J Evol Econ (2003) 13: 319–345
DOI: 10.1007/s00191-003-0156-x

Journal of Evolutionary Economics

Sophisticated play by idiosyncratic agents*

David P. Myatt and Chris Wallace

Department of Economics, Oxford University, Manor Road, Oxford OX1 3UQ, UK
(e-mail: {david.myatt,christopher.wallace}@economics.ox.ac.uk)

Abstract. Agents are drawn from a large population and matched to play a symmetric 2×2 coordination game, the payoffs of which are perturbed by agent-specific heterogeneity. Individuals observe a (possibly sampled) history of play, which forms the initial hypothesis for an opponent's behaviour. Using this hypothesis as a starting point, the agents iteratively reason toward a Bayesian Nash equilibrium. When sampling is complete and the noise becomes vanishingly small, a single equilibrium is played almost all the time. A necessary and sufficient condition for selection, shown to be closely related (but not identical) to risk-dominance, is derived. When sampling is sufficiently incomplete, the risk-dominant equilibrium is played irrespective of the history observed.

Key words: Sampling – Risk-dominance – Sophisticated play – Idiosyncrasy – Anticipation

JEL Classification: C72, C73

1 Introduction

Many games possess multiple Nash equilibria. Absent a suitable refinement concept, many theorists encourage the examination of the *context* in which a game is played.[1] In particular, the environment may provide a guideline for players. If this guideline specifies a particular Nash equilibrium, an agent can do no better than to play their part in it. This leaves open the question of how such an environment might arise.

Players may look to history to inform their decisions and this can, therefore, provide a suitable context. Of particular relevance are the actions taken previously

* The authors thank Tom Norman, Kevin Roberts, Hyun Shin, Peyton Young, the editor and an anonymous referee for helpful comments.

Correspondence to: C. Wallace

[1] See for example Binmore (1994) or Schelling (1960).

by others. Observation of these actions allows an agent to infer the typical play of a game and to select a strategy contingent on observation. The realised action becomes part of history and history thus evolves. Jointly modelling the interdependence of history and action choice allows selection between equilibria. This is the approach of the adaptive learning literature epitomised by the work of Kandori, Mailath, and Rob (1993, KMR). They specify a dynamic in which entrants respond to an existing population's strategy frequency. Such a dynamic is path-dependent – initial conditions determine long-run behaviour. In response the authors achieve ergodicity via the introduction of mutations (agents make mistakes when choosing their strategies). The invariant distribution of the resulting process is examined as mutations vanish and an equilibrium is selected irrespective of initial conditions.

The adaptive learning approach is subject to a number of critiques. Players form beliefs naïvely and also fail to choose optimal actions given those beliefs. The latter objection is highlighted by Myatt and Wallace (2002a,b) in which agents *differ* rather than *err*.[2] The focus in the present paper will be on the former objection. If a player repeatedly interacts with randomly selected members of the population, then adopting a best response to the incumbent strategy frequency may not be unreasonable. However, when an agent expects a single play against an opponent who simultaneously selects a strategy, this is somewhat suspect. Young (1993a) specifies a model where pairs of entrants, following a (sampled) observation of history, independently choose strategies and play a single game. In this situation, an agent might be expected to reason more carefully.

This approach is taken here. Observed frequencies no longer provide a theoretical opponent. Rather, context *seeds the beliefs* of a player in the following way: An agent initially conjectures that their opponent will act optimally given their observation. This yields a new hypothesis for an opponent's play. The agent, bearing this in mind, calculates a new best response. Stopping here would yield an *anticipatory* dynamic in the sense of Selten (1991). In his model, agents think "one step ahead" of some baseline mode of behaviour.[3] But what if agents anticipate such anticipation? Selten (1991) says:

> "Nevertheless, it is of interest to ask the question whether a learning process with anticipated anticipations has different stability properties [. . .] the change in strategies from one period to the next may be decomposed into a first-order effect due to the observed strategy [. . .] and a second-order effect due to the anticipated change to the opponent's strategy. To these effects one may add a third-order effect due to the anticipated effect of the opponent's anticipation . . . "

[2] The former paper (Myatt and Wallace 2002a) is referred to as MW henceforth. Ergodicity of the adaptive learning process requires occasional contrarian behaviour, where entrants choose against an established convention. An explanation for such behaviour is *idiosyncrasy* on the part of individuals, and this is the approach employed in the current paper.

[3] Selten (1991, p. 118) refers to a baseline mode of play as a "preliminary model" of behaviour. This preliminary model is not necessarily best-response play, but simply assumes that players choose a relatively better response. Following the "one step ahead" reasoning of agents, an *anticipatory learning process* is obtained. For a sufficiently slow adjustment process, equilibria are locally stable.

Following this idea, players in the current model anticipate the anticipations of their opponents. Anticipating this "third-order effect," an agent constructs a best response to this anticipation, and so on. Thus, entrants engage in an iterative best response process the limit of which depends upon the starting point. The starting point is provided by history. Such reflection upon the environment and the reasoning procedure of others is referred to as "sophisticated play." Stahl (1993) takes a related approach. In his terminology, a player who thinks n-steps ahead is "smart$_n$."[4,5] In the present paper, players reason indefinitely with a starting point determined by history.

Full rationality might seem a more appropriate response to the naïveté inherent in the adaptive learning literature. Unfortunately, it is not. A rational player would ignore payoff-irrelevant information and proceed directly to an equilibrium,[6] leaving the modeller unable to select between equally "rational" equilibria. The suggestion here is that agents *can* use observations of past play to coordinate on an equilibrium via the use of some reasonable thought experiment. This is motivated by a desire to model the way in which individuals might think strategically. Unlike full rationality, sophisticated play is a behavioural postulate – a common feature of adaptive and evolutionary research. But do agents really have the computational ability suggested in the model? Although players reason indefinitely, the intention here is to capture a situation in which agents think *to some extent* about their opponent's behaviour.[7]

When sampling is complete, agents observe identical strategy frequencies and this is common knowledge. The iterative reasoning process converges to a (history dependent) Bayesian Nash equilibrium of the underlying stage game. As strategies in such Bayesian equilibria are trigger rules, either action may be realised. Allowing the individual-specific heterogeneity to vanish, these equilibria correspond to the pure Nash equilibria of the unperturbed game. With vanishing noise, only one of the actions is played almost all of the time. If an equilibrium is both *risk-dominant* (Harsanyi and Selten, 1988) and would remain so if the payoffs were normalised by their relative variances, then it is selected. If this is not the case, then the risk-dominated equilibrium may be selected. This tends to occur when the individual-specific payoff heterogeneity is relatively high in the risk-dominant equilibrium. To illustrate this idea, an application to a public-goods contribution game is considered.

---

4 Stahl (1993) follows Selten (1991) by focusing on local stability, rather than equilibrium selection. He shows that, for sufficiently slow adjustment speeds, rationalizable strategies will be locally stable. Furthermore, evolution will "weed out" excessive smartness in the long run, since "being right is just as good as being smart" (Stahl 1993, p. 614).

5 Sáez-Martí and Weibull (1999) present a model in which the agents are "clever." They play a bargaining game after Young (1993b), rather than a coordination game. Some of the agents play simple myopic best responses, but some are clever in the sense that they play a best response to this myopic best response. Matros (2001) extends this idea to general two player games. The difference here is that the agents continue this thought process by considering best responses to best responses, and so on.

6 It is not, of course, quite as simple as that. Aumann and Brandenburger (1995) provide a detailed account of the conditions necessary for rational players to adopt Nash equilibrium strategies.

7 In fact, many of the conclusions are largely unaffected by allowing less sophistication on the part of the players – as long as the process is allowed to iterate sufficiently many times. Agents are allowed to reason indefinitely purely for mathematical convenience.

When sampling is incomplete, selection results may radically differ. Agents no longer observe common histories, and hence they must carefully consider not only their opponent's action but also their opponent's beliefs. An infection argument along the lines of Morris, Rob, and Shin (1995) applies. An agent initially conjectures a best response to the observed strategy frequency. They know that, with high probability, an opponent will observe a different frequency. Such an observation may generate a different strategy choice and hence the initial player might wish to switch strategy. With sufficiently many iterations of reasoning, a single strategy may be adopted by all entrants irrespective of their observations. In fact, when sampling is sufficiently incomplete all players adopt the risk-dominant equilibrium *regardless* of history. Moreover, this result holds without resorting to noise.

Section 2 outlines the model. The analysis with full sampling takes place in Section 3 and with incomplete sampling in Section 4. Section 5 contains some concluding remarks.

2 The model

The model is based upon a symmetric 2×2 coordination game with generic payoffs,

$$\begin{array}{c|c|c|} & 1 & 2 \\ \hline 1 & a,\ a & b,\ c \\ \hline 2 & c,\ b & d,\ d \\ \hline \end{array} \quad \text{or equivalently} \quad \begin{array}{c|c|c|} & 1 & 2 \\ \hline 1 & a-c,\ a-c & 0,\ 0 \\ \hline 2 & 0,\ 0 & d-b,\ d-b \\ \hline \end{array}\ , \tag{1}$$

where $a > c$ and $d > b$ ensure that the game has two pure strategy Nash equilibria $(1, 1)$ and $(2, 2)$.[8] Without loss of generality, it is assumed that $a - c > d - b$, ensuring that the equilibrium $(1, 1)$ is risk dominant (Harsanyi and Selten, 1988). Players care only about the difference in expected payoffs when making a choice, and hence the coordination game is strategically equivalent to the pure coordination game on the right hand side of Equation (1). It is further without loss of generality to set $b = c = 0$ throughout. In this formulation, the mixed strategy Nash equilibrium entails mixing probabilities of $[x^*, 1 - x^*]$ where $x^* = d/(a + d) < 1/2$ since $a > d$ by assumption.

The payoffs a and d represent mean utilities. To generate a Bayesian game, any individual player has idiosyncratic payoffs $\tilde{a}$ and $\tilde{d}$, generated by the addition of normally and independently distributed Harsanyian (1973) payoff trembles:

$$\begin{aligned} \tilde{a} &= a + \sigma\varepsilon_a \\ \tilde{d} &= d + \sigma\varepsilon_d \end{aligned} \quad \text{where} \quad \begin{bmatrix} \varepsilon_a \\ \varepsilon_d \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \xi_a^2 & \gamma\xi_a\xi_d \\ \gamma\xi_a\xi_d & \xi_d^2 \end{bmatrix} \right).$$

The parameters ξ_a and ξ_d allow the variance of the trembles to be strategy profile specific. Section 6 of Blume (1999) offers a similar "random utility" approach. In

[8] The genericity requirement is for convenience only, helping to eliminate integer problems and hence simplify the exposition. Such integer problems are avoided when $dn/(a + d) \notin \mathbb{Z}$ for $n \in \mathbb{Z}$. For this to hold it is sufficient to assume that a/d is an irrational number.

his model any payoff noise is added directly to the expected payoff difference of the two pure strategies, rather than to the payoffs from particular strategy profiles. Such a specification is equivalent to setting $\gamma = -1$ and $\xi_a = \xi_d$. The same observation can also be made of related papers by Brock and Durlauf (2001) and Blume and Durlauf (2001). The parameter σ is a common scaling factor which is allowed to vanish for the limiting results of Section 3. The normal distribution proves convenient for the subsequent analysis.[9] Its crucial property, however, is the unboundedness of the support, allowing either strategy to be dominant with some probability.

The game is played by a finite population of n players. At the beginning of a period each player simultaneously updates their strategy. Alternatively, this can be interpreted as a new group of players replacing last period's entire population. Each new entrant (or player updating) observes the strategy frequency of the population in a sample of size $s \leq n$ taken from last period's play. The agent considers playing a game with a prospective opponent who also observes a (possibly different) sample of size s. In particular, the iterative procedure outlined below takes place in the agent's mind, and a strategy is selected. This agent, along with their new strategy, then becomes part of the updated population, and the whole process repeats itself. Denote the number of individuals playing strategy 1 as z, a member of the finite state space $Z = \{0, ..., n\}$. Computation of the probability that an agent chooses either strategy yields a Markov chain on Z.

In the dynamic described above, all players simultaneously update. This coincides with the approach of KMR (1993). If agents were to update alone and enter a population that was essentially static in strategy frequency, it might be more reasonable for them to simply play a best response to the present state, as in KMR (1993) and Young (1993a). This is the methodology of MW (2002a) with players updating sequentially. When there is simultaneous updating, however, it may be more reasonable to assume that players would think very carefully about their planned action. In particular they would be concerned that a prospective opponent might also base their decision upon observations of history. To some extent, Young (1993a) adopts such an idea. Taking this a step further, a *sophisticated* player might act in the way described below.

Agents begin (as in previous work) by hypothesising a best reply to the strategy frequency of their sample. The agent conjectures a similar response from their opponent. Having calculated this, the agent can formulate a best response to *this* postulated behaviour.[10] An analogous thought experiment is anticipated for their opponent. Bearing in mind the possibly different sample of a prospective opponent an agent will iteratively reason until convergence. More formally, agents play a static Bayesian game where their private information is both the realised payoffs

[9] For $\sigma \to 0$ the key features are the asymptotic properties of the densities and hazard rates of the disturbances. Thus any other distribution sharing these features will lead to similar results. In fact, the logistic distribution produces identical results – see Appendix A. It is also worth noting that full generality of trembles, particularly allowing trembles to vary by state, as Bergin and Lipman (1996) have shown, would lead to inconclusive results.

[10] In Sáez-Martí and Weibull (1999) and Matros (2001), some of the agents conduct only one iteration, while some play simple myopic best responses.

and the observed sample. A myopic strategy profile initialises an iterative best response process which leads to a Bayesian Nash equilibrium.[11] This equilibrium generates the actual behaviour of the agent.

3 Analysis with complete sampling

3.1 Entrant response and Bayesian Nash equilibria

When selecting a strategy a player is concerned with the strategy choice of an opponent. A sufficient statistic for a player's beliefs about an opponent's behaviour is x, the probability with which strategy 1 is played. The expected payoffs from strategies 1 and 2 when an opponent plays 1 with probability x are respectively: $x(a+\sigma\varepsilon_a)$ and $(1-x)(d+\sigma\varepsilon_d)$.

If a player believes an opponent will play strategy 1 with probability x then it is optimal to reply with the same whenever the former is bigger than the latter – a natural trigger rule. This occurs whenever:

$$\varepsilon_d(1-x)-\varepsilon_a x<\frac{ax-d(1-x)}{\sigma}$$

The left hand side is normally distributed with zero mean and variance: $x^2\xi_a^2+(1-x)^2\xi_d^2-2x(1-x)\gamma\xi_a\xi_d$. The best response to a conjecture of x will therefore be strategy 1 with probability:

$$\Pr\left[1|x\right]=\Phi\left(\frac{ax-d(1-x)}{\sigma\sqrt{x^2\xi_a^2+(1-x)^2\xi_d^2-2x(1-x)\gamma\xi_a\xi_d}}\right)$$

where Φ is the standard Gaussian distribution. Notation is simplified by the following:

Definition 1. *Define $\kappa(x)$ as:*

$$\kappa(x)=\frac{ax-d(1-x)}{\sqrt{x^2\xi_a^2+(1-x)^2\xi_d^2-2x(1-x)\gamma\xi_a\xi_d}} \tag{2}$$

Using this notation, the above can be summarised in a convenient Lemma:

Lemma 1. *If an agent plays strategy 1 with probability x, then with probability $\rho\left(x,\sigma\right)=\Phi\left(\kappa\left(x\right)/\sigma\right)$ the optimal response will be strategy 1.*

[11] This kind of iterative thought process may not be too distant from the way in which individuals actually reason. Kinderman, Dunbar, and Bentall (1998) have shown that four levels of such reasoning is not uncommon among adults. Certain professions are conducive to an even greater degree of recursive empathy. Novelists may be able to achieve five or more levels of interactive reasoning, as they place themselves in the minds of their characters. For further details see the survey of Dunbar (1996).

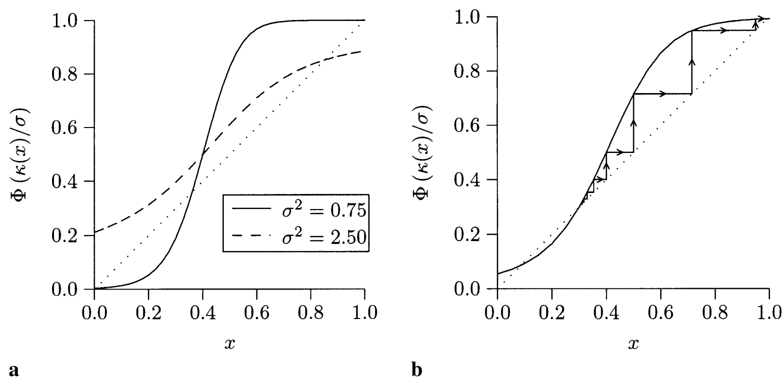

Fig. 1a,b. Bayesian Nash equilibria. **a** Equilibria for various σ. **b** Convergence to a Stable BNE

Ignoring for the moment the observations of agents, consider the game in isolation. The static game is one of incomplete information with uncertainty over the payoffs. In such a game, the Bayesian Nash equilibria correspond to fixed points of the mapping $x \mapsto \rho(x, \sigma)$. Since these equilibria play an important rôle in the following analysis a short examination of their properties is necessary. In particular, interest lies in the case where the Harsanyian perturbations are small ($\sigma \to 0$). Lemma 2 investigates (the proof is in Appendix B):

Lemma 2. *For σ sufficiently small there are three Bayesian Nash equilibria of the game, corresponding to fixed points of $x \mapsto \rho(x, \sigma)$ in the neighbourhood of* 0, x^* *and* 1. *They converge to these points as $\sigma \downarrow 0$.*

This lemma is illustrated in Figure 1a. As $\sigma \to 0$, $\Phi(\kappa(x)/\sigma)$ crosses the 45 degree line at three points in the neighbourhood of 0, x^* and 1.[12]

3.2 The reasoning process with full sampling

Lemma 2 establishes the existence and number of Bayesian Nash equilibria in the game. To select between them, an iterative best response procedure is constructed with the *sophisticated* nature of the players in mind. Note there is still no private information other than payoffs at this stage since agents obtain the same (complete) sample of the population. Hence a symmetric strategy profile can be summarised by the probability x that an individual plays strategy 1. Suppose an opponent initially is conjectured to play strategy 1 with probability x_0. The probability strategy 1 is played as a best response to this is $x_1 = \rho(x_0, \sigma)$ by Lemma 1. The sophisticated play paradigm suggests an iterative process with $x_t = \rho(x_{t-1}, \sigma)$. The agent iteratively calculates best responses to previous best responses until the process converges, yielding a sequence of probabilities x_t.

[12] All figures are evaluated from the following example: $a = 3$, $d = 2$, and so $x^* = 2/5$. Two specifications for ξ_a and ξ_d are used. The first is $\xi_a = \xi_d = 1$, the second is $\xi_a = 2$ and $\xi_d = 3/4$. For both $\gamma = 0$.

Consider now the limiting behaviour of such a procedure. Further suppose that σ is small enough such that there are three Bayesian Nash equilibria in the neighbourhood of 0, x^* and 1 by Lemma 2. Label these fixed points x_L, x_M and x_H respectively.

Lemma 3. *The limiting behaviour of the best response process satisfies:*

$$x_t \to \begin{cases} x_L & x_0 < x_M \\ x_M & x_0 = x_M \\ x_H & x_0 > x_M \end{cases}$$

Proof. Take $x_0 \in (x_M, x_H)$. In this range, $x_H > \rho(x, \sigma) > x$. So $x_t = \rho(x_{t-1}, \sigma) > x_{t-1}$. $\rho(\cdot, \sigma)$ is strictly increasing and continuous. $\{x_t\}_{t=0}^{\infty}$ is an increasing sequence bounded by x_H. By continuity $\lim_{t\to\infty} x_t = x_H$. A similar argument holds for $x_0 \notin (x_M, x_H)$. □

The process is illustrated in Figure 1b. For a given starting point the process converges to one of the two stable Bayesian Nash equilibria. This starting point is given by the agents' observations of the current population state. With full sampling this entails an identical observation of $s = n$ individuals' actions.

Lemma 4. *If agents observe i players out of n playing strategy 1 then, for sufficiently small σ, the sophisticated reasoning process converges to x_L for $i < \lceil nx^* \rceil$ and x_H for $i \geq \lceil nx^* \rceil$.*[13]

Proof. For σ small enough, x_M is sufficiently close to x^* to avoid discretisation problem. Otherwise follows from Lemma 3. □

When an agent observes a reasonably high proportion of strategy 1 players, $x_0 \in (x_M, x_H)$, they know their prospective opponent will also have observed a high proportion of strategy 1 players. They initially conjecture that their opponent will play a myopic best response to the population frequency. Knowing this, their opponent considers it *more* likely that the agent will play strategy 1, and so on. Once the fixed point is reached (x_H), the agent best responds with strategy 1 with probability exactly equal to x_H from the perspective of their opponent, and hence the process comes to a halt.

3.3 The ergodic distribution and equilibrium selection

Continuing under the assumption that there is no private information to a player aside from their payoffs ($s = n$), Lemma 4 shows that the sophisticated reasoning process induces each agent to play the same Bayesian Nash equilibria. The particular equilibrium played is contingent upon the (common) history observed. Furthermore, the probability distribution over an agent's actions is governed by

[13] The notation $\lceil u \rceil$ indicates the smallest integer above u. More formally, $\lceil u \rceil = \min\{k \in \mathbb{Z} : k \geq u\}$ and, for later reference, $\lfloor u \rfloor = \max\{k \in \mathbb{Z} : k \leq u\}$; $u \notin \mathbb{Z} \Rightarrow \lceil u \rceil > \lfloor u \rfloor$.

the strategy frequencies of the equilibrium. Hence the population state in any period determines which of the equilibria is played. This determines the state in the following period via the distribution over strategies implied by the equilibrium. Throughout, σ is assumed small enough for the lemmas of Sections 3.1 and 3.2 to apply, so that below (respectively above) some point agents play a Bayesian Nash equilibrium corresponding to x_L (respectively x_H).

In a 2×2 coordination game the state space can therefore be reduced to a two-state model. Throughout the remainder, these two states will be referred to (rather loosely) as "basins". Recall the original state space $Z = \{0, 1, ..., n\}$ where $z \in Z$ is the number of agents playing strategy 1.

Definition 2. *For the simultaneous updating dynamic define the Markov state space as* $Z^* = \{L, H\}$ *with generic element* z^*. *The transition probabilities are given by:* $p_{ij} = \Pr\left[z^*_{t+1} = j \mid z^*_t = i\right]$. *The associated Markov transition matrix is:*

$$P = \begin{bmatrix} p_{LL} & p_{LH} \\ p_{HL} & p_{HH} \end{bmatrix}$$

This reduced form is equivalent to $L = \{0, 1, ..., \lceil nx^* \rceil - 1\} \subset Z$ and $H = Z - L$. Further:

$$z^* = \begin{cases} L & z < \lceil nx^* \rceil \text{ and } z \in Z \\ H & z \geq \lceil nx^* \rceil \text{ and } z \in Z \end{cases}$$

Given $z^*_t = H$, $z_t \sim \text{Bin}\,(x_H, n)$. It follows that:

Lemma 5. *The reduced form Markov transition probabilities satisfy:*

$$p_{LH} = \sum_{i=\lceil nx^* \rceil}^{n} \binom{n}{i} x_L^i \left(1 - x_L\right)^{n-i} \qquad p_{LL} = 1 - p_{LH}$$

$$p_{HL} = \sum_{i=0}^{\lceil nx^* \rceil - 1} \binom{n}{i} x_H^i \left(1 - x_H\right)^{n-i} \qquad p_{HH} = 1 - p_{HL}$$

Proof. State H is reached if $z_{t+1} \geq \lceil nx^* \rceil$. Strategy 1 is played with probability x_L in state L. Applying the binomial distribution, obtain p_{LH}. Find the remaining terms similarly. □

This is a simple Markov process with only two states and $p_{ij} > 0\,\forall i, j$. In such cases the ergodic distribution satisfies:

$$\mu_L = \frac{p_{HL}}{p_{HL} + p_{LH}} \text{ and } \mu_H = \frac{p_{LH}}{p_{HL} + p_{LH}}$$

μ_i is the probability that the Bayesian Nash equilibrium corresponding to x_i is played in the long run. Interest lies in the relative frequencies of these two equilibria. Applying Lemma 5:

Lemma 6. *The relative frequency of states H and L in the ergodic distribution satisfies:*

$$\frac{\mu_H}{\mu_L} = \frac{\sum_{i=\lceil nx^*\rceil}^{n} \binom{n}{i} x_L^i \left(1 - x_L\right)^{n-i}}{\sum_{i=0}^{\lceil nx^*\rceil - 1} \binom{n}{i} x_H^i \left(1 - x_H\right)^{n-i}} \tag{3}$$

Note that both x_L and x_H depend upon σ. The next task is to examine the relative ergodic frequency as $\sigma \to 0$, thus enabling selection between the two available equilibria.

From Lemma 2, $x_L \to 0$ and $x_H \to 1$ as $\sigma \to 0$. Thus both numerator and denominator in Equation (3) tend to zero as σ vanishes. Investigation of the limiting behaviour begins with the following lemma.

Lemma 7. *The ratio in Equation (3) as $\sigma \to 0$ becomes:*

$$\lim_{\sigma \to 0} \frac{\mu_H}{\mu_L} = \frac{\lceil nx^*\rceil + 1}{n - \lceil nx^*\rceil} \times \lim_{\sigma \to 0} \frac{x_L^{\lceil nx^*\rceil}}{\left(1 - x_H\right)^{n - \lceil nx^*\rceil + 1}} \tag{4}$$

Proof. Multiplying and dividing the numerator of Equation (3) by $x_L^{\lceil nx^*\rceil}$:

$$\sum_{i=\lceil nx^*\rceil}^{n} \binom{n}{i} x_L^i \left(1 - x_L\right)^{n-i} = x_L^{\lceil nx^*\rceil} \sum_{i=\lceil nx^*\rceil}^{n} \binom{n}{i} x_L^{i - \lceil nx^*\rceil} \left(1 - x_L\right)^{n-i}$$

$$= x_L^{\lceil nx^*\rceil} \left\{ \binom{n}{\lceil nx^*\rceil} (1 - x_L)^{n - \lceil nx^*\rceil} + \sum_{i=\lceil nx^*\rceil + 1}^{n} \binom{n}{i} x_L^{i - \lceil nx^*\rceil} (1 - x_L)^{n-i} \right\}$$

As $\sigma \to 0$, $x_L \to 0$ and hence:

$$\sum_{i=\lceil nx^*\rceil + 1}^{n} \binom{n}{i} x_L^{i - \lceil nx^*\rceil} \left(1 - x_L\right)^{n-i} \to 0 \text{ and } \binom{n}{\lceil nx^*\rceil} \left(1 - x_L\right)^{n - \lceil nx^*\rceil} \to \binom{n}{\lceil nx^*\rceil}$$

Performing a similar operation on the denominator to obtain:

$$\lim_{\sigma \to 0} \frac{\mu_H}{\mu_L} = \frac{n!}{(n - \lceil nx^*\rceil)!\, \lceil nx^*\rceil!} \frac{(n - \lceil nx^*\rceil - 1)!\, (\lceil nx^*\rceil + 1)!}{n!}$$
$$\times \lim_{\sigma \to 0} \frac{x_L^{\lceil nx^*\rceil}}{\left(1 - x_H\right)^{n - \lceil nx^*\rceil + 1}} = \frac{\lceil nx^*\rceil + 1}{n - \lceil nx^*\rceil} \times \lim_{\sigma \to 0} \frac{x_L^{\lceil nx^*\rceil}}{\left(1 - x_H\right)^{n - \lceil nx^*\rceil + 1}}$$

Which is the desired result. □

When μ_H/μ_L is large, the Bayesian Nash equilibrium involving probabilities x_H and $1 - x_H$ is played most of the time. For vanishing σ agents play strategy 1 with high probability. The next definition formalises this idea in the limit case as $\sigma \to 0$.

Definition 3. *Strategy* 1 dominates for vanishing heterogeneity *if* $\lim_{\sigma\to 0}\mu_H/\mu_L = +\infty$. *Strategy* 2 dominates for vanishing heterogeneity *if* $\lim_{\sigma\to 0}\mu_H/\mu_L = 0$.

The following critical proposition establishes exactly which equilibrium will dominate for vanishing heterogeneity in terms of the parameters of the model.

Proposition 1. *If* $(n - \lceil nx^* \rceil + 1)\,\kappa(1)^2 > \lceil nx^* \rceil\,\kappa(0)^2$ *then strategy* 1 *dominates for vanishing idiosyncrasy. If the reverse holds then strategy* 2 *dominates.*

Proof. Recall that x_i satisfies the equality $x_i = \rho(x_i,\sigma) = \Phi(\kappa(x_i)/\sigma)$. Substituting into the second term on the right hand side of Equation (4):

$$\lim_{\sigma\to 0}\frac{x_L^{\lceil nx^*\rceil}}{(1-x_H)^{n-\lceil nx^*\rceil+1}} = \lim_{\sigma\to 0}\frac{\Phi(\kappa(x_L)/\sigma)^{\lceil nx^*\rceil}}{(1-\Phi(\kappa(x_H)/\sigma))^{n-\lceil nx^*\rceil+1}}$$

The first term is just a constant and hence is irrelevant. Separating the above into normal densities and hazards rates the following obtains:

$$\lim_{\sigma\to 0}\left(\frac{\phi(\kappa(x_H)/\sigma)}{1-\Phi(\kappa(x_H)/\sigma)}\right)^{n-\lceil nx^*\rceil+1}\left(\frac{\Phi(\kappa(x_L)/\sigma)}{\phi(\kappa(x_L)/\sigma)}\right)^{\lceil nx^*\rceil}\frac{\phi(\kappa(x_L)/\sigma)^{\lceil nx^*\rceil}}{\phi(\kappa(x_H)/\sigma)^{n-\lceil nx^*\rceil+1}}$$

Now as $\sigma\to 0$, $\kappa(x_L)\to\kappa(0)<0$ and $\kappa(x_H)\to\kappa(1)>0$. The limit becomes:

$$\lim_{\sigma\to 0}\left(\frac{\phi(\kappa(1)/\sigma)}{1-\Phi(\kappa(1)/\sigma)}\right)^{n-\lceil nx^*\rceil+1}\left(\frac{\Phi(\kappa(0)/\sigma)}{\phi(\kappa(0)/\sigma)}\right)^{\lceil nx^*\rceil}\frac{\phi(\kappa(0)/\sigma)^{\lceil nx^*\rceil}}{\phi(\kappa(1)/\sigma)^{n-\lceil nx^*\rceil+1}}$$

Note that $\kappa(0)/\sigma\to-\infty$ and $\kappa(1)/\sigma\to+\infty$ as $\sigma\to 0$. The hazard rate of the normal is asymptotically linear. Therefore the first two terms in the above expression are asymptotically polynomial. The third term is exponential however and dominates in the limit. Examining this crucial term:

$$\frac{\phi(\kappa(0)/\sigma)^{\lceil nx^*\rceil}}{\phi(\kappa(1)/\sigma)^{n-\lceil nx^*\rceil+1}} \propto \exp\left\{-\frac{\lceil nx^*\rceil\,\kappa(0)^2-(n-\lceil nx^*\rceil+1)\,\kappa(1)^2}{2\sigma^2}\right\}$$

It follows that:

$$\lim_{\sigma\to 0}\frac{\phi(\kappa(0)/\sigma)^{\lceil nx^*\rceil}}{\phi(\kappa(1)/\sigma)^{n-\lceil nx^*\rceil+1}} = +\infty \Leftrightarrow (n-\lceil nx^*\rceil+1)\,\kappa(1)^2 > \lceil nx^*\rceil\,\kappa(0)^2$$

i.e. strategy 1 dominates for vanishing idiosyncrasy; the desired result. □

Proposition 1 holds when the strategy choices of revising players are probit realisations. Would a similar result hold under alternative choice technologies? It is natural to consider the logit, as in Blume (1999), since the probit and logit are the leading binary choice models used by economists. As Appendix A shows, identical results are available for such a model.

Mistakes do not arise in this model. However, agents have idiosyncratic preferences and so it may appear to the modeller as if mistakes take place. Mutations, in

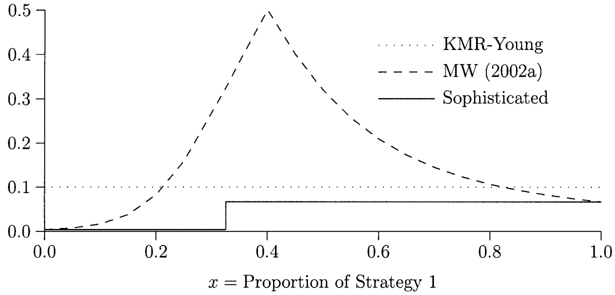

Fig. 2. Endogenously generated "mutations"

this sense, are observed when agents act against the flow of play. That is, an entrant takes a contrarian action relative to an entrant with mean payoffs. Since all agents are playing one of two Bayesian Nash equilibria this will occur with probability x_L in any state $z \in \{0, 1, ..., \lceil nx^* \rceil - 1\}$ and $1 - x_H$ in any state $z \in \{\lceil nx^* \rceil, ..., n\}$, (see Definition 2 and Lemma 5). Figure 2 illustrates these probabilities across the state space.[14]

Mutations are *state dependent* when they vanish to zero at different rates. Here, mutations arise endogenously and are a function of idiosyncrasy, σ, which vanishes to zero independently of the state. Nevertheless this model generates state dependence as Corollary 1 (which follows directly from Proposition 1) below shows.

Corollary 1. *The model endogenously generates state dependent mutations.*

Proof. Consider the limit of the ratio of "mutation" rates:

$$\lim_{\sigma \to 0} \frac{x_L}{1 - x_H} = \lim_{\sigma \to 0} \frac{\Phi\left(\kappa\left(x_L\right)/\sigma\right)}{1 - \Phi\left(\kappa\left(x_H\right)/\sigma\right)} = \lim_{\sigma \to 0} \exp\left\{-\frac{\kappa\left(0\right)^2 - \kappa\left(1\right)^2}{2\sigma^2}\right\}$$

The first equality follows from the fact that x_i satisfies $x_i = \rho\left(x_i, \sigma\right) = \Phi\left(\kappa\left(x_i\right)/\sigma\right)$, and the second from an analogous argument involving normal densities and hazards to that in Proposition 1. Finally note that this limit is either zero or infinity since $\kappa(0) \neq \kappa(1)$. □

Although Proposition 1 yields a precise statement of the selection result, it is helpful to recast this idea in terms of risk-dominance. Bergin and Lipman (1996) show that state dependent mutations can result in the selection of *any* Nash equilibrium – selection need not focus on risk-dominance. They conclude that underlying models of mutation generation are required to provide a convincing defense of any given equilibrium concept. Nevertheless, risk-dominance is surprisingly robust.

The following corollary establishes the relationship between risk-dominance and the above result. If an equilibrium is risk-dominant, and remains so when the payoffs are normalised by their variances, then it is selected. This corroborates the findings of MW (2002a).

[14] Figure 2 uses the second specification for ξ_a and ξ_d. It also illustrates equivalent probabilities (or mutation rates) for the same σ in the MW (2002a) case, and for the KMR (1993) dynamic – which of course is some constant (ε) across all the states and not, as a result, state-dependent.

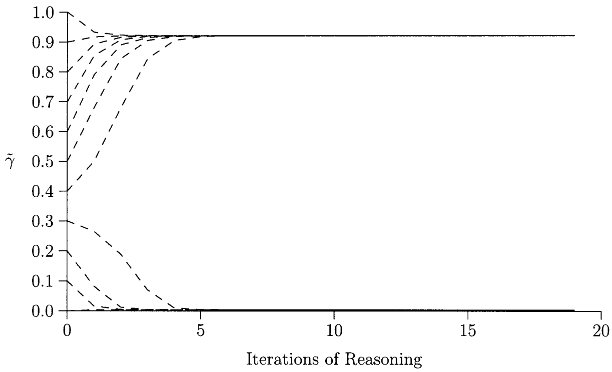

Fig. 3. Full sampling $s = 10$

Corollary 2. *If strategy 1 is risk-dominant* $(a > d)$ *and* $a/\xi_a > d/\xi_d$ *then it dominates for vanishing idiosyncrasy.*

Proof. Assume strategy 1 is risk-dominant and $a/\xi_a > d/\xi_d$ (the proof for strategy 2 is analogous). Using Definition 1:

$$\kappa(0)^2 = \frac{(-d)^2}{\xi_d^2} \text{ and } \kappa(1)^2 = \frac{a^2}{\xi_a^2}$$

If $a/\xi_a > d/\xi_d$ then $\kappa(1)^2 > \kappa(0)^2$. If strategy 1 is risk-dominant then $x^* < \frac{1}{2}$, so $n(1 - x^*) > nx^*$ and hence $(n - \lceil nx^* \rceil + 1) > \lceil nx^* \rceil$. Therefore by Proposition 1 strategy 1 dominates for vanishing idiosyncrasy. □

Given some $\sigma > 0$ players will reason their way to one of the two stable Bayesian Nash equilibria. Figure 3 illustrates this process. Which of the equilibria they play depends upon their starting point – provided by a common observation of history. In the figure, players rather quickly converge given any possible initial conjecture.[15]

Hence the two "basins" are characterised by a different uniform rate of mutation (see Fig. 2). In this diagram, the first basin – $[0, x_M)$ – is narrower, but has a much smaller "mistake probability", determined by the chance of playing Strategy 1 in the Bayesian Nash equilibrium which involves playing Strategy 2 with high probability. The second basin – $(x_M, 1]$ – is wider and has a relatively high probability of mutation. The depth and width of the basins determine selection.[16] Having played

[15] The vertical axis represents the belief probability that an opponent will play strategy 1. This is initialised by the sample observation and then updated according to the "sophisticated" reasoning process. For each possible sample observation this process is plotted. It only takes five iterations for convergence to the Bayesian Nash equilibrium probabilities. The "play strategy 1 with high probability" equilibrium is reached if more than 40% of the sample is seen to play strategy 1.

[16] The depth and width are determined by the parameters of the game. In fact, the widths are determined by a and d, whilst the depths are determined by a/ξ_a and d/ξ_d. The selection result of Proposition 1 essentially shows that it is the equilibrium associated with the greater overall basin "volume" that is selected. Corollary 2 can be interpreted as showing that a basin which is both *wider* and *deeper* clearly has a greater volume and hence results in selection.

their part in a Bayesian Nash equilibrium the agents' actions become part of history and the whole process repeats itself. As $\sigma \to 0$, the mutations vanish *at different rates*, the Bayesian equilibria become the two pure strategy Nash equilibria, and selection is achieved.

Interestingly, the results presented here contrast with those of KMR (1993) and Young (1993a). The reason is that, in those papers, the "mutations" were state-independent, and hence the basin of attraction for any particular equilibrium has a constant depth. Thus, it is *only* the width of a basin of attraction that determines selection. The present model suggests that the *depth* of a basin is also of critical importance. A general measure of basin depth would index the difficulty, as noise vanishes, of taking a step away from a particular equilibrium. In related work, Myatt and Wallace (2002b) develop a notion of *exponential cost* which formalises this notion.[17]

3.4 Application to the private provision of a public good

Proposition 1 leads to a prediction that may differ from the usual selection criterion of risk-dominance. Of course, risk-dominance continues to play a role. If a strategy is risk-dominant then the number of "mutations" needed to escape its basin of attraction is relatively large. Of equal importance is the likelihood with which such mutations take place. In the model presented here, the probabilities of such mutations are determined by $\kappa(0)$ and $\kappa(1)$. These "basin depths" are affected by the variances of payoffs as well as their means. The role of such variances can be easily seen in the context of a simple example.

Consider two players who must decide whether or not to contribute to the production of a public good. In a complete-information game, a contribution costs $k > 0$. This generates a benefit of $v > k$ to each player if and only if both choose to contribute. This generates a coordination problem: A player will wish to contribute whenever the opponent is expected to contribute. Using C for "contribute" and D for "don't" the strategic form is:

	C	D
C	$v-k$, $v-k$	$-k$, 0
D	0, $-k$	0, 0

By inspection, this game has two pure strategy Nash equilibria, (C,C) and (D,D). There is also a mixed strategy Nash equilibrium in which each player plays C with

[17] The incorporation of basin depth helps to explain the similarities between the results here and those of MW (2002a). Firstly, a risk-dominant equilibrium which remains so after the payoffs are normalised by their variances will be selected for vanishing heterogeneity. Secondly, when $\sigma > 0$, MW (2002a) conclude that the modes of the ergodic distribution correspond to the Bayesian Nash equilibria of the underlying stage game. Here, this fact is trivial. Agents reason their way to Bayesian Nash equilibria via the sophisticated deliberation process. The ergodic distribution consists of 2 atoms, one at each equilibrium.

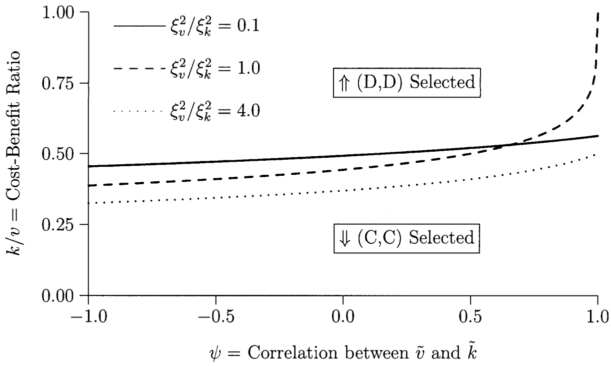

Fig. 4. Equilibrium selection in the public goods contribution game

probability $x^* = k/v$. When $x^* < \frac{1}{2}$, or equivalently $k < \frac{v}{2}$, the equilibrium (C,C) is risk-dominant.

The complete-information game described above abstracts from a more general specification in which individual players may differ in their preferences. A player's actual payoffs $\tilde{k}$ and $\tilde{v}$ may differ from the expectations k and v. For instance, a player might, on occasion, gain some direct private benefit from a contribution, so that $\tilde{k} < 0$. To capture this notion, suppose that $\tilde{k}$ and $\tilde{v}$ are drawn from a bivariate normal distribution with var$[\tilde{k}] = \sigma^2 \xi_k^2$ and var$[\tilde{v}] = \sigma^2 \xi_v^2$, and cov$[\tilde{k}, \tilde{v}] = \sigma^2 \psi \xi_k \xi_v$. The game is equivalent to the pure coordination game described in Section 2, where

$$\begin{bmatrix} \tilde{a} \\ \tilde{d} \end{bmatrix} = \begin{bmatrix} \tilde{v} - \tilde{k} \\ \tilde{k} \end{bmatrix} \sim N\left(\begin{bmatrix} v - k \\ k \end{bmatrix}, \sigma^2 \times \begin{bmatrix} \xi_v^2 + \xi_k^2 - 2\psi\xi_v\xi_k & \psi\xi_k\xi_v - \xi_k^2 \\ \psi\xi_k\xi_v - \xi_k^2 & \xi_k^2 \end{bmatrix} \right).$$

It is now straightforward to calculate $\kappa(1)$ and $\kappa(0)$:

$$\kappa(1) = \frac{v - k}{\sqrt{\xi_v^2 + \xi_k^2 - 2\psi\xi_v\xi_k}} \quad \text{and} \quad \kappa(0) = \frac{k}{\xi_k}$$

Suppose, for instance, that the process begins in a state in which no individual is contributing. A revising player will contribute only if $\tilde{k} \leq 0$, which happens with probability $1 - \Phi(\kappa(0)/\sigma)$. Thus $\kappa(0)$ indexes the difficulty of obtaining a voluntary contribution. It depends on the average contribution cost k and, more interestingly, the variability of this cost ξ_k. Similarly, in a state in which everyone contributes a revising player will stop contributing only when $\tilde{v} < \tilde{k}$. This happens with probability $1 - \Phi(\kappa(1)/\sigma)$. Thus $\kappa(1)$ indexes the robustness of the "all contribute" state. Notice that $\kappa(1)$ depends on the correlation between $\tilde{v}$ and $\tilde{k}$ as well as the variances. For instance, when $\psi < 0$, $\tilde{k}$ and $\tilde{v}$ are negatively correlated. This means that a revising player who finds it costly to contribute will also place little value on the public good. This increases the probability that a player will choose to cease contributing.

Combining these observations, Proposition 1 may be used to select an equilibrium in the long run. To escape from the state in which everyone contributes a proportion $1 - x^*$ of the population must stop contributing. From the discussion above, each such step has a "difficulty" of $\kappa(1)^2$. Similarly, to escape from the state in which nobody contributes, a proportion x^* of the population must begin contributing. Since each such contribution has a "difficulty" of $\kappa(0)^2$, the equilibrium (C,C) will be selected if and only if $(1-x^*)\kappa(1)^2 > x^*\kappa(0)^2$, where integer issues have been ignored. This reduces to

$$\left(\frac{v-k}{k}\right)^3 > 1 + \frac{\xi_v^2}{\xi_k^2} - 2\psi\frac{\xi_v}{\xi_k}$$

This selection condition is illustrated in Figure 4. When the cost-to-benefit ratio k/v is sufficiently small, the (C,C) equilibrium is selected. The required ratio for this to happen is increasing in ψ. This is because, for larger ψ, costs and benefits are positively correlated, and hence there is less chance of observing an individual with a high cost and low benefit. The effect of ξ_v^2/ξ_k^2 is ambiguous. Although high valuation variance tends to destabilise the (C,C) equilibrium for moderate values of ψ, this is not the case when ψ is large. For ψ close to 1, an increase in ξ_v^2 helps to exploit the positive correlation between value and cost.

4 Analysis with incomplete sampling

4.1 The reasoning process with incomplete sampling

With incomplete sampling, players may observe different strategy frequencies from the population. All agents obtain samples of size s. Hence there are $s+1$ possible observations. Index these by $i \in S = \{0, 1, ..., s\}$, where the generic element i is the number of individuals seen to be playing strategy 1. An agent's strategy is a mapping from their payoffs and sample to an action in the set $\{1, 2\}$.

Given their beliefs about an opponent's play, agents are restricted to consider only best responses. In Section 3.1 the optimal response of an agent who believes an opponent plays strategy 1 with probability x was shown to entail playing strategy 1 with probability $\rho(x, \sigma)$. Since payoffs are independently distributed, beliefs are allowed to depend only on the sample observed.[18] $\rho(x, \sigma)$ is the probability with which strategy 1 is a best response. All that remains is to specify the beliefs of the agents contingent on the sample they observe.

Represent a player's beliefs as a vector $\gamma \in [0, 1]^{s+1}$, where the ith element, $\gamma_i \in [0, 1]$, is the probability with which an opponent is believed to play strategy 1 given the player has observed i out of s individuals playing strategy 1. This forms the state variable for the reasoning process. Therefore, given a belief vector, a player observing a sample of $i \in S$ will play strategy 1 with probability $\rho(\gamma_i, \sigma)$. γ is determined by the sophisticated play paradigm: Players anticipate best responses in the population and construct their belief profiles iteratively.

Definition 4. *q_{ij} is the probability of an opponent observing a sample of $j \in S$ given the agent has observed a sample of $i \in S$.*

[18] Beliefs are not contingent on a players' identity; agents only have an "identity" in as far as they have idiosyncratic payoffs and samples.

q_{ij} is determined by the type of sampling procedure the agents use and their prior beliefs over the population states.[19] Two important examples of sampling procedure are considered in Section 4.3; uniform sampling with and without replacement. Of course, there are other possible procedures and for this reason the specification of q_{ij} is left open at this stage. The following Lemma is immediate.

Lemma 8. *If agents initially hold belief γ, then after one iteration of best response they will hold updated belief $\widetilde{\gamma}$, where:*

$$\widetilde{\gamma}_i = \sum_{j=0}^{s} q_{ij}\rho\left(\gamma_j, \sigma\right)$$

Assembling into vector notation, define $\rho\left(\gamma, \sigma\right) \in [0,1]^{s+1}$ as the vector with ith element $\rho\left(\gamma_i, \sigma\right)$ and Q as the matrix with (i, j)th element q_{ij}.

Definition 5. *The iterative reasoning mapping is $\widetilde{\gamma} : [0,1]^{s+1} \longmapsto [0,1]^{s+1}$, where:*

$$\widetilde{\gamma}\left(\gamma\right) = Q\rho\left(\gamma, \sigma\right)$$

The iterative reasoning mapping extends the best response process of Section 3.2. The difference is that play may be contingent on an agent's observation, and because of this players take into account the fact that others may have observed samples at odds with their own. "Sophisticated play" suggests that (*i*) the process should be iterated until it converges and (*ii*) the starting point for the process should be determined by the observations. Formally:

Definition 6. *The sophisticated reasoning process $\{\gamma^t\}$ is constructed with $\gamma^{t+1} = \widetilde{\gamma}\left(\gamma^t\right)$, where the starting point, γ^0, satisfies:*

$$\gamma_i^0 = \rho\left(\frac{i}{s}, \sigma\right)$$

Summarising, the idea is this: Agent A observes i players out of their sample of s playing strategy 1. They know that a prospective opponent (agent B) may have observed a different number of strategy 1 players, say j, with probability q_{ij}. With this in mind, A calculates the probability that B will play strategy 1, $\rho(j/s, \sigma)$. Thus, summing over all possible j, the total probability that B will play strategy 1 given A's observation of i is $\sum q_{ij}\rho(j/s, \sigma)$. B can also calculate this updated belief and hence the probability that A will play strategy 1 as a result of an observation of i. B can then calculate a best response to their *own* observation, again using the probabilities in Q, and these updated probabilities of A playing strategy 1. The process continues in this manner until convergence.

[19] Inevitably the prior is important insofar as it affects q_{ij}. Extreme beliefs are ruled out in the following, all population states are at least possible in the prior – a Beta distribution is assumed later for concreteness. A full discussion of priors is beyond the scope of the current work.

4.2 Equilibrium selection with incomplete sampling

The limiting behaviour of the sophisticated reasoning process determines the equilibrium to be played. The main proposition in this section gives sufficient conditions on the sampling procedure (and hence Q) for selection of the risk-dominant equilibrium.

If the sophisticated reasoning process converges, so that $\lim_{t\to\infty}\gamma^t = \gamma^*$, then the agents play a particular Bayesian Nash equilibrium of the trembled stage game with sampling. In such an equilibrium, an updating agent plays strategy 1 with probability $\sum_{i=0}^{s} \Pr[i \mid z]\,\gamma_i^*$, where $\Pr[i \mid z]$ is the probability of observing a sample $i \in S$ given the population state $z \in Z$. Hence the appropriate Markov transition probabilities can be constructed. Notice, however, that if $\gamma_i^* = \gamma_j^*$ for all $i \neq j$, then the updating agents play the same Bayesian Nash equilibrium regardless of their samples, or indeed the *population state* itself. Hence conditions are sought under which this may occur.

Since interest lies in the case of vanishing idiosyncrasy, the analysis below centres upon $\sigma = 0$. With incomplete sampling, equilibrium selection in this sense does not require an examination of the perturbed stage game. The sophisticated reasoning process itself is enough to select a unique equilibrium.

Condition 1. When a player observes $i < \lceil sx^* \rceil$ their belief that an opponent has observed $j > i$ must be greater than x^*. That is:

$$x^* < \min_{i<\lceil sx^*\rceil} \left\{ \sum_{j=i+1}^{s} q_{ij} \right\} \tag{5}$$

This is the substantive condition that is used in Proposition 2. The right hand side of Equation (5) is a measure of how "dispersed" the Q matrix is. This in turn reflects the agents' beliefs about opponents' observations given their own. The interpretation of Condition 4.2 is as follows. Take a player whose sample contains a minority playing strategy 1. Such a player must place a probability in excess of x^* on the event that an opponent's sample contains a larger number of strategy 1 players.

This condition is particularly weak. To see this, consider the relative probabilities of an opponent receiving a sample containing strictly more strategy 1 players to a sample containing strictly fewer strategy 1 players. Abusing notation slightly, this is the odds ratio $\Pr[j > i \,|\, i]/\Pr[j < i \,|\, i]$. When $i < 1/2$, which is certainly true when $i < x^*$, then for reasonable priors and sampling procedures this ratio will exceed one. In other words, a player who sees only a few strategy 1 players will find it more likely that an opponent will see more than less. As long as the probability of an opponent observing exactly the same number of strategy 1 players (i.e. $\Pr[j = i \,|\, i]$) is small (this will occur for sufficiently large sample sizes) then Condition 4.2 will hold. In summary: Players must believe that it is relatively more likely that opponents have observed samples that are less extreme than their own.[20]

[20] With a uniform prior, and a binomial sampling procedure, Condition 1 implies Condition 2, and Condition 1 holds for any $x^* < \frac{1}{2}$ so long as the sample size is sufficiently large. See Appendix B.

A second technical condition is also employed in the proof of Proposition 2:

Condition 2. The sampling procedure must satisfy a minimal first order stochastic dominance property, or more weakly:

$$x^* < \min_{i \geq \lceil sx^* \rceil} \left\{ \sum_{j=\lceil sx^* \rceil}^{s} q_{ij} \right\} \tag{6}$$

Proposition 2. *If Conditions 1 and 2 hold then for all* $i \in S$, $\lim_{t\to\infty} \gamma_i^t = 1$.

Proof. A best response to a belief of $\gamma_i \geq x^*$ is to play strategy 1. So, write $\rho(x, 0) = I(x \geq x^*)$ where I is an indicator function. The proof proceeds by showing that $\gamma_i^t \geq x^*$ for $j \geq \lceil sx^* \rceil - t$ and $t \leq \lceil sx^* \rceil$. Recall that $\gamma_i^0 = \rho\left(\frac{i}{s}, 0\right)$, then $\gamma_i^0 = 1$ for $i \geq \lceil sx^* \rceil$. Hence the hypothesis holds for $t = 0$, yielding an induction basis. Now suppose that the hypothesis holds for some $\tau < \lceil sx^* \rceil$, so that $\gamma_i^\tau \geq x^*$ for $i \geq \lceil sx^* \rceil - \tau$. Notice that:

$$\gamma_i^{\tau+1} = \sum_{j=0}^{s} q_{ij}\rho\left(\gamma_j^\tau, 0\right) = \sum_{j=0}^{s} q_{ij} I\left(\gamma_j^\tau \geq x^*\right) \geq \sum_{j=\lceil sx^* \rceil - \tau}^{s} q_{ij}$$

First consider $\lceil sx^* \rceil - \tau - 1 \leq i < \lceil sx^* \rceil$. Then:

$$\gamma_i^{\tau+1} \geq \sum_{j=\lceil sx^* \rceil - \tau}^{s} q_{ij} \geq \sum_{j=i+1}^{s} q_{ij} \geq \min_{i < \lceil sx^* \rceil} \left\{ \sum_{j=i+1}^{s} q_{ij} \right\} > x^*$$

Where the last inequality holds by Condition 1's Equation (5). For $i \geq \lceil sx^* \rceil$:

$$\gamma_i^{\tau+1} \geq \sum_{j=\lceil sx^* \rceil - \tau}^{s} q_{ij} \geq \sum_{j=\lceil sx^* \rceil}^{s} q_{ij} \geq \min_{i \geq \lceil sx^* \rceil} \left\{ \sum_{j=\lceil sx^* \rceil}^{s} q_{ij} \right\} > x^*$$

By Equation (6) in Condition 2. Thus, $\gamma_i^{\tau+1} > x^*$ for $i \geq \lceil sx^* \rceil - \tau - 1$. By the principle of induction $\gamma_i^t \geq x^*$ for $i \geq \lceil sx^* \rceil - t$ and $t \leq \lceil sx^* \rceil$. In particular, this holds for $\widehat{t} = \lceil sx^* \rceil$. So $\gamma_i^{\lceil sx^* \rceil} \geq x^*$. But then, for all $t > \widehat{t}$, $\gamma_i^t = 1$ for all i. Thus, $\lim_{t\to\infty} \gamma_i^t = 1$ for all $i \in S$. □

The proof employs an infection argument analogous to that in Morris, Rob, and Shin (1995). This point is returned to later. Two important remarks follow:

Remark 1. If strategy 1 is sufficiently risk-dominant then all agents will play strategy 1 regardless of the samples they observe.

In other words, for (almost) any sampling procedure, there is always an x^* small enough such that the Conditions 1 and 2 hold and hence Proposition 2 applies. With sufficiently incomplete sampling there is no need to rely upon evolutionary game theoretic arguments. Equilibrium selection is obtained without recourse to limiting results and the introduction of "mutations" is unnecessary. Fixing $\sigma = 0$, the sophisticated thought process alone allows agents to reason their way to the risk-dominant equilibrium. Convergence takes place instantaneously. Of course, for reasonable assumptions on the sampling procedure the conditions can hold for any $x^* < \frac{1}{2}$, as will be shown in the next section.

Remark 2. $\lceil sx^* \rceil$ is an upper bound on the number of levels of reasoning required to converge to the equilibrium.

As mentioned earlier (see Sect. 2), psychological research has focused on the number of iterations humans are capable of – and it is not many. Here, there is no need to assume the mathematical nicety of infinite reasoning capacity. In the incomplete sampling case equilibrium is reached in finite time.[21]

Conditions 1 and 2 determine precisely *how* incomplete sampling needs to be for any particular game and prior. Agents must place sufficiently high weight on the possibility that others have observed quite different samples from themselves. Given that they observe a sample of $i < \lceil sx^* \rceil$ agents playing the risk-dominant strategy, they must believe it more than likely that another agent has observed $j > i$.[22] Since this is true for all players, they can infer that each agent will believe it more than likely that the others have observed a $k > j$, and so on. Through this consideration of the deliberations of others, agents will eventually reach a point at which they consider it sufficiently likely that their prospective opponent will play the risk-dominant strategy to make their optimal response the same (with high probability). Once this point is reached both players begin (through the iterative thought process) to believe it more and more likely that their opponent will play the risk-dominant strategy. Condition 2 guarantees this. Eventually (and in finite iterations), all players optimally play the risk-dominant strategy with probability 1, see Figure 5.[23]

Notice that initial observations are ignored. As soon as the iterative procedure begins, an infection process much like that of Morris, Rob, and Shin (1995) takes the posteriors to the belief that the prospective opponent observed all risk-dominant players. Proposition 2 employs this infection argument formally.

4.3 Sampling procedures

If Conditions 1 and 2 are satisfied then all players adopt the same strategy. Note that no evolutionary considerations are needed – reasoning alone enables selection. A sufficiently risk-dominant strategy will be selected by the sophisticated reasoning process. More formally, fix a sampling procedure and associated matrix Q, then

[21] This time could be a very short period indeed. $\lceil sx^* \rceil$ can be very small with reasonable sample sizes, and in any case it is a rather *loose* upper bound.

[22] In the sense of Condition 1.

[23] In this diagram $\sigma > 0$ for illustrative purposes (the equilibrium does not involve a probability 1 belief that the opponent will play strategy 1). However, this does demonstrate that the results will go through in general when there are "mutations", although this case is of less interest. If σ is sufficiently small it is straightforward that the conclusions hold. For σ sufficiently large recall there is only one Bayesian Nash equilibrium, and hence selection is of no importance.

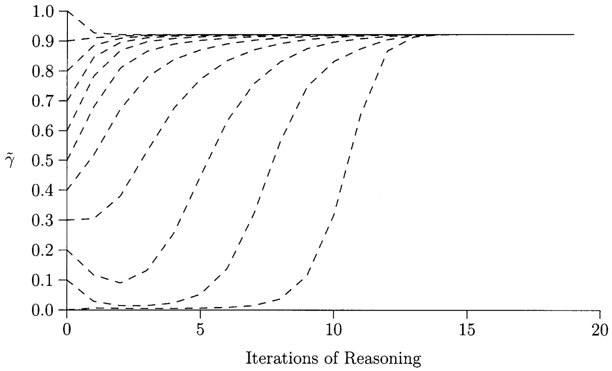

Fig. 5. Effects of sampling $s = 10$

for x^* sufficiently small, strategy 1 will be selected. A related question is raised here: Fixing x^*, what properties of the sampling procedure are required to result in the selection of the risk-dominant equilibrium? Conditions 1 and 2 are sufficient. Here the focus is on whether two reasonable sampling procedures (with/without replacement) satisfy these requirements.

First, suppose that agents observe a sample of size s, with replacement. The probability of observing an agent playing strategy 1 is $z/n = p$.[24] Therefore the number of agents playing strategy 1 in an individuals sample is distributed binomially with parameters p and s. The probability they observe exactly i strategy 1 players is:

$$\Pr[i \mid p] = \binom{i}{s} p^i (1-p)^{s-i}$$

Agents have priors over the state p (or z) , $g(p)$, with distribution function $G(p)$. So:

$$\Pr[i] = \int_p \Pr[i \mid p]\, dG(p)$$

Calculating the elements of Q, the q_{ij} are given by:

$$q_{ij} = \frac{\Pr[i \cap j]}{\Pr[i]} = \frac{\int_p \Pr[i \mid p] \Pr[j \mid p]\, dG(p)}{\int_p \Pr[i \mid p]\, dG(p)}$$

For concreteness assume the prior is Beta distributed with parameters β_1 and β_2:[25]

$$g(p) = \frac{\Gamma(\beta_1 + \beta_2)}{\Gamma(\beta_1)\Gamma(\beta_2)} p^{\beta_1 - 1} (1-p)^{\beta_2 - 1}$$

The Beta is technically convenient as well as simplifying nicely to a symmetric prior ($\beta_1 = \beta_2$) and, as a special case, the uniform ($\beta_1 = \beta_2 = 1$). A closed form can easily be obtained for the elements of the Q matrix – see Appendix B for the proof.

[24] p will be the standard notation throughout this section, simply for convenience.

[25] The Beta is a continuous distribution and hence n needs to be large enough for this to be a reasonable approximation. This is assumed throughout, although it is stated in Proposition 3.

Proposition 3. *With a Beta prior (and n sufficiently large) the elements of Q are:*

$$q_{ij} = \frac{\Gamma(s+1)}{\Gamma(j+1)\,\Gamma(s-j+1)} \frac{\Gamma(\beta_1+\beta_2+s)}{\Gamma(\beta_1+i)\,\Gamma(\beta_2+s-i)} \times \frac{\Gamma(\beta_1+i+j)\,\Gamma(\beta_2+2s-i-j)}{\Gamma(\beta_1+\beta_2+2s)}$$

Therefore, with the uniform distribution the elements reduce to:

$$q_{ij} = \frac{s+1}{2s+1}\binom{s}{i}\binom{s}{j}\Big/\binom{2s}{i+j} \tag{7}$$

Using this formulation Appendix B reports the results of numerical calculations to find (for a range of sample sizes, s) a lower bound on the maximum value of x^* for which Conditions 1 and 2 hold. This lower bound rapidly approaches $1/2$ as s increases. As x^* is increased a higher sample size s is required to satisfy the conditions. For a given x^*, the larger the sample size the more likely the risk-dominant equilibrium is selected immediately. This, combined with Appendix B, demonstrates that sampling with replacement is one procedure that can satisfy the conditions of Section 4.2.

An alternative to the procedure is the more intuitive case of sampling without replacement. This is the scenario Young (1993a) investigates. Players observe a random sample of size s consisting of different individuals' strategy choices. In a population of size n there are z agents playing strategy 1. The number of strategy 1 observations a player makes is hypergeometrically distributed with parameters z, n and s. The probability of observing exactly i agents playing strategy 1 is:

$$\Pr[i \mid z] = \binom{z}{i}\binom{n-z}{s-i}\Big/\binom{n}{s}$$

Retaining the assumptions above concerning the prior once again yields q_{ij}. Hypergeometric probabilities are well approximated by Binomial probabilities for sufficiently large n/s. Hence, by Proposition 3 and the numerical results of Appendix B, with n, n/s and s sufficiently large Conditions 1 and 2 are satisfied for a given $x^* < 1/2$. n/s is the only additional concern for sampling without replacement. The condition that this ratio be large is reminiscent of Young (1993a). Sample and population sizes must be large, but the number of individuals in the population must still dwarf the number observed.

5 Concluding remarks

Focal points often provide a way to select between strict equilibria. However, they can be arbitrary and game specific. The focal point here is provided by history – the *context* in which the game is played. Players do not imitate history or play a naïve best response. Rather they make an initial conjecture – that their opponent will play a best response – and consider what to do in such an event. Nor does their deliberative process stop there, they continue to make iterative conjectures in a "sophisticated" manner. This procedure leads them to an action which then becomes part of history and history evolves.

With such a framework in mind, the conclusions unsurprisingly rest upon the completeness of information available to agents. When each individual observes the entire population in their sample, an exact equilibrium selection criterion is found. Using the sophisticated play paradigm expounded above, agents reason their way to Bayesian Nash equilibria. In the classic coordination game, which equilibrium strategy is played depends on their initial observation of history. As the idiosyncratic nature of the population is reduced to zero, Bayesian equilibria become Nash and selection takes place. If a strategy is both risk-dominant and remains so when the payoffs are normalised by their variances then it is selected.

Agents are more likely to obtain only incomplete information concerning their environment. A complete characterisation is not available in such a scenario. Nevertheless, much can be deduced. There is no need for the introduction of artificial "mutations" to enable selection. In fact, with sufficiently incomplete sampling – and hence only partial information – the risk-dominant equilibrium is selected immediately. Players ignore history altogether in an effort to coordinate with their opponents. Two conditions on the game and sampling procedure suffice to ensure this result. For two common sampling procedures with reasonable constraints on the population and sample sizes, these conditions hold. Further, if a strategy is sufficiently risk-dominant it is selected. Finally, it only takes an agent a finite (and small) number of iterations of the form "I believe that you believe that I believe..." to find it optimal to play the risk-dominant strategy. Moreover, thought processes of this type are *all* that is required for selection to take place – mutations, evolution, and observations of history are rendered irrelevant.

Appendix A. Extension to logistic noise

The analysis in the main paper easily extends to the case of logistic noise. Suppose, for instance, that ε_a/ξ_a and ε_d/ξ_d are independent and that

$$\Pr\left[\frac{\varepsilon_i}{\xi_i} \leq x\right] = \frac{1}{1+e^{-x}} \quad \Rightarrow \quad \text{var}[\tilde{a}] = \frac{\sigma^2\xi_a^2\pi^2}{3} \quad \text{and} \quad \text{var}[\tilde{d}] = \frac{\sigma^2\xi_d^2\pi^2}{3}.$$

Suppose that the entire population is currently playing strategy 1, and hence $x = 1$. The probability that a best-response is strategy 2 is given by

$$\Pr[a + \sigma\varepsilon_a \leq 0] = \Pr\left[\frac{\varepsilon_a}{\xi_a} \leq -\frac{a}{\sigma\xi_a}\right] = e^{-a/\sigma\xi_a} \times \frac{e^{a/\sigma\xi_a}}{1+e^{a/\sigma\xi_a}}.$$

Thus, as $\sigma \to 0$, this probability behaves as $e^{-a/\sigma\xi_a}$. Similarly, when $x = 0$, the probability that a best-response is strategy 1 is approximately $e^{-d/\sigma\xi_d}$. It is straightforward to check that appropriate variants of Lemmas 1–7 continue to hold. Modifying the definition of κ such that

$$\kappa(0)^2 = \frac{d}{\xi_d} \quad \text{and} \quad \kappa(1)^2 = \frac{a}{\xi_a},$$

Proposition 1 and Corollaries 1–2 continue to apply.

6 Appendix B. Omitted proofs and numerical calculations

Section B.1 and B.2 provide the proofs of Lemma 2 and Proposition 3 respectively. Section B.3 contains the numerical calculations referred to throughout Section 4.3.

B.1 Proof of Lemma 2

Fixed points of $\rho(x, \sigma)$ correspond to roots of:

$$f(x) = \Phi\left(\frac{\kappa(x)}{\sigma}\right) - x$$

Notice that $f'(x) = \phi(\kappa(x)/\sigma)\kappa'(x)/\sigma - 1$. When $\sigma \to 0$, $f(x) \to 1 - x$ if $x > x^*$, and $f(x) \to -x$ if $x < x^*$, so there cannot be a fixed point unless it is local to $\{0, x^*, 1\}$. Consider the interval $0 \leq x \leq \varepsilon$. For sufficiently small σ, $f(x)$ is decreasing in this interval. Moreover, $f(0) \geq 0$ and $f(\varepsilon) < 0$. Therefore there is exactly one root in this interval. It is immediate that the fixed point converges to 0 in the limit. A similar argument applies to $1 - \varepsilon \leq x \leq 1$. Now consider $x^* - \varepsilon \leq x \leq x^* + \varepsilon$. Then $f(x^* - \varepsilon) < 0$ and $f(x^* + \varepsilon) > 0$. Again there is at least one root in this interval. $\Phi(\kappa(x)/\sigma)$ is strictly increasing. A fixed point of $\Phi(\kappa(x)/\sigma)$ corresponds to a fixed point of its inverse. Local to x^* the derivative of the inverse is less than one. This locality expands as σ gets small. Within this region there can be only one fixed point of the inverse and hence in this interval the root of $f(x)$ is unique. Convergence for $\sigma \to 0$ is again immediate.

B.2 Proof of Proposition 3

Allowing n sufficiently large for the prior to be well approximated by a Beta distribution, note that:

$$\begin{aligned}\Pr(p \mid i) = \frac{g(p)\Pr(i \mid p)}{\int_0^1 \Pr(i \mid x)\, g(x)\, dx} &\propto g(p)\Pr(i \mid p) \\ &\propto p^{\beta_1 - 1}(1-p)^{\beta_2 - 1} p^i (1-p)^{s-i} \\ &\propto p^{\beta_1 + i - 1}(1-p)^{\beta_2 + s - i - 1}\end{aligned}$$

Which is Beta with parameters $\beta_1 + i$ and $\beta_2 + s - i$. The multiplicative factor to ensure this probability integrates to one can be inserted to give an exact form for $\Pr(p \mid i)$. So:

$$\Pr(p \mid i) = \frac{\Gamma(\beta_1 + \beta_2 + s)}{\Gamma(\beta_1 + i)\,\Gamma(\beta_2 + s - i)} p^{\beta_1 + i - 1}(1-p)^{\beta_2 + s - i - 1}$$

Now the elements of Q are:

$$
\begin{aligned}
q_{ij} &= \Pr(j \mid i) = \int_0^1 \Pr(j \mid p)\Pr(p \mid i)\,dp \\
&= \frac{s!}{j!\,(s-j)!}\frac{\Gamma(\beta_1+\beta_2+s)}{\Gamma(\beta_1+i)\,\Gamma(\beta_2+s-i)}\int_0^1 p^{\beta_1+i+j-1}(1-p)^{\beta_2+2s-i-j-1}\,dp \\
&= \frac{s!}{j!\,(s-j)!}\frac{\Gamma(\beta_1+\beta_2+s)}{\Gamma(\beta_1+i)\,\Gamma(\beta_2+s-i)}\frac{\Gamma(\beta_1+i+j)\,\Gamma(\beta_2+2s-i-j)}{\Gamma(\beta_1+\beta_2+2s)} \\
&\times \int_0^1 \frac{\Gamma(\beta_1+\beta_2+2s)}{\Gamma(\beta_1+i+j)\,\Gamma(\beta_2+2s-i-j)}p^{\beta_1+i+j-1}(1-p)^{\beta_2+2s-i-j-1}\,dp
\end{aligned}
$$

The second term in the last line is a Beta distribution with parameters β_1+i+j and $\beta_2+2s-i-j$ which integrates to one. Hence, as required:

$$
\begin{aligned}
q_{ij} &= \frac{\Gamma(s+1)}{\Gamma(j+1)\,\Gamma(s-j+1)} \\
&\times \frac{\Gamma(\beta_1+\beta_2+s)}{\Gamma(\beta_1+i)\,\Gamma(\beta_2+s-i)}\frac{\Gamma(\beta_1+i+j)\,\Gamma(\beta_2+2s-i-j)}{\Gamma(\beta_1+\beta_2+2s)}
\end{aligned}
$$

B.3 Numerical calculations

Setting $\beta_1=\beta_2=1$ yields a uniformly distributed prior. Then q_{ij} simplifies to the expression in Equation (7). Condition 1 provides a lower bound for maximum values of x^* for which it is satisfied. Since $x^*<1/2$:

$$
\min_{i<\lceil sx^*\rceil}\left\{\sum_{j=i+1}^{s} q_{ij}\right\} \geq \min_{i<\lceil s/2\rceil}\left\{\sum_{j=i+1}^{s} q_{ij}\right\}
$$

The second expression can be numerically calculated for a range of s, using Equation (7). If x^* is less than this value, Condition 1 is satisfied. Figure 6 illustrates these lower bounds. Finally, Condition 1 implies Condition 2 for such q_{ij} as shown in the following:

Lemma 9. *If $j>i$ then $q_{ij}<q_{i+1j}$ and if $j\leq i$ then $q_{ij}\geq q_{i+1j}$, where:*

$$
q_{ij} = \frac{s+1}{2s+1}\binom{s}{i}\binom{s}{j}\Big/\binom{2s}{i+j}
$$

Proof. Consider the ratio q_{ij}/q_{i+1j}:

$$
\frac{q_{ij}}{q_{i+1j}} = \frac{s+1}{2s+1}\binom{s}{i}\binom{s}{j}\binom{2s}{i+1+j}\Big/\frac{s+1}{2s+1}\binom{2s}{i+j}\binom{s}{i+1}\binom{s}{j}
$$

Expanding the terms of this fraction leaves:

$$
\begin{aligned}
\frac{q_{ij}}{q_{i+1j}} &= \frac{(2s-i-j)!\,(i+j)!\,(s-i-1)!\,(i+1)!}{(s-i)!i!\,(2s-i-1-j)!\,(i+1+j)!} \\
&= \frac{(i+1)(s-i)+(i+1)(s-j)}{(i+1)(s-i)+j(s-i)}
\end{aligned}
$$

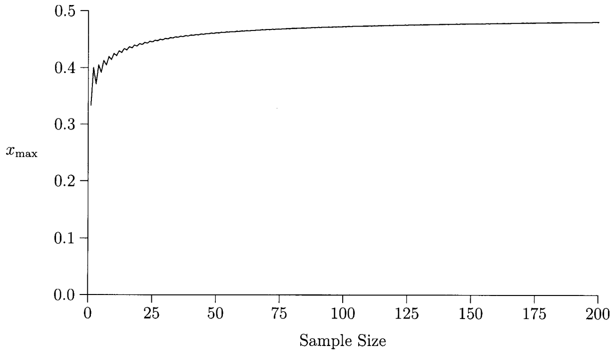

Fig. 6. Lower Bounds on the maximum value of x^* for Condition 1 to hold

If $j > i$ then (since i and j are integers and lie in $[0, s]$) $s\,(j-i) > s-j$ and hence $j\,(s-i) > (i+1)\,(s-j)$, so $q_{ij}/q_{i+1j} < 1$ and the result holds. Likewise for $j \leq i$. □

Lemma 10. *Condition 1 implies Condition 2 for such* q_{ij}.

Proof. Consider the bound in Condition 2:

$$\min_{i\geq\lceil sx^*\rceil}\left\{\sum_{j=\lceil sx^*\rceil}^{s} q_{ij}\right\} = \min_{i\geq\lceil sx^*\rceil}\left\{1-\sum_{j=0}^{\lceil sx^*\rceil-1} q_{ij}\right\} = 1 - \max_{i\geq\lceil sx^*\rceil}\left\{\sum_{j=0}^{\lceil sx^*\rceil-1} q_{ij}\right\}$$

Notice that in the last term $j < i$ throughout the sum. Bringing the results of Lemma 9 to bear, $q_{ij} \geq q_{i+1j}$ for all such $j < i$. Hence, the maximum takes place at the minimum i available, i.e. $i = \lceil sx^*\rceil$. So, the equality becomes:

$$\min_{i\geq\lceil sx^*\rceil}\left\{\sum_{j=\lceil sx^*\rceil}^{s} q_{ij}\right\} = 1 - \left\{\sum_{j=0}^{\lceil sx^*\rceil-1} q_{\lceil sx^*\rceil j}\right\} = \sum_{j=\lceil sx^*\rceil}^{s} q_{\lceil sx^*\rceil j}$$

But if Condition 1 holds then:

$$x^* < \min_{i<\lceil sx^*\rceil}\left\{\sum_{j=i+1}^{s} q_{ij}\right\} \leq \sum_{j=\lceil sx^*\rceil}^{s} q_{\lceil sx^*\rceil-1j} < \sum_{j=\lceil sx^*\rceil}^{s} q_{\lceil sx^*\rceil j}$$

Where the first inequality follows simply because the minimum is being taken over all values of $i < \lceil sx^*\rceil$, which includes $i = \lceil sx^*\rceil - 1$. The second inequality again follows from Lemma 9. $i = \lceil sx^*\rceil - 1 < j$ and hence $q_{ij} < q_{i+1j}$. That is, $q_{\lceil sx^*\rceil-1j} < q_{\lceil sx^*\rceil j}$ for all such j. The bound in Condition 2 is met once Condition 1 is satisfied. □

Figure 6 therefore illustrates lower bounds on the maximum value of x^* for *both* Conditions to hold. From the numerical calculations, any value of $x^* < 1/2$ can satisfy the Conditions given a large enough s. Even for relatively small sample sizes, x^* can be very close to $1/2$.

References

Aumann R, Brandenburger, AM (1995) Epistemic conditions for Nash equilibrium. Econometrica 63(4): 1161–1180

Bergin J, Lipman B (1996) Evolution with state-dependent mutations. Econometrica 64(4): 943–956

Binmore K (1994) Game theory and the social contract, vol 1. Playing fair. MIT Press, Cambridge, MA

Blume LE (1999) How noise matters. Mimeo, SFI Working Paper

Blume LE, Durlauf SN (2001) The interactions based approach to socioeconomic behaviour. In: Durlauf SN, Young HP (eds) Social dynamics, pp 15–44. MIT Press, Cambridge, MA

Brock WA, Durlauf N (2001) Discrete choice with social interactions. The Review of Economic Studies 68(2): 235–260

Dunbar R (1996) Grooming, gossip and the evolution of language. Faber and Faber, London

Harsanyi J, Selten R (1988) A general theory of equilibrium selection in games. MIT Press, Cambridge MA

Harsanyi JC (1973) Games with randomly disturbed payoffs: a new rationale for mixed-strategy equilibrium points. International Journal of Game Theory 2: 1–23

Kandori M, Mailath GJ, Rob R (1993) Learning, mutation and long-run equilibria in games. Econometrica 61(1): 29–56

Kinderman P, Dunbar R, Bentall R (1998) Theory of mind deficits and causal attributions. British Journal of Psychology 89: 191–204

Matros A (2001) Clever agents in adaptive learning. The Journal of Economic Theory (forthcoming)

Morris S, Rob R, Shin HS (1995) p-dominance and belief potential. Econometrica 63(1): 145–157

Myatt DP, Wallace C (2002a) Adaptive play by idiosyncratic agents. Oxford University, Department of Economics Discussion Paper, 89

Myatt DP, Wallace C (2002b) A multinomial probit model of stochastic evolution. The Journal of Economic Theory (forthcoming)

Sáez-Martí M, Weibull JW (1999) Clever agents in Young's evolutionary bargaining model. The Journal of Economic Theory 86: 268–279

Schelling TC (1960) The strategy of conflict. Harvard University Press, Cambridge MA

Selten R (1991) Anticipatory learning in games. In: Selten R (ed) Game equilibrium models I, chap 3. Springer-Verlag, New York

Stahl DO (1993) Evolution of Smart$_n$ players. Games and Economic Behavior 5(4): 604–617

Young HP (1993a) The evolution of conventions. Econometrica 61(1): 57–84

Young HP (1993b) An evolutionary model of bargaining. The Journal of Economic Theory 59: 145–168