Scenario Sampling in Large Games

Econometric Methods for Social Spillovers and Networks

Universität St.Gallen, October 7 to 11, 2024

Bryan S. Graham ⓡ

University of California - Berkeley

Andrin Pelican

Peer Effects with Binary Actions

There are $t = 1, \ldots, T$ players, each of whom choose binary action $Y_t \in \{0, 1\}$.

Preferences:

$$\upsilon_t(\mathbf{y}; \mathbf{X}, \mathbf{U}, \theta) = y_t \left(X_t' \beta + \delta s(\mathbf{y}_{-t}) - U_t \right)$$

with

$$s(\mathbf{y}_{-t}) = \frac{1}{T-1} \sum_{s \neq t} y_s.$$

Agents prefer to take action (e.g., 'smoke') when more of their peers do so as well ($\delta \geq 0$).

X_t is a vector of observed agent attributes; U_t a random utility term.

Peer Effects with Binary Actions: Equilibria

$\mathbf{Y} = (Y_1, \ldots, Y_T)'$ is a NE in pure strategies if

$$Y_t = \mathbf{1}\left(X_t'\beta + \delta s\left(\mathbf{Y}_{-t}\right) \geq U_t\right)$$

simultaneously for all $t = 1, \ldots, T$.

When $\delta \geq 0$ there exists, for all $\mathbf{U} \in \mathbb{U}^T$, *at least* one NE in pure strategies (Tarski, 1955).

Policy implications of $\delta > 0$ are profound.

A System of Nonlinear Simultaneous Equations

If $U_t | \mathbf{X} \overset{iid}{\sim} \mathcal{N}(0,1)$,then we have a T simultaneous equations 'probit' model (e.g., Heckman, 1978, Maddala, 1983).

Model exhibits both 'simultaneity' and 'completeness' issues.

Simultaneity: Y_t enters the decision rule for player $s \Rightarrow U_t$ and Y_s covary, since Y_s is a component of $s(\mathbf{Y}_{-t})$, $\Rightarrow U_t$ and $s(\mathbf{Y}_{-t})$ will covary as well.

Incompleteness: There may be multiple NE and the model is silent on which one is selected ($\Rightarrow$ distribution of $\mathbf{Y}|\mathbf{X}$ not fully defined).

$\therefore$ a probit fit of Y_t onto X_t and $s(\mathbf{Y}_{-t})$ does not consistently estimate β and/or δ.

This Paper

Studies simulated maximum likelihood (SML) estimation of a class of supermodular games (of which the peer effects game is a particularly simple example).

T players each take M binary actions with TM large (e.g., in the tens of thousands).

Existing approaches to the econometrics of games ill-suited to our setting

We make an explicit equilibrium selection assumption and focus on likelihood evaluation.

General Setting: Preferences

$i = 1, \ldots, N$ indexes a random sample of games ($N \geq 1$).

Each game consists of $t = 1, \ldots, T$ players, each taking $m = 1, \ldots, M$ binary actions, $Y_{itm} \in \{0, 1\}$.

Player t's utility from game outcome $\mathbf{y} \in \{0,1\}^{TM}$ equals

$$\begin{aligned} \upsilon_t(\mathbf{y}; \mathbf{X}, \mathbf{U}, \theta) =& \upsilon(y_t, \mathbf{y}_{-t}; X_t, U_t, \theta) \\ \stackrel{def}{\equiv}& \sum_{m=1}^{M} y_{tm} \left(X'_{tm}\beta_m + s_m \left(y_{t,-m}, \mathbf{y}_{-t} \right)' \delta_m - U_{tm} \right). \end{aligned}$$

Preferences (continued)

Agent utility varies with

1. X_{tm}, agent-by-action specific attributes;

2. $s_m\left(y_{t,-m}, \mathbf{y}_{-t}\right)$ a known function of

   (a) player t's actions other than the m^{th} one and

   (b) player t's peers' actions;

3. U_{tm}, a cost-of-action shock, iid across actions & players with known distribution.

Supermodularity

We assume that $v(y_t, \mathbf{y}_{-t}; X_t, U_t, \theta)$ is a supermodular function of y_t (Topkis, 1998).

The $s_m\left(y_{t,-m}, \mathbf{y}_{-t}\right)$ term allows for player t's marginal utility from action m to depend on what other actions *she* chooses to take.

We assume that for any configuration of peer actions $s_m\left(y_{t,-m}, \mathbf{y}_{-t}\right)' \delta$ is weakly increasing in $y_{t,-m}$ (for $m = 1, \ldots, M$).

This means that the M individual actions are complementary.

Supermodularity (continued)

The $s_m\left(y_{t,-m}, \mathbf{y}_{-t}\right)$ term also captures how the choices of *other* players alter the utility player t attaches to action m.

We assume that for any configuration of own actions $s_m\left(y_{t,-m}, \mathbf{y}_{-t}\right)' \delta$ is weakly increasing in peers' actions $\mathbf{y}_{-t}$ (for $m = 1, \ldots, M$).

These properties ensure that player reaction functions are (weakly) increasing in own and peers' actions.

$\upsilon\left(y_t, \mathbf{y}_{-t}; X_t, U_t, \theta\right)$ has increasing differences in y_t and $\mathbf{y}_{-t}$.

The Game

Simultaneous move, complete information game.

Supermodular in the sense of Milgrom and Roberts (1990).

Can use Tarski's (1955) Theorem to show the existence of a Nash Equilibrium (NE) in pure strategies.

There exist two extremal NE in pure strategies: minimal and maximal (all rationalizable strategy profiles are bounded by these two extremal NE).

We assume that the minimal equilibrium is the one observed in the data.

Example: Network Effects

Single binary action: $Y_t \in \mathbb{Y}_t = \{0, 1\}$, learn random set theory $(Y_t = 1)$ or not $(Y_t = 0)$.

Payoff function is increasing in the number of other adopters:

$$\upsilon_t (\mathbf{y}; \mathbf{x}, \mathbf{u}, \theta) \overset{def}{\equiv} y_t \left(x_t' \beta + \delta s (\mathbf{y}_{-t}) - u_t \right) \tag{1}$$

with $s(\mathbf{y}_{-t}) = \frac{1}{T-1} \sum_{s \neq t} y_s$.

Stylized version of many studies of technology adoption with "network effects" or peer effects with binary actions.

$(N \to \infty$, T fixed, M fixed$)$.

cf., Manski (1993), Brock and Durlauf (2001), Krauth (2006), Soetevent and Koorman (2007).

Example: Strategic Network Formation

Agents decide whether to direct a link to each of the $T-1$ other agents ($Y_{ts}=1$) or not ($Y_{ts}=0$).

In this example there are $M = T-1$ actions per player $\Rightarrow 2^{TM} = 2^{T(T-1)}$ possible pure strategy combinations!

Payoff for directing a link is increasing in the number of "friends in common" (transitivity)

$$\upsilon_t\left(\mathbf{y};\mathbf{x},\mathbf{u},\theta\right) \overset{def}{\equiv} \sum_{s\neq t} y_{ts}\left(x'_{ts}\beta + \delta s\left(y_{t,-m},\mathbf{y}_{-t}\right) - u_{ts}\right). \tag{2}$$

with $s\left(y_{t,-m},\mathbf{y}_{-t}\right) = \sum_{r=1}^{T} y_{tr}y_{rs}$.

(N fixed, $M=T-1$, $T\to\infty$).

This Paper

Since (i) the distribution of $\mathbf{U}$ is parametrically specified and (ii) we make an equilibrium selection assumption, the likelihood function is well-defined.

Unfortunately it is difficult (impossible?) to write down and evaluate when T and/or M is moderately large.

We show how to approximate the log-likelihood function (and its derivatives) by simulation.

Ademaro, Brunhilde and the EDM Concert

Ademaro $(t = 1)$ and Brunhilde $(t = 2)$ are close friends deciding whether to attend, $y_t \in \{0, 1\}$, a local electronic dance music (EDM) concert.

Utility equals

$$\upsilon\left(y_t, y_{-t}; x_t, u_t, \theta\right) = y_t \left(x_t'\beta + \delta y_{-t} - u_t\right). \tag{3}$$

The payoff from attendance depends on peer behavior.

It's more enjoyable to attend the concert with a friend: $\delta > 0$.

Buckets

We can use the utility function and possible peer behaviors to partition the support of U_t in *buckets:*

$$\mathbb{R} = \left(-\infty, X_t'\beta\right] \cup \left(X_t'\beta, X_t'\beta + \delta\right] \cup \left(X_t'\beta + \delta, \infty\right)$$

Bucket boundaries coincide with possible values of the deterministic return to attendance.

Any draw $U_t \sim F_U$ will fall into one, and only one, bucket.

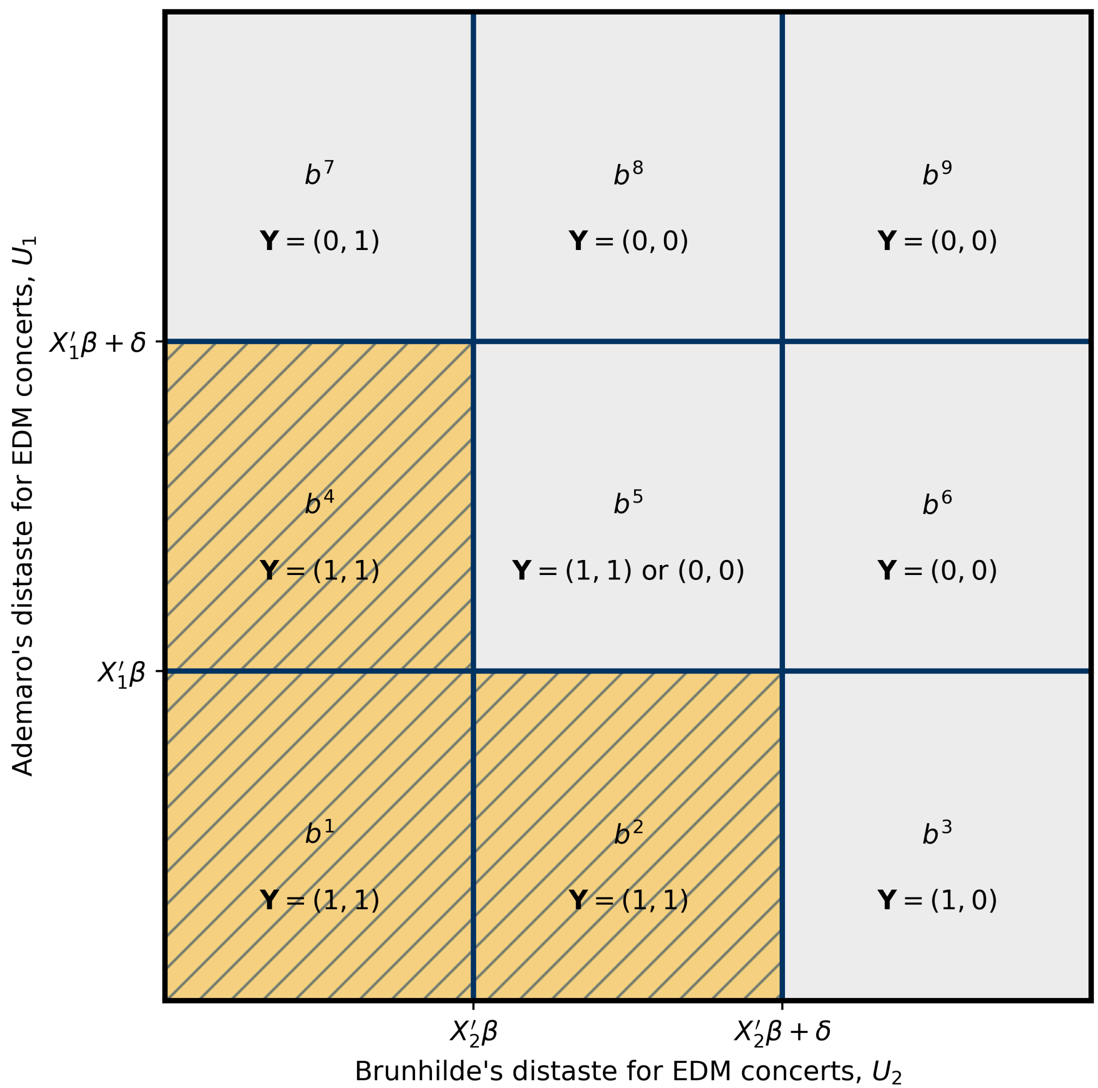
Scenarios

In a similar manner, the support of $\mathbf{U} = (U_1, U_2)'$ can be partitioned into a set of rectangles (e.g., Bresnahan and Reiss, 1991).

$$\mathbb{R}^2 = b^1 \cup b^2 \cup \cdots b^9.$$

We can these rectangles *scenarios*:

$$\begin{aligned} b^2 &= \left(-\infty, X_1'\beta\right] \times \left(X_2'\beta, X_2'\beta + \delta\right] \\ &= \left(\underline{b}_1^2, \bar{b}_1^2\right] \times \left(\underline{b}_2^2, \bar{b}_2^2\right]. \end{aligned}$$



b^7
$\mathbf{Y} = (0, 1)$
b^8
$\mathbf{Y} = (0, 0)$
b^9
$\mathbf{Y} = (0, 0)$
b^4
$\mathbf{Y} = (1, 1)$
b^5
$\mathbf{Y} = (1, 1)$ or $(0, 0)$
b^6
$\mathbf{Y} = (0, 0)$
b^1
$\mathbf{Y} = (1, 1)$
b^2
$\mathbf{Y} = (1, 1)$
b^3
$\mathbf{Y} = (1, 0)$
$X_1'\beta + \delta$
$X_1'\beta$
$X_2'\beta$
$X_2'\beta + \delta$
Ademaro's distaste for EDM concerts, U_1
Brunhilde's distaste for EDM concerts, U_2

Equilibrium Selection

For all $\mathbf{U} \in b^2$ Ademaro will go to the EDM concert "no matter what", while Brunhilde is on the fence and only wants to go if Ademaro does.

The NE in this case is $\mathbf{y} = (1, 1)$; they both go.

For all $\mathbf{U} \in b^5$ Ademaro's (Brunhilde's) utility/cost shock is such that he (she) would prefer to attend the concert if Brunhilde (Ademaro) does as well; but would prefer not to attend if Brunhilde (Ademaro) also decides to stay at home.

Both and $\mathbf{y} = (0, 0)$ and $\mathbf{y} = (1, 1)$ are NE in this case.

We assume the *minimal* equilibrium is always selected.

Likelihood

With an equilibrium selection assumption in hand, the probability of any game outcome $\mathbf{Y} = \mathbf{y} = (y_1, y_2)'$ corresponds to the probability that $\mathbf{U} = (U_1, U_2)'$ falls into one of the scenarios in which $\mathbf{Y} = \mathbf{y}$ is the (selected) NE.

The probability of observing $\mathbf{Y} = (1, 1)'$, for example, corresponds to the ex ante chance that a pair of random utility shocks falls into one of the three cross-hatched scenarios.

For $\mathbf{y} = (1, 1)'$ we have $\mathbb{B}_{\mathbf{y}} = \{b_1, b_2, b_4\}$.

Likelihood (continued)

For $\mathbf{y} = (1,1)'$ we integrate $f_{\mathbf{U}}(\mathbf{u}) = f(u_1)\, f(u_2)$ over the three cross-hatched scenarios.

$$
\begin{aligned}
\Pr(\mathbf{Y} = \mathbf{y} | \mathbf{X}; \theta) &= \sum_{b \in \mathbb{B}_{\mathbf{y}}} \int_{\mathbf{u} \in b} f_{\mathbf{U}}(\mathbf{u})\, \mathrm{d}\mathbf{u} \\
&= \int_{\mathbf{u} \in b^1} f_{\mathbf{U}}(\mathbf{u})\, \mathrm{d}\mathbf{u} + \int_{\mathbf{u} \in b^2} f_{\mathbf{U}}(\mathbf{u})\, \mathrm{d}\mathbf{u} + \int_{\mathbf{u} \in b^4} f_{\mathbf{U}}(\mathbf{u})\, \mathrm{d}\mathbf{u} \\
&= \sum_{j=1,2,4} \left[F\left(\bar{b}_1^j\right) - F\left(\underline{b}_1^j\right) \right] \left[F\left(\bar{b}_2^j\right) - F\left(\underline{b}_2^j\right) \right] \\
&= F\left(X_1'\beta\right) F\left(X_2'\beta\right) + F\left(X_1'\beta\right) \left[F\left(X_2'\beta + \delta\right) - F\left(X_2'\beta\right) \right] \\
&\quad + \left[F\left(X_1'\beta + \delta\right) - F\left(X_1'\beta\right) \right] F\left(X_2'\beta\right). \qquad (4)
\end{aligned}
$$

Likelihood (continued)

With $T = 3$, we would have 4 buckets, for 64 different scenarios.

In general, the number of scenarios is exponential in the number of players/binary actions.

Summing over relevant scenarios to evaluate the likelihood is not feasible in large games.

Simulated Likelihood

The probability that a random draw of $\mathbf{U} = (U_1, U_2)'$ lies in scenario b is simply

$$\zeta(b;\theta) \stackrel{def}{\equiv} \int_{\mathbf{u}\in b} f_{\mathbf{U}}(\mathbf{u})\,\mathrm{d}\mathbf{u}, \tag{5}$$

where we suppress the role of covariates, $\mathbf{X}$.

Let θ_0 denote the population parameter, $\zeta(b;\theta_0)$ gives the probability that the players (in a randomly sampled game) find themselves in scenario b.

$\zeta(b;\theta)$ is a pmf for scenarios with support $\mathbb{B}$.

Simulated Likelihood (continued)

An accept/reject Monte Carlo ("dartboard") simulation estimate is

$$\hat{\Pr}\left(\mathbf{Y}=\mathbf{y}|\,\mathbf{X};\theta\right)=\frac{1}{S}\sum_{s=1}^{S}\mathbf{1}\left(B^{(s)}\in\mathbb{B}_{\mathbf{y}}\right). \tag{6}$$

with $B^{(s)}$ now a random draw from $\mathbb{B}$ with distribution $\zeta\left(b;\theta\right)$.

It is easy to generate random draws from $\zeta\left(b;\theta\right)$ because the population distribution over $\mathbb{B}$ is induced by the one for the random utility shifters $\mathbf{U}$ (which are easy to simulate).

$\hat{\Pr}\left(\mathbf{Y}=\mathbf{y}|\,\mathbf{X};\theta\right)=\frac{1}{S}\sum_{s=1}^{S}\mathbf{1}\left(\mathbf{y}\text{ is } \textit{the} \text{ NE at } \mathbf{U}^{(s)}\right)$.

Simulated Likelihood (continued)

Unfortunately in large games we will have $\mathbf{1}\left(B^{(s)} \in \mathbb{B}_{\mathbf{y}}\right) = 0$ with very high probability.

This means infeasibly many simulation draws would be required to accurately estimate the likelihood.

Importance Sampling Scenarios

Let $\lambda_{\mathbf{y}}(b;\theta)$ be a function which assigns probabilities to the elements of $\mathbb{B}_{\mathbf{y}}$.

We require that

1. $\lambda_{\mathbf{y}}(b;\theta)$ be strictly greater than zero for any $b \in \mathbb{B}_{\mathbf{y}}$ and zero otherwise (i.e., $b \in \mathbb{B} \setminus \mathbb{B}_{\mathbf{y}}$);

2. satisfy the adding up condition $\sum_{b \in \mathbb{B}_{\mathbf{y}}} \lambda_{\mathbf{y}}(b;\theta) = 1$.

Importance Sampling Scenarios (continued)

Rewrite the likelihood function as an *average* over those scenarios in the set $\mathbb{B}_{\mathbf{y}}$.

Let $\theta^{(0)}$ be some (fixed) value for the parameter; we have that

$$\begin{aligned} \Pr\left(\mathbf{Y}=\mathbf{y}\middle|\,\mathbf{X};\theta\right) &= \sum_{b\in\mathbb{B}_{\mathbf{y}}} \zeta\left(b;\theta\right) \\ &= \sum_{b\in\mathbb{B}_{\mathbf{y}}} \frac{\zeta\left(b;\theta\right)}{\lambda_{\mathbf{y}}\left(b;\theta^{(0)}\right)}\lambda_{\mathbf{y}}\left(b;\theta^{(0)}\right) \\ &= \mathbb{E}_{\tilde{B}}\left[\frac{\zeta\left(\tilde{B};\mathbf{X},\theta\right)}{\lambda_{\mathbf{y}}\left(\tilde{B};\theta^{(0)}\right)}\right], \end{aligned} \qquad (7)$$

where $\tilde{B}$ denotes a random draw from $\lambda_{\mathbf{y}}\left(b;\theta^{(0)}\right)$.

Importance Sampling Scenarios (continued)

Let $\tilde{B}^{(s)}$ be $s = 1, \ldots, S$ independent draws from $\lambda_{\mathbf{y}}\left(b; \theta^{(0)}\right)$.

An importance sampling Monte Carlo estimate of the likelihood function is:

$$\hat{\Pr}\left(\mathbf{Y} = \mathbf{y} \middle| \mathbf{X}; \theta\right) = \frac{1}{S} \sum_{s=1}^{S} \frac{\zeta\left(\tilde{B}^{(s)}; \theta\right)}{\lambda_{\mathbf{y}}\left(\tilde{B}^{(s)}; \theta^{(0)}\right)}. \tag{8}$$

This estimate, because the cardinality of $\mathbb{B}_{\mathbf{y}}$ is finite, is consistent as $S \to \infty$.

All summands in (8) are non-zero.

This Paper

Develops an algorithm for sampling scenarios from $\mathbb{B}_{\mathbf{y}}$.

Allows for SML estimation of a class of supermodular games where T players take M binary actions each.

The analyst observes $N \geq 1$ games.

The space of action profiles $\mathbb{Y}$ for each game has cardinality 2^{TM}.

Can easily handle examples with TM in the tens of thousands.

Key Idea

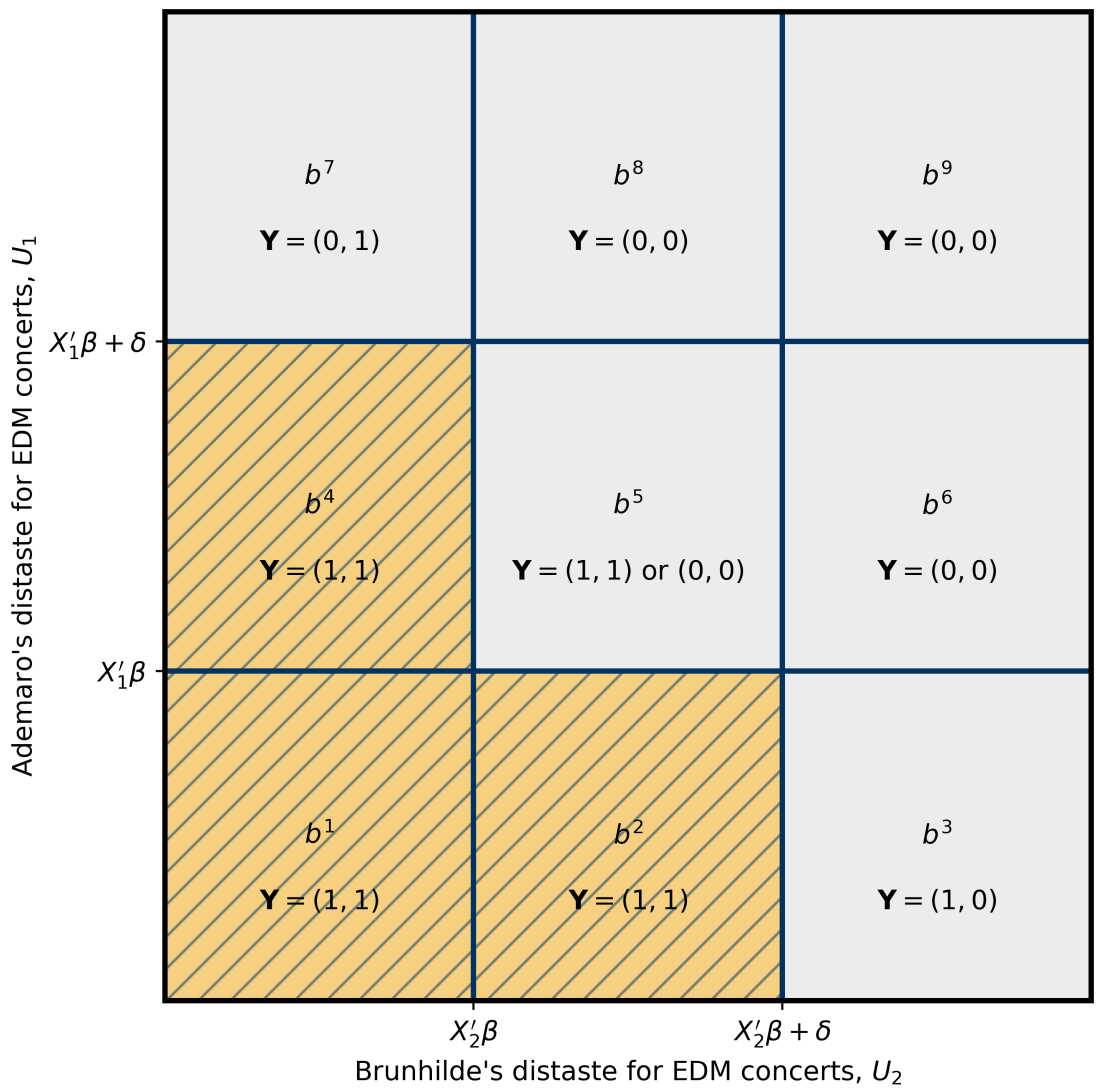
We proceed by drawing $\mathbf{U}$ such that $\mathbf{U} \in \tilde{B}$, $\tilde{B} \in \mathbb{B}_{\mathbf{y}}$ with probability one.

If we draw the elements of $\mathbf{U} = (U_1, \ldots, U_T)'$ *independently*, then $\mathbf{U} \in B$, but $B \in \mathbb{B}_{\mathbf{y}}$ with negligible probability.

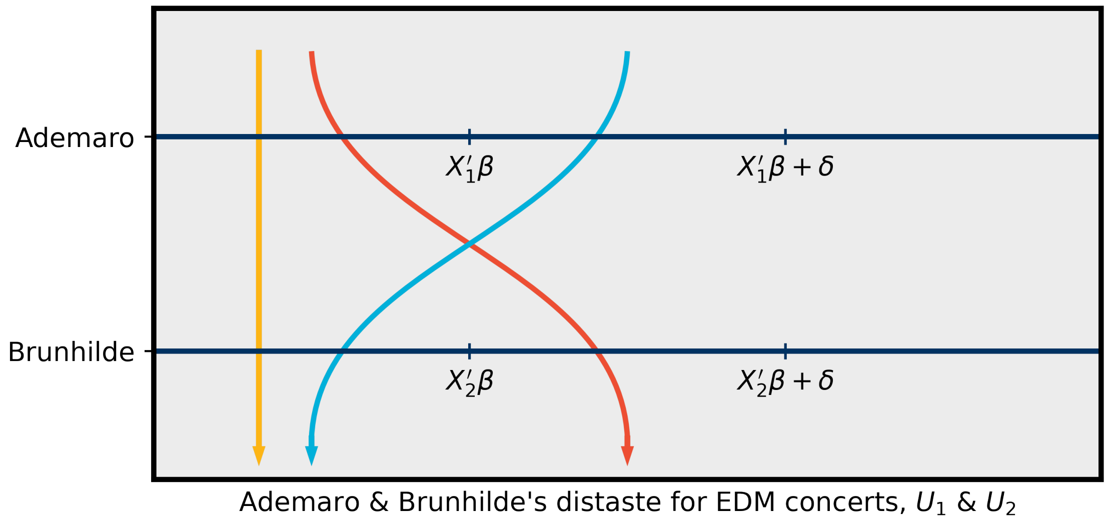
Instead we draw $U_1, U_2, \ldots$ *sequentially*.

The support of U_t will depend on the realizations of U_s for $s < t$. We vary the support such that, in the end, $\mathbf{U} \in \tilde{B}$, $\tilde{B} \in \mathbb{B}_{\mathbf{y}}$ with probability one.

The logic of NE (and supermodularity) allows us to find the correct support for each draw.



b^7
$\mathbf{Y} = (0, 1)$
b^8
$\mathbf{Y} = (0, 0)$
b^9
$\mathbf{Y} = (0, 0)$
b^4
$\mathbf{Y} = (1, 1)$
b^5
$\mathbf{Y} = (1, 1)$ or $(0, 0)$
b^6
$\mathbf{Y} = (0, 0)$
b^1
$\mathbf{Y} = (1, 1)$
b^2
$\mathbf{Y} = (1, 1)$
b^3
$\mathbf{Y} = (1, 0)$
$X_1'\beta + \delta$
$X_1'\beta$
$X_2'\beta$
$X_2'\beta + \delta$
Ademaro's distaste for EDM concerts, U_1
Brunhilde's distaste for EDM concerts, U_2


Ademaro
Brunhilde
$X_1'\beta$
$X_1'\beta + \delta$
$X_2'\beta$
$X_2'\beta + \delta$
Ademaro & Brunhilde's distaste for EDM concerts, U_1 & U_2

Our Paper (continued)

Our method utilizes a new *importance sampling* algorithm (cf. McFadden, 1989; Krauth 2006, Ackerberg, 2009; Bajari, Hong and Ryan, 2010).

1. We can compute SML estimates in models with tens of thousands binary decisions (TM) and hundreds of parameters with a pocket calculator;

2. Method produces simulation estimates of both the log-likelihood function as well as its score;

3. For some classes of models further computational speed-ups are available (I will ignore this today).

Our Paper (continued)

Some related work:

1. Supermodular games: Jia (2008), Nishida (2015), Uetake and Watanabe (2013), Xu and Lee (2015), Miyauchi (2016);

2. Simulation: McFadden (1989), Krauth (2006), Soetevent and Koorman (2007), Ackerberg (2009), Bajari, Hong and Ryan (2010).

This is work in progress (cf. Graham and Pelican, 2020; Pelican and Graham, 2021).

Paper (so far) is about computation only.

Simulation Algorithm

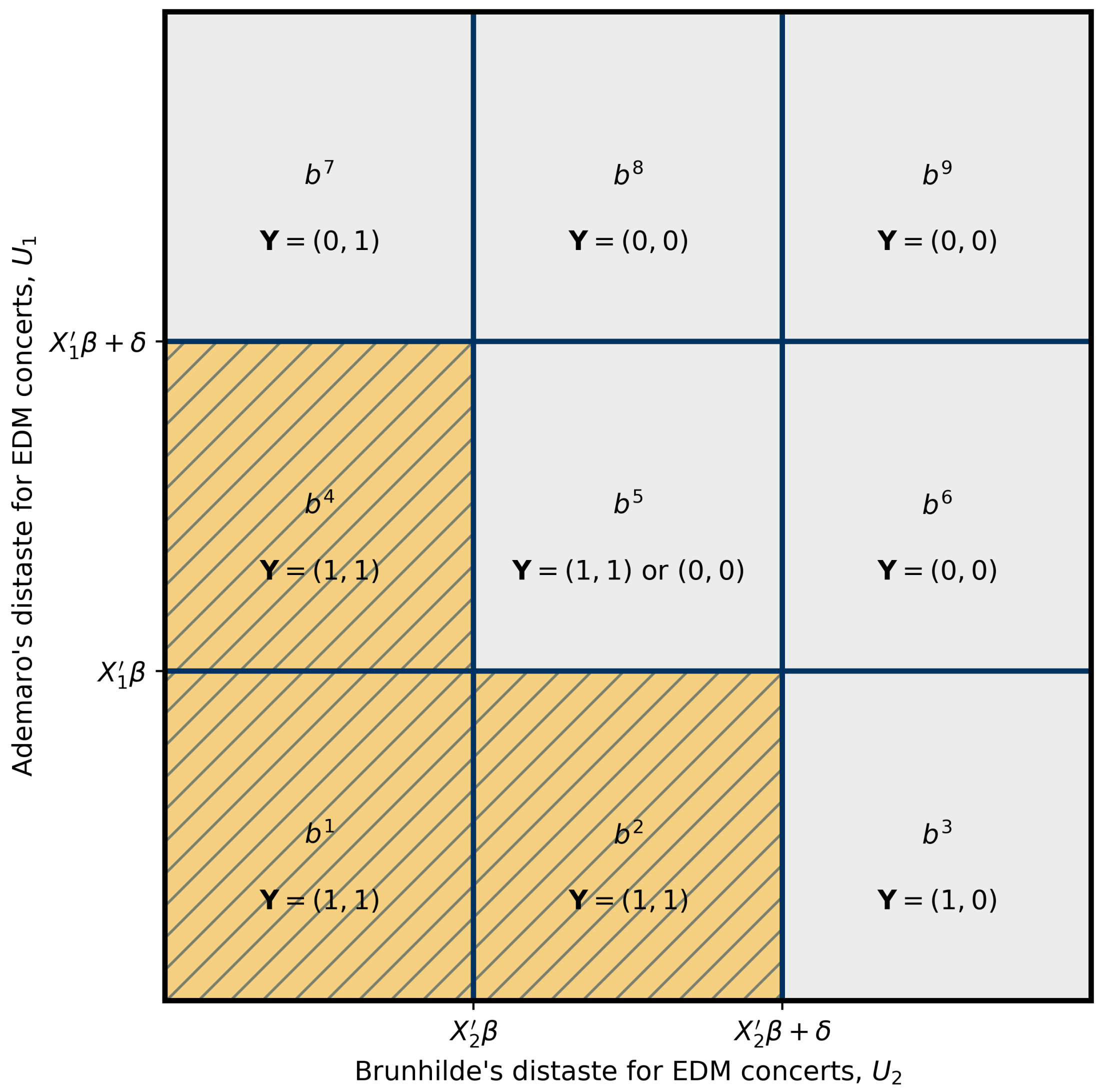
1. $\mathbf{y}$ is target NE. We want $\mathbf{U} \in \tilde{B}$ with $\tilde{B} \in \mathbb{B}_{\mathbf{y}}$.

2. Start with $y_t = 0$ cases: draw $U_t \in \left(X_t' \beta + s\left(\mathbf{y}_{-t}\right)' \delta, \infty \right)$.

3. Go through $y_t = 1$ cases one at a time and

   (a) check how many "defections" would occur if t – contrary to fact – doesn't take action ($\Rightarrow$ new NE with $\tilde{\mathbf{y}} \leq \mathbf{y}$);

   (b) get threshold $\bar{h}_t \in \left(X_t' \beta, X_t' \beta + s\left(\mathbf{y}_{-t}\right)' \delta \right]$ such that if $U_t \leq \bar{h}_t$ our sequence "stays on track."


b^7
$\mathbf{Y} = (0, 1)$
b^8
$\mathbf{Y} = (0, 0)$
b^9
$\mathbf{Y} = (0, 0)$
b^4
$\mathbf{Y} = (1, 1)$
b^5
$\mathbf{Y} = (1, 1)$ or $(0, 0)$
b^6
$\mathbf{Y} = (0, 0)$
b^1
$\mathbf{Y} = (1, 1)$
b^2
$\mathbf{Y} = (1, 1)$
b^3
$\mathbf{Y} = (1, 0)$
$X_1'\beta + \delta$
$X_1'\beta$
$X_2'\beta$
$X_2'\beta + \delta$
Ademaro's distaste for EDM concerts, U_1
Brunhilde's distaste for EDM concerts, U_2

Random Utility Draws for $y_t = 1$ Cases

Finding the appropriate range restriction on U_t for the $y_t = 1$ cases is key.

1. Since $s(\mathbf{y}_{-t})' \delta \geq 0$, if $U_t \in (-\infty, X_t' \beta]$ the action will be taken (strictly dominant strategy).

2. Also possible that a draw of $U_t \in \left(X_t' \beta, X_t' \beta + s(\mathbf{y}_{-t})' \delta\right]$ is sufficiently low such that agent t would still choose to take the action.

3. If $U_t \in \left(X_t' \beta + s(\mathbf{y}_{-t})' \delta, \infty\right)$ agent t will not take the action (no matter what other agents do).

Random Utility Draws for $y_t = 1$ Cases (continued)

We can conclude that there exists an agent-by-action-specific *threshold* $\bar{h}_t \in \Big(X_t'\beta, X_t'\beta + s$
such that

- if $U_t \leq \bar{h}_t$, then it is possible to construct subsequent draws such that, in the end, $\mathbf{U} \in \tilde{B}$ with $\tilde{B} \in \mathbb{B}_{\mathbf{y}}$ (as needed),

- whereas if $U_t > \bar{h}_t$, it will not be possible.

Algorithm 1: Scenario sampler

Inputs: $\mathbf{z} = (\mathbf{X}, \mathbf{y})$, θ (i.e., a target pure strategy combination and a utility/payoff function)

1. Initialize $\mathbf{U} = (U_1, \ldots, U_T)' = \underline{0}_T$.

2. For $t = 1, \ldots, T$

   (a) If $y_t = 0$, then sample $U_t \in \left[X_t'\beta + s\left(\mathbf{y}_{-t}\right)'\delta, \infty\right)$ from the conditional density $\frac{f(u)}{1-F\left(X_t'\beta + s(\mathbf{y}_{-t})'\delta\right)} \stackrel{def}{\equiv} \omega_t f\left(u\right)$.

3. For $t = 1, \ldots, T$

(a) If $y_t = 1$, then

i. determine $\bar{h}_t$ using $\textsc{Threshold}$($\mathbf{z}$, θ, $\mathbf{U}$, t);

ii. sample $U_t \in \left(-\infty, \bar{h}_t\right]$ from the conditional density $\frac{f(u)}{F\left(\bar{h}_t\right)} \overset{def}{\equiv} \omega_t f(u)$.

4. Find $\tilde{B} \in \mathbb{B}_{\mathbf{y}}$ such that $\mathbf{U} \in \tilde{B}$.

Outputs: The $T \times 1$ weight vector $\underline{\omega} = (\omega_1, \ldots, \omega_T)'$, the vector of utility shifters $\mathbf{U}$ and a (random) scenario $\tilde{B} \in \mathbb{B}_{\mathbf{y}}$.

Algorithm 2: Threshold finder

Inputs: $\mathbf{z} = (\mathbf{X}, \mathbf{y})$, θ, $\mathbf{U}$, t

1. For $t' = 1, \ldots, T$

   (a) if $y_{t'} = 0$, then set $\tilde{U}_{t'} = U_{t'}$;

   (b) if $y_{t'} = 1$, then

      i. if $t' < t$, then set $\tilde{U}_{t'} = U_{t'}$ ($\bar{h}_{t'}$ already found)

      ii. if $t' > t$, then set $\tilde{U}_{t'} = X_{t'}'\beta - 1$ ($\bar{h}_{t'}$ not already found; force $\tilde{Y}_{t'} = 1$)

2. Set $\tilde{U}_t = X_t'\beta + s(\mathbf{y}_{-t})'\delta + 1$ (ensures that player t *will not* want to choose $\tilde{Y}_t = 1$ in Step 3 below)

3. Find the minimal NE, $\mathbf{Y}$, associated with $\tilde{\mathbf{U}}$. Set $\bar{h}_t = X_t' \beta + s\left(\mathbf{Y}_{-t}\right)' \delta$

Output: The threshold, $\bar{h}_t$.

Threshold finder (intuition)

By forcing player t to not take the action (Step 2), some players – for whom we have already simulated utility shocks ($t' < t$) – will choose to also now not take action (even thought $y_{t'} = 1$). This induces a new NE (step 3) with $\mathbf{Y} \leq \mathbf{y}$ (supermodularity).

$\bar{h}_t$ is the maximal value of U_t such that the "defections" in $\tilde{\mathbf{Y}}$ don't occur,

If $U_t \in \left(-\infty, \bar{h}_t\right]$, then player t will take the action as desired, and those players $t' < t$ which "defected" in $\tilde{\mathbf{Y}}$ will also take the action.

OTH, if $U_t > \bar{h}_t$, then player t not taking the action, and some subset of players $t' < t$ also not taking action, yields a minimal NE ($\tilde{\mathbf{Y}}$) below the target.

Monte Carlo Experiments

Peer effects on networks example.

$$\upsilon_t\left(\mathbf{y};\mathbf{x},\mathbf{u},\theta\right) \stackrel{def}{\equiv} y_t\left(x_t'\beta+\delta\left(\sum_{s\neq t} d_{ts}y_s\right)-u_t\right).$$

Friendships generated by a random geometric network. Four covariates, two discrete, two continuous.

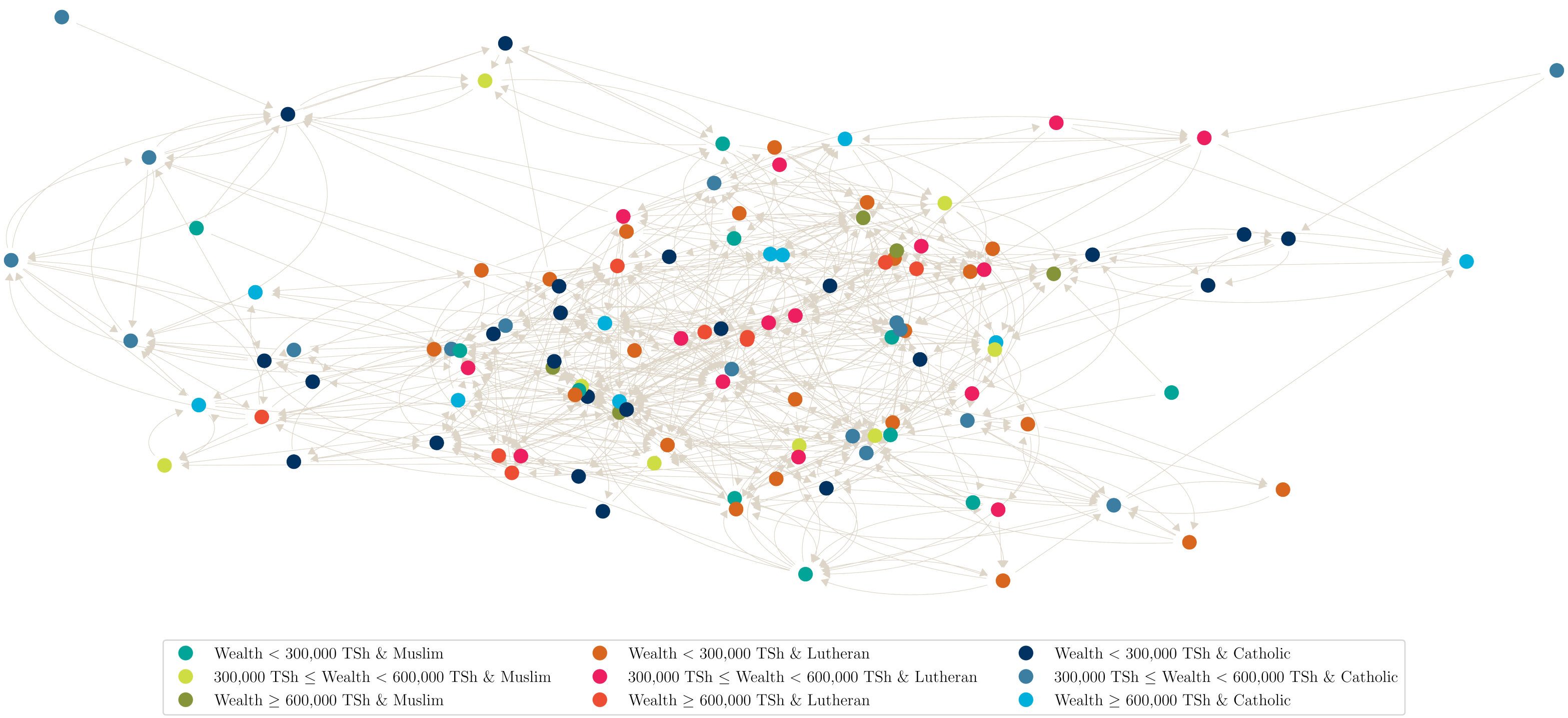
Two cases:

1. 2000 agents in 100 distinct friendship networks;

2. 500 agents in a single friendship network.

Monte Carlo Experiments (continued)

	Panel A Many Games ($N = 100$)			Panel B Single Game ($N = 1$)		
Number of players per game, T	20	20	20	500	500	500
Number of scenario draws, S	1	10	100	1	10	100
Number Monte Carlo replications	500	500	500	500	500	500
Mean of $\hat{\delta}$	0.208	0.200	0.197	0.206	0.199	0.197
Std. Dev. of $\hat{\delta}$	0.020	0.030	0.033	0.030	0.043	0.051
Likelihood Ratio test size ($H_0 : \delta = \delta_0, \alpha = 0.05$)	0.072	0.050	0.034	0.060	0.040	0.072
Confidence interval coverage ($1 - \alpha = 0.95$)	0.942	0.952	0.936	0.958	0.896	0.872


Wealth < 300,000 TSh & Muslim
300,000 TSh ≤ Wealth < 600,000 TSh & Muslim
Wealth ≥ 600,000 TSh & Muslim
Wealth < 300,000 TSh & Lutheran
300,000 TSh ≤ Wealth < 600,000 TSh & Lutheran
Wealth ≥ 600,000 TSh & Lutheran
Wealth < 300,000 TSh & Catholic
300,000 TSh ≤ Wealth < 600,000 TSh & Catholic
Wealth ≥ 600,000 TSh & Catholic

"Regressor"	Probit	SMLE ($S = 1$)	SMLE ($S = 10$)	SMLE ($S = 100$)
Support	0.183	0.166	0.127	0.146
($\sum_{r=1}^{T} y_{rt} y_{rs}$)	(0.031)	(0.015)	(0.014)	(0.031)
Parents, children	1.485	1.511	1.509	1.510)
and siblings	(0.116)	(0.113)	(0.113)	(0.117)
Nephews, nieces, uncles, aunts,	0.919	0.897	0.921	0.929)
cousins, grandparents, grandchildren	(0.128)	(0.127)	(0.127)	(0.128)
Other blood relative	0.697	0.691	0.714	0.702
	(0.102)	(0.100)	(0.100)	(0.101)
Distance (km)	-1.375	-1.396	-1.420	-1.394
	(0.100)	(0.099)	(0.099)	(0.101)
(and more!)				

Recap

Our importance sampling approach:

1. Makes SML estimation feasible in supermodular games with many agents (T) and/or many actions (M).

2. Because we can also construct score estimates, we can fit high dimensional models (i.e., don't need to rely on grid searches).

3. Opens up a wide variety of large games to formal/structural empirical analysis.