Behavioral Economics:
Theory and Evidence on Adaptive Learning

University of Oxford, Michaelmas Term 2013

and

Asia Summer Institute in Behavioral Economics
National University of Singapore, 22 July-2 August 2013

Vincent P. Crawford, University of Oxford, All Souls College, and University of California, San Diego

(with thanks to Colin Camerer, Teck-Hua Ho, and Juin-Kuan Chong)

Revised 7 July 2013; minor corrections 29 October 2013

Introduction

Adaptive learning models describe how players adjust their decisions over time in response to their experience with analogous games.

The learning process is usually modeled as repetition of a fixed “stage game”, so the analogies are perfect; some recent work relaxes that.

Players view their decisions in the stage game as the objects of choice, and the dynamics of their decisions are modeled either directly, or indirectly in terms of their beliefs with decisions best responding.

In either case, decisions are usually allowed to be “noisy” in the sense that the model is taken to describe the mean decision, with errors.

I focus on *adaptive* learning models because they appear to be the most useful for understanding behavior.

But I will also discuss how adaptive learning models relate to other models of learning, such as “rational learning” or “long-run equilibrium”.

An adaptive learning model has two main components:

● A model of players' interaction patterns
● A model of how players adjust decisions in response to experience

Models of interaction patterns follow evolutionary game theory—the main reason evolutionary models are interesting for economics, in my view.

I first discuss evolutionary models, illustrating the effects of interaction patterns under the standard assumption that the population frequencies of players' decisions increase with their payoffs in the current population.

(That assumption is at least approximately true for the leading adaptive learning models, but it is only a proxy for a detailed model of learning.)

I next discuss more detailed adaptive learning models of how players adjust their decisions, given the population's interaction pattern.

I conclude with an analysis of some especially informative experiments.

“Evolutionary” models of players’ interaction patterns

In evolutionary models, a population or populations of players repeatedly play a given “stage game”, without or with distinguished player roles. E.g.

- A two-person stage game is played repeatedly by pairs of players randomly drawn from a single population, or with asymmetric player roles randomly filled from separate, distinguished populations.
- Or an *n*-person stage game is played repeatedly by an entire population or (with distinguished player roles) populations.

Players’ payoffs in a given stage are assumed to be determined by their own actions and the population action frequencies.

(This is implied by random pairing, assuming expected-payoff maximization; but it is restrictive for other interaction patterns.)

Players with a given observable role label are identical but for actions.

For simplicity, players are usually assumed to play only pure actions.

In evolutionary models, players' actions are not chosen but inherited.

In adaptive learning models actions are chosen, but not fully rationally in the sense of equilibrium in the game that describes the entire learning process ("rational learning"); instead action choices are shaped by their payoffs in the current population.

Other important differences from traditional game theory include:

- The population is the unit of analysis, rather than the individual.
- Labeling of actions and/or populations has substantive implications, because labels are the "language" in which players (implicitly or explicitly) recognize analogies between current and previous games, hence the language in which they encode their experience.

A final important difference pertains to the scientific "cultures":

- Traditional game theory grew out of von Neumann's analysis of games of pure conflict, and is at its best in zero-sum two-person games.
- Evolutionary game theory grew out of analyses of coordination in games like Battle of the Sexes (the "Hawk-Dove" game), and this gives it and adaptive learning models advantages in analyzing coordination.

In evolutionary game theory the law of motion of the population action frequencies is derived, with a functional form known as the "replicator dynamics", from the assumption that players inherit their actions from parents who reproduce at rates ("fitnesses") equal to current payoffs.

Put another way, if we describe game outcomes as fitnesses, in simple evolutionary models the replicator dynamics are like accounting identities, so that the action frequencies must follow them by definition.

In evolutionary game theory and simple adaptive learning models, dynamics other than the replicator dynamics are allowed, but the population action frequencies usually respond to payoff differences in a way that is qualitatively similar to the replicator dynamics.

(The distinction between replicator and non-replicator dynamics can be important for some purposes, but it will not matter much here.)

The goal of an evolutionary analysis is usually taken to be identifying the locally stable steady states of the replicator dynamics.

(The notion of "evolutionary stability" is an imperfect static proxy for this.)

If the dynamics converge, they must converge to a steady state in which the actions that persist are optimal in the stage game, given the limiting action frequencies; thus, the limiting frequencies are in Nash equilibrium.

Remarkably, even though players' actions are not rationally chosen—not even chosen!—the population collectively "learns" the equilibrium as its frequencies evolve, with selection doing the work of strategic thinking.

Note that although the limiting action frequencies are in Nash equilibrium, individuals' actions need not be.

If for example individuals play only pure strategies in a game with a mixed-strategy equilibrium, then the limiting population action frequencies mimic a purified "equilibrium in beliefs".

This link between steady states of the dynamics and Nash equilibria has a counterpart for adaptive learning dynamics:

In the simplest models, adaptive learning cannot converge to population frequencies that are not in Nash equilibrium.

And though there are no general convergence results, adaptive learning dynamics have a strong tendency to converge to *some* Nash equilibrium.

The interesting question in most applications is, Which equilibrium?

The answer depends both on players' interaction patterns, and on the details of the learning rules by which players adjust their decisions.

Identifying the "right" evolutionary model of how players interact goes a long way toward analyzing the dynamics, as I illustrate next.

But the details of learning also matter, as I will illustrate later.

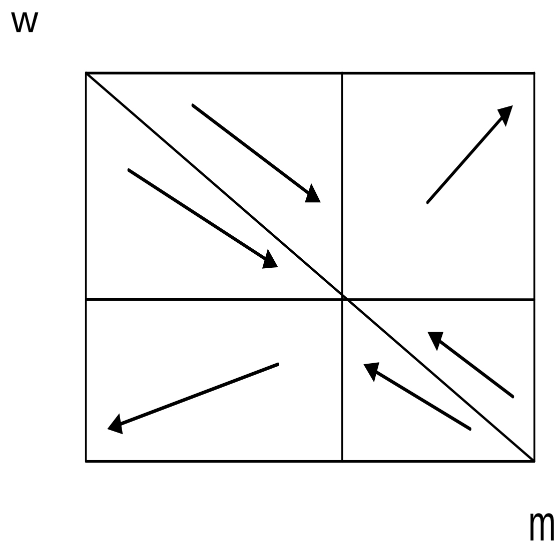
Evolutionary dynamics

First imagine a large population of men and women repeatedly and anonymously paired (heterosexually, and with gender observable so they can base their actions on it) to play Battle of the Sexes ("BoS").

	Fights	Ballet
Fights	2, 1	0, 0
Ballet	0, 0	1, 2

Battle of the Sexes

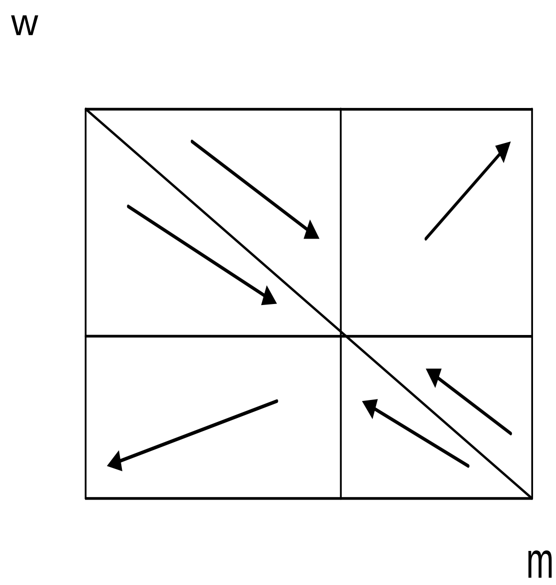
Draw a differential equation phase diagram with the population frequency of men playing Fights, *m*, on the horizontal axis and the frequency of women playing Fights, *w*, on the vertical axis.



The diagram allows a simple analysis of adaptive learning dynamics.

There are four regions of the state space:

$m > 2/3$, $w > 1/3$
$m > 2/3$, $w < 1/3$
$m < 2/3$, $w > 1/3$
$m < 2/3$, $w < 1/3$.

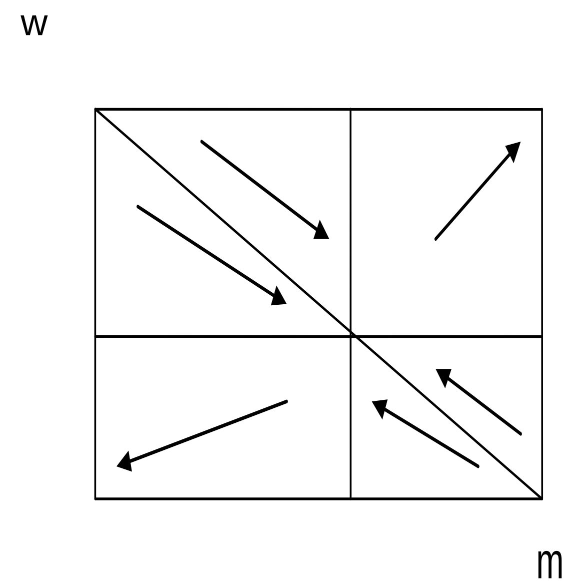


For men the expected payoff of Fights is higher than Ballet whenever $w > 1/3$ ($2w > 1 - w$).

For women the expected payoff of Fights is higher than Ballet whenever $m > 2/3$ ($m > 2(1 - m)$).

Thus when ($m > 2/3$, $w > 1/3$), m and w rise, so $m \rightarrow 1$ and $w \rightarrow 1$.

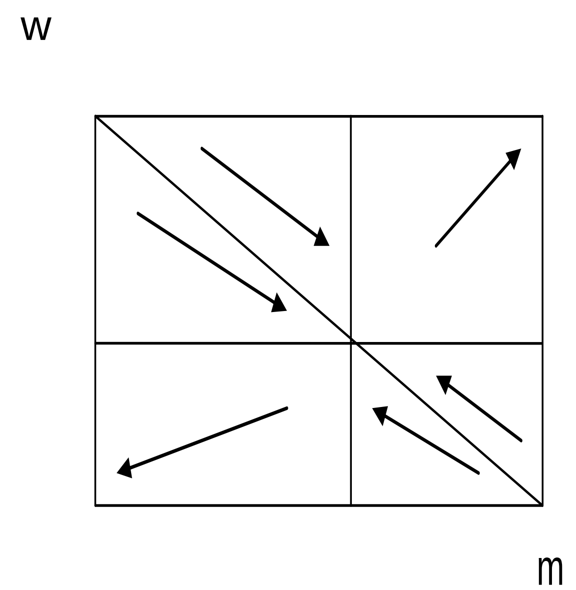
And when ($m < 2/3$, $w < 1/3$), m and w fall, so $m \rightarrow 0$ and $w \rightarrow 0$.



When (m > 2/3, w < 1/3), m falls and w rises; and when (m < 2/3, w > 1/3), m rises and w falls.

In these cases, if the initial condition is above the diagonal—if $m + w > 1$—the system enters (m > 2/3, w > 1/3), $m \to 1$, and $w \to 1$; if it's below the diagonal, the system enters (m < 2/3, w < 1/3), $m \to 0$, and $w \to 0$.

In all four cases the limiting outcome "is" one of BoS's two pure-strategy equilibria ("is" only in that the game that describes the populations' interactions is not BoS; but there is a simple correspondence).



In each case, in the limiting outcome people follow a convention based on labels, although labels are assumed irrelevant in traditional theory.

In deterministic evolutionary dynamics, which convention people follow is completely determined by whether the initial sum of the frequencies of arrogant men and wimpy women, $m + w > 1$.

Now consider a closely related model in which players are repeatedly and anonymously randomly paired from a large “unisex” population to play the same game with two pure-strategy equilibria, one favored by one player and the other by the other; but with no observable labeling.

Even without labels, players can distinguish actions via payoffs.

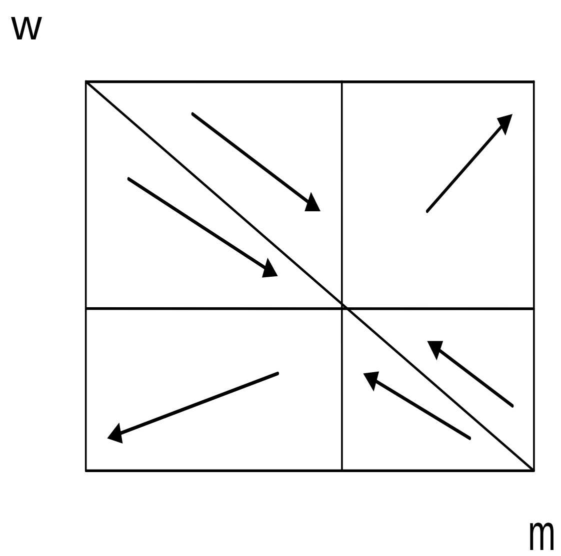
Call the action that *could* yield a player’s best outcome “Hawk”—in BoS, Fights for men and Ballet for women—and call the other action “Dove”.

With these labels BoS has symmetric player roles for men and women, an important convenience in modeling the difficulty of coordination:

	Hawk	**Dove**
Hawk	0, 0	2, 1
Dove	1, 2	0, 0

Hawk-Dove Game

One way to analyze the dynamics recycles the BoS phase diagram, adding the restriction that in a single, unisex population, the frequencies of Hawk in each player role are equal, like the previous frequencies *m* of men playing Fights or 1 – *w* of women playing Ballet; so that *m* + *w* = 1. This limits the dynamics to the northwest/southeast diagonal.



If it were possible to go off the diagonal, the symmetric mixed-strategy equilibrium of the game at Pr{Hawk} = 2/3 would be unstable, and players would converge to one of the pure-strategy equilibria.

But the dynamics must now converge to the intersection in the center, the symmetric mixed-strategy equilibrium of the game at Pr{Hawk} = 2/3.

The lack of observable labels disables the gender- or decision-based convention the population used before to break the symmetry of roles as needed for efficient coordination, completely changing the outcome.

With no observable labeling, the off-diagonal states, including the pure-strategy equilibria, are not even in the state space.

Off-diagonal states are ruled out, even though they could occur with positive probability, because when players are randomly paired from a single population with no observable labeling, there can be no *systematic* difference in the frequencies of Hawk in each player role.

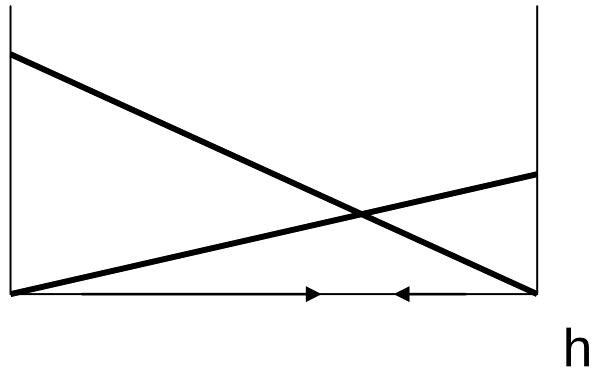
This illustrates how evolutionary game theory models the difficulty of coordination, using a careful account of what labels players observe to rule out “magical” coordination usually allowed in traditional analyses.

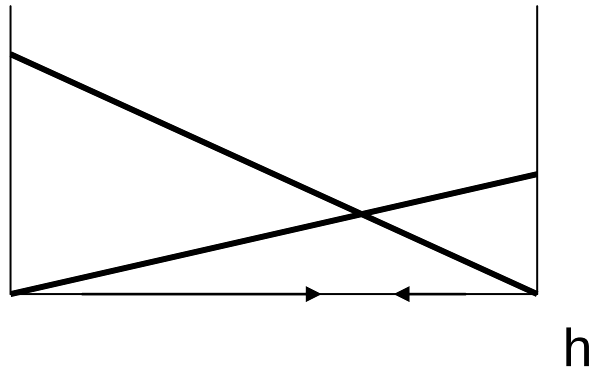
That evolutionary game theory has machinery to do this in a systematic way is an important advantage—although such magical coordination is now often ruled out in traditional analyses on an ad hoc, intuitive basis.

Another, equivalent way to analyze the dynamics graphs the expected payoffs of Hawk and Dove in either role against the frequency of Hawk *h*.

This builds in the restriction that with unisex labeling, the frequency of players playing Hawk must be the same in both roles.

This restriction allows us to represent the dynamics in a one-dimensional phase diagram, with the expected payoffs of Hawk and Dove on the vertical axis and the population frequency of Hawk on the horizontal axis.





When the frequency of Hawk is low, Hawk has higher payoff than Dove; and vice versa.

Thus the dynamics follow the arrows on the horizontal axis, converging to the frequency of Hawk where the payoff lines cross, which is Pr{Hawk} = 2/3, again as in the game's mixed-strategy Nash equilibrium.

These simple examples illustrate the methods of evolutionary analyses, assuming only that the population frequencies of players' decisions increase with the decisions' payoffs against the current population.

I will adapt those methods to analyze Van Huyck et al.'s coordination experiments, where the details of learning play an essential role; but first a brief introduction to the leading adaptive learning models.

Adaptive learning models

Adaptive learning models are inherently "irrational" from the viewpoint of Nash equilibrium (*not* necessarily from the viewpoint of decision theory).

They implicitly assume that a player thinks other players' decisions won't adjust, even though the player knows that he himself is adjusting.

And they allow a player's adjustments to be only partial.

I will show that something like these "irrational" (though sensible) features are needed to elucidate observed behavior.

Adaptive learning models do share the rationality-based feature that a player's probability of making a decision increases with the payoff that decision yielded, or would have yielded, in the current population.

There are three leading classes of adaptive learning models:

- ● Reinforcement learning
- ● Beliefs-based learning
- ● Experience-weighted attraction (“EWA”) learning

(Beyond the scope of these lectures are interesting and largely orthogonal refinements such as Camerer, Ho, and Chong’s 2002 *JET* “strategic teaching”, Stahl’s 1996 *GEB* “rule learning”, and Selten’s 1991 and Conlisk’s 1993ab *JEBO* “sophisticated learning”.)

All three leading models can be described by assuming that strategies have numerical “attractions,” which determine their choice probabilities in the same way for all three models.

Given this, specifying an adaptive learning model requires specifying each player’s initial attractions, how he updates them in response to his experience, and how his choice probabilities depend on the attractions.

Initial attractions are the subject of strategic thinking, discussed later; I focus here on updating and choice probabilities.

Reinforcement (or “stimulus-response” or “rote”) learning

Reinforcement learning was originally developed to describe the behavior of (mostly non-human) laboratory animals in simple settings.

(Snapshots here and below from Camerer and Ho 1999 *Econometrica.*)

The initial reinforcement level of strategy s_i^j of player i is $R_i^j(0)$. These initial reinforcements can be assumed a priori (based on a theory of first-period play) or estimated from the data. Reinforcements are updated according to two principles:

$$(2.3) \qquad R_i^j(t) = \begin{cases} \phi \cdot R_i^j(t-1) + \pi_i(s_i^j, s_{-i}(t)) & \text{if} \quad s_i^j = s_i(t), \\ \phi \cdot R_i^j(t-1) & \text{if} \quad s_i^j \neq s_i(t). \end{cases}$$

In reinforcement and other models, attractions determine choice probabilities as follows (with attractions A(·) generalizing reinforcements R(·) to other models):

Attractions must determine probabilities of choosing strategies in some way. $P_i^j(t)$ should be monotonically increasing in $A_i^j(t)$ and decreasing in $A_i^k(t)$ (where $k \neq j$). Three forms have been used in previous research: Exponential (logit), power, and normal (probit). In estimation reported below we use the logit function, which is commonly used in studies of choice under risk and uncertainty, brand choice, etc. (Ben-Akiva and Lerman (1985), Anderson, Palma, and Thisse (1992)), and is given by

$$(2.11) \qquad P_i^j(t+1) = \frac{e^{\lambda \cdot A_i^j(t)}}{\sum_{k=1}^{m_i} e^{\lambda \cdot A_i^k(t)}}.$$

The parameter λ measures sensitivity of players to attractions. Sensitivity could vary due to the psychophysics of perception or whether subjects are highly motivated or not. In this probability function, the exponent in the numerator is just the weighted effect of strategy s_i^j's attraction, $\lambda \cdot A_i^j(t)$, on the probability of choosing strategy s_i^j. Models in which cross-effects of attractions on other

In reinforcement learning, strategies' attractions are updated only according to the payoffs received for the strategies actually played:

The initial reinforcement level of strategy s_i^j of player i is $R_i^j(0)$. These initial reinforcements can be assumed a priori (based on a theory of first-period play) or estimated from the data. Reinforcements are updated according to two principles:

$$(2.3) \qquad R_i^j(t) = \begin{cases} \phi \cdot R_i^j(t-1) + \pi_i(s_i^j, s_{-i}(t)) & \text{if} \quad s_i^j = s_i(t), \\ \phi \cdot R_i^j(t-1) & \text{if} \quad s_i^j \neq s_i(t). \end{cases}$$

(The choice probabilities of other strategies can still change via (2.11).)

Reinforcement learning is defined without reference to the structure of the game, or even to whether players know they are playing a game.

Thus there can be no hypothetical "what if I had done that instead?" thought experiments of the kind humans often seem to engage in when trying to learn from experience. (Such thinking depends on the structure.)

Despite these limitations, reinforcement learning fits observed human behavior well in some games, such as simple matrix games with mixed-strategy equilibria.

Reinforcement learning tends to fit best when subjects have little information about the structure, but it fits well in some games even when subjects know the structure.

But reinforcement learning often fits poorly in games with larger strategy spaces; and it often adjusts far too slowly to describe human behavior, as it would in Van Huyck et al.'s experiments discussed below.

(One might argue, uncharitably, that reinforcement learning fits well in settings that leave human subjects bored or confused.)

Beliefs-based learning

In beliefs-based models such as "fictitious play", players are assumed, by contrast, to know the structure of the game.

Players update probabilistic beliefs about others' likely strategies from past experience, calculate the expected payoffs of their own strategies, and choose strategies that have higher payoffs with higher probabilities.

Fictitious play assumes others' strategies are drawn from a stationary distribution, so estimated beliefs are simple averages of past history.

At the other extreme is "Cournot dynamics" (best responding to others' last-period play), which discounts all history but the most recent.

But beliefs-based learning includes many more sophisticated rules, which weight past experience in more plausible ways.

For example, in Camerer and Ho's (1999) treatment of beliefs,

We also allow past experience to be depreciated or discounted by a factor ρ (presumably between zero and one). Formally, the prior beliefs for player i about choices of others are specified by a vector of relative frequencies of choices of strategies s^k_{-i}, denoted $N^k_{-i}(0)$. Call the sum of those frequencies (dropping the player subscript for simplicity) $N(t) = \sum_{k=1}^{m_{-i}} N^k_{-i}(t)$. Then the initial prior $B^k_{-i}(0)$ is

$$(2.6) \qquad B^k_{-i}(0) = \frac{N^k_{-i}(0)}{N(0)},$$

with $N^k_{-i}(0) \geq 0$ and $N(0) > 0$. Beliefs are updated by depreciating the previous counts by ρ, and adding one for the strategy combination actually chosen by the other players. That is,

$$(2.7) \qquad B^k_{-i}(t) = \frac{\rho \cdot N^k_{-i}(t-1) + I(s^k_{-i}, s_{-i}(t))}{\sum_{h=1}^{m_{-i}} \left[\rho \cdot N^h_{-i}(t-1) + I(s^h_{-i}, s_{-i}(t)) \right]}.$$

EWA (experience-weighted attraction) learning

In EWA learning, just as in reinforcement or beliefs-based learning, strategies' attractions reflect prior predispositions; and attractions determine choice probabilities according to a rule like the logit in (2.11).

EWA nests beliefs-based and reinforcement models of learning.

The key idea is "hypothetical reinforcement" of strategies that were not chosen, according to the payoff they would have yielded.

Hypothetical reinforcement is allowed to have a different weight than reinforcement and beliefs-based models' reinforcement of the strategies that were actually chosen.

In Camerer and Ho (1999) (with parts that are also pertinent to more flexible specifications of beliefs-based models other than EWA):

The core of the EWA model is two variables which are updated after each round. The first variable is $N(t)$, which we interpret as the number of 'observation-equivalents' of past experience. The second variable is $A_i^j(t)$, player i's attraction of strategy s_i^j *after* period t has taken place.

The variables $N(t)$ and $A_i^j(t)$ begin with some prior values, $N(0)$ and $A_i^j(0)$.These prior values can be thought of as reflecting pregame experience, either due to learning transferred from different games or due to introspection. (Then $N(0)$ can be interpreted as the number of periods of actual experience, which is equivalent in attraction impact to the pregame thinking.)

Updating is governed by two rules. First,

$$(2.1) \qquad N(t) = \rho \cdot N(t-1) + 1, \quad t \geq 1.$$

The parameter ρ is a depreciation rate or retrospective discount factor that measures the fractional impact of previous experience, compared to one new period.

The second rule updates the level of attraction. A key component of the updating is the payoff that a strategy either yielded, or would have yielded, in a period. The model weights hypothetical payoffs that unchosen strategies would have earned by a parameter δ, and weights payoffs actually received, from chosen strategy $s_i(t)$, by an additional $1 - \delta$ (so they receive a total weight of 1). Using an indicator function $I(x, y)$ that equals 1 if $x = y$ and 0 if $x \neq y$, the weighted payoff can be written as $[\delta + (1 - \delta) \cdot I(s_i^j, s_i(t))] \cdot \pi_i(s_i^j, s_{-i}(t))$.

The rule for updating attraction sets $A_i^j(t)$ to be the sum of a depreciated, experience-weighted previous attraction $A_i^j(t-1)$ plus the (weighted) payoff from period t, normalized by the updated experience weight:

$$(2.2) \qquad A_i^j(t) = \frac{\phi \cdot N(t-1) \cdot A_i^j(t-1) + [\delta + (1-\delta) \cdot I(s_i^j, s_i(t))] \cdot \pi_i(s_i^j, s_{-i}(t))}{N(t)}.$$

The factor ϕ is a discount factor or decay rate, which depreciates previous attraction.

Thus the strength of EWA's hypothetical reinforcement is tuned by a parameter δ; δ = 0 is reinforcement learning and δ = 1 is a class of weighted fictitious play beliefs-based learning models.

The decay parameters φ and ρ in (2.1)-(2.2) depreciate attractions and the amount of experience; φ = ρ is belief-based; ρ = 0 is reinforcement.

EWA combines the best features of reinforcement and beliefs-based models: allowing attractions to begin and grow flexibly as reinforcement does, but reinforcing strategies not chosen as belief-based models do.

In applications typical estimates of δ are around .50, of φ around 1, and of ρ from 0 to φ, showing that the generalization to EWA has bite.

Because EWA is nonlinear in δ, the resulting models are (much) more than a linear combination of beliefs-based and reinforcement models; and in applications EWA often fits better than both extremes.

See for example Camerer and Ho's 1998 *J. Mathematical Psychology* EWA analysis of Van Huyck et al.'s coordination experiments, discussed below from Crawford's 1995 *Econometrica* beliefs-based point of view.

Reinforcement, beliefs-based, and EWA learning models have similar convergence properties.

Although general results are elusive, and convergence may depend on the details of the game and the stochastic structure, these models all normally converge to some steady state pattern in the stage game.

Moreover, when the stage game is a normal-form game, these models can converge only to a Nash equilibrium of that game (with qualifications regarding equilibrium in beliefs for mixed-strategy equilibria).

When the stage game is a nontrivial extensive-form game, convergence may be to a generalization known as self-confirming equilibrium, which allows deviations from subgame-perfect or sequential equilibrium in parts of the game tree where play does not test strategies' performance.

Just as it is surprising that evolutionary dynamics converge to Nash equilibrium, it may be surprising that reinforcement learning usually converges to Nash equilibrium, because reinforcement learners lack much of the information about the game that its equilibria depend on.

But Hopkins 2002 *Econometrica* shows that the *expected* motions of particular versions of stochastic fictitious play and reinforcement learning with experimentation can both be written as a perturbed form of the evolutionary replicator dynamics. This result probably holds for EWA too.

Reinforcement learning yields the same adjustments *on average* over time as stochastic fictitious play, even though fictitious play's adjustment each period depends on details that reinforcement learning ignores.

This suggests that the details of learning rules matter mainly in how they affect probabilities of "tunneling": jumping between basins of attraction.

Tunneling probabilities do vary with the details; more on this below.

Van Huyck, Battalio, and Beil's 1990 *AER*, 1991 *QJE* experiments

Repeated play of symmetric coordination games in populations of subjects, interacting all at once in "large groups", or with random pairing.

Subjects chose simultaneously among 7 efforts, with payoffs and optimal choices determined by their own efforts and the median or minimum effort in large groups, or the current pair's minimum with random pairing.

Explicit communication was prohibited, the median or minimum was publicly announced after each play (random pairs told only their own pair's minimum), and the structure was publicly announced at the start.

There were five leading treatments, using the minimum in 1990 and the median in 1991, varying the size of the population and their interaction patterns in ways that make the results much more informative.

(Here I focus on the "C_d" treatment with repeatedly random pairing, but contrast with the "C_f" treatment's initially random but then fixed pairing.)

The stage games have seven strict, symmetric, Pareto-ranked equilibria.

The stage games are like a meeting that everyone would prefer to start on time, but which can't start until a given quorum is achieved—100% in the large-group minimum game, 50% in the large-group median games.

Although there is an "obviously" right way to play, the Pareto-superior equilibrium is intuitively more fragile, the larger the quorum or the group.

In the experiments—very surprisingly at the time, because the experiments were far ahead of theory—coordination was less efficient, the more fragile was the Pareto-superior equilibrium.

Traditional equilibrium analysis and refinements (with exceptions like Harsanyi and Selten's 1988 risk-dominance) don't address this issue; and more recent refinements come only slightly closer.

PAYOFF TABLE Γ

		Median value of X chosen						
		7	6	5	4	3	2	1
Your choice of X	7	1.30	1.15	0.90	0.55	0.10	−0.45	−1.10
	6	1.25	1.20	1.05	0.80	0.45	0.00	−0.55
	5	1.10	1.15	1.10	0.95	0.70	0.35	−0.10
	4	0.85	1.00	1.05	1.00	0.85	0.60	0.25
	3	0.50	0.75	0.90	0.95	0.90	0.75	0.50
	2	0.05	0.40	0.65	0.80	0.85	0.80	0.65
	1	−0.50	−0.05	0.30	0.55	0.70	0.75	0.70

PAYOFF TABLE A

		Smallest Value of X Chosen						
		7	6	5	4	3	2	1
Your Choice of X	7	1.30	1.10	0.90	0.70	0.50	0.30	0.10
	6	–	1.20	1.00	0.80	0.60	0.40	0.20
	5	–	–	1.10	0.90	0.70	0.50	0.30
	4	–	–	–	1.00	0.80	0.60	0.40
	3	–	–	–	–	0.90	0.70	0.50
	2	–	–	–	–	–	0.80	0.60
	1	–	–	–	–	–	–	0.70

VHBB's Leading Median and Minimum Payoff Tables

Results

The five leading treatments all evoked similar initial responses (table from Crawford 1991 *Games and Economic Behavior*, p. 55)).

TABLE I

		Minimum treatment				
		A (%)	*B* (%)	*A′* (%)	C_d (%)	C_f (%)
Subject's initial effort	7	33 (31)	76 (84)	23 (25)	11 (37)	13 (42)
	6	10 (9)	1 (1)	1 (1)	1 (3)	0 (0)
	5	34 (32)	2 (2)	2 (2)	2 (7)	6 (19)
	4	18 (17)	5 (5)	7 (8)	5 (17)	2 (6)
	3	5 (5)	1 (1)	7 (8)	3 (10)	1 (3)
	2	5 (5)	1 (1)	17 (19)	1 (3)	1 (3)
	1	2 (2)	5 (5)	34 (37)	7 (23)	8 (26)
Totals		107 (101)	91 (99)	91 (100)	30 (100)	31 (99)

		Median treatment		
		Γ, Γdm (%)	Ω (%)	Φ (%)
Subject's initial effort	7	8 (15)	14 (52)	2 (7)
	6	4 (7)	1 (4)	3 (11)
	5	15 (28)	9 (33)	9 (33)
	4	19 (35)	3 (11)	11 (41)
	3	8 (15)	0 (0)	2 (7)
	2	0 (0)	0 (0)	0 (0)
	1	0 (0)	0 (0)	0 (0)
Totals		54 (100)	27 (100)	27 (99)

Subjects almost always converged to *some* stage-game equilibrium.

But the dynamics and limiting outcomes varied with the order statistic, interaction pattern, and group size, with very large differences in drift, history-dependence, and rate of convergence:

- In 12 out of 12 large-group median runs, there was "lock-in" on the initial median, although it varied across runs and was usually inefficient

- In 9 out of 9 large-group minimum runs, there was strong downward drift, with subjects always approaching the least efficient equilibrium

- In 2 out of 2 random-pairing minimum runs, there was very slow convergence, no discernible drift, and moderate inefficiency

Comparing the first two reveals an "fragility" effect: coordination is less efficient, the smaller the groups that can disrupt efficient equilibria.

Comparing the last two reveals a "group size" effect: holding the order statistic constant, coordination is less efficient in larger groups.

TABLE III

MEDIAN CHOICE FOR THE FIRST TEN PERIODS OF ALL EXPERIMENTS

	Period									
Treatment	1	2	3	4	5	6	7	8	9	10
Gamma										
Exp. 1	4	4	4	4	4	4	4*	4	4*	4*
Exp. 2	5	5	5	5	5	5	5	5	5	5
Exp. 3	5	5	5	5	5	5	5	5	5	5*
Gammadm										
Exp. 4	4	4	4	4	4	4*	4*	4*	4*	4*
Exp. 5	4	4	4	4*	4*	4*	4*	4*	4*	4*
Exp. 6	5	5	5	5	5	5	5	5*	5*	5*
Omega										
Exp. 7	7	7	7	7*	7*	7*	7*	7*	7*	7*
Exp. 8	5	5	5	5	5*	5*	5*	5*	5*	5*
Exp. 9	7	7	7*	7*	7*	7*	7*	7*	7*	7*
Phi										
Exp. 10	4	4	4	4	4*	4*	4*	4*	4*	4*
Exp. 11	5	5	5	5*	5*	5*	5*	5*	5*	5*
Exp. 12	5	5	5	5*	5*	5*	5*	5*	5*	5*

Notes. Exp. = experiment. * = indicates a mutual best response outcome.

TABLE 2—EXPERIMENTAL RESULTS FOR TREATMENT *A*

	Period									
	1	2	3	4	5	6	7	8	9	10
Experiment 1										
No. of 7's	8	1	1	0	0	0	0	0	0	1
No. of 6's	3	2	1	0	0	0	0	0	0	0
No. of 5's	2	3	2	1	0	0	1	0	0	0
No. of 4's	1	6	5	4	1	1	1	0	0	0
No. of 3's	1	2	5	5	4	1	1	1	0	1
No. of 2's	1	2	2	4	8	7	8	6	4	1
No. of 1's	0	0	0	2	3	7	5	9	12	13
Minimum	2	2	2	1	1	1	1	1	1	1
Experiment 2										
No. of 7's	4	0	1	0	0	0	0	0	0	1
No. of 6's	1	0	1	0	0	1	0	0	0	0
No. of 5's	3	3	2	1	0	0	1	1	0	1
No. of 4's	4	6	2	3	3	0	0	0	0	0
No. of 3's	1	4	2	5	0	1	1	0	1	0
No. of 2's	3	2	6	5	5	9	3	4	3	1
No. of 1's	0	1	2	2	8	5	11	11	12	13
Minimum	2	1	1	1	1	1	1	1	1	1
Experiment 3										
No. of 7's	4	4	1	0	1	1	1	0	0	2
No. of 6's	2	0	2	0	0	0	0	0	0	0
No. of 5's	5	6	1	1	1	0	0	0	0	0
No. of 4's	3	3	2	1	2	1	0	0	0	1
No. of 3's	0	0	7	6	0	2	3	0	0	0
No. of 2's	0	1	1	4	5	3	6	3	2	2
No. of 1's	0	0	0	2	5	7	4	11	12	9
Minimum	4	2	2	1	1	1	1	1	1	1
Experiment 4										
No. of 7's	6	0	1	1	0	0	1	0	0	0
No. of 6's	0	6	2	0	0	1	0	0	0	0
No. of 5's	8	5	5	5	0	1	0	0	0	0
No. of 4's	1	1	4	6	7	1	2	1	1	0
No. of 3's	0	2	3	2	4	3	2	2	1	0
No. of 2's	0	1	0	0	2	3	7	4	2	2
No. of 1's	0	0	0	1	2	6	3	8	11	13
Minimum	4	2	3	1	1	1	1	1	1	1

TABLE 2—EXPERIMENTAL RESULTS FOR TREATMENT *A*, Continued

	Period									
	1	2	3	4	5	6	7	8	9	10
Experiment 5										
No. of 7's	2	2	3	1	1	1	1	0	0	0
No. of 6's	1	3	1	0	0	0	0	0	0	0
No. of 5's	9	3	0	4	1	0	2	0	0	0
No. of 4's	3	4	6	2	1	2	0	2	1	1
No. of 3's	1	2	2	4	6	0	0	0	0	1
No. of 2's	0	2	2	3	4	6	5	2	5	3
No. of 1's	0	0	2	2	3	7	8	12	10	11
Minimum	3	2	1	1	1	1	1	1	1	1
Experiment 6										
No. of 7's	5	3	1	1	1	1	2	2	2	3
No. of 6's	2	0	0	0	1	0	0	0	0	0
No. of 5's	5	1	0	0	0	1	0	0	0	0
No. of 4's	2	3	4	0	0	0	0	0	0	0
No. of 3's	1	5	4	2	2	2	1	0	2	0
No. of 2's	0	2	4	5	3	3	6	4	5	5
No. of 1's	1	2	3	8	9	9	7	10	7	8
Minimum	1	1	1	1	1	1	1	1	1	1
Experiment 7										
No. of 7's	4	3	1	1	1	1	1	1	1	1
No. of 6's	1	0	0	0	0	0	0	0	0	0
No. of 5's	2	3	0	0	0	0	0	0	0	0
No. of 4's	4	0	1	2	1	0	0	0	0	0
No. of 3's	1	3	2	1	1	0	0	0	0	0
No. of 2's	1	3	2	2	4	4	4	4	5	3
No. of 1's	1	2	8	8	7	9	9	9	8	10
Minimum	1	1	1	1	1	1	1	1	1	1

TABLE 5—DISTRIBUTION OF ACTIONS FOR TREATMENT *C*: RANDOM PAIRINGS

	Period				
	21	22	23	24	25
Experiment 6					
No. of 7's	5	5	4	10	8
No. of 6's	0	1	3	0	0
No. of 5's	2	5	3	3	4
No. of 4's	3	1	1	1	1
No. of 3's	1	1	1	0	0
No. of 2's	1	1	2	2	2
No. of 1's	4	2	2	0	1
Experiment 7					
No. of 7's	–	–	6	5	5
No. of 6's	–	–	1	0	1
No. of 5's	–	–	0	3	0
No. of 4's	–	–	2	1	4
No. of 3's	–	–	2	0	0
No. of 2's	–	–	0	0	1
No. of 1's	–	–	3	5	3

Aside

In case you are wondering, here are the results for the C_f treatment's initially random but then fixed pairing: radically different from C_d's results.

TABLE 4—EXPERIMENTAL RESULTS FOR TREATMENT C:
FIXED PAIRINGS

	Period						
	21	22	23	24	25	26	27
Experiment 5							
Pair 1							
Subject 1	7	7	7	7	7	7	7
Subject 16	7	7	7	7	7	7	7
Minimum	7*	7*	7*	7*	7*	7*	7*
Pair 2							
Subject 2	7	2	7	7	7	7	7
Subject 15	1	7	3	6	7	7	7
Minimum	1	2	7	7	7	7	7
Pair 3							
Subject 3	1	1	1	1	1	1	1
Subject 14	1	1	7	1	1	1	7
Minimum	1*	1*	1	1*	1*	1*	1
Pair 4							
Subject 4	1	7	7	7	7	7	7
Subject 13	7	2	5	7	7	7	7
Minimum	1	2	5	7*	7*	7*	7*
Pair 5							
Subject 5	1	7	4	7	7	7	7
Subject 12	1	4	7	7	7	7	7
Minimum	1	4	4	7*	7*	7*	7*
Pair 6							
Subject 6	5	7	7	7	7	7	7
Subject 11	7	7	7	7	7	7	7
Minimum	5	7*	7*	7*	7*	7*	7*
Pair 7							
Subject 7	1	7	6	7	7	7	7
Subject 10	5	3	6	7	7	7	7
Minimum	1	3	6*	7*	7*	7*	7*

Pair 8							
Subject 8	7	6	6	7	7	7	7
Subject 9	3	5	7	7	7	7	7
Minimum	3	5	6	7*	7*	7*	7*
Experiment 6							
Pair 1							
Subject 2	7	7	4	5	6	6	7
Subject 15	2	3	6	6	7	7	7
Minimum	2	3	4	5	6	6	7*
Pair 2							
Subject 3	5	7	7	7	7	7	7
Subject 14	7	7	7	7	7	7	7
Minimum	5	7*	7*	7*	7*	7*	7*
Pair 3							
Subject 4	1	1	1	1	4	4	1
Subject 13	7	1	1	3	1	1	2
Minimum	1	1*	1*	1	1	1	1
Pair 4							
Subject 5	5	7	7	7	7	7	7
Subject 12	7	7	7	7	7	7	7
Minimum	5	7*	7*	7*	7*	7*	7*

TABLE 4—FIXED PAIRINGS, Continued

	Period						
	21	22	23	24	25	26	27
Pair 5							
Subject 6	4	5	7	7	7	7	7
Subject 11	4	5	7	7	7	7	7
Minimum	4*	5*	7*	7*	7*	7*	7*
Pair 6							
Subject 7	5	7	7	7	7	7	7
Subject 10	5	7	7	7	7	7	7
Minimum	5*	7*	7*	7*	7*	7*	7*

* – Denotes a mutual best-response outcome.

There is clear evidence of “strategic teaching” (Camerer, Ho, and Chong 2002 *JET*), with 12 of 14 pairs reaching the most efficient equilibrium.

Subjects seemed to understand that with C_d’s random pairing, teaching is pointless, because it’s costly but others reap almost all of the benefits.

But with C_f’s fixed pairing most subjects saw the point of teaching, and devised repeated-game strategies that used it to get efficient outcomes.

These outcomes cannot be modeled taking stage-game strategies as the objects of choice, because teaching must look beyond current payoffs.

But since subjects played the repeated game that describes their entire interaction only once, it’s not clear how to model these outcomes taking repeated-game strategies as the objects of choice either.

I return to Van Huyck et al.’s main treatments, which *can* be modeled with subjects thinking of stage-game strategies as the objects of choice.

End of aside

Explaining the results of Van Huyck et al.’s main treatments

Rational learning?

In rational learning models, players’ decisions in the stage game are determined by an equilibrium in the repeated game that describes the entire learning process, sometimes with a particular selection.

Any pattern of perfectly coordinated jumping from one pure-strategy equilibrium to another over time is a rational learning equilibrium, in any of Van Huyck et al.’s treatments. (And there are many other equilibria.)

Thus rational learning does not even try to explain the results of Van Huyck et al.’s treatments, beyond being possibly consistent with them.

Quantal response equilibrium (“QRE”) in the stage game can be viewed as a variant of rational learning, with time-varying precision used to describe a learning process.

QRE addresses some but not all of the issues raised by these results.

Deterministic evolutionary dynamics?

Inexperienced subjects' initial strategic thinking didn't react strongly to order statistic or group size.

TABLE I

		Minimum treatment				
		A (%)	B (%)	A′ (%)	C_d (%)	C_f (%)
Subject's	7	33 (31)	76 (84)	23 (25)	11 (37)	13 (42)
initial	6	10 (9)	1 (1)	1 (1)	1 (3)	0 (0)
effort	5	34 (32)	2 (2)	2 (2)	2 (7)	6 (19)
	4	18 (17)	5 (5)	7 (8)	5 (17)	2 (6)
	3	5 (5)	1 (1)	7 (8)	3 (10)	1 (3)
	2	5 (5)	1 (1)	17 (19)	1 (3)	1 (3)
	1	2 (2)	5 (5)	34 (37)	7 (23)	8 (26)
Totals		107 (101)	91 (99)	91 (100)	30 (100)	31 (99)

		Median treatment		
		Γ, Γdm (%)	Ω (%)	Φ (%)
Subject's	7	8 (15)	14 (52)	2 (7)
initial	6	4 (7)	1 (4)	3 (11)
effort	5	15 (28)	9 (33)	9 (33)
	4	19 (35)	3 (11)	11 (41)
	3	8 (15)	0 (0)	2 (7)
	2	0 (0)	0 (0)	0 (0)
	1	0 (0)	0 (0)	0 (0)
Totals		54 (100)	27 (100)	27 (99)

Thus the strong treatment effects must be due to dynamics of learning.

Deterministic evolutionary dynamics are the simplest such models.

Such dynamics give a simple account of history-dependent equilibrium selection, in which the population always converges to the equilibrium whose basin of attraction includes its initial state.

To build intuition I start with simplified, two-effort versions of Van Huyck et al.'s treatments, and then generalize.

The random-pairing and large-group minimum games are larger versions of two-effort Stag Hunt games (Rousseau's *Discourse on* Inequality); and the large-group median games can also be simplified this way.

Other Player

	Stag	**Rabbit**
Stag	2 2	1 0
Rabbit	0 1	1 1

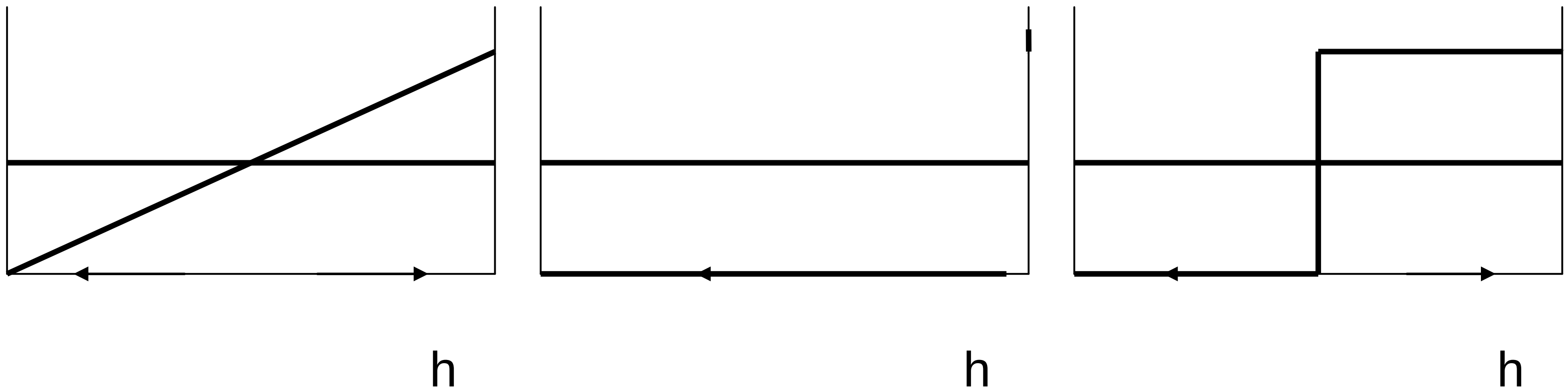
Two-Person Stag Hunt

All Other Players

	All-Stag	**Not All-Stag**
Stag	2	0
Rabbit	1	1

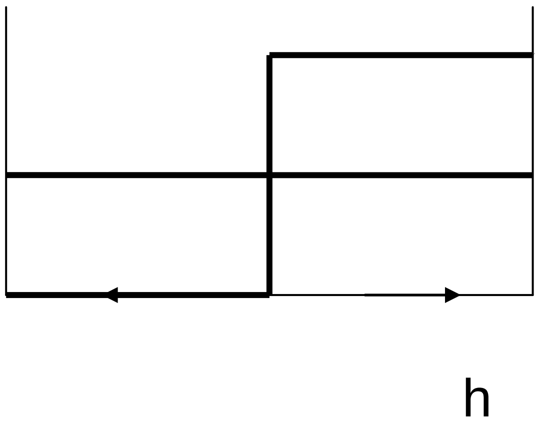
***n*-Person Stag Hunt**

Graph the expected payoffs of high (Stag) and low (Rabbit; expected payoffs always constant) effort against the population frequency of high effort h in the random pairing and large-group minimum games, and in the large-group median game.



Random-pairing minimum Large-group minimum Large-group median

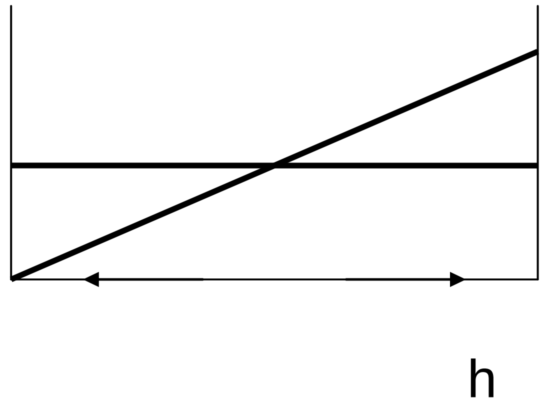
In the large-group median game, the all-Stag and all-Rabbit equilibria are both locally stable.



With random initial conditions, the population is equally likely to converge to all-Stag or all-Rabbit.

If the initial conditions favor one equilibrium, say via strategic thinking, then that equilibrium’s probability of being selected is higher.

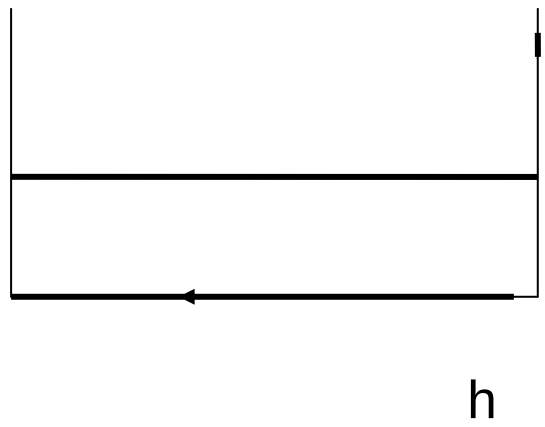
In the random-pairing minimum game, the all-Stag and all-Rabbit equilibria are again both locally stable.



With random initial conditions, the population is equally likely to converge to all-Stag or all-Rabbit.

If the initial conditions favor one equilibrium, say via strategic thinking, then that equilibrium’s probability of being selected is higher.

In the large-group minimum game, the all-Rabbit equilibrium is locally stable but all-Stag is locally unstable.



With random initial conditions, the population is almost certain to converge to all-Rabbit.

So, not bad; but what about the actual, seven-effort treatments?

Crawford's 1995 *Econometrica* model nests deterministic and stochastic evolutionary dynamics and beliefs-based adaptive learning.

Learning is characterized in the style of the engineering adaptive control literature, with beliefs represented by the optimal choices they imply.

(This style is nonstandard in economics and game theory, but in this respect the model is close to Selten's "learning direction theory".)

Players ignore their influences on the order statistic, as the data suggest.

Learning is purely beliefs-based, in that adjustments are influenced equally by what happened and by what would have happened if a player had done something else. (Camerer and Ho 1998 *J. Math. Psychology* show that EWA fits somewhat better than a purely beliefs-based model.)

Recall that the stage games in all five of Van Huyck et al.'s leading treatments have seven strict, symmetric, Pareto-ranked equilibria.

Players' best responses are always an order statistic of the population effort distribution.

This is true even in the random pairing minimum treatment, where for algebraic reasons a player's best response equals the population median; Crawford 1995, p. 110, footnote 10.

Players' beliefs are represented by the optimal efforts they imply, x_{it}, assumed continuously variable.

(This can be relaxed, with the x_{it} as latent variables in an ordered probit.)

The order statistic y_t that determines players' best responses is a continuous function of the x_{it}:

(1) $$y_t \equiv f(x_{it},...,x_{nt}),$$

where for any $x_{1t},...x_{nt}$ and constants a and $b \geq 0$,

(2) $$f(a+bx_{1t},...,a+bx_{nt}) \equiv a+bf(x_{1t},...,x_{nt}).$$

The initial x_{it} are i.i.d. draws, with mean α_0 and shocks ζ_{i0}:

(3) $$x_{i0} = \alpha_0 + \zeta_{i0}.$$

The later x_{it} adjust toward the value suggested by the most recent y_{t-1}:

(4) $x_{it} = \alpha_t + \beta_t y_{t-1} + (1-\beta_t)x_{it-1} + \zeta_{it}$, t = 1,…; $0 < \beta_t \leq 1$; and $\alpha_t \to 0$ as $t \to \infty$.

The adjustment rule (4) includes fictitious play and best-response learning.

Although (4) suggests partial adjustment, think of it as full adjustment to players' estimates of their optimal efforts, which respond less than fully to new y_t observations because they are only part of players' information.

The i.i.d. shocks ζ_{it} represent differences in players' initial beliefs and interpretations of new observations, with mean 0 and variances $\sigma^2_{\zeta t}$.

The model's recursive structure and i.i.d. shocks rule out unmodeled coordination: Coordination can occur only via player's independent responses to common observations of the order statistic.

In (4), deterministic evolutionary dynamics have $\sigma^2_{\zeta 0} > 0$, $\zeta_{it} \equiv \sigma^2_{\zeta t} \equiv \alpha_t \equiv 0$ for all i and $t = 1, \ldots$. This allows unlimited initial heterogeneity but no subsequent differences in players' interpretations of new observations.

Proposition 1 gives a general account of the history-dependence of deterministic evolutionary dynamics.

Proposition 1: Suppose $\alpha_t = 0$ and $\beta_t \in (0,1]$ for all $t = 1, \ldots$, and that there is a $T \geq 1$ such that $\zeta_{it} \equiv \sigma^2_{\zeta t} \equiv 0$ for all $t = T, \ldots$. Then for all i, $x_{it} \rightarrow y_{T-1}$ monotonically, without overshooting, and $y_t = y_T$ for all $t = T, \ldots$, independent of the number of players n and the order statistic $f(\cdot)$.

The proof uses the fact that with no subsequent differences in players' interpretations of new observations, y_t changes only if more players overshoot it in one direction than in the other. By (4) $x_{iT} - y_{T-1}$ then has the same sign as $x_{iT-1} - y_{T-1}$, with x_{iT} closer to y_{T-1} than x_{iT-1} was, and the x_{it} collapse mechanically on the current value of the order statistic, independent of the number of players n and the order statistic $f(\cdot)$.

Proposition 1 shows that deterministic evolutionary dynamics imply some of the history-dependent equilibrium selection found in the experiments.

Deterministic evolutionary dynamics also indirectly capture some effects of strategic uncertainty, in that the treatment variables affect the sizes of basins of attraction of equilibria.

But deterministic evolutionary dynamics rule out the “tunneling” across basins of attraction that sometimes occurs in the experiments.

E.g. Proposition 1 shows that in the large-group minimum game, deterministic dynamics always make the population converge monotonically to the initial minimum, without ever changing it. But in the experiments the initial minimum was above one in five out of seven sessions, but it always converged quickly down to one. For example:

Table 2—Experimental Results for Treatment *A*

	Period									
	1	2	3	4	5	6	7	8	9	10
Experiment 1										
No. of 7's	8	1	1	0	0	0	0	0	0	1
No. of 6's	3	2	1	0	0	0	0	0	0	0
No. of 5's	2	3	2	1	0	0	1	0	0	0
No. of 4's	1	6	5	4	1	1	1	0	0	0
No. of 3's	1	2	5	5	4	1	1	1	0	1
No. of 2's	1	2	2	4	8	7	8	6	4	1
No. of 1's	0	0	0	2	3	7	5	9	12	13
Minimum	2	2	2	1	1	1	1	1	1	1
Experiment 2										
No. of 7's	4	0	1	0	0	0	0	0	0	1
No. of 6's	1	0	1	0	0	1	0	0	0	0
No. of 5's	3	3	2	1	0	0	1	1	0	1
No. of 4's	4	6	2	3	3	0	0	0	0	0
No. of 3's	1	4	2	5	0	1	1	0	1	0
No. of 2's	3	2	6	5	5	9	3	4	3	1
No. of 1's	0	1	2	2	8	5	11	11	12	13
Minimum	2	1	1	1	1	1	1	1	1	1
Experiment 3										
No. of 7's	4	4	1	0	1	1	1	0	0	2
No. of 6's	2	0	2	0	0	0	0	0	0	0
No. of 5's	5	6	1	1	1	0	0	0	0	0
No. of 4's	3	3	2	1	2	1	0	0	0	1
No. of 3's	0	0	7	6	0	2	3	0	0	0
No. of 2's	0	1	1	4	5	3	6	3	2	2
No. of 1's	0	0	0	2	5	7	4	11	12	9
Minimum	4	2	2	1	1	1	1	1	1	1
Experiment 4										
No. of 7's	6	0	1	1	0	0	1	0	0	0
No. of 6's	0	6	2	0	0	1	0	0	0	0
No. of 5's	8	5	5	5	0	1	0	0	0	0
No. of 4's	1	1	4	6	7	1	2	1	1	0
No. of 3's	0	2	3	2	4	3	2	2	1	0
No. of 2's	0	1	0	0	2	3	7	4	2	2
No. of 1's	0	0	0	1	2	6	3	8	11	13
Minimum	4	2	3	1	1	1	1	1	1	1

Long-run equilibria of stochastic evolutionary dynamics?

Deterministic evolutionary dynamics may have many steady states, and which one the population converges to depends on the initial state and is hard to predict without knowing the history and players' learning rules.

A popular way to address this difficulty is via analyses of "long-run equilibria" of stochastic evolutionary dynamics (Kandori, Mailath, and Rob 1993 *Econometrica* and Young 1993 *Econometrica*).

Stochastic evolutionary dynamics allow players' strategy adjustments to be subject to random "mutations", whose probability is constant over time and independent of the state.

In (4), this means $\sigma_{\zeta t}^{2} \equiv \varepsilon > 0$ and $\alpha_t \equiv 0$ and $\beta_t \equiv \beta$ for all t = 1,. . . .

(4) $x_{it} = \alpha_t + \beta_t y_{t-1} + (1 - \beta_t) x_{it-1} + \zeta_{it}$, t = 1,…; $0 < \beta_t \leq 1$; and $\alpha_t \to 0$ as $t \to \infty$.

The resulting dynamics are ergodic, losing the influence of initial conditions and all but the most recent history.

In the long run the process cycles perpetually among steady states of the dynamics without mutations, with “tunneling” across basins of attraction.

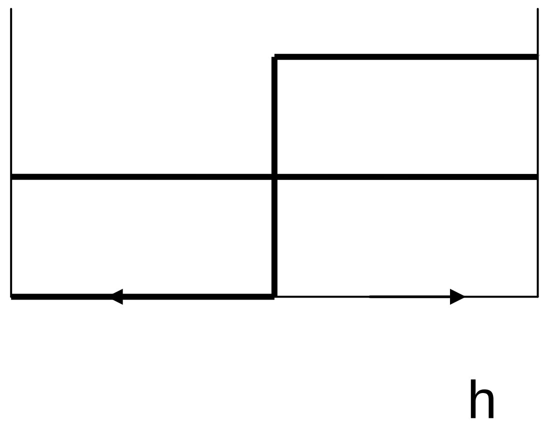
The ergodic distribution of the steady states depends on the probability of mutations and is hard to characterize in general.

But when the probability approaches zero the ergodic distribution approaches a “long-run equilibrium”, which usually puts probability approaching one on one steady state of the dynamics without mutations.

Which steady state is determined by the relative difficulties of moving from alternative steady states to the basins of attraction of other steady states, and can be characterized by counting the number of simultaneous mutations it takes for the population to “tunnel” from one equilibrium to the edge of the basin of attraction of another. (Once at the edge, the deterministic dynamics take over with high probability.)

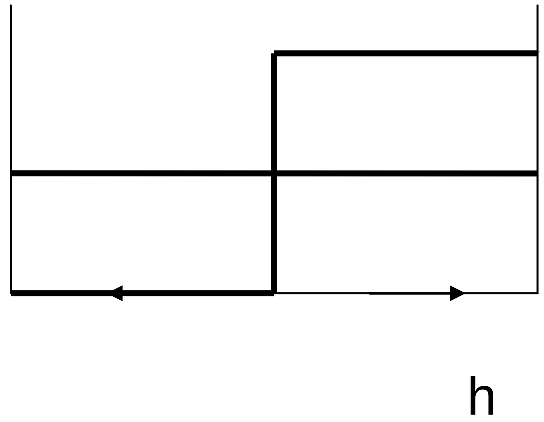
To build intuition I again start with two-effort versions of Van Huyck et al.'s treatments, and then generalize.

Recall the large-group median Stag Hunt game, but contemplate alternative order statistics.



When the order statistic is below the median, the discontinuous drop in effort 2's expected payoff occurs in the right half of the horizontal axis.

Transitions between (symmetric pure-strategy) equilibria occur if more players cross the order statistic from below than above, or vice versa.



When the basin of attraction of the low-effort equilibrium at the left end is larger than that of the high-effort equilibrium at the right end, fewer mutations are needed to go from the high-effort equilibrium to the edge of the basin of attraction of the low-effort equilibrium than vice versa.

A noninfinitesimal mutation probability therefore makes the probability of tunneling leftward across the boundary between basins of attraction higher than the probability of tunneling rightward, so that the ergodic distribution assigns higher probability to the low-effort equilibrium.

As the mutation probability approaches zero, the ratio of the two tunneling probabilities approaches infinity, and the probability of the low-effort equilibrium in the ergodic distribution approaches one.

Thus, a long-run equilibrium analysis discriminates among strict equilibria and obtains unique predictions in most of VHBB's treatments.

These predictions are obtained without modeling initial responses or using empirical information, by studying ergodic dynamics and passing to the limit as the mutation probability approaches zero.

What about the actual, seven-effort treatments?

Proposition 3: In Van Huyck et al.'s 1990, 1991 games, the long-run equilibrium assigns probability one to the equilibrium with lowest (highest) effort whenever the order statistic is below (above) the median, and assigns positive probability to every equilibrium when the order statistic is the median. In each case the long-run equilibrium is independent of the number of players and the order statistic, as long as it remains below (or above) the median.

The proof is a simple mutation-counting argument like the one given in Robles 1997 *JET*.

Proposition 3 shows that a long-run equilibrium analysis discriminates among equilibria in ways that are qualitatively generally consistent with the variations across treatments Van Huyck et al. observed.

But the long-run equilibrium is completely, and unrealistically, independent of the number of players and the order statistic, as long as the latter remains below (or above) the median.

By limiting the effects of history, a long-run equilibrium analysis eliminates much of the information about the effects of changes in the environment an analysis of VHBB’s results could provide.

Beliefs-based adaptive learning?

Consider adaptive learning models, with $\sigma^2_{\zeta 0} > 0,\ \sigma^2_{\zeta t} \to 0,$ and $\alpha_t \to 0$ (or $\alpha_t \equiv 0$ for all *t* = 1, . . .)

(3) $$x_{i0} = \alpha_0 + \zeta_{i0}.$$

(4) $x_{it} = \alpha_t + \beta_t y_{t-1} + (1-\beta_t)x_{it-1} + \zeta_{it}$, *t* = 1,…; $0 < \beta_t \le 1$; and $\alpha_t \to 0$ as $t \to \infty$.

The model is a Markov process with nonstationary transition probabilities (the expected motion of the learning dynamics plus shocks whose stochastic structure is like the nonstationary structure assumed in the strong law of large numbers, with nontrivial interactions), whose long-run steady states coincide with pure-strategy stage-game equilibria.

Unless $\sigma^2_{\zeta t} \to 0$ very slowly, Crawford 1995 *Econometrica* shows that the learning dynamics converge, with probability 1, to a symmetric equilibrium of the stage game. (The variance condition is just as one would expect from the strong law of large numbers.)

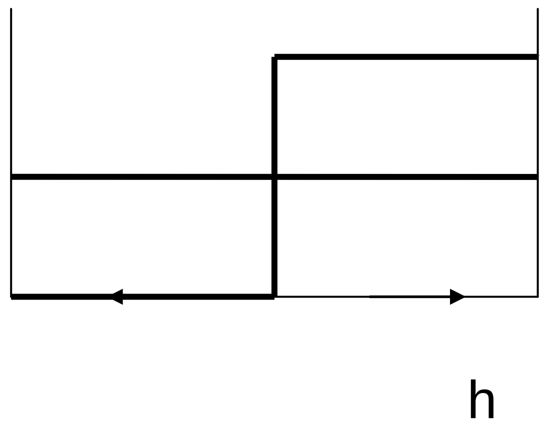
The key difference between adaptive learning models and stochastic evolutionary dynamics is that the heterogeneity of players' beliefs (modeled as i.i.d. random perturbations about a common mean) converges to zero over time ($\sigma_{\zeta t}^{2} \to 0$) rather than remaining constant.

This is what gives adaptive learning dynamics nonstationary transition probabilities and makes it nonergodic.

This is what enables the history-dependence seen in the data, in which the dynamics lock in on a particular equilibrium in the stage game.

Again, to build intuition I start with two-effort versions of Van Huyck et al.'s treatments, and then generalize.

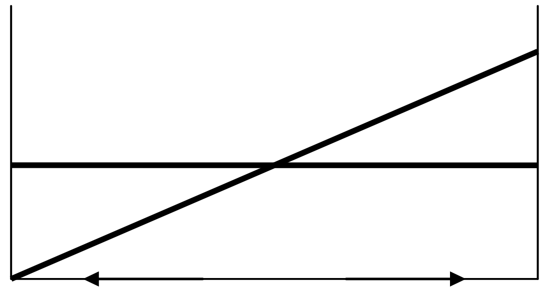
In the large-group median game, the all-Stag and all-Rabbit equilibria are both locally stable.



By symmetry, random shocks are neutral, equally likely to flip the population (via "tunneling") from all-Stag to all-Rabbit or vice versa, so the learning dynamics have no trend.

The population is therefore likely to lock in on the initial median, all-Stag or all-Rabbit, roughly as it did in Van Huyck et al.'s median experiments.

In the random-pairing minimum game, the all-Stag and all-Rabbit equilibria are again both locally stable.

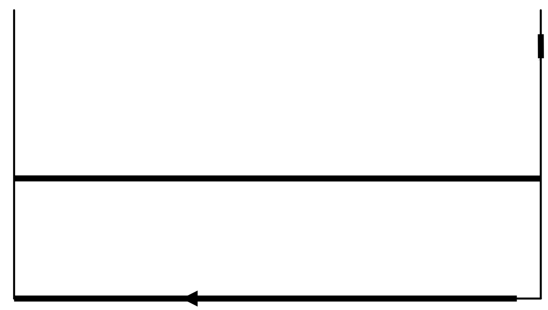


Random shocks are again neutral, so the learning dynamics have no trend.

The population is therefore likely to lock in on the initial median, all-Stag or all-Rabbit, roughly as it did in Van Huyck et al.'s median experiments.

But with random pairing a subject samples only a fraction of the population effort distribution each period (his current partner's effort is an estimate of the population median, but a noisy one), so convergence will be much slower, as it was in the experiments.

In the large-group minimum game, the all-Rabbit equilibrium is locally stable but all-Stag is locally unstable.



h

Random shocks are not neutral, and the learning dynamics have a strong downward trend, yielding strong convergence to all-Rabbit, much as what was observed in Van Huyck et al.'s experiments.

What about the actual, seven-effort treatments?

Consider Crawford's 1995 *Econometrica* adaptive learning model, with $\sigma_{\zeta 0}^2 > 0,\ \sigma_{\zeta t}^2 \to 0,$ and $\alpha_t \to 0$ (or $\alpha_t \equiv 0$ for all t = 1, . . .)

(3) $$x_{i0} = \alpha_0 + \zeta_{i0}.$$

(4) $x_{it} = \alpha_t + \beta_t y_{t-1} + (1-\beta_t)x_{it-1} + \zeta_{it},$ t = 1,…; $0 < \beta_t \leq 1$; and $\alpha_t \to 0$ as $t \to \infty$.

The learning process is like a random walk in the aggregate, but with declining variances and nonzero drift.

The limiting outcome is determined by the cumulative drift before learning eliminates strategic uncertainty and the process locks in on an equilibrium.

The model's implications for equilibrium selection can be summarized by the prior probability distribution of the limiting equilibrium, normally nondegenerate due to the persistent effects of strategic uncertainty.

Proposition 4 shows how the outcome is built up period by period from the shocks that represent the differences in players' beliefs.

Proposition 4: The unique solution of (3) and (4), for all *i* and *t*, is

(5) $$x_{it} = \alpha_0 + \sum_{s=0}^{t-1} \beta f_s + z_{it}$$

and

(6) $$y_t = \alpha_0 + \sum_{s=0}^{t-1} \beta f_s + f_t,$$

where

(7) $$z_{it} \equiv \sum_{s=0}^{t} (1-\beta)^{t-s} \zeta_{is} \quad \text{and} \quad f_t \equiv f(z_{1t},\ldots,z_{nt}).$$

The proof is immediate by induction on *t*. The solution is constructed using the scaling property of *f(·)*in (2) and the linearity of the average adjustment rule to pass the common elements of the x_{it} through *f(·)*.

In the learning dynamics, players' beliefs and efforts differ in response to their different experiences and interpretations of their experiences.

But as they learn from their common observations of y_t, their beliefs and efforts become progressively more correlated over time.

This correlation would normally make analysis difficult.

But viewing players as ex ante i.i.d. draws from a common distribution and using Proposition 4's closed-form solution of the dynamics in terms of the shocks yields simple expressions for Ex_{it} and Ey_t in terms of behavioral and statistical parameters and treatment variables.

Let σ_{zt}^2 denote the common variance of the z_{it} in (7).

(7) $z_{it} \equiv \sum_{s=0}^{t} (1-\beta)^{t\text{-}s} \zeta_{is}$ and $f_t \equiv f(z_{1t},\ldots,z_{nt})$.

(7) implies that $\sigma_{zt}^2 \equiv \sum_{s=0}^{t} [(1-\beta)^{t-s}]^2 \sigma_{zs}^2$.

Define $\mu_t \equiv Ef(z_{1t}/\sigma_{zt},\ldots,z_{nt}/\sigma_{zt})$. Because the z_{it}/σ_{zt} are standardized, with mean 0 and variance 1, μ_t is completely determined by n, $f(\cdot)$, and the joint distribution of the z_{it}/σ_{zt}.

Proposition 5: The ex ante means of y_t and the x_{it}, for all i and t, are

(11) $$Ex_{it} = \alpha_0 + \beta \sum_{s=0}^{t-1} \sigma_{zs}\mu_s$$ and

(12) $$Ey_t = \alpha_0 + \beta \sum_{s=0}^{t-1} \sigma_{zs}\mu_s + \sigma_{zt}\mu_t.$$

The proof takes expectations in (5) and (6), whose shock terms are known functions of the z_{it}, ex ante i.i.d. across i with 0 means, using (7) and (13) $Ef(z_{1s},\ldots,z_{ns}) \equiv E[\sigma_{zs} f(z_{1s}/\sigma_{zs},\ldots,z_{ns}/\sigma_{zs})] \equiv \sigma_{zs}\mu_s$.

Proposition 5 shows how the drift that strategic uncertainty imparts to the dynamics depends on the variances that represent the dispersion of players' beliefs, behavioral parameters, and statistical parameters that reflect the influence of the treatment variables.

Suppose, by way of approximation, that the z_{it}/σ_{zt} are normal, so that $\mu_t \equiv \mu$; and that $\sum_{s=0}^{t} \sigma_{zs} \to S$ as $t \to \infty$. Then Ey_t and $Ex_{it} \to \alpha_0 + \mu\beta S$.

This formula shows how the mean coordination outcome is determined by the behavioral parameters; the number of players; the order statistic, via μ; and the initial dispersion of beliefs and the rate at which it is eliminated by learning, via S.

By symmetry $\mu = 0$ for VHBB's median and random-pairing minimum treatments, so there is no drift and Ey_t, $Ex_{it} \to \alpha_0$.

But $\mu = -1.74$ for the large-group minimum treatment, where the approximate common limit of Ey_t, Ex_{it}, $\alpha_0 + \mu\beta S$, is < 1.10. As intuition suggests, the downward trend in the large-group minimum treatment is stronger, the larger the group or the quorum.

Overall, the analysis yields the following conclusions:

- The perfect history-dependence in VHBB's 1991 median treatments is due to no drift and small variance; but convergence to initial median in 12 of 12 trials may overstate history-dependence: initial median "explains" 46-81% of variance of final median.

- The lack of history-dependence in VHBB's 1990 large-group minimum treatment is due to strong downward drift, which yields convergence to lower bound with very high probability; but convergence in 9 of 9 trials may understate the difficulty of coordination: in simulations it occurred in 500 of 500 trials.

- The slow convergence, weak history-dependence, and lack of trend in VHBB's 1990 random-pairing minimum treatment are due to no drift and subjects' observation of only their current pair's minimum, which is a very noisy estimate of the population median that determined their best responses.

The analysis also yields qualitative comparative dynamics conclusions about the direct effects of changes in treatment variables, holding the behavioral parameters constant:

- Coordination is less efficient the lower the order statistic (the smaller the subsets of the population that can adversely affect the outcome), because small numbers of deviations are more likely than large numbers.

- Coordination is less efficient in larger groups (holding the order statistic constant, measured from the bottom) because it requires coherence among more independent decisions. (This is not an up-down asymmetry!)

The dependence of the dynamics and limiting outcomes on empirical parameters is eliminated in other approaches.

In equilibrium or rational learning analyses this is done by ruling out significant strategic uncertainty.

In evolutionary long-run equilibrium analyses this is done by ruling out any persistent effect of strategic uncertainty, by using ergodic dynamics to model learning and letting the probability of mutations go to 0.

But real-world learning processes are almost always history-dependent.

Realistic models of learning and equilibrium selection must come to grips with empirical behavioral parameters, and ideally provide a framework within which to estimate them.

Afterword: Van Huyck, Battalio, and Beil’s 1993 design and results

Van Huyck et al.’s 1993 *GEB* design was the same as their 1991 *QJE* design, with repeated play of one of the 1991 median games, but with the right to play auctioned each period to the highest 9 bidders in a population of 18 (an English clock auction, with the same price paid by all winning bidders).

The market-clearing price was publicly announced after each period’s auction, the median was publicly announced after each period’s play, and the structure was publicly announced at the start.

The stage game has a range of symmetric equilibria, in which all bid the payoff of some equilibrium of the median game and play that equilibrium, unless others bid differently.

In 8 of 8 trials, subjects quickly bid the price to a level that could only be recouped in the most efficient equilibrium and then converged to that equilibrium: strong, precise selection among a wide range of equilibria.

Auctioning the right to play had a strong efficiency-enhancing effect via focusing subjects’ beliefs on more efficient ways to coordinate—a new and potentially important mechanism by which competition promotes efficiency.

TABLE V

DISTRIBUTION OF ACTIONS FOR GAME Γ(9): EC AUCTION

	Period														
	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Exp. 10															
Price	1.24	1.24	1.28	1.29	1.30	1.30	1.30	1.30	1.30	1.30	—	—	—	—	—
Undom. actions	≥6	≥6	7	7	7	7	7	7	7	7	—	—	—	—	—
# of 7s	7	8	9	8	9	9	9	9	9	9	—	—	—	—	—
# of 6s	2	1	0	0	0	0	0	0	0	0	—	—	—	—	—
# of 5s	0	0	0	0	0	0	0	0	0	0	—	—	—	—	—
# of 4s	0	0	0	0	0	0	0	0	0	0	—	—	—	—	—
# of 3s	0	0	0	0	0	0	0	0	0	0	—	—	—	—	—
# of 2s	0	0	0	0	0	0	0	0	0	0	—	—	—	—	—
# of 1s	0	0	0	1	0	0	0	0	0	0	—	—	—	—	—
Median	7	7	7*	7	7*	7*	7*	7*	7*	7*	—	—	—	—	—
Exp. 11															
Price	1.00	1.20	1.29	1.30	1.29	1.30	1.29	1.29	1.30	1.30	—	—	—	—	—
Undom. actions	≥4	≥6	7	7	7	7	7	7	7	7	—	—	—	—	—
# of 7s	4	5	9	9	9	9	9	9	9	9	—	—	—	—	—
# of 6s	1	3	0	0	0	0	0	0	0	0	—	—	—	—	—
# of 5s	2	0	0	0	0	0	0	0	0	0	—	—	—	—	—
# of 4s	2	1	0	0	0	0	0	0	0	0	—	—	—	—	—
# of 3s	0	0	0	0	0	0	0	0	0	0	—	—	—	—	—
# of 2s	0	0	0	0	0	0	0	0	0	0	—	—	—	—	—
# of 1s	0	0	0	0	0	0	0	0	0	0	—	—	—	—	—
Median	6	7	7*	7*	7*	7*	7*	7*	7*	7*	—	—	—	—	—
Exp. 12															
Price	.95	1.04	1.08	1.10	1.15	1.20	1.25	1.25	1.30	1.30	1.30	1.30	1.30	1.30	1.30
Undom. actions	≥3	≥4	≥5	≥5	≥5	≥6	≥6	≥6	7	7	7	7	7	7	7
# of 7s	1	0	1	3	2	5	8	9	9	9	9	9	9	9	9
# of 6s	0	3	5	6	7	4	0	0	0	0	0	0	0	0	0
# of 5s	6	2	2	0	0	0	0	0	0	0	0	0	0	0	0
# of 4s	1	4	1	0	0	0	0	0	0	0	0	0	0	0	0
# of 3s	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
# of 2s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
# of 1s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Median	5	5	6	6	6	7	7	7	7*	7*	7*	7*	7*	7*	7*
Exp. 13															
Price	1.05	1.14	1.18	1.25	1.29	1.25	1.25	1.30	1.25	1.30	1.30	1.30	1.30	1.30	1.30
Undom. actions	≥4	≥5	≥6	≥6	7	≥6	≥6	7	≥6	7	7	7	7	7	7
# of 7s	2	2	4	6	9	9	9	9	9	9	9	9	9	9	9
# of 6s	1	6	5	3	0	0	0	0	0	0	0	0	0	0	0
# of 5s	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0
# of 4s	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0
# of 3s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
# of 2s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
# of 1s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Median	5	6	6	7	7*	7*	7*	7*	7*	7*	7*	7*	7*	7*	7*
Exp. 14															
Price	1.05	1.15	1.27	1.25	1.25	1.30	1.30	1.25	1.30	1.30	1.25	1.30	1.30	1.30	1.30
Undom. actions	≥4	≥5	7	≥6	≥6	7	7	≥6	7	7	7	7	7	7	7
# of 7s	0	5	8	8	9	9	9	9	9	9	9	9	9	9	9
# of 6s	7	4	0	1	0	0	0	0	0	0	0	0	0	0	0
# of 5s	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
# of 4s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
# of 3s	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
# of 2s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
# of 1s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Median	6	7	7	7	7*	7*	7*	7*	7*	7*	7*	7*	7*	7*	7*

TABLE VI

Distribution of Actions for Game Γ(9): EC Auction and Experienced Subjects

	Period														
	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Exp. 7 ($M = 5$)															
Price	1.09	1.09	1.10	1.19	1.29	1.29	1.30	1.29	1.30	1.30	1.30	1.29	1.30	1.25	1.29
Undom. actions	≥5	≥5	≥5	≥6	7	7	7	7	7	7	7	7	7	≥6	7
# of 7s	0	0	2	5	9	9	9	9	9	9	9	9	9	5	9
# of 6s	2	1	5	4	0	0	0	0	0	0	0	0	0	0	0
# of 5s	6	8	2	0	0	0	0	0	0	0	0	0	0	0	0
# of 4s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
# of 3s	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
# of 2s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
# of 1s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Median	5	5	6	7	7*	7*	7*	7*	7*	7*	7*	7*	7*	7*	7*
Exp. 8 ($M = 5$)															
Price	1.09	1.25	1.28	1.29	1.30	1.29	1.30	1.30	1.29	1.30	1.29	1.30	1.29	1.30	1.30
Undom. actions	≥5	≥6	7	7	7	7	7	7	7	7	7	7	7	7	7
# of 7s	3	7	9	9	9	9	9	9	9	9	9	9	9	9	9
# of 6s	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0
# of 5s	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
# of 4s	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
# of 3s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
# of 2s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
# of 1s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Median	6	7	7*	7*	7*	7*	7*	7*	7*	7*	7*	7*	7*	7*	7*
Exp. 9 ($M = 6$)															
Price	1.15	1.21	1.29	1.29	1.29	1.29	1.29	1.29	1.29	1.29	1.29	1.29	1.29	1.29	1.29
Undom. actions	≥5	≥6	7	7	7	7	7	7	7	7	7	7	7	7	7
# of 7s	0	7	9	9	9	9	9	9	9	9	9	9	9	9	9
# of 6s	8	1	0	0	0	0	0	0	0	0	0	0	0	0	0
# of 5s	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
# of 4s	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
# of 3s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
# of 2s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
# of 1s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Median	6	7	7*	7*	7*	7*	7*	7*	7*	7*	7*	7*	7*	7*	7*

Notes. * indicates mutual best response outcome. ___ Partitions actions into $FI(P)$ and the complement of $FI(P)$.

Explaining Van Huyck et al.'s 1993 results

Crawford and Broseta 1998 *AER*, following Crawford 1995 *Econometrica* and Broseta 2000 *GEB*, show that this effect can be understood as following from effects that formalize "order statistic," "optimistic subjects," and "forward induction" intuitions.

The optimistic subjects and order statistic effects together have approximately the same magnitude in VHBB's environment (where the right to play a nine-person median game was auctioned in a group of 18) as the order statistic effect in an 18-person coordination game without auctions in which payoffs and best responses are determined by the fifth highest (the median of the nine highest) of all 18 players' efforts.

Auctioning the right to play a 9-person median game in a group of 18 effectively turns the game into a "75^{th} percentile" game (0.75 = 13.5/18), whose order statistic effect contributes a large upward drift, as the previous analyses suggest there would have been in such a game without auctions.

Crawford and Broseta's analysis attributes the other half of the efficiency-enhancing effect of auctions in VHBB's environment to a strong forward induction effect.

The analysis shows that coordination is more efficient with more intense competition for the right to play, because it yields higher prices for a given level of dispersion in bidding strategies, and it increases the optimistic subjects effect.

This effect should extend to related environments, but may not always yield full efficiency.