

Empirical strategy-proofness

Alexander L. Brown and Rodrigo A. Velez

Texas A&M University

CRETE
July 2022

Strategy-proofness

- ▶ Central property in market/mechanism design.
- ▶ $g : \Theta_1 \times \cdots \times \Theta_n \rightarrow X$
- ▶ An scf is strategy proof (SP) if it gives agents a weak incentive to reveal private information.
- ▶ Incentives for truthful revelation are usually weak.

DS hypothesis

- ▶ Dominant Strategy hypothesis: people will play a dominant strategy when it is available;
- ▶ **not supported by data.** (Coppinger et al. 1980 EI...Li, 2017 AER).
 - ▶ Data in Cason et al (2006, GEB), Andreoni et al (2006, GEB), Healy (2006, JET) supports convergence to equilibria that are not intended by the scf.
- ▶ SP is not a guarantee of performance.

Traditional (popular) equilibrium refinements are silent about this phenomenon

- ▶ **Tremble-based refinements** (Selten 1975; Myerson 1978;...Milgrom and Mollner 2017,2018; Fudenberg and He 2018); **stability** (Kohlberg and Mertens 1986); van Damme (1991)...
- ▶ **implicitly or explicitly** assume “admissibility” (weakly dominated behavior implausible).
- ▶ This paper: propose a new path to refine Nash equilibrium; show that it allows us to understand this phenomenon (what are the equil. of a SP game that should concern us).

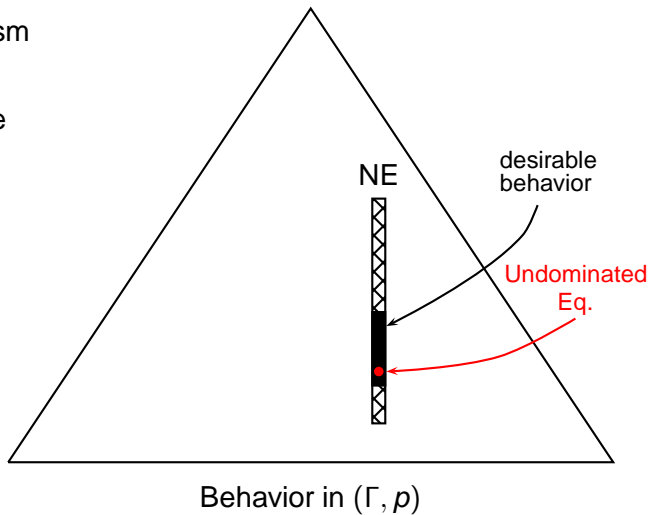
Our approach

Θ Type space

Γ mechanism

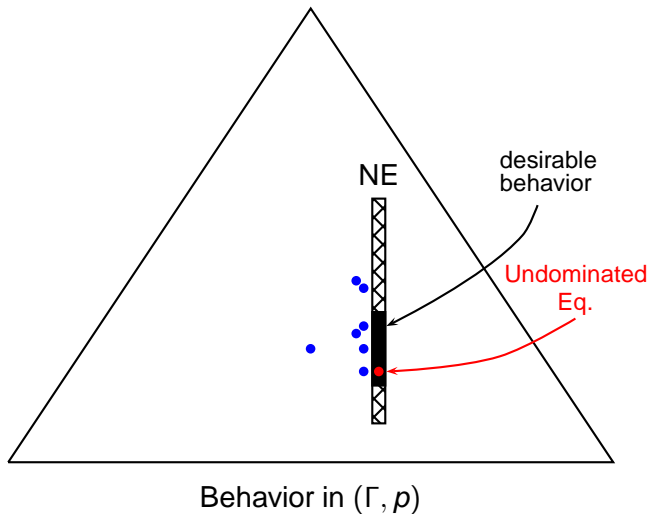
$p \in \Delta(\Theta)$

(Γ, p) game



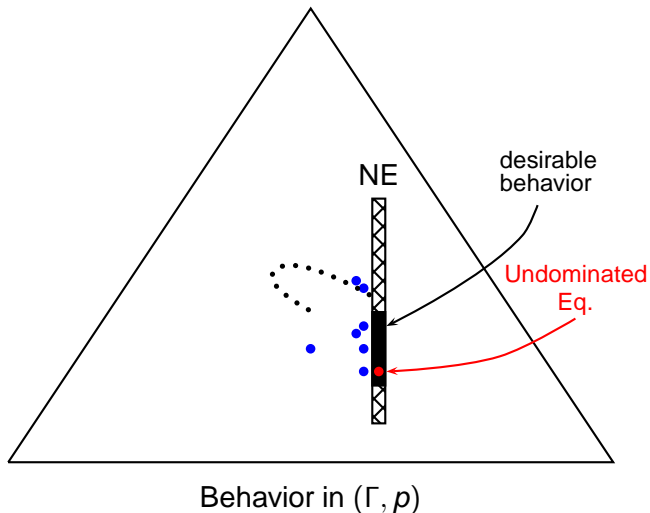
Experimental evidence

Experimental data reveals: NE is too pessimistic, UD is too optimistic.



Behavior in experiments is not random

McKelvey and Palfrey (1995): Weakly dominated behavior can be the limit of increasingly sophisticated Logistic QRE.

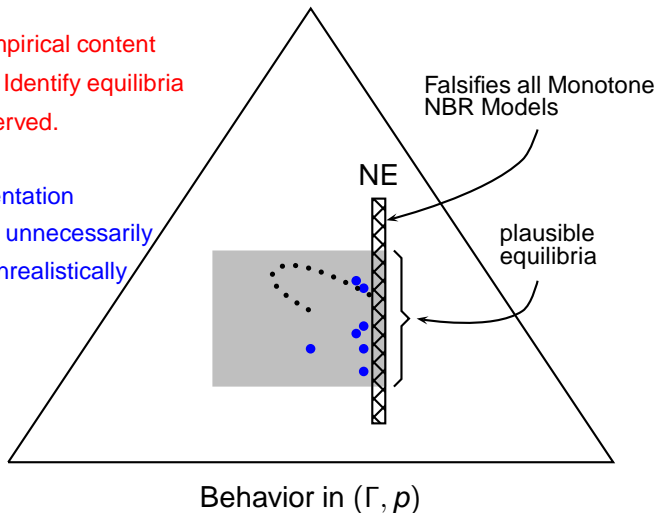


Noisy best response models

Monotone noisy best response (NBR) models do a good job at rationalizing experimental data when behavior has the chance to converge

Idea: calculate empirical content of these theories. Identify equilibria that won't be observed.

New Full implementation paradigm: neither unnecessarily pessimistic, nor unrealistically optimistic.

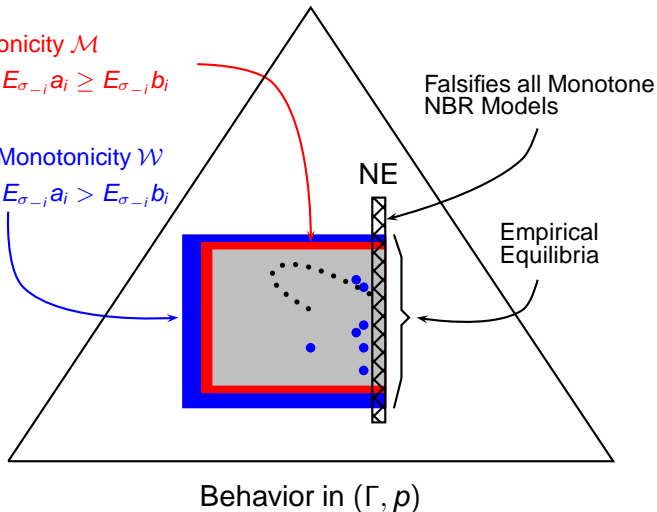


Which Monotone NBR theory should we base our analysis on?

Logistic QRE, monotone structural QRE, control costs, regular QRE,
Harsanyi/van Damme monotone RDPM,...?

Payoff monotonicity \mathcal{M}
 $\sigma_i(a_i) \geq \sigma_i(b_i) \Leftrightarrow E_{\sigma_{-i}} a_i \geq E_{\sigma_{-i}} b_i$

Weak Payoff Monotonicity \mathcal{W}
 $\sigma_i(a_i) > \sigma_i(b_i) \Rightarrow E_{\sigma_{-i}} a_i > E_{\sigma_{-i}} b_i$



Empirical equilibrium

Some notation: a_i , $u_i(a_1, \dots, a_n)$, $\sigma_i(a_i)$
 (u, σ) both observable

Definition 1 ($\mathcal{W}(u)$)

σ is **weakly payoff monotone** for u if for each $i \in N$, each $a_i, b_i \in A_i$, if $\sigma_i(a_i) > \sigma_i(b_i)$ we have that $E_{u_i}(a_i | \sigma_{-i}) > E_{u_i}(b_i | \sigma_{-i})$.

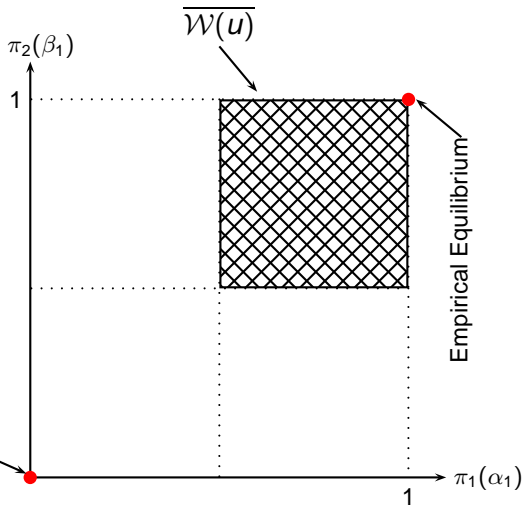
Definition 2 ($EE(u) = NE(u) \cap \overline{\mathcal{W}(u)}$)

An **empirical equilibrium** of u is a Nash equilibrium that is the limit of weakly payoff monotone distributions for u .

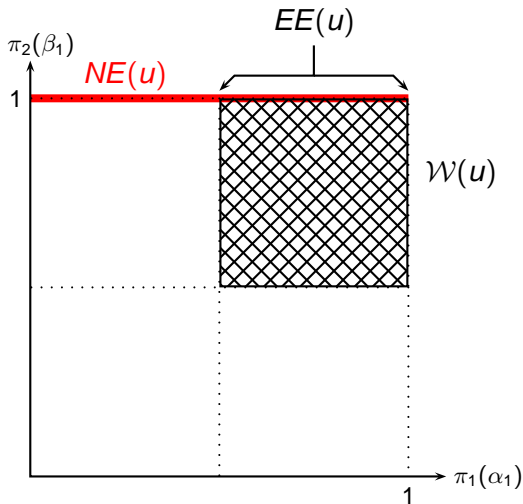
Empirical equilibria exist in each finite game. They may admit weakly dominated behavior (McK-P 1995).

	β_1	β_2
α_1	1, 1	0, 0
α_2	0, 0	0, 0

Implausible: would reject all monotone NBR theories



	β_1	β_2
α_1	2, 2	2, 1
α_2	2, 3	0, 0



$EE \not\subseteq$ Undominated E; Proper, Perfect $\not\subseteq$ EE.

EE differentiates between SP mechanisms

For each common prior info structure all EE are truthful equivalent	
TTC	+
Second Price Auction	-

Why is this? Why do we have + in TTC, and - in SPA? For which info structures is SPA failing?

Model: detail free implementation

All Finite (private values):

- ▶ $N \equiv \{1, \dots, n\}$ agents; X outcome space
- ▶ Θ_i payoff type space, $\theta_i \in \Theta_i$, $\Theta \equiv \Theta_1 \times \dots \times \Theta_n$
- ▶ Common prior $p \in \Delta(\Theta)$ (can be generalized).
- ▶ Scf $g : \Theta \rightarrow X$
- ▶ Direct revelation mechanism of g , (Θ, g)
- ▶ Induces Bayesian game (Θ, g, p)
- ▶ Behavior strategy: $\theta_i \in \Theta_i \mapsto \sigma_i(\cdot | \theta_i) \in \Delta(\Theta_i)$
- ▶ σ is weakly payoff monotone for (Θ, g, p) if for each $i \in N$, each $\theta_i \in \Theta_i$, $\sigma_i(\tau_i | \theta_i) > \sigma_i(\tau_i' | \theta_i)$ implies $E_{u_i(\cdot | \theta_i)}(\tau_i | \sigma_{-i}, p) > E_{u_i(\cdot | \theta_i)}(\tau_i' | \sigma_{-i}, p)$.
- ▶ An **empirical equilibrium** of (Θ, g, p) is one of its Bayesian NE for which there is a sequence of weakly payoff monotone distributions converging to it.
- ▶ An equilibrium of (Θ, g, p) **implements g** if for each state θ in the support of p it obtains $g(\theta)$.

Robust full implementation in EE

Theorem 3

Let g be an scf. The following are equivalent:

- ▶ For each common prior p , each empirical equilibrium of (Θ, g, p) implements g .
- ▶ g is *strategy-proof* and *non-bossy*.

Robust interior full implementation in EE: Information Structure Matters

Theorem 4

Let g be an scf. The following are equivalent:

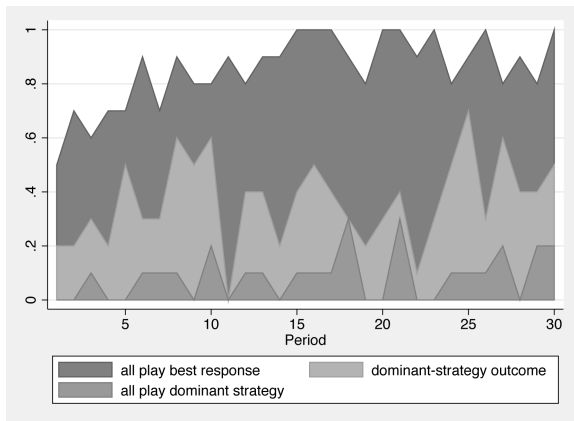
- ▶ *For each interior common prior p , each empirical equilibrium of (Θ, g, p) implements g .*
- ▶ *g is **strategy-proof** and **has no bossy dominant strategy**.*

Conclusions

- ▶ It is safe (for the long term performance of a mechanism) to operate a SP scf if one can expect truthful behavior is frequent enough and
 - ▶ SCF is non-bossy.
 - ▶ SCF has no bossy dominant strategy and there is enough uncertainty.
- ▶ It may not be safe to operate a SP scf in which an agent can be bossy where there is no enough uncertainty.
- ▶ Performance of SP mechanisms may depend on the information structure in which they are operated.
- ▶ It is possible to make a meaningful selection of weakly dominated NE based on regularities observed in experiments on simultaneous-move mechanisms.

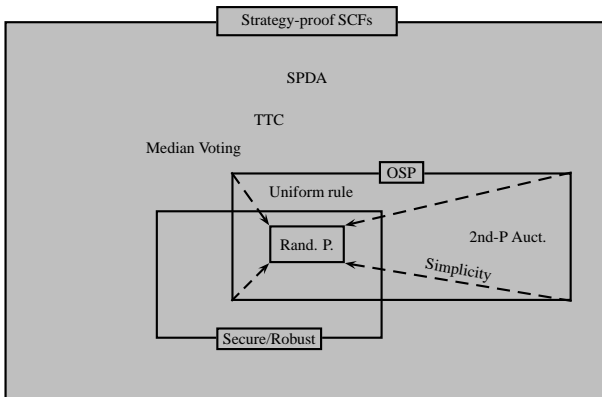
Thanks for your attention.

Andreoni et al (2006): Second price auction



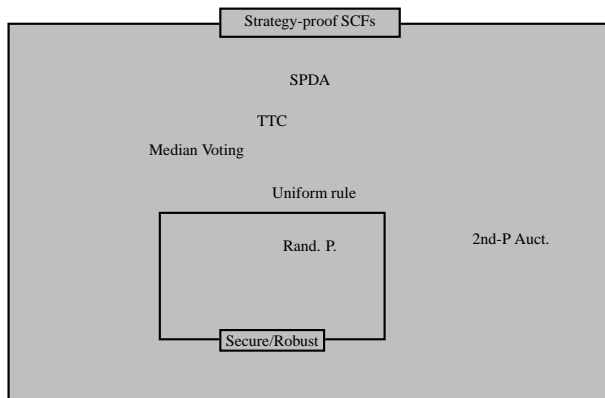
What SP scfs, and in which conditions, admit this type of behavior?

Related literature: not all sp scfs work the same



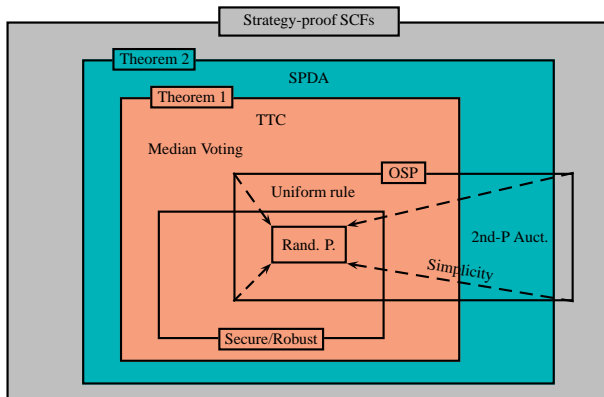
- ▶ Secure/robust implementation: All Nash equilibria are optimal (Saijo et al. 2007; Bochet and Sakai 2010; Fujinaka Wakayamama 2011; Adachi 2014)
- ▶ Obvious Strategy-proofness/Simplicity: recognizing truthful report as a best response is simpler (Li 2017; Bade and Gonczarowski 2017; Ashlagi and Gonczarowski 2018; Arribillaga et al 2019; Pycia and Troyan 2019).

What if mechanism is not secure/robust/simple?



- ▶ Not all NE are plausible.
- ▶ Our objective: understand the conditions under which it is plausible that behavior can approximate NE that produce suboptimal outcomes in SP games.

Our classification in the context of literature



Proof of results

Our proof allows us to strengthen the policy implications of our theorems.

1. It is enough that truthful behavior be frequent enough (at least uniform random play) so behavior in a mechanism satisfying the properties in our theorems will never approach a bad Nash equilibrium.
2. If the conditions in the theorems are violated, one can construct bad empirical equilibria that are approached by distributions generated by some of the most popular noisy best response models.

Empirical support

Experiments that provide empirical support:

- ▶ evaluate whether subject play is baseline-truthful in strategy-proof games;
- ▶ compare two, largely equivalent direct-revelation games with complete information, one under a strategy-proof scf that is bossy and one that is non-bossy;
- ▶ compare mechanisms with no bossy dominant strategies under interior and non-interior information;
- ▶ evaluate whether subject play obeys our assumption of weak payoff monotonicity and thus when data approximates mutual best responses it is consistent with empirical equilibrium.