Available online at www.sciencedirect.com
SCIENCE DIRECT
Journal of Economic Theory 113 (2003) 286–301
JOURNAL OF Economic Theory
http://www.elsevier.com/locate/jet

A multinomial probit model of stochastic evolution ☆

David P. Myatt and Chris Wallace*

Department of Economics, Oxford University, Oxford OX1 3UQ, UK

Received 2 August 2001; final version received 7 August 2002

Abstract

A strategy revision process in symmetric normal form games is proposed. Following Kandori et al. (Econometrica 61 (1993) 29), members of a population periodically revise their strategy choice, and choose a myopic best response to currently observed play. Their payoffs are perturbed by normally distributed Harsanyian trembles, so that strategies are chosen according to multinomial probit probabilities. As the variance of payoffs is allowed to vanish, the graph theoretic methods of the earlier literature continue to apply. The distributional assumption enables a convenient closed form characterisation for the weights of the rooted trees. An illustration of the approach is offered, via a consideration of the role of dominated strategies in equilibrium selection.

JEL classification: C72; C73

Keywords: Multinomial probit; Stochastic evolution; Idiosyncrasy; Dominated strategies

1. Strategy revision processes with noise

Recent research has been active in tackling the selection of equilibria using evolutionary methods. Kandori, Mailath and Rob [8, KMR] and Young [14] examined the long run behaviour of strategy revision processes with noise (a term

☆The authors would like to thank Peyton Young and Kevin Roberts for helpful discussion. They also thank the associate editor and anonymous referees for their comments and suggestions. Elements of this paper previously appeared under the title "Dominated Strategies and Equilibrium Selection."

*Corresponding author. Fax: +44-1865-279902.

E-mail addresses: david.myatt@economics.ox.ac.uk (D.P. Myatt), christopher.wallace@economics.ox.ac.uk (C. Wallace).

0022-0531/03/$ - see front matter
doi:10.1016/S0022-0531(03)00069-3

due to Blume [2]). The approach is now familiar. An underlying dynamic is specified in which revising players choose a best response to either current or historical strategy choices.[1] Since such processes are often path dependent (becoming "locked in" to pure strategy Nash equilibria), noise is added to yield ergodicity. This ensures that the process generates a unique long-run distribution of play. KMR [8] and Young [14] employed the graph-theoretic methods of Freidlin and Wentzell [4] to characterise this distribution as noise is allowed to vanish. The distribution places all weight on play corresponding to a single pure strategy Nash equilibrium.

Unfortunately, the selection results obtained by this approach depend critically upon the specification of noise. Early contributions adopted a "mutation" framework. Revising players fail to choose a best response with some fixed probability, independent of the state of play. Bergin and Lipman [1] demonstrated that any equilibrium may be selected via an appropriate choice of state-dependent mutations. This indicates the need for an examination of a wider variety of noise processes. Blume [2] constructed a stochastic evolution model in which the expected payoff differences between two strategies are perturbed by logit noise.[2] He found conditions under which the original selection results (focusing on the risk-dominance of equilibria) are maintained.[3]

An alternative approach, and the one taken here, is the adoption of payoff idiosyncrasy. The preferences of revising players may be different from those in the underlying stage game. Although players hold myopic (and hence boundedly rational) beliefs they do not fail to optimise, since any "contrarian" behaviour is driven by idiosyncrasy. Instead of adding noise to the expected payoff differences, normally distributed Harsanyian [6] trembles are added directly to payoffs. It follows that a best response to an observed strategy frequency is characterised by a multinomial probit. The analogue of vanishing mutations is a purification procedure whereby the variances of trembles tend to zero.

This approach permits convenient closed form results and the continued application of the "least cost rooted tree" technique, pioneered by KMR [8] and Young [14]. The tree weights are shown to have a quadratic form in the expected payoff differences. These can be interpreted as the "difficulty" of taking steps away from a given equilibrium. Hence, in contrast to the earlier literature, selection depends not only on the number of mutations required to move between equilibria, but also on the likelihood of such mutations.

[1] Such processes can be more general. For instance, Vega-Redondo [13] studied players who imitate the best performing strategy observed, but occasionally experiment by choosing a strategy at random.

[2] For symmetric 2×2 games, Blume [2] allowed the probability of a mutation to depend upon the expected payoff difference between the strategies. He used the logit specification as an example of this. Such a specification arises when the expected payoff difference between the two strategies is subject to a logistically distributed disturbance. If the expected payoff advantage of the best response is $\triangle$, then this strategy is chosen with probability proportional to $\exp(\beta\triangle)$ where $\beta \to \infty$ as the variance of the logistic noise is allowed to vanish. The logit and probit are leading econometric models of discrete choice.

[3] More generally, he found that risk-dominant equilibria are selected when the noise process is "skew-symmetric." This means that the probability of a mutation depends only upon the absolute difference in payoffs. Myatt and Wallace [11] discuss exactly *how* skew-symmetric the noise process needs to be.

One way to test the robustness of a selection criterion is to compare the predictions of different noise processes. Perhaps unsurprisingly, the results from mistake-driven and idiosyncrasy-driven models may sometimes differ. This is particularly likely to be true when the number of mutations needed to move between equilibria, and the likelihood of such mutations depend upon the properties of two different subsets of strategies. An application of this idea is provided in which the presence of a strictly dominated strategy prevents the selection of a 1/2-dominant equilibrium.

2. A model of stochastic evolution

Consider a two-player symmetric strategic form game with m actions and generic payoffs:[4]

	1	2		m
1	a_{11}, a_{11}	a_{12}, a_{21}	$\cdots$	a_{1m}, a_{m1}
2	a_{21}, a_{12}	a_{22}, a_{22}	$\cdots$	a_{2m}, a_{m2}
	$\vdots$	$\vdots$	$\ddots$	$\vdots$
m	a_{m1}, a_{1m}	a_{m2}, a_{2m}	$\cdots$	a_{mm}, a_{mm}

The payoffs $[a_{ij}]$ will be viewed as the *expected* payoffs for a player. Fixing the payoffs of a static game of complete information is doubtless a simplification. Individual players will have *idiosyncratic* payoffs. These are modelled via Harsanyian [6] *payoff trembles*. Each payoff is subject to an independent and normally distributed perturbation

$$\tilde{a}_{ij} = a_{ij} + \varepsilon_{ij} \quad \text{where} \quad \varepsilon_{ij} \sim N(0, \sigma^2) \quad \text{and} \quad E[\varepsilon_{ij}\varepsilon_{kl}] = 0 \quad \text{for } ij \neq kl.$$

Notice that the payoff disturbances have a fully parametric form. Importantly, this specification obtains clear closed-form results. It is also possible to think of this assumption as an appropriate representation of differing payoffs across players. Such differences could arise from the sum of many individual idiosyncratic factors, yielding the normal distribution as a natural approximation. In any case, the use of a

[4] Although the specification is symmetric, the results of Section 3 may be used in more general settings.

fully general model is precluded by the work of Bergin and Lipman [1]. They show that full generality of trembles, particularly allowing trembles to vary by state, leads to ambiguous results. The state-dependent mutations generated by this model have a natural form and avoid the inconclusiveness associated with full generality.

The game is played by a population of n players. The population evolves according to the following strategy revision process. During a period each player repeatedly plays a fixed strategy against randomly selected opponents from the remaining $n-1$ players. At the end of each period, a randomly selected member of the population leaves, and is replaced by another player with newly trembled payoffs $\tilde{a}_{ij}$. This player observes the strategy distribution among the incumbents, prior to the exit of the leaving player, and selects a best response to this frequency.[5] The process repeats itself.

3. Strategy choice with vanishing noise

The analysis begins with the strategy choice of idiosyncratic revising players. Allowing the variance of payoff perturbations to vanish, the limiting choice probabilities are shown to have a convenient closed form. Begin by using $x \in \mathbb{R}^m$ to denote a strategy frequency vector, satisfying $x_i \geqslant 0$ and $\sum_{i=1}^m x_i = 1$. The following definition will be useful.

Definition 1. The *normalised mean payoff* of strategy i facing frequency x is

$$\mu_i(x) = \frac{\sum_{j=1}^m x_j a_{ij}}{\sqrt{\sum_{j=1}^m x_j^2}}.$$

Define the *normalised mean payoff advantage* of i over j as $\delta_{ij}(x) = \mu_i(x) - \mu_j(x)$.

Thus strategy i is, on average, a strictly better response than j to an observed strategy frequency x whenever $\delta_{ij}(x) > 0$. Of course, the actual best response of a revising player depends upon the player's payoffs, and is stochastic from the perspective of the analyst. The probability that a revising player finds i to be a best response is as follows.[6]

Lemma 1. *An entrant facing strategy frequency x adopts strategy i with probability*

$$\rho_i(x;\sigma) = \int_{-\infty}^{\infty} \left\{ \prod_{j \neq i} \Phi\left(z + \frac{\delta_{ij}(x)}{\sigma} \right) \right\} \phi(z)\, \mathrm{d}z,$$

where Φ and ϕ denote the Gaussian distribution and density functions, respectively.

[5] The dynamic might be specified such that observation takes place post-exit. However, this is immaterial to the results that follow and would only complicate notation. A dynamic in which all members of the population simultaneously revise their strategies would also generate similar results.

[6] The proof to this, and all subsequent results, can be found in Appendix A.

Thus the strategy choice of a revising player is the realisation of a homoskedastic multinomial probit. Unfortunately, the choice probabilities are not available in closed analytic form—numeric evaluation of multiple integrals is required. Subsequent selection analysis, however, will place great interest in the behaviour of these probabilities as payoff idiosyncrasy vanishes. Clearly, if $\mu_i > \mu_j \forall j \neq i$ then $\rho_i \to 1$ and $\rho_j \to 0$ as $\sigma \to 0$. Interest will focus, however, on the *rate* at which these latter probabilities vanish. Although analytic expressions are unavailable for fixed σ, these probabilities become parametric as idiosyncrasy vanishes. This idea is formalised using the following definition:

Definition 2. $f(\sigma) > 0$ has *exponential cost* $c > 0$ if for arbitrarily small $\xi > 0$:

$$\lim_{\sigma \to 0} f(\sigma) \exp\left(\frac{c+\xi}{2\sigma^2}\right) = \infty \quad \text{and} \quad \lim_{\sigma \to 0} f(\sigma) \exp\left(\frac{c-\xi}{2\sigma^2}\right) = 0.$$

This property is denoted $f(\sigma) = \tilde{o}(c)$ or alternatively $c(f(\cdot)) = c$.

Thus a function has exponential cost c if it behaves as $\exp(-c/2\sigma^2)$ does when σ vanishes. The exponential cost property has parallels with the standard $o(\cdot)$ and $O(\cdot)$ notation familiar from the asymptotic behaviour of functions and sequences. The main difference is that the behaviour of $f(\sigma)\exp(c/2\sigma^2)$ is undefined. Familiar properties hold.

Lemma 2. *Exponential cost has the following properties*:

$$\prod_{i=1}^{m} \tilde{o}(c_i) = \tilde{o}\left(\sum_{i=1}^{m} c_i\right), \quad \sum_{i=1}^{m} \tilde{o}(c_i) = \tilde{o}\left(\min_{1 \leqslant i \leqslant m} c_i\right) \quad \textit{and} \quad a \times \tilde{o}(c_i) = \tilde{o}(c_i).$$

Further, taking ratios of functions of σ: $c_i > c_j \Rightarrow \lim_{\sigma \to 0} [\tilde{o}(c_i)/\tilde{o}(c_j)] = 0$.

The following proposition determines the exponential cost of the probit probabilities.[7]

Proposition 1. *The multinomial probit* ρ_i *has exponential cost*

$$c = J_i \times \operatorname*{var}_{j:\delta_{ij} \leqslant 0} (\delta_{ij}) = \sum_{j:\delta_{ij} \leqslant 0} \delta_{ij}^2 - \frac{\left(\sum_{j:\delta_{ij} \leqslant 0} \delta_{ij}\right)^2}{J_i},$$

where $J_i = \sum_j \mathbb{I}(\delta_{ij} \leqslant 0)$ *and* $\mathbb{I}(\cdot)$ *represents the indicator function.*

Notice immediately that if strategy i is the best response in mean payoffs, then $c = 0$, since $\lim_{\sigma \to 0} \rho_i = 1$. More generally, if there are J_i strategies that are weakly better than i in mean payoffs, then the exponential cost is J_i times the variance of the

[7] In unpublished work, Ruud [12] takes a similar approach in an econometric setting.

mean payoff advantage across this set of strategies. This convenient formulation follows from the judicious choice of payoff tremble distribution.

Intuition is built by consideration of the $m = 2$ case. Suppose that $a_{11} > a_{21}$ so that $(1, 1)$ is a pure strategy Nash equilibrium, and suppose that $x_1 = 1$. Omitting dependence on x, it is clear that $\delta_{12} = a_{11} - a_{21} > 0$, and furthermore $\tilde{a}_{11} - \tilde{a}_{21} \sim N(\delta_{12}, 2\sigma^2)$. In this simple two strategy case choice probabilities are simple binomial probits. Clearly,

$$\rho_2 = 1 - \Phi\left(\frac{\delta_{12}}{\sigma\sqrt{2}}\right) = \frac{1 - \Phi(\delta_{12}/\sigma\sqrt{2})}{\phi(\delta_{12}/\sigma\sqrt{2})} \times \phi\left(\frac{\delta_{12}}{\sigma\sqrt{2}}\right).$$

Since $\delta_{12} > 0$ it is clear that $\rho_2 \to 0$ as $\sigma \to 0$. Examining the right-hand side of this equation, the first term is the (inverted) hazard rate of the normal distribution. The hazard rate $\phi(z)/(1 - \Phi(z))$ is asymptotically linear as $z \to \infty$. It follows that the first term is dominated by the second (exponential) term in the limit. Hence

$$\rho_2 \to \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\delta_{12}^2}{4\sigma^2}\right) = \tilde{o}\left(\frac{\delta_{12}^2}{2}\right). \tag{1}$$

The probability of a non-best response depends upon the square of the payoff difference between the two strategies. Proposition 1 extends this to the multinomial case.

4. Long-run behaviour

The current state of play in the population is characterised by the number of agents playing each strategy i, yielding the state space $S = \{s \in \mathbb{Z}_+^m\colon \sum_{i=1}^m s_i = n\}$. The behaviour of revising players is determined solely by the strategies employed by incumbents. It follows that the strategy revision process described in Section 2 is a Markov chain on S. Since revisions occur one at a time, each Markov transition entails a single step. To describe such transitions, it is useful to introduce the following notation. Denote by e_i a $m \times 1$ vector, with zero elements except for the ith element, which is set to 1.

Lemma 3. *Suppressing σ in $\rho_i(\cdot;\sigma)$, the Markov transition probabilities satisfy*

$$p_{ss'} = \Pr[s_{t+1} = s' \mid s_t = s] = \begin{cases} \dfrac{s_k}{n}\rho_i\left(\dfrac{s}{n}\right) & s' = s - e_k + e_i, \\ \sum_{i=1}^m \dfrac{s_i}{n}\rho_i\left(\dfrac{s}{n}\right) & s' = s, \\ 0 & s' \neq s - e_k + e_i \quad \forall i, k. \end{cases}$$

Thus a player switches from strategy k to i with probability $(s_k/n)\rho_i(s/n)$. The term s_k/n is a multiplicative factor independent of the payoff idiosyncrasy σ. Limiting behaviour as $\sigma \to 0$ is determined entirely by the multinomial choice probability $\rho_i(s/n)$, and its corresponding exponential cost. Proposition 1 yields the following:

Corollary 1. *For $s \neq s' = s - e_k + e_i$ the transition $p_{ss'}$ has exponential cost $J_i \times \mathrm{var}_{j:\delta_{ij} \leqslant 0}(\delta_{ij})$.*

Inspection of Lemma 3 reveals that there is positive probability of moving between any two states in a finite number of steps, and hence the Markov chain is irreducible. Furthermore, since there is positive probability of remaining in a state, established cycles cannot occur and the chain is aperiodic. These features combine [5] to yield ergodicity. There exists a unique ergodic distribution $[\pi_s]_{s \in S}$ satisfying $\pi_s = \lim_{t \to \infty} \Pr[s_t = s]$ independent of any initial conditions. For 2×2 games, the state space may be reduced to $\{0, 1, \ldots, n\}$, the strategy revision process becomes a simple birth–death chain, and the ergodic distribution is easy to characterise.[8] For $m > 2$, however, the analysis is more complex. Following KMR [8] and Young [14], the graph-theoretic approach of Freidlin and Wentzell [4] is employed.

The Freidlin and Wentzell [4] technique constructs a directed graph on the state space S with edge weights corresponding to Markov transition probabilities. The directed edge set $E \subseteq S \times S$ has weights $p : E \mapsto \mathbb{R}_+$, where the first and second coordinates represent source and target nodes, respectively. A *tree* rooted at s is a set of edges $h \subseteq E$ such that each node $s' \neq s$ has a unique successor. All sequences of edges lead to s, which has no successor. The collection of trees rooted at s is H_s. The *weight* of such a tree h is w_h and the combined weight of all trees rooted at s is q_s where

$$w_h = \prod_{(r,r') \in h} p_{rr'} \quad \text{and} \quad q_s = \sum_{h \in H_s} w_h.$$

At each step of the Markov chain, a route opens from each node to another. This yields a directed edge set on the state space. Restricting to rooted trees gives route sets which eventually lead to a node s. From Freidlin and Wentzell [4, Chapter 6, Lemma 3.1]:

Lemma 4. *The ergodic distribution μ satisfies*

$$\mu_s = \frac{q_s}{\sum_{s' \in S} q_{s'}} = \frac{\sum_{h \in H_s} \prod_{(r,r') \in h} p_{rr'}}{\sum_{s' \in S} \sum_{h \in H_{s'}} \prod_{(r,r') \in h} p_{rr'}}.$$

This lemma provides an immediate closed form for the invariant distribution. The relative weights of any two states in this distribution may be assessed by considering the ratio $q_s/q_{s'}$. Notice that the weight of a tree is the product of transition

[8] Both Blume [2] and Myatt and Wallace [11] consider such 2×2 games, and fully characterise the ergodic distribution.

probabilities. The exponential cost of this product may be calculated by summing the component costs.

Lemma 5. *Denote* $c_{ss'} = c(p_{ss'})$. *The exponential cost of a tree* h *is* $c(w_h) = \sum_{(s,s')\in h} c_{ss'}$.

Combining this result with the characterisation of the ergodic distribution from Lemma 4 permits the examination of long-run behaviour as $\sigma \to 0$. First, only the least cost rooted trees for each state s are of importance in the limit. To see this, use the properties of exponential cost from Lemma 2: $c(q_s) = \min_{h\in H_s} c(w_h)$. Second, if state s has a lower cost tree than the lowest cost tree from state s', then state s has infinitely more weight in the limit. Again from Lemma 2: $c(q_s) < c(q_{s'}) \Rightarrow \lim_{\sigma\to 0}[q_{s'}/q_s] = 0$. Attention is focused, therefore, on the set of states S^* with least cost rooted trees

$$S^* = \left\{ s \in S\colon \min_{h\in H_s} \{c(w_h)\} \leqslant \min_{s'\in S} \min_{h'\in H_{s'}} \{c(w_{h'})\} \right\}.$$

Allowing $\sigma \to 0$, these states are "selected" in the ergodic distribution. Formally:

Proposition 2. *States in* S^* *attract all weight in the limit*: $\lim_{\sigma\to 0} \sum_{s\in S^*} \mu_s = 1$.

Proposition 2 allows equilibrium selection to take place. If the set S^* consists of a single state corresponding to a pure strategy Nash equilibrium, then (for sufficiently small σ) that equilibrium will be played almost always in the long run. Finding such an equilibrium involves a search for the least cost rooted tree—the exact approach taken by KMR [8] and Young [14]. The difference here is that the cost of each non-best response transition is no longer a constant, but rather a function of the expected payoffs for each available pure strategy.

5. An application to dominated strategies

The simplest possible context for the application of strategy revision processes is the class of symmetric 2×2 coordination games, where $m = 2$, $a_{11} > a_{21}$ and $a_{22} > a_{12}$. Without loss of generality set $a_{11} - a_{21} > a_{22} - a_{12}$ so that the pure strategy Nash equilibrium $(1,1)$ is risk-dominant [7]. Such a game is strategically equivalent to the pure coordination game with payoffs

$$[a_{ij}] = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix} \quad \text{where} \quad a_{11} > a_{22} \quad \Rightarrow \quad \frac{a_{22}}{a_{11}+a_{22}} < \frac{1}{2}. \tag{2}$$

In this case, the state space for a single-population strategy revision process is $S = \{s \in \mathbb{Z}^2_+\colon s_1 + s_2 = n\}$. A best response dynamic with mistakes [8,14] selects the equilibrium $(1,1)$. To see why, first note that the process becomes "locked in" to both $s = (n,0)$ and $s = (0,n)$, which correspond to the pure strategy Nash equilibria.

It is a best response to choose strategy 1 whenever $s_1/n > a_{22}/(a_{11}+a_{22})$. Hence $\lceil na_{22}/(a_{11}+a_{22})\rceil$ mutations are required to escape from the equilibrium $(2,2)$.[9] On the other hand $\lceil na_{11}/(a_{11}+a_{22})\rceil$ mutations are required to escape from the equilibrium $(1,1)$. Hence it is more difficult to move from $(1,1)$ to $(2,2)$ than vice versa.

This selection result continues to hold under the idiosyncratic framework described here. The steps required to escape from the two equilibria are exactly as calculated above. In addition, the cost (or difficulty) of taking such steps depends upon the square of the expected payoff differences. For example, the exponential cost of a move away from state $s=(n,0)$ is $a_{11}^2/2$ (this follows from Eq. (1)). Similarly, the exponential cost of a move away from state $s=(0,n)$ is $a_{22}^2/2 < a_{11}^2/2$. Intuitively, the two effects reinforce each other to confirm the selection result.[10] Notice that *both* effects require a comparison of the relative performance of strategies 1 and 2.

In larger games these two effects need no longer operate in the same direction. This may be illustrated by the addition of a dominated strategy:

$$[a_{ij}] = \begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & a_{22} \\ a_{11}-\varepsilon & -\varepsilon & -\varepsilon \end{bmatrix}. \tag{3}$$

Notice that strategy 3 is strictly dominated by strategy 1.[11] A best response to strategy 3 however, is strategy 2. This dominated strategy provides an additional route via which the population can move from equilibrium $(1,1)$ to equilibrium $(2,2)$.

A mistake-driven process continues to select equilibrium $(1,1)$. This is because the equilibrium remains 1/2-dominant in the sense of Morris et al. [10]. A fraction $a_{11}/(a_{11}+a_{22}) > 1/2$ of the population must switch strategies to escape from equilibrium $(1,1)$ and a fraction $a_{22}/(a_{11}+a_{22}) < 1/2$ must switch to escape from $(2,2)$. 1/2-dominant equilibria (when they exist) continue to be selected in larger games, as shown by Ellison [3] and Maruta [9]. Importantly, the number of mutations required to move between equilibria involves a comparison between the relative performance of strategies 1 and 2 in different population states. Strategy 3 is dominated, and hence never selected. Its presence may, however, influence the identity of a best response.[12]

[9] The notation $\lceil y \rceil$ indicates the smallest integer that is weakly greater than y.

[10] To calculate the exact selection criterion interior states where both strategies are used in the population must also be considered. For a more detailed discussion of the relevant issues and a proof of this selection result, see [11].

[11] More accurately, strategy 3 is strictly dominated in mean payoffs. There is always some probability that a revising player finds it optimal to play this strategy.

[12] It is well known that the addition of a dominated strategy can affect the selection outcome from a strategy revision process. The difference here is that the dominated strategy does not affect the 1/2-dominance of equilibrium $(1,1)$.

This selection result does not hold in the present framework. Suppose the population state is currently $s = (n, 0, 0)$. An idiosyncratic tremble of ε is all that is required to convince a revising player to switch to strategy 3, since this strategy is only a little bit suboptimal relative to strategy 1. This transition has exponential cost $\varepsilon^2/2$. For small ε this step is easy to take. In contrast, any step away from the state $(0, n, 0)$ has an exponential cost of at least $a_{22}^2/2$. Importantly, the exponential cost of an escape from $(1, 1)$ involves a comparison between the relative performance of strategies 2 and 3. This is a different combination than that determining the number of steps required. It is only possible to make this distinction when there are more than two strategies—hence the need for an examination of larger games. In conclusion, as long as ε is small *enough*, the relative ease of taking steps away from equilibrium $(1, 1)$ more than compensates for the larger number of steps required. Formally:

Proposition 3. *Take a* 2×2 *symmetric coordination game where strategy* 1 *is risk-dominant and add a third strategy. The additional strategy may be constructed so that* (i) *the strategy is strictly dominated in mean payoffs*, (ii) *strategy* 1 *remains* $1/2$-*dominant and* (iii) *strategy* 2 *is selected as* $\sigma \to 0$.

The proof of Proposition 3 constructs a dominated strategy in the manner described in Eq. (3), where ε is chosen to be sufficiently small. An upper bound on ε is available, which establishes the required proximity of strategy 3 to strategy 1. The precise expression is available in the proof to Proposition 3.

Appendix A. Omitted proofs

Proof of Lemma 1. Facing x, the payoff from strategy i is $\sum_{j=1}^{m} x_j \tilde{a}_{ij}$. It is chosen when

$$\pi_i = \sum_{j=1}^{m} x_j \tilde{a}_{ij} \geqslant \pi_k = \sum_{j=1}^{m} x_j \tilde{a}_{kj} \quad \forall k \neq i.$$

Recalling the assumptions of payoff idiosyncrasy

$$\pi_i = \sum_{j=1}^{m} x_j a_{ij} + \sum_{j=1}^{m} x_j \varepsilon_{ij} \backsim N\left(\sum_{j=1}^{m} x_j a_{ij}, \sigma^2 \sum_{j=1}^{m} x_j^2\right).$$

The homoskedasticity assumption simplifies the analysis. The consequence is that each strategy yields a payoff with a common variance. Notice

$$\pi_i \geqslant \pi_k \quad \Leftrightarrow \quad \frac{\pi_i}{\sqrt{\sum_{j=1}^{m} x_j^2}} \geqslant \frac{\pi_k}{\sqrt{\sum_{j=1}^{m} x_j^2}}.$$

Denoting this normalised payoff as y_i:

$$y_i \backsim N\left(\frac{\sum_{j=1}^m x_j a_{ij}}{\sqrt{\sum_{j=1}^m x_j^2}}, \sigma^2\right).$$

Strategy selection is thus a realisation of a homoskedastic multinomial probit model where option i has expectation $\mu_i(x)$. The probability of selection is thus

$$\begin{aligned}\rho_i(x;\sigma) = \Pr[y_i \geqslant y_j \ \ \forall j \neq i] &= E_{\varepsilon_i}\left[\Pr\left[\frac{\mu_i - \mu_j + \varepsilon_i}{\sigma} \geqslant \frac{\varepsilon_j}{\sigma} \quad \forall j \neq i | \varepsilon_i\right]\right] \\ &= \int_{-\infty}^{\infty}\left\{\prod_{j\neq i}\Phi\left(z + \frac{\delta_{ij}(x)}{\sigma}\right)\right\}\phi(z)\,\mathrm{d}z,\end{aligned}$$

which is a standard multinomial probit choice probability. □

Proof of Lemma 2. Consider m functions $f_i(\sigma)$ with exponential costs $\{c_i\}$. For ξ arbitrarily small, $m\xi$ is arbitrarily small. Hence

$$\exp\left(\frac{m\xi + \sum_{i=1}^m c_i}{2\sigma^2}\right)\prod_{i=1}^m f_i(\sigma) = \prod_{i=1}^m \exp\left(\frac{\xi + c_i}{2\sigma^2}\right) f_i(\sigma).$$

From this, the first property of Lemma 2 follows easily. The remaining properties follow in a similar fashion. □

Proof of Proposition 1. Recall from Lemma 1 that

$$\rho_i(\cdot;\sigma) = \int_{-\infty}^{\infty}\left\{\prod_{j\neq i}\Phi\left(z + \frac{\delta_{ij}}{\sigma}\right)\right\}\phi(z)dz. \tag{A.1}$$

Write the product of cumulative distributions as a product of densities and hazards

$$\begin{aligned}&\prod_{j\neq i}\Phi\left(z+\frac{\delta_{ij}}{\sigma}\right) \\ &\quad = \underbrace{\prod_{j:\delta_{ij}>0}\Phi\left(z+\frac{\delta_{ij}}{\sigma}\right)}_{\text{cdfs}\to 1} \times \underbrace{\prod_{j\neq i:\delta_{ij}\leqslant 0}\frac{\Phi(z+\sigma^{-1}\delta_{ij})}{\phi(z+\sigma^{-1}\delta_{ij})}}_{\text{hazards}} \times \underbrace{\prod_{j\neq i:\delta_{ij}\leqslant 0}\phi\left(z+\frac{\delta_{ij}}{\sigma}\right)}_{\text{densities}}.\end{aligned}$$

The product of densities is combined with $\phi(z)$ to obtain

$$\phi(z)\prod_{j\neq i:\delta_{ij}\leqslant 0}\phi\left(z+\frac{\delta_{ij}}{\sigma}\right) = \frac{1}{(2\pi)^{J_i/2}}\exp\left(-\frac{z^2 + \sum_{j\neq i:\delta_{ij}\leqslant 0}(z+\sigma^{-1}\delta_{ij})^2}{2}\right),$$

where $J_i = 1 + \sum_{j \neq i} \mathbb{I}(\delta_{ij} \leqslant 0) = \sum_j \mathbb{I}(\delta_{ij} \leqslant 0)$. Completing the square yields

$$\begin{aligned} z^2 + \sum_{j \neq i : \delta_{ij} \leqslant 0} \left(z + \frac{\delta_{ij}}{\sigma} \right)^2 &= \left(z\sqrt{J_i} + \frac{\sum_{j \neq i:\delta_{ij} \leqslant 0} \delta_{ij}}{\sigma \sqrt{J_i}} \right)^2 + \frac{\sum_{j \neq i:\delta_{ij} \leqslant 0} \delta_{ij}^2}{\sigma^2} - \frac{(\sum_{j \neq i:\delta_{ij} \leqslant 0} \delta_{ij})^2}{J_i \sigma^2} \\ &= \left(z\sqrt{J_i} + \frac{\sum_{j \neq i:\delta_{ij} \leqslant 0} \delta_{ij}}{\sigma \sqrt{J_i}} \right)^2 + \frac{\kappa^2}{\sigma^2}, \end{aligned}$$

where κ^2 denotes

$$\kappa^2 = \sum_{j:\delta_{ij} \leqslant 0} \delta_{ij}^2 - \frac{(\sum_{j:\delta_{ij} \leqslant 0} \delta_{ij})^2}{J_i} = J_i \times \underset{j:\delta_{ij} \leqslant 0}{\mathrm{var}} (\delta_{ij}).$$

Notice that in the above summation $j : \delta_{ij} \leqslant 0$ includes $j = i$. Adding the term $\delta_{ii} = 0$ does not affect the summations, and allows the variance interpretation on the right-hand side. It is convenient to economise notation as follows

$$\prod_{j:\delta_{ij} > 0} \Phi_i = \prod_{j:\delta_{ij} > 0} \Phi\left(z + \frac{\delta_{ij}}{\sigma} \right) \quad \text{and} \quad \tilde{\phi}(z) = \phi\left(z\sqrt{J_i} + \frac{\sum_{j \neq i:\delta_{ij} \leqslant 0} \delta_{ij}}{\sigma\sqrt{J_i}} \right).$$

The choice probability (Eq. (A.1)) is now

$$\rho_i = \frac{\exp(-\kappa/2\sigma^2)}{(2\pi)^{(J_i-1)/2}} \int_{-\infty}^{\infty} \tilde{\phi}(z) \prod_{j:\delta_{ij} > 0} \Phi_i \prod_{j:\delta_{ij} \leqslant 0} \frac{\Phi(z + \sigma^{-1}\delta_{ij})}{\phi(z + \sigma^{-1}\delta_{ij})} \, dz. \tag{A.2}$$

Consider $c < \kappa^2$. In this case

$$\frac{\exp(-\kappa^2/2\sigma^2)}{(2\pi)^{(J_i-1)/2}} \exp\left(\frac{c}{2\sigma^2} \right) = \frac{1}{(2\pi)^{(J_i-1)/2}} \exp\left(-\frac{(\kappa^2 - c)}{2\sigma^2} \right) \to 0.$$

In addition, the integrand of Eq. (A.2) tends to zero and hence

$$\exp\left(\frac{c}{2\sigma^2} \right) \times \int_{-\infty}^{\infty} \left\{ \prod_{j \neq i} \Phi\left(z + \frac{\delta_{ij}}{\sigma} \right) \right\} \phi(z) \, dz \to 0.$$

Next consider $c > \kappa$. In this case

$$\frac{\exp(-\kappa^2/2\sigma^2)}{(2\pi)^{(J_i-1)/2}} \exp\left(\frac{c}{2\sigma^2} \right) = \frac{1}{(2\pi)^{(J_i-1)/2}} \exp\left(\frac{(c - \kappa^2)}{2\sigma^2} \right) \to \infty. \tag{A.3}$$

This expression diverges at an exponential rate towards $+\infty$. The integral of (A.2) vanishes to zero, however. The rest of the proof constructs a lower bound on this integral. It will be shown that this lower bound is polynomial in σ^{-1}. Exponential terms dominate polynomials in the limit, and hence the divergent exponential part of (A.3) will dominate. First bound the integral of (A.2) by integrating over a subset of its range. To do this, first find the best alternative to strategy i and denote its advantage over i by δ_H. Hence

$$\delta_H = \max_{j \neq i} \delta_{ji} > 0.$$

Bound the integral as follows:

$$\int_{-\infty}^{\infty} \tilde{\phi}(z) \prod_{j:\delta_{ij}>0} \Phi_i \prod_{j \neq i:\delta_{ij} \leqslant 0} \frac{\Phi(z+\sigma^{-1}\delta_{ij})}{\phi(z+\sigma^{-1}\delta_{ij})}\, dz$$
$$\geqslant \int_{0}^{\sigma^{-1}\delta_H} \tilde{\phi}(z) \prod_{j:\delta_{ij}>0} \Phi_i \prod_{j \neq i:\delta_{ij} \leqslant 0} \frac{\Phi(z+\sigma^{-1}\delta_{ij})}{\phi(z+\sigma^{-1}\delta_{ij})}\, dz.$$

Next bounds are sought on each term in the integrand. First, consider the product of distribution functions

$$\prod_{j:\delta_{ij}>0} \Phi\left(z + \frac{\delta_{ij}}{\sigma}\right) \geqslant [\Phi(z)]^{m-J_i}.$$

This achieves a minimum at the lower limit of integration $z=0$, yielding

$$\prod_{j:\delta_{ij}>0} \Phi\left(z + \frac{\delta_{ij}}{\sigma}\right) \geqslant \frac{1}{2^{m-J_i}} \quad \text{for} \quad \frac{\delta_H}{\sigma} \geqslant z \geqslant 0.$$

Next recall that the hazard ratio ϕ/Φ is decreasing in its argument. Hence

$$\prod_{j \neq i:\delta_{ij} \leqslant 0} \frac{\phi(z+\sigma^{-1}\delta_{ij})}{\Phi(z+\sigma^{-1}\delta_{ij})} \leqslant \left[\max_{j \neq i:\delta_{ij} \leqslant 0} \frac{\phi(z+\sigma^{-1}\delta_{ij})}{\Phi(z+\sigma^{-1}\delta_{ij})}\right]^{J_i-1} = \left[\frac{\phi(z-\sigma^{-1}\delta_H)}{\Phi(z-\sigma^{-1}\delta_H)}\right]^{J_i-1}.$$

Furthermore, on the range of integration this achieves a maximum at $z=0$, yielding

$$\prod_{j \neq i:\delta_{ij} \leqslant 0} \frac{\phi(z+\sigma^{-1}\delta_{ij})}{\Phi(z+\sigma^{-1}\delta_{ij})} \leqslant \left[\frac{\phi(-\sigma^{-1}\delta_H)}{\Phi(-\sigma^{-1}\delta_H)}\right]^{J_i-1} \quad \text{for } \frac{\delta_H}{\sigma} \geqslant z \geqslant 0.$$

On taking the reciprocal, the inequality is reversed, giving a lower bound. Finally, note

$$\int_0^{\sigma^{-1}\delta_H} \tilde{\phi}(z)\, dz = \int_0^{\sigma^{-1}\delta_H} \phi\left(z\sqrt{J_i} + \frac{\sum_{j \neq i:\delta_{ij} \leqslant 0} \delta_{ij}}{\sigma\sqrt{J_i}}\right) dz$$
$$= \frac{1}{\sqrt{J_i}} \int_0^{\sigma^{-1}\delta_H} d\Phi\left(z\sqrt{J_i} + \frac{\sum_{j \neq i:\delta_{ij} \leqslant 0} \delta_{ij}}{\sigma\sqrt{J_i}}\right)$$
$$= \frac{1}{\sqrt{J_i}}\left\{\Phi\left(\frac{\sqrt{J_i}}{\sigma}\left[\delta_H + \frac{1}{J_i}\sum_{j \neq i:\delta_{ij} \leqslant 0} \delta_{ij}\right]\right) - \Phi\left(\frac{1}{\sigma\sqrt{J_i}} \sum_{j \neq i:\delta_{ij} \leqslant 0} \delta_{ij}\right)\right\}.$$

The second term vanishes

$$\lim_{\sigma\to 0} \Phi\left(\frac{1}{\sigma\sqrt{J_i}}\left[\sum_{j\neq i:\delta_{ij}\leqslant 0} \delta_{ij}\right]\right) = 0$$

since $\delta_{ij} < 0$ for some j. Next notice that

$$\delta_H > \frac{(J_i - 1)\delta_H}{J_i} > \frac{\sum_{j\neq i:\delta_{ij}\leqslant 0} \delta_{ji}}{J_i} = -\frac{\sum_{j\neq i:\delta_{ij}\leqslant 0} \delta_{ij}}{J_i}$$
$$\Rightarrow \lim_{\sigma\to 0} \Phi\left(\frac{\sqrt{J_i}}{\sigma}\left[\delta_H + \frac{1}{J_i}\sum_{j\neq i:\delta_{ij}\leqslant 0} \delta_{ij}\right]\right) = 1.$$

Having obtained a bound for $\int_0^{\sigma^{-1}\delta_H} \tilde{\phi}(z)\,dz$, the bounding components are assembled

$$\begin{aligned}
&\int_{-\infty}^{\infty} \tilde{\phi}(z) \prod_{j:\delta_{ij}>0} \Phi_i \prod_{j\neq i:\delta_{ij}\leqslant 0} \frac{\Phi(z+\sigma^{-1}\delta_{ij})}{\phi(z+\sigma^{-1}\delta_{ij})}\,dz \\
&\geqslant \int_0^{\sigma^{-1}\delta_H} \tilde{\phi}(z) \prod_{j:\delta_{ij}>0} \Phi_i \prod_{j\neq i:\delta_{ij}\leqslant 0} \frac{\Phi(z+\sigma^{-1}\delta_{ij})}{\phi(z+\sigma^{-1}\delta_{ij})}\,dz \\
&\geqslant \frac{1}{2^{m-J_i}} \int_0^{\sigma^{-1}\delta_H} \tilde{\phi}(z) \prod_{j\neq i:\delta_{ij}\leqslant 0} \frac{\Phi(z+\sigma^{-1}\delta_{ij})}{\phi(z+\sigma^{-1}\delta_{ij})}\,dz \\
&\geqslant \frac{1}{2^{m-J_i}} \left[\frac{\Phi(-\sigma^{-1}\delta_H)}{\phi(-\sigma^{-1}\delta_H)}\right]^{J_i-1} \int_0^{\sigma^{-1}\delta_H} \tilde{\phi}(z)\,dz \\
&\to \frac{\sigma^{J_i-1}}{2^{m-J_i}\sqrt{J_i}\delta_H^{J_i-1}},
\end{aligned}$$

where the last step employs the asymptotic linearity of normal hazards.[13] This is polynomial in σ^{-1} and dominated asymptotically by the exponential term, so

$$\int_{-\infty}^{\infty} \left\{\prod_{j\neq i} \Phi\left(z+\frac{\delta_{ij}}{\sigma}\right)\right\}\phi(z)dz \times \exp\left(\frac{c}{2\sigma^2}\right) \to \infty,$$

which completes the proof. □

Proof of Lemma 3. Start in state s. The population updates one player at a time, and hence must lose a strategy and gain a (possibly identical) strategy. It loses strategy k with probability s_k/n, and the entrant adopts strategy $i\neq k$ with probability $\rho_i(s/n)$. Alternatively, the process may not shift, requiring any strategy to be lost and replaced by the same, yielding the summation. □

Proof of Lemma 5. Follows directly from Lemma 2. □

[13] $\phi(x)/(1-\Phi(x))$ is asymptotically linear as $x\to\infty$, following from an application of l'Hôpital's rule.

Proof of Lemma 4. From [4]. □

Proof of Proposition 2. Define $c_{\min} = \min_{s' \in S} \min_{h' \in H_{s'}} \{c(w_{h'})\}$. This is the weight of the least cost rooted tree across all nodes. Take a state $s \in S$ and $s' \notin S$. Using Lemma 2 the cost of q_s is $c(q_s) = \min_{h \in H_s} c(w_h) = c_{\min}$. Similarly, the cost of $q_{s'}$ is $c(q_{s'}) = \min_{h' \in H_{s'}} c(w_{h'}) > c_{\min}$. Hence, again using Lemma 2 it must be the case that $\lim_{\sigma \to 0} [q_{s'}/q_s] = 0 \Rightarrow \lim_{\sigma \to 0} \mu_{s'} = 0$. Thus all states outside S have zero weight in the limit, which corresponds to the desired result. □

Proof of Proposition 3. Take a 2×2 coordination game as required. This may be normalised to yield a pure coordination game as given in Eq. (2), where $(1,1)$ is risk-dominant. Add a third dominated strategy, as shown in Eq. (3). It is first necessary to check that the equilibrium $(1,1)$ is 1/2-dominant. Strategy 1 must be a best response to any mixed strategy satisfying $x_1 \geqslant 1/2$. In this case, the payoff from strategy 1 is $x_1 a_{11}$. The payoff from strategy 2 is $(x_2 + x_3)a_{22} = (1 - x_1)a_{22} < x_1 a_{11}$, following from the risk dominance of $(1,1)$ in the original 2×2 game. It remains to show that strategy 2 is selected as $\sigma \to 0$.

Begin by considering a tree rooted at state $s = (n, 0, 0)$. This requires at least one step away from state $s' = (0, n, 0)$. A best response to state s' is strategy 2, and hence a step away requires a revising player to choose either strategy 1 or strategy 3. It is easiest for such a player to choose strategy 1. Both strategies 1 and 2 are weakly better responses than strategy 1, and hence $J_1 = 2$. Application of Proposition 1 shows that the cost of this transition is $a_{22}^2/2$. This provides a lower bound for the weight of any tree rooted at s. Next, consider a tree rooted at s'. Such a tree must provide an "escape route" from s. Starting at s, construct a sequence of $\lceil n a_{11}/(a_{11} + a_{22}) \rceil$ transitions where at each step a revising player chooses strategy 3 rather than strategy 1. At each step on this route, strategy 1 is a better response than strategy 3. The difference in expected payoffs is ε, and hence the cost of each step is at most $\varepsilon^2/2$.[14] At the end of this route, it is a best response to choose strategy 2. Hence all other states can be mapped to a best response with zero exponential cost. This yields a tree with an exponential cost of at most $\varepsilon^2 \lceil n a_{11}/(a_{11} + a_{22}) \rceil /2$. This tree has a lower cost than any tree rooted at s if $\varepsilon^2 \lceil n a_{11}/(a_{11} + a_{22}) \rceil < a_{22}^2$. This inequality will hold for ε sufficiently small. In fact, a tighter bound is available. An application of Propositions 1 and 2 reveals that a sufficient condition is

$$\varepsilon^2 < \left[\sum_{k=0}^{\lfloor n a_{22}/(a_{11}+a_{22}) \rfloor} \frac{[k a_{11} - (n-k) a_{22}]^2}{k^2 + (n-k)^2} \right] \Bigg/ \left[\sum_{k=0}^{\lfloor n a_{11}/(a_{11}+a_{22}) \rfloor} \frac{n^2}{k^2 + (n-k)^2} \right],$$

where $\lfloor y \rfloor$ is the greatest integer weakly below y. □

[14] As the number of strategy 3 players increases it may eventually become easier for a revising player to jump directly to strategy 2. This can only lower the cost of such a transition.

References

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

[2] L.E. Blume, How noise matters, Working Paper, Santa Fe Institute, 1999.

[3] G. Ellison, Basins of attraction, long run stochastic stability and the speed of step-by-step evolution, Rev. Econom. Stud. 67 (2000) 17–46.

[4] M.I. Freidlin, A.D. Wentzell, Random Perturbations of Dynamical Systems, Springer, Berlin, 1984.

[5] G.R. Grimmett, D.R. Stirzaker, Probability and Random Processes, 3rd Edition, Oxford University Press, Oxford, 2001.

[6] J.C. Harsanyi, Games with randomly disturbed payoffs: a new rationale for mixed-strategy equilibrium points, Int. J. Game Theory 2 (1973) 1–23.

[7] J.C. Harsanyi, R. Selten, A General Theory of Equilibrium Selection in Games, MIT Press, Cambridge, MA, 1988.

[8] M. Kandori, G.J. Mailath, R. Rob, Learning, mutation and long-run equilibria in games, Econometrica 61 (1993) 29–56.

[9] T. Maruta, On the relationship between risk-dominance and stochastic stability, Games Econom. Behav. 19 (1997) 221–234.

[10] S. Morris, R. Rob, H.S. Shin, Dominance and belief potential, Econometrica 63 (1995) 145–157.

[11] D.P. Myatt, C. Wallace, Adaptive play by idiosyncratic agents, Discussion Paper 89, Department of Economics, Oxford University, 2002.

[12] P.A. Ruud, Approximation and simulation of the multinomial probit model: An analysis of covariance matrix estimation, Working Paper, University of California, Berkeley, 1996.

[13] F. Vega-Redondo, The evolution of Walrasian behavior, Econometrica 65 (1997) 375–384.

[14] H.P. Young, The evolution of conventions, Econometrica 61 (1993) 57–84.