Motivation: To be credible, economic applications should use only Nash equilibria that are viable.

Simple dual indices:
formation index, $F(\pi) =$ the # of players that “can form $\pi$,”
defection index, $D(\pi) =$ the # of defectors that “$\pi$ can sustain.”

Surprisingly, these simple indices:

1. predict the performance of Nash equilibria in social systems and lab experiments, i.e. assess viability,

2. they also uncover new properties of Nash equilibria that have eluded game theory refinements.
Viable Nash Equilibria: Formation & Defection
Ehud Kalai

Motivation: To be credible, economic analysis should restrict itself to use of only those Nash equilibria that are viable.

Viable Nash Equilibria: Formation & Defection

The broader objective:
Develop theoretical tools to address behavioral questions.

Similar to
$E(X)$ & $SD(X)$ that assess viability of investments $X$,
$F(\pi)$ & $D(\pi)$ assess viability of equilibria $\pi$.

Duality: $F(\pi) + D(\pi) = n$ the # of players.

Simple minimal departure from Nash:
• Stay with anonymous ordinal defections.
• Only replace Nash’s assumption that “no opponents defect” by $\delta = “the # of potential defectors.”$
:: attribute all new observations to $\delta$.

Avoid game theory refinements for:
• broad applicability.
• simple best-response computations.
Motivation: Viable Nash equilibria provide good understanding of functioning social systems, whereas unviable Nash equilibria are often contrived, and have the appearance of useless theory.

Viable Nash Equilibria

Ehud Kalai

I study simple dual indices to assess equilibrium viability: a formation index, \( F(\pi) \), and a deterrence index, \( D(\pi) \).

Surprisingly, despite their simplicity these indices:
1. identify new properties of Nash equilibria and stability issues beyond game theory refinements and
2. provide insights for the viability of Nash equil in functioning social systems and lab experiments.

Related earlier work

- **In theory:** \( D(\pi) \) = the level of subgame perfection in play with revisions, as in Kalai & Neme (1992).
- **In applications:** \( D \) originated in *distributive computing*, adopted to *implementation theory* by Eliaz (2002), Abraham et al. (2006), and Gradwohl and Reingold (2014).
- \( F \) is new.
- This paper deals with theory and applications of *more extensive properties of \( D \) and the new index \( F \).*
The game theory story:

- Rational players play an equilibrium \( \pi \), but concerned about defections by opponents.

Examples:

- Faulty opponents, irrational, unpredictable, see Eliaz (2002).
- Defections by coalitions of rational opponents.
- Incomplete game specifications: e.g., threats, bribes, reputation for future play, misspecified payoffs,…

A strategy is “highly viable,” if it is “dominant” against “many” defections. Strong condition, yet observed in social systems; more manageable than optimal Bayesian response.
Lecture topics:
Definitions and properties (1)
Subjective viability assessments (1)
Behavioral observations (1)
Incomplete info game (1)
Comparisons with standard GT (1)
Rational coalitional defections (1/2)
Forming/switching equilibrium (1/4)
Implementations with faulty players (1/4)
Viability in network games (1/4)
Future research (1/2)
Proposed experiment (1/4)
Definitions and properties

\(\pi\) - a fixed strategy profile of an \(n\)-person strategic game \(\Gamma\).

\(D(\pi)\) and \(F(\pi)\) below are dual/equivalent, nevertheless, each is the natural primitive in its own set of applications.
Definitions and properties

\( \pi \) - a fixed strategy profile of an \( n \)-person strategic game \( \Gamma \).

**Definition:**

\( D(\pi) \equiv \) the minimal number of defections from \( \pi \) needed to create a profile \( \pi' \) to which \( \pi \) is not a best response.

**D(\pi) measures the confidence in individual strategies:**

With any # of defections < \( D(\pi) \), everybody’s \( \pi_i \) is optimal.

**Equivalent definition:** resilience in Abraham et al (2006)

\( D(\pi) \equiv \) the maximal “# of potential defectors \( d \) that \( \pi \) deters.”

That is: in any game \( D \) played by \( d \) potential defectors, \( \pi_i \) is a dominant strategy for each \( i \in D \).

(Assuming that the remaining \( n-d \) players in \( D^c \) are \( \pi \) loyalists).

**Beyond Nash’s deterrence of individual defectors,**

\( \pi \) strongly deters defection of any group of size \( \leq D(\pi) \).
Definitions and properties

\[ D(\pi) \equiv \text{the maximal } d \text{ s.t. in any game } D \text{ played by } d \text{ potential defectors (while the remaining } n-d \text{ players are fixed at their } \pi_j \text{s), } \pi_i \text{ is a dominant strategy for each } i \in D. \]

Next: A dual restatement, stated through the number of loyalists:

(Dual) Definition of the formation index
\[ F(\pi) \equiv \text{The minimal } l \text{ s.t. for any group of } l\text{-loyalists, } \pi_i \text{ is a dominant strategy for any player } i \text{ outside the group.} \]

Any \( F(\pi) \) loyalists strongly induce the play of \( \pi \) on the rest.

\begin{tcolorbox}[colback=green!5!white, colframe=green!50!black]
\textbf{Duality}

\[ D(\pi) + F(\pi) = n, \text{ each useful in different applications.} \]
\end{tcolorbox}
Definitions and properties

Equivalent to an earlier definition of Eliaz (2002):

Tolerance of $D(\pi) - 1$ faulty players: If $D(\pi) - 1$ players are faulty (irrational and unpredictable), $\pi$ remains a Nash equil of the non-faulty players.

Thus, a new natural concept:

Nash critical mass of $\pi \equiv F(\pi) + 1$. bounded “Small worlds”
If a group $G$ has at least $F(\pi) + 1$ players, then $\pi$ is a Nash equilibrium for $G$ no matter what the $G$-outsiders play.

Nash/dominance complementarity:
$\pi$ is a Nash eq of any group with $c + 1$ or more players iff $\pi$ is a dominant strategy for any player in a complementary group with $n - c$ or fewer players.
Definitions and properties

Nash equilibria are rungs on a ladder of sustainability/dominance

$D(\pi) = 0, 1, \ldots, n$ partitions all stgy profiles $\pi$ of n player games

$D(\pi) = n$ iff $\pi$ is a dominant strategy equilibrium.
$D(\pi) = n-1$ iff $\pi_i$ is a dominant stgy for each player $i$, when conditioning on “at least 1 opponent playing $\pi$".

$D(\pi) = 2$ iff $\pi$ deters up to 2 defectors
$D(\pi) = 1$ iff $\pi$ deters only single defectors
$D(\pi) = 0$ iff $\pi$ is not a Nash equilibrum
Example: Asymmetric game

The Party Line game.
3 Democrats and 5 Republicans, each chooses $E$ or $F$.
Payoffs: # opposite-party players you mismatch.
Divisive equil: Dems choose $F$; Reps choose $E$.

$$D(\text{Div}) = \min(2,3) = 2$$  
Why?

$$F(\text{Div}) = 8 - 2 = 6$$
Lecture topics:
Definitions and properties (1)
Subjective viability assessments (1)
Behavioral observations (1)
Incomplete info game (1)
Comparisons with standard GT (1)
Rational coalitional defections (1/2)
Forming/switching equilibrium (1/4)
Implementations with faulty players (1/4)
Viability in network games (1/4)
Future research (1/2)
Proposed experiment (1/4)
$D$ assesses equil sustainability

**High sustainability: Language Matching**

$n = 200M$ players, choose a language.  
**Payoff:** # of opponents you match.  
**Equil:** everybody chooses English, $eE$.  
$$D(eE) = 100M.$$  

**Low sustainability: All or Lose (AOL)**

$n = 200M$ players, each can play or not.  
**Player payoff** = \[
\begin{cases}
1, & \text{iff all } 200M \text{ play}, \\
-1, & \text{otherwise}. \\
\end{cases}
\]
**Non-player’s** = 0  
**Equil:** all 200M play, $aP$.  
$$D(aP) = 1.$$
### Viability assessment

<table>
<thead>
<tr>
<th>Sustainability</th>
<th>(D(\pi))</th>
</tr>
</thead>
<tbody>
<tr>
<td>200M</td>
<td></td>
</tr>
<tr>
<td>100M</td>
<td>all English</td>
</tr>
<tr>
<td></td>
<td>all play AOLose</td>
</tr>
</tbody>
</table>

- From 200M: All play AOLose
- From 100M: All play AOLose
- From English: All play AOLose
Difficult formation Language Matching

$n = 200M$ players, choose a language.
Payoff: # of opponents you match.
Equil: everybody chooses Mandarin, eM

Defection deterrence: $D(\text{eM}) = 100M$.
Formation difficulty: $F(\text{eM}) = 200M - 100M = 100M$

Easy formation: New Communication Network

$n = 200M$ players subscribe or not, subscription costs $9.99$
Subscriber’s payoff: # (of other subscribers) $- 9.99$ .
Non-subscriber’s payoff: 0.
Equil: all subscribe, aSub.
Formation difficulty: $F(\text{aSub}) = 10$.
$D(\text{aSub}) = 200M - 10$. 

Sustainable but formation requires 100M loyalists.
Sustainable formation requires only 10 loyalists.
# Viability assessment

<table>
<thead>
<tr>
<th>Formation difficulty</th>
<th>Sustainability</th>
<th>Subjective Viability Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F(\pi)$</td>
<td>$D(\pi)$</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>200M</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>200M-10</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100M</td>
<td>100M English</td>
<td></td>
</tr>
<tr>
<td>200M</td>
<td>all play AOL</td>
<td></td>
</tr>
</tbody>
</table>

*All play AOL.*
<table>
<thead>
<tr>
<th>Formation difficulty</th>
<th>Sustainability</th>
<th>Subjective Viability Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$F(\pi)$</td>
<td>$D(\pi)$</td>
</tr>
<tr>
<td>all subscribe</td>
<td>0</td>
<td>200M</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>200M-10</td>
</tr>
<tr>
<td>all Mandarin</td>
<td>100M</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>all English</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>all play AOL</td>
</tr>
<tr>
<td></td>
<td>200M</td>
<td></td>
</tr>
</tbody>
</table>

Viability assessment
### Viability assessment

<table>
<thead>
<tr>
<th>Formation difficulty</th>
<th>Sustainability</th>
<th>Subjective Viability Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F(\pi)$</td>
<td>$D(\pi)$</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>10</th>
<th>200M</th>
<th>200M-10</th>
<th>all subscribe</th>
<th>all subscribe</th>
</tr>
</thead>
<tbody>
<tr>
<td>all</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>all subscribe</td>
<td></td>
</tr>
<tr>
<td>all</td>
<td>100M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>all Mandarin/English</td>
<td>English/Mandarin</td>
</tr>
<tr>
<td>all</td>
<td></td>
<td>2</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td>all play AOL</td>
<td>all play AOL</td>
</tr>
<tr>
<td>all</td>
<td>200M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Viability assessment

<table>
<thead>
<tr>
<th>Formation difficulty</th>
<th>Sustainability</th>
<th>Subjective Viability Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F(\pi)$</td>
<td>$D(\pi)$</td>
<td></td>
</tr>
<tr>
<td>all subscribe</td>
<td>200M</td>
<td>High viability</td>
</tr>
<tr>
<td>Mandarin/English</td>
<td>200M-10</td>
<td></td>
</tr>
<tr>
<td>all play AOL</td>
<td>100M</td>
<td>?</td>
</tr>
<tr>
<td>English/Mandarin</td>
<td>100M-10</td>
<td>Low viability</td>
</tr>
<tr>
<td>all play AOL</td>
<td>200M</td>
<td></td>
</tr>
<tr>
<td>all subscribe</td>
<td>200M-10</td>
<td></td>
</tr>
</tbody>
</table>
Viability assessment

<table>
<thead>
<tr>
<th>Formation difficulty $F(\pi)$</th>
<th>Sustainability $D(\pi)$</th>
<th>Subjective Viability Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>all subscribe</td>
<td>High viability</td>
<td>Viable if $\pi$ already exists</td>
</tr>
<tr>
<td>Mandarin/English</td>
<td>Viable if $\pi$ already exists</td>
<td>Low viability</td>
</tr>
<tr>
<td>all play AOL</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

200M-10

100M

100M

200M

0

1

2

10
## Viability assessment

<table>
<thead>
<tr>
<th>Formation difficulty</th>
<th>Sustainability</th>
<th>Subjective Viability Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F(\pi)$</td>
<td>$D(\pi)$</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>200M</td>
<td>High viability</td>
</tr>
<tr>
<td>1</td>
<td>all subscribe</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>200M-10</td>
<td></td>
</tr>
<tr>
<td>.</td>
<td>all subscribe</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **all subscribe**
- **all Mandarin/English**
- **all play AOL**

- **all 100M English/English**
- **all 100M English/Mandarin**
- **all 200M Mandarin**
- **all 200M-10 Mandarin**
- **all 200M play AOL**
- **all 200M-10 play AOL**
Lecture topics:
Definitions and properties (1)
Subjective viability assessments (1)
Behavioral observations (1)
Incomplete info game (1)
Comparisons with standard GT (1)
Rational coalitional defections (1/2)
Forming/switching equilibrium (1/4)
Implementations with faulty players (1/4)
Viability in network games (1/4)
Future research (1/2)
Proposed experiment (1/4)
Viability in functioning systems

Recall Language-matching game

\[ n = 200M \] symmetric players.
Strategy: choose a language.
Payoff: number of opponents you match.
Equil: everybody chooses English, \( eE \).
Def deterrence: \( D(eE) = 100M \).

Many similar social systems

- All English, all Spanish, all Mandarin...
- All use dollars, all use euros, ...
- All use the metric system, all use the U.S. measurement system...
- All use the Qwerty keyboard, all drive on the right, ...
- Subscribers to comm’n networks: Facebook, Twitter, Zoom,...
Social avoidance of unviable equilibria

All judges submit zero score in the Beauty-Contest game.
See experimental results in Nagel (1995) and follow up papers.

All reporting to work in a simple production line.

Centralized trade, Centralized communication...
Papers in finance and in network communication design.

Mixed strategies equilibrium in coordination game.
See O'Neill (1987) and follow up papers for experimental results.

Nobody confess in the Confession game.
Non-disclosure agreements in bio and high tech companies.
Lecture topics:
Definitions and properties (1)
Subjective viability assessments (1)
Behavioral observations (1)
Incomplete info game (1)
Comparisons with standard GT (1)
Rational coalitional defections (1/2)
Forming/switching equilibrium (1/4)
Implementations with faulty players (1/4)
Viability in network games (1/4)
Future research (1/2)
Proposed experiment (1/4)
Example with Incomplete Information: Duplicating signals

Why members of the same political party repeat the same talking points again and again?
A signaling game

\[ S = \{\alpha, \beta\}; \text{two pos. states, recommendations, and actions.} \]

**Players:** 1000 DMs; 3 recommenders: 1 Malicious and 2 Honest.

Knowing the state \( \theta \in S \); each recommender recommends \( R_i \in S \).

Knowing the majority recommendation, each DM select \( A_i \in S \).

**Payoff:** DMs and Honest recommenders = \( \#(A_i = \theta) \).

Malicious recommender = \( \#(A_i \neq \theta) \).

Equil \( \pi \): \( R_H = \theta; R_M = \theta^c \); DMs follows the majority recom’n.

Sustainability: \( D(\pi) = 1 \). Why? What if \( R_M = \theta \)?

Altered game with 4 Honest recom’ers has \( D(\pi) = 2 \) (>1).

The game and the altered game are “common-knowledge equivalent” but duplicating senders yields higher \( D \)-values.

Political parties duplicate recommenders with identical talking points to maximize (or to minimize) the viability of equilibria.

Similarly, why do losing voters bother to vote?
Lecture topics:
Definitions and properties (1)
Subjective viability assessments (1)
Behavioral observations (1)
Incomplete info game (1)
Comparisons with standard GT (1)
Rational coalitional defections (1/2)
Forming/switching equilibrium (1/4)
Implementations with faulty players (1/4)
Viability in network games (1/4)
Future research (1/2)
Proposed experiment (1/4)
Comparisons with gt refinements in a confession game

A Confession Game

$n = 36$ crime participants, confess or not. nobody confesses $\rightarrow$ all go free.
some confess $\rightarrow$ each gets 10yrs jail, but confessor only 3yrs.

Equil: nobody confesses, $nC$.

Defection deterrence: $D(nC) = 1$. 
Game theory laureates: \( nC \) is Perfect, Proper, Strong Nash.

\( nC \) is coalition proof

But mafias say \( nC \) is not viable.

Despite passing the GT refinements, \( D(nC) = 1 \rightarrow nC \) is barely sustainable.

Somebody will confess, if just for the fear that others would.

Killing confessors should make \( nC \) a dominant strategy.
Comparison with stochastic stability in match the boss game.
The boss $B$ and $n$ subordinates, each chooses a language. $B$’s payoff: 1 if he chooses $E$, 0 otherwise. Subordinate’s payoff: 1 if she matches $B$, 0 otherwise.

All choose $E$, $eE$, is \textit{stochastically stable} a la Young (1993), KMR (1993), and other basin-of-attraction arguments.

But still, $eE$, is \textit{only minimally sustainable}, $D(eE) = 1$,

\textbf{Why?} The subordinates depend on the choice of one player, B, who may be unreliable.
Improvement: have a committee of three bosses, all prefer $E$, and $n$ subordinates, wishing to match most bosses, now:

$$D(eE) = 2 (> 1).$$

Evaluation committees, not deans
Replace dictators by politburos,…

But still, $eE$, is only minimally sustainable, $D(eE) = 1$,

Why? The subordinates depend on the choice of one player, B, who may be unreliable.
Lecture topics:
Definitions and properties (1)
Subjective viability assessments (1)
Behavioral observations (1)
Incomplete info game (1)
Comparisons with standard GT (1)
Rational coalitional defections (1/2)
Forming/switching equilibrium (1/4)
Implementations with faulty players (1/4)
Viability in network games (1/4)
Future research (1/2)
Proposed experiment (1/4)
Rideshare game

8 players, each can (1) take a taxi for $80, or (2) ride the bus. The Bus costs, $180, will be shared equally by the bus choosers.

Equilibrium: Everybody takes a taxi, \( eT \). \( D(eT) = 2 \)

How is the cost of the taxi related to \( D(eT) \)?

Dominant

\[ D(eT) = 0, 1, \ldots, 8, \text{ are all possible.} \]
When taxi cost = $80, $D(e_T) = 2$, why?

A 2 player defection is a loss, cost go up from $80 to $90; but a 3 player defection may be a gain, cost go down from $80 to $60.

As we vary the taxi cost, we obtain all values $D(e_T) = 0, ..., 8$. 

<table>
<thead>
<tr>
<th># of bus riders</th>
<th>cost/rider $x$</th>
<th>cost of taxi $c$</th>
<th>Deterrence $D(e_T)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>180</td>
<td>$180 &lt; c$</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>90</td>
<td>$90 &lt; c \leq 180$</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>60</td>
<td>$45 &lt; c \leq 60$</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>45</td>
<td>$36 &lt; c \leq 45$</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>36</td>
<td>$30 &lt; c \leq 36$</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>30</td>
<td>$25.7 &lt; c \leq 30$</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>25.7</td>
<td>$22.5 &lt; c \leq 25.7$</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>22.5</td>
<td>$c \leq 22.5$</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8</td>
</tr>
</tbody>
</table>
These are **rational incentives** to defect, simple enough for bus riders, politicians,…

<table>
<thead>
<tr>
<th># of bus riders</th>
<th>cost/rider $180/x</th>
<th>cost of taxi $c</th>
<th>Deterrence $D(eT))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>180 $&lt; c$</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>180</td>
<td>90 $&lt; c$ $\leq$ 180</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>90</td>
<td>60 $&lt; c$ $\leq$ 90</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>60</td>
<td>45 $&lt; c$ $\leq$ 60</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>45</td>
<td>36 $&lt; c$ $\leq$ 45</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>36</td>
<td>30 $&lt; c$ $\leq$ 36</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>30</td>
<td>25.7 $&lt; c$ $\leq$ 30</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>25.7</td>
<td>22.5 $&lt; c$ $\leq$ 25.7</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>22.5</td>
<td>$c$ $\leq$ 22.5</td>
<td>8</td>
</tr>
</tbody>
</table>

Vote against Kavanaugh?
Lecture topics:
Definitions and properties (1)
Subjective viability assessments (1)
Behavioral observations (1)
Incomplete info game (1)
Comparisons with standard GT (1)
Rational coalitional defections (1/2)
Forming/switching equilibrium (1/4)
Implementations with faulty players (1/4)
Viability in network games (1/4)
Future research (1/2)
Proposed experiment (1/4)
The Ridesharing game: 8 players, each can (1) take a taxi, for $80; or (2) ride the bus. The bus costs, $180, will be shared equally among the bus choosers.

Everybody bus, eB, has $F(eB) = 2$.

The small $F(eB)$ is likely to cause a switch to the bus:

**EX.1** The bus company guarantees that the first 2 bus choosers will pay at most $75 each. Two will take it and make the bus the dominant strategy for the remaining players, all paying $22.5/rider.

**EX.2** Seeing the possible savings, 2 players initiate the switch to the bus on their own, without waiting for the bus company...
Unlikely switch

Switching to everybody choosing Mandarin, $eM$, in the language-choice game is unlikely. $n=200M$ symmetric players choosing a language. $D(eM) = 100M$, highly sustainable.

But you need to first convince $F(eM) = (200M - 100M) = 100M$ players.

Switching to ‘everybody chooses Mandarin` is unlikely, it requires the commitment of $100M$ players to make it dominant for the rest.
No Switch: A Sad Example

 Attempts to switch the US out of its measurement system to metric keep failing.

• In 1866, an act of the US Congress made it "lawful throughout the United States of America to employ the weights and measures of the metric system...“
  
  
  
• In 1975 Congress enacted the Metric Conversion Act but left the conversion voluntary. This attempt failed too, and the resistance to Metric continues. The high formation index may require a legal action to make people switch.
Lecture topics:
Definitions and properties (1)
Subjective viability assessments (1)
Behavioral observations (1)
Incomplete info game (1)
Comparisons with standard GT (1)
Rational coalitional defections (1/2)
Forming/switching equilibrium (1/4)
Implementations with faulty players (1/4)
Viability in network games (1/4)
Future research (1/2)
Proposed experiment (1/4)
Equilibrium with “faulty players”


Gradwohl and Reingold, ”Fault Tolerance in Large Games,” *GEB*, 2014.
Implementation with \( k \) faulty players

Eliaz (2002): \( n \)-player implementation method, when \( k \) unknown players may be faulty, i.e., irrational and unpredictable. Uses “\( k \)-Fault Tolerant Nash Eq.”, \( k \)-FTNE: where the \( n - k \) rational players have the Nash incentives to play an equilibrium \( \pi \), regardless of the strategies of the \( k \) faulty unknown players.

In our terminology, the \# rational players \( \geq \) Nash critical mass, or:

\[
\pi \text{ is a } k\text{-FTNE iff } n - k \geq \text{Nash c.m. } = F(\pi) + 1 = n - D(\pi) + 1
\]

\[
\therefore \pi \text{ facilitates Eliaz } k\text{-faulty-player implementation iff } D(\pi) > k.
\]

**DIFFERENCE** Our potential defectors may be **faulty** as in Eliaz, but may also be **rational**. Allows **new applications** such as: bribing and/or threatening rational players, defections by rational coalitions, equilibrium formation/switching by rational players, ...
Lecture topics:
Definitions and properties (1)
Subjective viability assessments (1)
Behavioral observations (1)
Incomplete info game (1)
Comparisons with standard GT (1)
Rational coalitional defections (1/2)
Forming/switching equilibrium (1/4)
Implementations with faulty players (1/4)
Viability in network games (1/4)
Future research (1/2)
Proposed experiment (1/4)
Graph matching games

Equilibrium viability in social networks.
Matching neighbors in a graph

Γ: a directed graph; V - set of n vertices; E - set of edges.
The set of (out) neighbors of ν, η(ν) ≡ {w ∈ V|(ν, w) ∈ E}.

The Γ matching game:
V - the set of players; each choose a language.
Payoff: the number of your neighbors you match.
Equilibrium: everybody choosing E, eE.

THEOREM  Let \( \hat{V} = \{v|\eta(v) \neq \emptyset\} \), the connected vertices.
\( D(eE) \): If \( \hat{V} = \emptyset \), \( D(eE) = n \).
If \( \hat{V} \neq \emptyset \), then \( D(eE) = \min_{v \in \hat{V}} \left[ \frac{|\eta(v)|}{2} \right] + 1 \).
= the strict majority of neighbors of the most vulnerable player, i.e., one with a minimal number of neighbors.
Decentralized equilibria are more sustainable

A currency match game with \( n \) tradors:
An equilibrium, everybody chooses \( D, eD \).

In a complete graph \( D(eD) \approx n/2 \)

Trading through middleman: \( D(eD) = 1 \).

Similarly for political, communications, supply chains, ...
Lecture topics:
Definitions and properties (1)
Subjective viability assessments (1)
Behavioral observations (1)
Incomplete info game (1)
Comparisons with standard GT (1)
Rational coalitional defections (1/2)
Forming/switching equilibrium (1/4)
Implementations with faulty players (1/4)
Viability in network games (1/4)
Future research (1/2)
Proposed experiment (1/4)
Summary and Future research

That two simple theoretic indices, $D$ and $F$ explain and predict observed strategic behavior, suggests an important optimistic research agenda:

_Develop game theory tools to predict/explain actual strategic behavior._

Some open problems

• Axiomatize $D$, $F$ (or other indices).
• Expand Nash existence theorem.
• Advance more refined indices.
• Non anonymous indices and equil formation.
• Sequential equilibrium formation.
Nash’s Theorem restated: In any finite $n$ person game there exists a profile $\pi$ with $D(\pi) \geq 1$.

For applications in which equilibria $\pi$ with $D(\pi)=1$ are not sufficiently reliable, we need:

Important open problem: For a finite $n$ person strategic game $\Gamma$, find sufficient conditions for the existence of equilibrium $\pi$ with $D(\pi) > 1$. 

Extending Nash’s existence theorem
Limitations of $D$ (all shared with Nash eq)

**Needed Continuity:**
for a $D(eE)$ that communication viability in the two populations.

$D = 1$

$D = \frac{n}{2}$

**Needed Discontinuity:**
Unlike the lhs game, a single defection in the rhs game totally destroys the equilibrium.

$D = 1$

**Non anonymous indices:** consider the identity of defectors and their position in the game.
Equil formation & non-anonymous indices

A coalition $G$ is a generator of $\pi$, if the play of $\pi$ by $G$ makes $\pi$ a dominant strategy for the rest of the players, i.e., $G$ forms $\pi$.

Notice
1. **Monotonicity**: If $G$ is a $\pi$-generator then so is $G'$, for $G' \supseteq G$. $N$ is the unique maximal $\pi$-generator under containment.
2. **Simple structure**: There is a finite number of minimal (by containment) generators, called roots of $\pi$. A coalition is a generator iff it contains a root.
3. **Non-anonymous formation**, (where you convince a root to play $\pi$) may be more efficient than the anonymous formation (where you convince $F(\pi)$ players to play $\pi$).
Illustration: the divisive equilibrium.

Recall: The divisive equilibrium in the Party Line game, \( Div \), in which 3 Ds choose F and 5 Rs choose E. Roots: The 3 Ds, The 5 Rs any group of 2Ds and 3Rs, \( \{2Ds&3Rs\} \).

Alternative ways to form \( Div \)
1. Convince the 3 Ds to choose F.
2. Convince the 5 Rs to choose E.
3. Convince any 2 Ds and any 3Rs to play \( Div \).
4. Convince any 6 players to play \( Div \).

(1, 2, and 3) target the roots.
(4) relies on the anonymous index \( F(Div) = 6 \).

Non-anonymous methods require more computations.
Illustration: the divisive equilibrium.

Recall: T divisive equilibrium in the Party Line game, \( Div \), in which the 3 Ds choose F and the 5 Rs choose E.

**Sequential equilibrium formation:** for example, convince two Ds, this will convince all the Rs, which in turn will convince the third D.
There are multiple paths for the order and choices of coalitions to convince, harder to compute ...

For example, the “genius of a dictator” is the skill to navigate a sequential formation process that lead to an equilibrium in which all the players obey her wishes.
Lecture topics:
Definitions and properties (1)
Subjective viability assessments (1)
Behavioral observations (1)
Incomplete info game (1)
Comparisons with standard GT (1)
Rational coalitional defections (1/2)
Forming/switching equilibrium (1/4)
Implementations with faulty players (1/4)
Viability in network games (1/4)
Future research (1/2)
Proposed experiment (1/4)
Proposed Experiment:
20-Player Participation Game with threshold $t$

$t$ - a non-negative integer, $t \geq 1$.
Simultaneously, each player chooses to participate or not.
Non-participant’s payoff = $10$.
Each participant’s payoff = \[
\begin{cases} 
$20$, if the # of participants $> t, \\
0$, otherwise.
\end{cases}
\]

Everybody participates, $eP$, is a Nash eq for $t = 1, \ldots, 19$, with formation difficulty index $F(eP) = t$.

**Conjecture:** Participation rates should decrease as $t$ increases.

**Question:** When do they $\approx 0$?

Experiment question: would the number of participants decrease as the threshold $t$ is increased?
Basin of attraction of an equilibrium $\pi$, with $D(\pi) = 1$, in a 3 player game.

A 1 dimensional cross in a 3 dimension strategy space.
Basin of attraction of an equilibrium $\pi$, with $D(\pi) = 2$, in a 3 player game.

A 2 dimensional cross in a 3 dimension strategy space.

domination of player 1
domination of player 2
domination of player 3
defections of Pl’s 1 and 2
defections of Pl’s 1 and 3
defections of Pl’s 2 and 3