estimator.

2. Then estimate game with NPL (or CMLE) estimator with starting values from (1).

3. If (2) converges, then report these.

4. If (2) does not converge, then report (1).

Take-away: tML should be avoided for the signaling games in IR – better and easier approaches exist.
Future Work

Where we go from here:

1. Expand everything here to more general error structures
2. Release an R package for the PL and NPL
3. Consider more even more general estimators that allow for changes in eq. selection over time
Thank you!
Start with $B$, and let $p_F$ be the probability that $A$ fights conditional on challenging

- $B$ resists only if

$$C_B < p_F(\bar{W}_B + \varepsilon_B) + (1 - p_F)V_B$$
Start with $B$, and let $p_F$ be the probability that $A$ fights conditional on challenging.

- $B$ resists only if

\[
C_B < p_F(\bar{W}_B + \varepsilon_B) + (1 - p_F)V_B
\]

\[
p_R = \Phi\left(\frac{p_F\bar{W}_B + (1 - p_F)V_B - C_B}{p_F}\right) := f(p_F; \theta)
\]
Strategies

A challenges when

\[ S_A < p_R \max\{\bar{a} + \varepsilon_a, \bar{W}_A + \varepsilon_A\} + (1 - p_R)V_A \]
Strategies

A challenges when

\[
S_A < p_R \max\{\bar{a} + \varepsilon_a, \bar{W}_A + \varepsilon_A\} + (1 - p_R)V_A
\]

\[
c^* < \max\{\bar{a} + \varepsilon_a, \bar{W}_A + \varepsilon_A\}
\]
A challenges when

\[ S_A < p_R \max\{\bar{a} + \varepsilon_a, \bar{W}_A + \varepsilon_A\} + (1 - p_R)V_A \]
\[ c^* < \max\{\bar{a} + \varepsilon_a, \bar{W}_A + \varepsilon_A\} \]
\[ p_C = 1 - \Pr (\varepsilon_A < c^* - \bar{W}_A\ \text{AND} \ \varepsilon_a < c^* - \bar{a}) \]
A challenges when

\[ S_A < p_R \max\{\bar{a} + \varepsilon_a, \bar{W}_A + \varepsilon_A\} + (1 - p_R)V_A \]
\[ c^* < \max\{\bar{a} + \varepsilon_a, \bar{W}_A + \varepsilon_A\} \]
\[ p_C = 1 - \Pr(\varepsilon_A < c^* - \bar{W}_A \text{ AND } \varepsilon_a < c^* - \bar{a}) \]
\[ = 1 - \Phi(c^* - \bar{W}_A)\Phi(c^* - \bar{a}) := g(p_R; \theta) \]
Recall that $p_F$ was the probability of $A$ fighting given that $A$ challenged

$$p_F = \Pr(\bar{W}_A + \varepsilon_A > \bar{a} + \varepsilon_a | \max\{\bar{a} + \varepsilon_a, \bar{W}_A + \varepsilon_A\} > c^*)$$
Recall that $p_F$ was the probability of $A$ fighting given that $A$ challenged

$$p_F = \Pr(\bar{W}_A + \varepsilon_A > \bar{a} + \varepsilon_a \mid \max\{\bar{a} + \varepsilon_a, \bar{W}_A + \varepsilon_A\} > c^*)$$

$$= \Pr(\varepsilon_a - \varepsilon_A < \bar{W}_A - \bar{a} \text{ AND } \varepsilon_A < \bar{W}_A - c^*)/p_C$$
Recall that $p_F$ was the probability of $A$ fighting given that $A$ challenged

$$p_F = \Pr(\bar{W}_A + \epsilon_A > \bar{a} + \epsilon_a | \max\{\bar{a} + \epsilon_a, \bar{W}_A + \epsilon_A\} > c^*)$$

$$= \Pr(\epsilon_a - \epsilon_A < \bar{W}_A - \bar{a} \text{ AND } \epsilon_A < \bar{W}_A - c^*)/p_C$$

$$= \Phi_2\left(\frac{\bar{W}_A - \bar{a}}{\sqrt{2}}, \bar{W}_A - c^*, \frac{1}{\sqrt{2}}\right) / g(p_R; \theta) := h(p_R; \theta)$$
The Problems of Eq. Selection

![Graph showing log-likelihood for 1 Dyad and 10 Dyads.](image)

- For 1 Dyad:
  - Log-likelihood ranges from -200 to -600.
  - The log-likelihood is relatively flat.

- For 10 Dyads:
  - Log-likelihood ranges from -2000 to -6000.
  - The log-likelihood shows a more pronounced variation.

The graph depicts the behavior of the log-likelihood with respect to \( \hat{\beta}_{WB}^1 \) for different numbers of dyads.
## Specification

<table>
<thead>
<tr>
<th>Variable</th>
<th>Utilities</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed to 0</td>
<td>$S_B, V_B$</td>
<td>Identification restriction</td>
<td>–</td>
</tr>
<tr>
<td>Constant</td>
<td>$V_A, C_B, \bar{W}_i, \bar{a}$</td>
<td>Omitted from $S_A$ for identification</td>
<td>–</td>
</tr>
<tr>
<td>Econ. Dep$_A$</td>
<td>$S_A, \bar{W}_A$</td>
<td>$A$’s economic dependence on $B$</td>
<td>TIES</td>
</tr>
<tr>
<td>Dem$_A$</td>
<td>$S_A, \bar{a}, \bar{W}_A$</td>
<td>$A$’s Polity2 score</td>
<td>Polity IV</td>
</tr>
<tr>
<td>Contiguity</td>
<td>$S_A, C_B$</td>
<td>Contiguity between the states</td>
<td>COW</td>
</tr>
<tr>
<td>Alliance$_B$</td>
<td>$V_A, C_B$</td>
<td>Alliance between the states (0/1)</td>
<td>COW</td>
</tr>
<tr>
<td>Costs$_A$</td>
<td>$V_A$</td>
<td>Anticipated costs to $A$</td>
<td>TIES</td>
</tr>
<tr>
<td>Econ. Dep$_B$</td>
<td>$C_B$</td>
<td>$B$’s economic dependence on $A$</td>
<td>TIES</td>
</tr>
<tr>
<td>Costs$_B$</td>
<td>$C_B$</td>
<td>Anticipated costs to $B$</td>
<td>TIES</td>
</tr>
<tr>
<td>Cap. Ratio</td>
<td>$\bar{W}_i$</td>
<td>(log) ratio of $A$’s capabilities to $B$</td>
<td>COW</td>
</tr>
<tr>
<td>Dem$_B$</td>
<td>$\bar{a}, \bar{W}_B$</td>
<td>$B$’s Polity2 score</td>
<td>Polity IV</td>
</tr>
</tbody>
</table>
Monte Carlo Results: Unique Equilibrium II

Bias in Signaling Estimators

<table>
<thead>
<tr>
<th>Estimator</th>
<th>CMLE</th>
<th>PL</th>
<th>tMLE</th>
<th>NPL</th>
</tr>
</thead>
</table>

Within–game observations

Median Absolute Bias

<table>
<thead>
<tr>
<th>5 25 50 100 200</th>
<th>5 25 50 100 200</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0 0.5 1.0 1.5</td>
<td>0.0 0.5 1.0 1.5</td>
</tr>
</tbody>
</table>

Monte Carlo Results
Monte Carlo Results: Multiple Equilibria II

Bias in Signaling Estimators

<table>
<thead>
<tr>
<th>Median Absolute Bias</th>
<th>25 Games</th>
<th>50 Games</th>
<th>100 Games</th>
<th>200 Games</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimator</td>
<td>CMLE</td>
<td>PL</td>
<td>tMLE</td>
<td>NPL</td>
</tr>
</tbody>
</table>

Within-game observations
Monte Carlo Results: Unstable Equilibria I

RMSE in Signaling Estimators

- Estimator CMLE PL tMLE NPL

Proportion Unstable vs. Log RMSE

Proportion Unstable

0.00 0.25 0.50 0.75 1.00
Monte Carlo Results: Unstable Equilibria II

Convergence Rates in Signaling Estimators

Proportion Converged vs. Proportion Unstable for different estimators:
- CMLE
- PL
- tMLE
- NPL

Estimator legend:
- CMLE
- PL
- tMLE
- NPL
Monte Carlo Results: CMLE Implementations I

Within-game observations
Logged RMSE
Estimator CMLE−AugLag PL CMLE−IPOPT NPL

- **25 Games**
- **50 Games**
- **100 Games**
- **200 Games**

Estimator:
- CMLE−AugLag
- PL
- CMLE−IPOPT
- NPL
Monte Carlo Results: CMLE Implementations II

Convergence Rates in Signaling Estimators

Estimator

- CMLE−AugLag
- PL
- CMLE−IPOPT
- NPL
Implementation: Optimizer

The CMLE is a large scale optimization problem that we solve using the open-source program IPOPT (Wächter and Biegler 2006).

- Interior point optimization does not require computing or satisfying the equilibrium at every iteration.
- Avoids the computational burden of repeatedly computing equilibrium at every iteration (Rust 1987).
- Time trials suggest that IPOPT does as well as proprietary software (KNITRO) and other CMLE methods.
Implementation: AD

Algorithmic differentiation (AD) allows for fast computation of derivatives which are necessary for precise estimation.

AD allows us to write down only the likelihood and constraint.

Repeated application of the chain rule lead to fast/precise computation of sparse 1st and 2nd order derivatives.


