The Evolution of Conflict under Inertia*

Thomas Norman
Nuffield College, Oxford

February 17, 2003

Abstract

In Norman (2003a), the introduction of individual strategy switching costs, and thus inertia, into stochastic evolutionary coordination games was found *inter alia* to strengthen the mixed-strategy equilibrium as a short- to medium-run equilibrium. This paper considers the impact of such switching costs on the conflict scenario of Hawk-Dove games. The "attractive" mixed-strategy equilibrium of Hawk-Dove games represents a far better candidate for *long-run* equilibrium than its unstable counterpart in coordination games, and yet robust selection results have proved elusive, with conditions on the selection dynamics generally being required. Such a condition remains a necessity in the switching cost model with state-independent mutations. However, a more realistic model of state-*dependent* mutations driven by stochastic switching costs overcomes this problem, and identifies a threshold mean switching cost, above which the mixed-strategy equilibrium is selected in the long run for a wide class of switching cost distributions.

*Discussions with my supervisors, David Myatt and Chris Wallace, as well as with Margaret Meyer, Kevin Roberts and seminar participants at the Department of Economics, Oxford University, are gratefully acknowledged. The usual disclaimer applies.

1 Introduction

"The truth is rarely pure, and never simple."
Oscar Wilde, *The Importance of Being Earnest* (1895) Act I

In Norman (2003a), the implications of player inertia in stochastic evolutionary coordination games were considered. The presence of "switching costs" to individual strategy changes in the Kandori, Mailath, and Rob (1993) (henceforth KMR) model was found to create a number of new short- to medium-run equilibria, centred around the mixed-strategy equilibrium. However, this unusual evolutionary justification for the controversial concept of mixed-strategy equilibrium did not extend to the long run, where risk-dominance continued to reign supreme. But if the "unattractive" mixed-strategy equilibrium of coordination games can be strengthened by inertia, it is natural to ask what will be the effect in games which have an "attractive" mixed-strategy equilibrium. This question motivates this paper's shift of focus to the conflict scenario of Hawk-Dove games.

In contrast to coordination games, the mixed-strategy equilibrium in Hawk-Dove games constitutes the unique *symmetric* Nash equilibrium. Yet despite this "attractiveness" of the Hawk-Dove mixed-strategy equilibrium, existing models (Kandori, Mailath, and Rob 1993, Robson and Vega-Redondo 1996) have not found it to be selected *robustly* as a long-run equilibrium. Instead, long-run behaviour in this "most problematic type of game"[1] depends crucially on the precise selection dynamic employed, and convergence to the mixed-strategy equilibrium generally requires some restrictive conditions to be placed on that dynamic. It is shown in section 3 that this unhappy situation is essentially unaltered by the presence of switching costs when the mutation rate is uniform; the same restrictive conditions must be placed on the selection dynamic in order for the mixed-strategy equilibrium to be selected, and in fact some dilution of this selection occurs due to the creation of competing mixed absorbing states.

However, in section 4, the model is altered to make the switching cost itself stochastic across players and time. This is a realistic modification; in reality, different players (with different priorities, abilities and constraints) will vary in the size of their switching costs, and any given player's switching cost will fluctuate in size over time (as his priorities, abilities and constraints change). Moreover, the introduction of this stochastic switching cost obviates the need to introduce the state-independent mutations of the KMR-style model, thus also serving to address the criticisms of Bergin and Lipman (1996) by building a model with endogenously generated "state-dependent mutations". In this setting, the problem of requiring the selection dynamic to meet a restrictive condition in order to select the mixed-strategy equilibrium of Hawk-Dove games is overcome. Instead, there is a threshold mean

[1] Robson and Vega-Redondo (1996), p. 67.

switching cost, above which the appropriate selection takes place for a wide class of switching cost distributions.

To begin with, however, the next section introduces the relevant literature.

2 Relevant Literature

There is a longstanding controversy in game theory over the validity of mixed-strategy equilibrium. Mixed-strategy equilibria are seen by many as dubious on a number of grounds, chiefly doubts over whether players randomise and whether knife-edge behaviour (playing with the exact required probabilities) based on indifference can be stable. Various responses have been offered to these criticisms, generally offering different interpretations of mixed-strategy equilibrium to that of individual-level randomisation taken literally.[2] The most notable of these is Harsanyi's (1973) purification argument, which implies that *all* mixed equilibria are approximations of strict, and hence evolutionarily stable, equilibria of games with slightly perturbed payoffs. However, the most interesting interpretation of mixed-strategy equilibrium from an evolutionary standpoint is as a steady state of an environment in which players act repeatedly and ignore any strategic link that may exist between plays.[3] This is precisely the stochastic evolutionary paradigm: statistical frequencies of current and/or past play form the basis of players' beliefs about the future behaviour of the other players, which in turn form the basis of their own play.

Stochastic evolutionary game theory was born in the papers of Foster and Young (1990), Kandori, Mailath, and Rob (1993), and Young (1993). Foster and Young (1990), drawing on the Markovian techniques of Freidlin and Wentzell (1984), introduced the concept of stochastic stability in continuous dynamical systems into evolutionary biology, and provided a method for the analytical computation of the stochastically stable set. Kandori, Mailath, and Rob (1993), and Young (1993) then brought the somewhat simpler discrete analysis into the realm of economics. These papers provided the basis of a new and fruitful direction for evolutionary game theory, appearing to offer a solution to the persistent problem of multiple equilibria and path dependence: the long-run equilibrium of a stochastic model was unique for generic games with strict Nash equilibria.

This "solution" was, however, soon questioned by Bergin and Lipman (1996), who highlight the arbitrariness of mutations occurring at a rate independent of the current state of the system. Such "state-independent mutations", embodied in the fixed mutation rate ε of KMR and others, imply that

[2]See Osborne and Rubinstein (1994), pp. 37-44, for an excellent survey of such interpretations.

[3]See Osborne and Rubinstein (1994), pp. 38-9.

players make mistakes (or experiment, etc.) with the same probability irrespective of the current strategy frequencies, and thus of the expected payoffs at stake. Bergin and Lipman demonstrate that, given any model of the effect of mutations, any invariant distribution of the "mutationless" process is close to an invariant distribution of the process with appropriately chosen small mutation rates. This implies that any strict Nash equilibrium of a strategic form game is selected under some suitably chosen mutation model. Bergin and Lipman's paper highlights the importance of developing models or other criteria to determine "reasonable" classes of "state-*dependent* mutations". Myatt and Wallace (1998) present a candidate for just such a "reasonable" mutation process with their model of state-dependent mutations driven by payoff heterogeneity (rather than mistakes, or experimentation). More generally, Blume (1999) alleviates the indeterminacy of the Bergin and Lipman critique with his finding that the known stochastic stability results are preserved for the (large) class of noise processes satisfying a certain symmetry condition.

On the basis of existing stochastic evolutionary work, however, there is little support for the population steady state interpretation of mixed-strategy equilibrium. In most stochastic evolutionary models, mixed-strategy equilibria are unstable knife-edges, and as such are comprehensively deselected in the long run. As will be seen in section 3 below, Kandori, Mailath, and Rob (1993) and Robson and Vega-Redondo (1996) do provide theorems for long-run selection of the mixed-strategy equilibrium in Hawk-Dove games, but even this requires the imposition of apparently ad hoc constraints on the selection dynamics. Oechssler (1997) also investigates evolutionary convergence to mixed-strategy equilibrium, but again for a particular form of selection dynamic; every period each player receives the opportunity to adjust his strategy with probability $\theta \in (0, 1)$. This variable speed dynamic guarantees eventual convergence to the mixed-strategy equilibrium in Hawk-Dove games.

Some models of learning and deterministic evolution have sought to justify mixed-strategy equilibrium in population contexts. In two related papers, Fudenberg and Kreps (1993) and Ellison and Fudenberg (2000) explore a potential justification for mixed-strategy equilibria based on the idea that an equilibrium distribution might arise in a large population as the result of a learning process in the style of fictitious play. This follows the earlier pessimistic result of Crawford (1985) that such a process is locally unstable at the mixed equilibrium for almost all games. Fudenberg and Kreps (1993) find play unlikely to converge to a mixed-strategy equilibrium in their model, until the stage game is perturbed in the manner of Harsanyi's purification theorem, at which point such convergence becomes natural. Their paper and several subsequent analyses show that, for the type of "smoothed" learning models studied, play in 2×2 games converges to the mixed equilibrium in games like matching pennies, while the seemingly unreasonable mixed equilibria of coordination games are unstable. Ellison and Fudenberg (2000), meanwhile, extend the model to more complicated settings, and find that - contrary to some previous suggestions - learning can sometimes provide a justification for complicated mixed equilibria. Whether an equilibrium is stable is found often to depend on the distribution of payoff perturbations.

On the evolutionary side, meanwhile, Eshel and Sansone (1995) also appeal to Harsanyi's purification argument in warning against a premature rejection of mixed strategies in an evolutionary context, arguing that small payoff perturbations and incomplete information about opponents' payoffs are inescapable in realistic evolutionary models. Binmore and Samuelson (2001) explore this idea further, seeking to reconcile Harsanyi's defence of mixed-strategy equilibrium with Selten's (1980) demonstration that *no* mixed equilibria are evolutionarily stable when players can condition their strategies on their roles in a game. They find that approximations of mixed equilibria are likely to persist when payoff perturbations are relatively important and role identification is relatively noisy, but are unlikely to persist when payoff perturbations are unimportant and role identification is precise.

However, despite all of this support in deterministic contexts (albeit perturbed), as far as selection in explicit stochastic models is concerned, the mixed-strategy equilibrium remains on very shaky ground. The contention of this paper is that this ground becomes somewhat firmer in the presence of switching costs.

The idea of inertia driven by switching costs in repeated game contexts is of course not a new one. Within the traditional perfectly rational paradigm, Klemperer's (1987a, 1987b) consumer switching costs[4] and Radner's (1980) ε-Nash equilibrium are two obvious examples. The closest parallel to the present paper in the perfectly rational literature, however, is Lipman and Wang (2000), who study the effects of switching costs in general repeated game contexts. They add small costs of changing actions and frequent repetition to finitely repeated games, and find that doing so makes credible certain commitments which then serve to overturn all the standard results for finitely repeated games.

By contrast with the perfectly rational literature, however, it would seem that inertia has for the most part been left unmodelled in evolutionary contexts. In reinforcement models, for example, the probability of taking an action in the present increases with the payoff that resulted from taking that action in the past.[5] Such models admit an intuitive role for inertia, but this inertia is assumed exogenous, its root causes left unmodelled. Meanwhile, inertia plays an explicit role in the KMR (1993) model, with their weakly

[4]On consumer switching costs see, for example, Klemperer (1995), Beggs and Klemperer (1992), Farrell and Klemperer (2001), Farrell (1987), Farrell and Shapiro (1988, 1989), and Padilla (1995).

[5]See Bush and Mosteller (1955), Suppes and Atkinson (1960), Arthur (1993), Roth and Erev (1995), Börgers and Sarin (1995, 1997).

monotonic selection dynamic capturing the idea that only some (as opposed to all) players need be adjusting their behaviour in any given period. But this ignores the important possibility that *no* player adjusts his behaviour in a given period, and moreover, there is again no endogenous determination of the inertia.

There are admittedly some evolutionary papers that model costly play of some sort. One example is Sethi (1998), whose model of "strategy-specific barriers to learning" in the replicator dynamics explores the consequences of strategies varying in the ease with which they can be learned. Another example is van Damme and Weibull's (1998) model of "mutations driven by control costs", which has mutation rates determined by individual mistake probabilities, which players can control at some cost. State-dependent mutations are thus based on an economically justified model here, in the manner suggested by Bergin and Lipman (1996). But whilst both of these papers model important ways in which strategy adoption might be costly, neither captures the idea that strategy *change* is costly compared to the (cost-free) status quo. The *switching costs* postulated in this paper, by contrast, draw attention to the learning and implementation costs incurred when a player switches to a new strategy - a plausible source of individual-level inertia.

3 Uniform Mutations

As has been seen, the unique symmetric Nash equilibrium in Hawk-Dove games is that in mixed strategies. However, despite this "attractiveness" of the mixed-strategy equilibrium in Hawk-Dove games, existing evolutionary models have not found it to be selected *robustly* as a long-run equilibrium. Indeed, Hawk-Dove games are seen by some as "the most problematic type of game"[6]. Applying their model to Hawk-Dove games, Kandori, Mailath, and Rob (1993) do derive a result selecting the mixed-strategy equilibrium in the long run, but only under a certain condition on the selection dynamics.[7] In effect, adjustment must be sufficiently slow to rule out overshooting of the mixed-strategy equilibrium. This condition seems somewhat ad hoc, and so this section explores whether introducing switching costs into the model can serve to robustify the result, as the findings of Norman (2003a) suggest it might. In fact, analysis reveals that it cannot do so to any satisfactory degree, a similar condition being required on the deterministic dynamic for long-run selection of the mixed-strategy equilibrium to hold in the presence of switching costs as in their absence. Whilst the condition on the selection dynamics is weaker under switching costs, the long-run selection of the mixed-strategy equilibrium is also *diluted* somewhat, given the concomitant

[6] Robson and Vega-Redondo (1996), p. 67.

[7] Robson and Vega-Redondo (1996) derive a similar result for a sufficiently large population in their model.

creation of competing mixed absorbing states. These findings highlight the limitations in the fixed-switching cost, state-independent mutations model employed.

3.1 The Model

To begin with then, the uniform mutation rate model of Norman (2003a) is employed. This is essentially the Kandori, Mailath, and Rob (1993) model in the presence of a *switching cost* c of changing strategies at the individual player level. Thus, there is a finite number N (N an even number) of players who are repeatedly matched to play the symmetric 2×2 stage game defined by the general payoff matrix

$$\Lambda = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \tag{1}$$

and who adjust their behaviour over time. Actions are taken in discrete time, $t = 1, 2, \ldots$, each player choosing his pure strategy for a period t at the beginning of that period. The number z^t of players playing strategy 1 at time t defines the state of the dynamical system, the state space being $Z = \{0, 1, \ldots, N\}$. The average payoff of a player with strategy i, $\pi_i(z^t)$, is assumed to be[8]

$$\pi_1(z) = \frac{(z-1)}{(N-1)} a_{11} + \frac{(N-z)}{(N-1)} a_{12}, \tag{2}$$

$$\pi_2(z) = \frac{z}{(N-1)} a_{21} + \frac{(N-z-1)}{(N-1)} a_{22} \tag{3}$$

These become expected payoffs under the assumption that players are myopic.

KMR's "Darwinian" selection dynamics - under which better strategies are better represented in the population in the next period - are once again modified to include a fixed switching cost c, incurred whenever a player switches strategies.

Definition 1 *A* 1-incumbent *(resp.* 2-incumbent*) in period t is a player who was a 1-strategist (resp. 2-strategist) in period $t-1$.*

Definition 2 *The* cost-adjusted payoff matrices *are*

$$\Lambda_1 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} - c & a_{22} - c \end{bmatrix} \tag{4}$$

[8]See KMR (1993), p. 37, for models of matching which generate these payoffs. Note that a player will ignore his own play in forming his beliefs from the current population strategy frequency.

and

$$\Lambda_2 = \begin{bmatrix} a_{11} - c & a_{12} - c \\ a_{21} & a_{22} \end{bmatrix} \tag{5}$$

for 1*-incumbents and* 2*-incumbents, respectively.*

Given the presence of this switching cost, KMR's underlying *deterministic dynamic*,

$$z^{t+1} = b(z^t) \tag{6}$$

now has the "modified Darwinian" property:[9,10]

$$(b(z) - z) \text{ is } \left\{ \begin{array}{ll} \text{strictly negative} & \text{iff } \pi_1(z) < \pi_2(z) - c \\ \text{nonnegative} & \text{iff } \pi_1(z) \geq \pi_2(z) - c \\ \text{nonpositive} & \text{iff } \pi_1(z) - c \leq \pi_2(z) \\ \text{strictly positive} & \text{iff } \pi_1(z) - c > \pi_2(z) \end{array} \right\} \text{ for } z \neq 0, N. \tag{7}$$

Or, more succinctly,[11]

$$(b(z) - z) \text{ is } \left\{ \begin{array}{ll} \text{strictly negative} & \text{iff } \pi_1(z) - \pi_2(z) < -c \\ 0 & \text{iff } -c \leq \pi_1(z) - \pi_2(z) \leq c \\ \text{strictly positive} & \text{iff } \pi_1(z) - \pi_2(z) > c \end{array} \right\} \text{ for } z \neq 0, N. \tag{8}$$

The model is now made stochastic by introducing some noise ("mutations") into the system. Following KMR, it is assumed that each player's strategy "flips" with probability ε in each period (i.i.d. across players and over time).[12] This yields the nonlinear stochastic difference equation

$$z^{t+1} = b(z^t) + x^t - y^t, \tag{9}$$

where x^t and y^t have the binomial distributions:

$$x^t \sim Bin(N - b(z^t), \varepsilon) \quad \text{and} \quad y^t \sim Bin(b(z^t), \varepsilon)$$

The dynamical system in equation (9) defines a Markov chain on the finite state space $Z = \{0, 1, \ldots N\}$. $P^\varepsilon = [p_{ij}]$ is the Markov matrix, with the transition probabilities given by

$$p_{ij} = \Pr(z^{t+1} = j \mid z^t = i) \tag{10}$$

[9] The strict and weak inequalities are assigned according to the assumption that indifference leads to inertia at the individual level. This assumption is unimportant for the results.

[10] For the extreme states, 0 and N, it is assumed that $b(0) > 0$ iff $\pi_1(0) - c > \pi_2(0)$, and $b(0) = 0$ iff $\pi_1(0) - c \leq \pi_2(0)$. Similarly, $b(N) < N$ iff $\pi_1(N) < \pi_2(N) - c$, and $b(N) = N$ iff $\pi_1(N) \geq \pi_2(N) - c$.

[11] For the extreme states, similarly, $b(0) > 0$ iff $\pi_1(0) - \pi_2(0) > c$,and $b(0) = 0$ iff $\pi_1(0) - \pi_2(0) \leq c$; $b(N) < N$ iff $\pi_2(N) - \pi_1(N) > c$, and $b(N) = N$ iff $\pi_2(N) - \pi_1(N) \leq c$.

[12] The usual stories of "experimentation" or of players dying (with probability 2ε) and being replaced by ignorant newcomers may be appealed to. For example, see KMR (1993), pp. 38-9.

All elements in the matrix P^{ε} are strictly positive under the above assumptions.

The long-run behaviour of the Markov chain in equation (9) is captured by the *stationarity equations* $\mu^{\varepsilon}P^{\varepsilon} = \mu^{\varepsilon}$, the solution of which is the distribution μ^{ε} (over states) that is stationary under P^{ε}. However, it is not immediately clear whether this solution is unique. If it is not, long-run behaviour will be sensitive to initial conditions, so that the path dependence of deterministic models remains. However, for an *ergodic* Markov process, the stationarity equations *will* have a unique solution and the long-run behaviour embodied in μ^{ε} will be independent of initial conditions. The unique invariant distribution μ^{ε} may then be used to explore the equilibrium selection issue by providing information on whether some outcomes are much more likely than others. Indeed, μ^{ε} can be interpreted as the proportion of time that the society spends in each state in the long run.

Lemma 1 *An irreducible and aperiodic Markov chain is ergodic.*

This is a standard result - see, for instance, Theorem 11.2 of Stokey, Lucas, and Prescott (1989).

Proposition 1 *The adaptive response dynamic defined by the transition probabilities p_{ij} in equation (10) is an irreducible, aperiodic Markov process on the finite state space Z. Consequently, it has a unique invariant (ergodic) distribution μ^{ε}.*

Proof. Since all elements in the Markov matrix P^{ε} are strictly positive, i.e. any state is accessible from any other state in a single period, it follows that the process is irreducible. Moreover, since in every state there is a positive probability of the system remaining in that state in the next period, the process is aperiodic. Lemma 1 then implies that the process has a unique invariant distribution. ■

The stage game Each time the players are matched to play the game Λ in (1), they will play a symmetric "Hawk-Dove" game. The class of symmetric Hawk-Dove games is the set of all games satisfying the conditions $a_{11} < a_{21}$ and $a_{22} < a_{12}$. This class of games has two pure-strategy Nash equilibria, $(1, 2)$ and $(2, 1)$, and one mixed-strategy Nash equilibrium which has strategy 1 being played with probability

$$\rho = \frac{a_{22} - a_{12}}{(a_{22} - a_{12}) + (a_{11} - a_{21})}$$

There is thus a set of three Nash equilibria, $\Theta^{NE} = \{(1, 2), (2, 1), (\rho, \rho)\}$, with no apparent way of selecting between them - the classic game-theoretic problem of multiple equilibria. However, the mixed-strategy equilibrium is

	Dove	Hawk
Dove	$\frac{1}{2}V$, $\frac{1}{2}V$	0 , V
Hawk	V , 0	$\frac{1}{2}V-K$, $\frac{1}{2}V-K$

Figure 1: Maynard Smith's (1982) "Hawk-Dove" game

immediately somewhat more appealing than the two pure-strategy equilibria, by virtue of being the unique *symmetric* Nash equilibrium.[13]

Example 1 *Maynard Smith's (1982) "Hawk-Dove" game in figure 1 modelled two animals contesting a resource worth $V > 0$. In this game's original biological context, the payoffs are interpreted as the incremental fitnesses that accrue to the two animals, where "fitness" is the property of an animal that determines its likelihood of survival/reproduction. Each animal can choose to be aggressive (playing Hawk) or to acquiesce (playing Dove). If one animal plays Hawk and the other animal plays Dove, then the aggressive animal takes the whole resource and the acquiescent animal gets nothing. If, meanwhile, both animals choose the same pure strategy, then each is equally likely to get the resource, but if the strategy played is Hawk, the resulting conflict causes each to pay a cost of $K \geq V$. The game has two pure-strategy Nash equilibria, $(Hawk, Dove)$ and $(Dove, Hawk)$, but a unique symmetric Nash equilibrium in mixed strategies which has each player using Hawk with probability $\frac{K-\frac{1}{2}V}{K}$.*

This sort of game has obvious applications to international conflicts, and the "Hawk-Dove" terminology is widely used in the international relations literature.

[13]It is worth noting that the principal refinement of *deterministic* evolutionary game theory, namely *evolutionary stability*, selects the mixed-strategy equilibrium as the unique *evolutionarily stable strategy*.

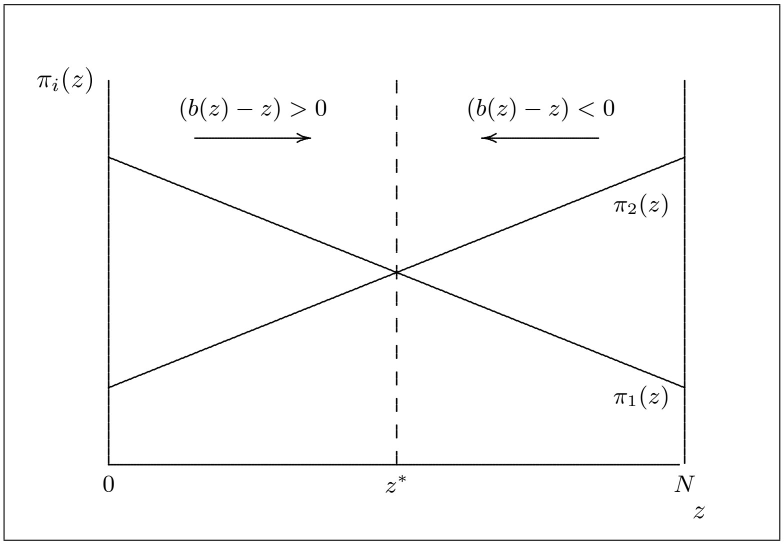


Figure 2: KMR (1993) model: Hawk-Dove games

3.2 Analysis

3.2.1 The KMR theorem

In figure 2, the average payoffs in equations (2) and (3) are graphed as linear functions of z,

$$\pi_1(z) = \frac{(a_{11} - a_{12})}{(N-1)} z + \frac{Na_{12} - a_{11}}{(N-1)} \tag{11}$$

$$\pi_2(z) = \frac{(a_{21} - a_{22})}{(N-1)} z + a_{22} \tag{12}$$

Note that the $\pi_2(z)$ line will be steeper than the $\pi_1(z)$ line if and only if

$$a_{11} - a_{21} < a_{12} - a_{22}$$

and that this holds necessarily for Hawk-Dove games given that $a_{11} < a_{21}$ and $a_{22} < a_{12}$.[14] Moreover, these same two defining conditions of Hawk-Dove games imply that $\pi_2(z) < \pi_1(z)$ for $z = 0$ and that $\pi_2(z) > \pi_1(z)$ for $z = N$ (as is clear from equations (2) and (3)), so that there will certainly exist a value $z^* \in Z$ where the two average payoff lines cross (though this z^* need not, of course, be an integer). This corresponds to the mixed-strategy equilibrium of the Hawk-Dove stage game, which puts probability

[14] There is, however, no implication that $\pi_2(z)$ need have a positive slope and $\pi_1(z)$ a negative slope.

$\rho = \frac{(a_{22}-a_{12})}{a_{11}-a_{21}+a_{22}-a_{12}}$ on strategy 1.[15] It is straightforward to solve for the value of z^*:

$$\begin{aligned} 0 &= \pi_1(z^*) - \pi_2(z^*) \\ \overset{(2),(3)}{\Longrightarrow} z^* &= \frac{N(a_{22}-a_{12}) + a_{11} - a_{22}}{\theta}, \end{aligned} \tag{13}$$

where $\theta = (a_{11} - a_{21} + a_{22} - a_{12})$ is the sum of the normalised stage game payoffs, $a_1 = a_{11} - a_{21}$ and $a_2 = a_{22} - a_{12}$.

Looking at figure 2, it is clear that the mixed-strategy equilibrium of Hawk-Dove games is "attractive" in the sense that the deterministic dynamic points towards z^* in all states. In the coordination game setting of KMR (1993) and Norman (2003a), by contrast, the configuration of the two average payoff lines was reversed, so that the average payoff line for strategy 1 was steeper than that for strategy 2, and the deterministic dynamic pointed *away* from the mixed-strategy equilibrium z^* in any given state.[16] The "attractiveness" of z^* in the Hawk-Dove setting accords with the fact that the unique symmetric Nash equilibrium of such games is that in mixed strategies.

In order to investigate whether this "attractiveness" of the mixed-strategy equilibrium leads it to be selected as the long-run equilibrium of the model, use is made of Young's (1993) method[17] for computing the stochastically stable states of a regular[18] perturbed Markov process. This method is based on the notion of *rooted trees* constructed on the set of the recurrent class(es) E_1, E_2, ..., E_K of the unperturbed process P^0. For a given pair of (distinct)

[15] In this population context, the *exact* analogue of the stage game mixed-strategy equilibrium is the state ρN (which again need not be an integer) where a fraction ρ of all players are playing strategy 1. Note that z^* will not be exactly equal to ρN because 1-incumbents and 2-incumbents face slightly different strategy distributions due to the finiteness of the population. The difference between z^* and ρN does, however, vanish as the population size becomes large.

[16] Note that the only other possible configuration of the two average payoff lines in figure 2 is for the lines not to cross within the state space. This then corresponds to the remaining class of symmetric 2×2 games - that where there is a dominant strategy (the strategy whose average payoff line lies entirely above that of the other strategy). Question-beggingly high switching costs are required to overthrow the dominant-strategy equilibrium as the unique long-run equilibrium of this class of games.

[17] See Young (1998), section 3.4, for more detail and illustrative examples.

[18] Let P^ε be a Markov process on Z for each ε in some interval $[0, \varepsilon^*]$. P^ε is described as a *regular* perturbed Markov process if P^ε is irreducible for every $\varepsilon \in (0, \varepsilon^*]$, and for every $i, j \in Z$, p^ε_{ij} approaches p^0_{ij} at an exponential rate, that is,

$$\lim_{\varepsilon \to 0} p^\varepsilon_{ij} = p^0_{ij},$$

and

$$\text{if } p^\varepsilon_{ij} > 0 \text{ for some } \varepsilon > 0, \text{ then } 0 < \lim_{\varepsilon \to 0} \frac{p^\varepsilon_{ij}}{\varepsilon^{r(i,j)}} < \infty \text{ for some } r(i,j) \geq 0.$$

The real number $r(i,j)$ is called the *resistance* of the transition $i \to j$. Note that transitions that can occur under P^0 have zero resistance (i.e. $p^0_{ij} > 0$ if and only if $r(i,j) = 0$).

recurrent classes E_i and E_j, an ij-path is a sequence of states that begins in E_i and ends in E_j. The *resistance* of this path is then the number of mutations required to transit from recurrent class E_i to recurrent class E_j along this path, and the minimum resistance over all possible ij-paths is denoted r_{ij}.[19] Construct a complete directed graph with K vertices (one for each recurrent class), and weight each directed edge $i \to j$ with the appropriate minimum resistance r_{ij}. A tree rooted at a particular recurrent class E_j's vertex j (a *j-tree*) is then a set of $(K-1)$ directed edges such that there is a unique directed path in the tree to j from every vertex (i.e. recurrent class) other than j. The *resistance of a rooted tree T* is the sum of the resistances r_{ij} on T's $(K-1)$ edges, and the minimum resistance over all trees rooted at j is called the *stochastic potential* γ_j of the recurrent class E_j.

Definition 3 *A state z is* stochastically stable *(Young 1993) if*

$$\lim_{\varepsilon \to 0} \mu^{\varepsilon}(z) > 0$$

The selection criterion of *stochastic stability* identifies which outcome(s) receive positive weight in the ergodic distribution as the amount of noise tends to zero. Intuitively, stochastically stable states are those that are most likely to be observed over the long run when noise is small, and they are thus sometimes referred to as the *long-run equilibria* of a system.

Lemma 2 (Young (1993)) *If P^{ε} is a regular perturbed Markov process, and μ^{ε} is its unique stationary distribution for each $\varepsilon > 0$, then $\lim_{\varepsilon \to 0} \mu^{\varepsilon}(z) = \mu^0$ exists, and the limiting distribution μ^0 is a stationary distribution of the unperturbed process P^0. The stochastically stable states are precisely those states contained in the recurrent class(es) of P^0 having minimum stochastic potential.*

The intuition for this result is that, for a small and positive noise level ε, the process is most likely to follow paths leading towards the recurrent classes having minimum potential.

Returning to the application at hand, for the coordination games of KMR (1993), the set of recurrent classes of the unperturbed process was invariant under the admitted class of "Darwinian" selection dynamics. Moreover, from any given state the underlying deterministic dynamic pointed in the direction of one and only one absorbing state. For states below the mixed-strategy equilibrium ($z < z^*$), the deterministic dynamic pointed towards the absorbing state 0, whilst for states above the mixed-strategy equilibrium ($z > z^*$), it pointed towards the absorbing state N. In no state did the deterministic dynamic point towards the mixed-strategy equilibrium. Thus, the speed of the dynamic was unimportant for long-run equilibrium;

[19] Note that other evolutionary models (such as KMR) employ the analogous notion of the *cost* of a transition path, rather than the *resistance*.

immediate best-response by all players would transport the system to the relevant extreme state within one period, but the same state would eventually be reached if the dynamic specified that only one player reviewed his strategy each period.

In Hawk-Dove games, by contrast, the deterministic dynamic in any given state points in the direction of a multitude of possible recurrent classes. This is because the set of recurrent classes of the unperturbed process changes according to the precise specification of the selection dynamic. The speed of adjustment thus becomes crucially important for both short- to medium- and long-run equilibrium. To illustrate, consider what happens if the speed of adjustment is as fast as it could possibly be - i.e. if the deterministic dynamic is the simultaneous-review best-response dynamic $B^N(z)$, under which all players best-respond every period. In this case, the unperturbed process P^0 has one recurrent class, $\{0, N\}$.[20] In any state $z \in [0, z^*)$, all players' best response is to play strategy 1 ($B^N(z) = N$), so that the system jumps to state N in the next period. In any state $z \in (z^*, N]$, meanwhile, all players' best response is to play strategy 2 ($B^N(z) = 0$), so that the system jumps to state 0 in the next period. Thus, once the recurrent class $\{0, N\}$ has been entered, the system oscillates between the two extreme states from one period to the next. Since this is the only recurrent class, it is (trivially) the recurrent class with minimum stochastic potential once noise is introduced, and (by Lemma 2, page 13) is thus selected with probability one by the long-run invariant distribution as $\varepsilon \to 0$ (with each of states 0 and N receiving probability $\frac{1}{2}$).

The results are very different, however, if the speed of adjustment is as slow as it could possibly be under KMR's "Darwinian" dynamics - i.e. if the deterministic dynamic is the single-revisions best-response dynamic $B^1(z)$, under which only one player best-responds each period, the rest retaining their previous strategy. In this case, the mixed-strategy equilibrium (or, more precisely, its neighbouring integers) is selected in the long run. There is again just one recurrent class, but this time it is the set $\{\alpha, \beta\}$ of integers surrounding the mixed-strategy equilibrium z^*. For $z \in [0, z^*)$, $B^1(z) = z+1$, whilst for $z \in (z^*, N]$, $B^1(z) = z-1$. Thus, $B^1(\alpha) = \beta$ and $B^1(\beta) = \alpha$, so that $\{\alpha, \beta\}$ is the unique recurrent class which, once entered, sees the system oscillate between the states α and β from one period to the next. Another trivial application of Lemma 2 shows that the limit distribution of the perturbed process puts probability $\frac{1}{2}$ on each of the states α and β.

In fact, the selection dynamic does not have to be as slow as this "one-step" single-revisions dynamic in order to deliver the mixed-strategy equilibrium as the unique long-run equilibrium. KMR's (1993) Theorem 5 provides a condition on the selection dynamics which, if satisfied, ensures the long-

[20] In the unlikely event that the mixed-strategy equilibrium z^* happens to be an integer, it will also form a (singleton) recurrent class, but an extremely unstable one.

run selection of $\{\alpha, \beta\}$ as above. Given a mixed-strategy equilibrium that places probability ρ on strategy 1 being played, KMR say that $b(z)$ is a *contraction relative to* ρN[21] if

$$\lceil |b(z) - \rho N| \rceil \leq \lceil |z - \rho N| \rceil \qquad \text{for all } z, \tag{14}$$

with the inequality holding strictly if $\rho N \in Z$ or $\lceil |z - \rho N| \rceil \geq 2$. If $b(z)$ is a contraction relative to ρN, and if N is sufficiently large that $1 \leq \rho N \leq N-1$, then the limit distribution mimics the mixed-strategy equilibrium in the way that it did for the one-step dynamic above.

By constraining the selection dynamic to behave in this way, the mixed-strategy equilibrium becomes inevitable, but it is not clear whether the "contraction" condition is a reasonable constraint to impose. For KMR, the condition is best viewed as an assumption on the length of a period: periods are sufficiently short that only a small fraction of the population can adjust. For example, it is satisfied by the above one-step dynamic. However, it is not satisfied by a "two-step" dynamic, nor indeed by any "n-step" dynamic where $n > 1$; a one-step dynamic is required to keep the system at the mixed-strategy equilibrium once it is reached, and any faster n-step dynamic will not select the mixed-strategy equilibrium in the long run. An unperturbed two-step best-response dynamic would have two recurrent classes, $\{(\beta - 1), \alpha\}$ and $\{\beta, (\alpha + 1)\}$, each of which would receive probability $\frac{1}{2}$ under the limit distribution of the perturbed process. Similarly, an unperturbed three-step best-response dynamic would have three recurrent classes, $\{(\beta - 2), \alpha\}$, $\{(\beta - 1), (\alpha + 1)\}$ and $\{\beta, (\alpha + 2)\}$, each of which would receive probability $\frac{1}{3}$ under the limit distribution of the perturbed process. This indeterminacy of the n-step dynamic will grow with n (giving n equally likely recurrent classes), reaching a maximum when $n = \min\{\alpha, N - \beta\}$; at this point, the limits of the state space start to reduce the number of recurrent classes, until it eventually falls to 1 again under the simultaneous-review best-response dynamic. These considerations reveal the stringent requirements imposed by KMR's "contraction" condition, and suggest the focality of the mixed-strategy equilibrium in stochastic evolutionary Hawk-Dove games to be far from robust.

Matters are helped little by moving to a dynamic where a constant fraction of those wishing to switch strategies may do so each period. What the "contraction" condition literally requires is that a sufficiently small fraction of those players whose current best response is to switch strategies actually do so, and that this fraction shrink the closer is the current state to the mixed-strategy equilibrium ρN. Thus, with a dynamic that specifies a constant fraction of would-be strategy-switchers as adjusting each period, the

[21] Recall that ρN is the *exact* mixed-strategy equilibrium analogue, by contrast with the close approximation z^* delivered by players ignoring their own play when forming beliefs (see note 15, page 12).

fraction required to hold the system around the mixed-strategy equilibrium once reached (i.e. $1/\max\{\alpha, N-\beta\}$) provides a (very low) upper bound on the permissible speed of review for the whole process, if the mixed-strategy equilibrium is to be selected in the long run. Indeed, the fastest *possible* review here occurs when $\alpha = N - \beta \approx \frac{N}{2}$, so that the permitted reviewing fraction is $\frac{2}{N}$.

All in all, one is left either with the requirement of very slow review for selection of the mixed-strategy equilibrium, or with trying to provide a foundation for a dynamic which sees the fraction of adjusting players *fall* (sufficiently quickly) as the system approaches the mixed-strategy equilibrium. The natural question to arise at this point, in view of the findings of Norman (2003a), is whether the introduction of switching costs to strategy changes could serve to strengthen the mixed-strategy equilibrium of Hawk-Dove games, and thus robustify its selection as the long-run equilibrium.

3.2.2 Introducing inertia

As in Norman (2003a), the introduction of switching costs into the model creates a region of new mixed absorbing states, born of player inertia when the switching cost outweighs any potential payoff gain from a strategy change.

Proposition 2 *In the presence of a switching cost $c > 0$, there exists a multiplicity of* mixed absorbing states[22] *of the unperturbed process P^0 in addition to the set of recurrent classes present when $c = 0$.*

Proof. The modified Darwinian property of the selection dynamics in equation (8) says that there is a range of the average payoff difference between the two strategies, $(\pi_1(z) - \pi_2(z)) \in [-c, c]$, for which $(b(z) - z) = 0$. Since the average payoff lines in figure 2 will certainly cross at some $z^* \in Z$ (as was shown above), it follows that there will be a range of z, $[z_L, z_H] \in Z$, for which $(b(z) - z) = 0$. Now, it is clear from the deterministic dynamic in equation (6) that any state for which $(b(z) - z) = 0$ is an absorbing state of the unperturbed process P^0 (since then $z^{t+1} = z^t$ for $\varepsilon = 0$). Thus, there is a range of $z \in \{Z \backslash 0, N\}$, each (integer) element of which is a (mixed) absorbing state. ■

Definition 4 *Given equation (8), the lower limit z_L of the range of z for*

[22] A *mixed* state is one where both strategies are being played by some strictly positive number of players (as opposed to the two *pure* states, 0 and N). This has also been termed a *polymorphic profile*(by contrast with a *monomorphic profile*) in the literature (e.g. Robson and Vega-Redondo (1996)).

which $(b(z) - z) = 0$ *is defined as*

$$
\begin{aligned}
c &= \pi_1(z_L) - \pi_2(z_L) \\
\overset{(2),(3)}{\Longrightarrow} z_L &= \frac{N(a_{22} - a_{12} + c) + a_{11} - a_{22} - c}{\theta},
\end{aligned} \tag{15}
$$

whilst the upper limit z_H *is defined as*

$$
\begin{aligned}
-c &= \pi_1(z_H) - \pi_2(z_H) \\
\overset{(2),(3)}{\Longrightarrow} z_H &= \frac{N(a_{22} - a_{12} - c) + a_{11} - a_{22} + c}{\theta}
\end{aligned} \tag{16}
$$

Definition 5 *As in KMR (1993), the integers around* z^* *are defined as*

$$
\begin{aligned}
\alpha &= \min\{z \in Z \mid \pi_2(z) > \pi_1(z)\} = \lceil z^* \rceil, \text{ and} \\
\beta &= \max\{z \in Z \mid \pi_2(z) < \pi_1(z)\} = \lfloor z^* \rfloor
\end{aligned}
$$

Similarly, the integers around z_L *are defined as*

$$
\begin{aligned}
\alpha_L &= \min\{z \in Z \mid \pi_1(z) - \pi_2(z) < c\} = \lceil z_L \rceil, \text{ and} \\
\beta_L &= \max\{z \in Z \mid \pi_1(z) - \pi_2(z) > c\} = \lfloor z_L \rfloor
\end{aligned}
$$

whilst those around z_H *are defined as*

$$
\begin{aligned}
\alpha_H &= \min\{z \in Z \mid \pi_1(z) - \pi_2(z) < -c\} = \lceil z_H \rceil, \text{ and} \\
\beta_H &= \max\{z \in Z \mid \pi_1(z) - \pi_2(z) > -c\} = \lfloor z_H \rfloor
\end{aligned}
$$

Figure 2 is thus modified to yield figure 3. The set E_M of new mixed absorbing states is then straightforwardly characterised as:

$$
\begin{aligned}
E_M &= \{\alpha_L, \alpha_L + 1, \ldots, \beta, \alpha, \ldots, \alpha_H, \beta_H\} \\
&= \{\alpha_L + j\}, \quad j = 0, 1, \ldots, (\beta_H - \alpha_L)
\end{aligned} \tag{17}
$$

The number m of these mixed absorbing states is clearly given by

$$
\begin{aligned}
m &= \beta_H - \alpha_L + 1 \\
&= \beta_H - \beta_L = \alpha_H - \alpha_L
\end{aligned} \tag{18}
$$

Defining $\zeta_L = \alpha_L - z_L$ and $\zeta_H = \alpha_H - z_H$, m can then be expressed as a function of the parameters of the model:

$$
\begin{aligned}
m &= (z_H + \zeta_H) - (z_L + \zeta_L) \\
&= -\frac{2(N-1)c}{\theta} + (\zeta_H - \zeta_L)
\end{aligned} \tag{19}
$$

Since $\theta < 0$ for Hawk-Dove games, the number of mixed absorbing states m is clearly increasing in N and c, but decreasing in the absolute value of the sum of the normalised stage game payoffs θ.

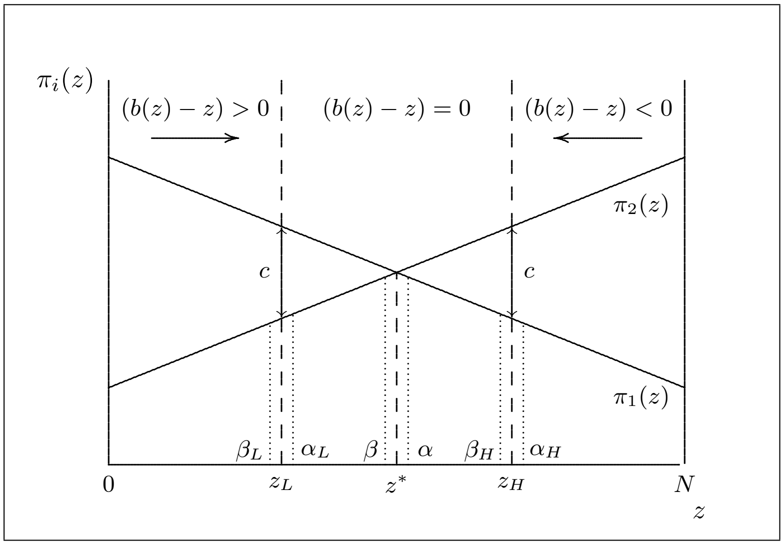


Figure 3: KMR Hawk-Dove games with switching costs

The existence of the new mixed absorbing states is independent of the precise selection dynamic employed, and thus hints at a new stability lacking in the above inertialess model. In fact, however, in order for any of the new mixed absorbing states to be selected as long-run equilibria, a "contraction" condition (albeit slightly modified) must once again be imposed on the selection dynamics. Moreover, the KMR result is in fact *weakened* in some sense by the introduction of switching costs, the mixed-strategy equilibrium losing its unique focality.

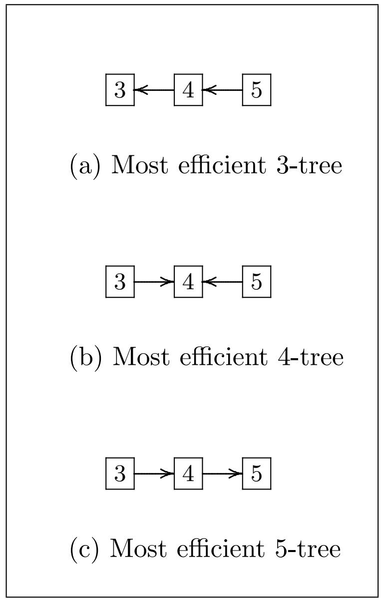
Definition 6 *Given a mixed-strategy equilibrium that places probability ρ on strategy 1 being played, say that $b(z)$ is a* contraction relative to E_M *if*

$$\lceil |b(z)-\rho N| \rceil \leq \lceil |z-\rho N| \rceil \qquad \textit{for all } z, \tag{20}$$

with the inequality holding strictly if $\lceil |z-\rho N| \rceil > (\frac{1}{2}m)$ (or $\lceil |z-\rho N| \rceil > (\frac{1}{2}m+1)$ for m odd).

Proposition 3 *If $b(z)$ is a contraction relative to E_M, then the limit distribution of the perturbed Hawk-Dove game in the presence of a fixed switching cost c puts probability $\frac{1}{m}$ on each of the m mixed absorbing states.*

Proof. The mixed absorbing states, E_M, constitute the only recurrent classes of the unperturbed process when the "modified contraction" condition of Definition 6 is satisfied. Each mixed absorbing state has an essentially identical most efficient (i.e., most probable) j-tree constituted by one-step jumps between absorbing states (see figure 4), each such jump requiring one



(a) Most efficient 3-tree

(b) Most efficient 4-tree

(c) Most efficient 5-tree

Figure 4: Most efficient j-trees, $m = 3$

mutation. All of the mixed absorbing states thus have the same stochastic potential, so that they are all selected as equally likely long-run equilibria by Lemma 2 (page 13). ∎

The only strengthening of the result to emerge with the addition of switching costs is that obtained by the weakening of the "contraction" condition in Definition 6; the inequality in equation (20) is now only required to hold strictly if the system currently lies outside the E_M region (compare equation (14), page 15). This weakening is possible because the deterministic dynamic does not act within the E_M region, so that it is only required to deliver the system into the E_M region, rather than all the way to the mixed-strategy equilibrium as before. As a result, a wider class of selection dynamics will now select the E_M states than would have selected the mixed-strategy equilibrium in the absence of inertia. Thus, for example, if there is a region of five mixed absorbing states in the presence of a given switching cost c, then a six-step dynamic satisfies the "modified contraction" condition and will deliver the selection results of Proposition 3. This, however, seems a hollow victory in the wake of the indeterminacy of the Proposition.

3.3 Discussion

Thus, contrary to initial expectation, the presence of switching cost-driven inertia in the evolutionary Hawk-Dove game studied does not serve to robustify the long-run selection of the “attractive” mixed-strategy equilibrium to any satisfactory degree. The direct effect of the inertia is confined to the region of the state space taken up by the new mixed absorbing states, so that a KMR-like “contraction” condition must be placed on the selection dynamics in order for convergence to that region to take place. The presence of the mixed absorbing states does serve to weaken the requirements of the “contraction” condition somewhat, but it also serves to dilute the focality of the mixed-strategy equilibrium by providing a number of equally attractive alternative equilibria.

These findings may signal that the KMR theorem is as strong a selection result as it is possible to derive for the mixed-strategy equilibrium here, and - in combination with the findings of Norman (2003a) - that the presence of switching costs cannot significantly affect long-run equilibrium. More likely, however, is that it points up the limitations of the model employed so far. In particular, there are two main unrealistic features of the model that may be distorting the results or hiding additional effects: the fact that the switching cost is uniform (across players) and fixed (over time); and the fact that the probability of mutation is uniform across states. Both problems can be addressed using a model of state-dependent mutations driven by stochastic switching costs, to which attention is now turned.

4 State-Dependent Mutations

Whilst it is highly realistic that individual players will face switching costs of changing their strategies, it has been argued that it is far less realistic that there will be a single time-invariant switching cost that is the same across all players, as was the case with c in the above KMR-style model. Moreover, the model is clearly subject to the Bergin and Lipman (1996) critique of models with state-independent mutations, the mutation rate ε remaining constant across states. Since mutations may be chosen such that any invariant distribution of the unperturbed process is the limiting ergodic distribution of a perturbed process, Bergin and Lipman argue that state-independent mutations are arbitrary, and that an economically justified model of state-*dependent* mutations should instead be employed.

Thus, in this section, the assumption of a fixed switching cost is dropped, and c is instead allowed to be stochastic, determined as the realisation of a random variable C. Each individual player takes a draw from C each period to determine his switching cost for that period, yielding the realistic feature of *idiosyncratic*, *time-varying* switching costs. Hence, this sort of *stochastic switching costs* model explicitly incorporates player heterogeneity, with players varying in their switching costs according to abilities, situations, priorities and so on. Indeed, it has parallels with the Myatt and Wallace (1998) model of adaptive play by idiosyncratic agents. The move to stochastic switching costs also has the potential to address the Bergin and Lipman (1996) critique by endogenising the mutation rate. The stochastic switching costs can provide the "error" necessary to yield an ergodic Markov chain, with this "error" now being interpreted as players behaving differently to what one would expect from the *cost-less* payoff matrix Λ in equation (1).

However, if - as the intuition for switching costs would at first suggest - the support of the random variable C were to be restricted to the positive real line $\mathbb{R}_+$, then the resulting Markov process would not be irreducible; "mutations" against the flow of the deterministic dynamic would have zero probability since, with a positive switching cost, no player will ever switch away from the strategy which currently has the higher expected payoff.[23] A reducible Markov chain remains subject to path-dependence, and does not yield an ergodic long-run distribution. To attain this, the state-independent mutation rate ε of the earlier chapters could be re-introduced, but this seems *ad hoc* and defeatist. Instead, the support of the switching cost random variable C could be extended to the whole real line $\mathbb{R}$, allowing the possibility of switching *benefits*. The existence of switching benefits in some players, grounded for example in an urge for creativity or nonconformity, is a realistic feature to incorporate within the model, and can be kept relatively improbable by assuming that C is distributed with positive mean. This will suffice to deliver an irreducible Markov chain - all states now being accessible from all others - and thus the desired ergodic long-run distribution over states.

Long-run selection of the mixed-strategy equilibrium is found to hold in the stochastic switching costs model under the *single-revisions* dynamic, which assumes that a *single* randomly selected member of the population has the opportunity to revise his strategy at the end of each period. However, this was to be expected, given that the "one-step" single-revisions dynamic satisfies both of the "contraction" conditions from the previous section. Generality thus requires the model to be extended to a KMR-style *simultaneous-revisions* framework, in which every player simultaneously has the opportunity to review his strategy each period. This allows the rate at which players review their strategies to be endogenised, so that the problems encountered in section 3 are overcome, and long-run convergence to the mixed-strategy equilibrium can be given a robust theoretical foundation. It turns out that there exists a threshold mean switching cost, above which the mixed-strategy equilibrium is selected in the long run for a wide class of switching cost distributions.

[23] This observation illustrates well the Bergin and Lipman (1996) criticism of the arbitrariness of state-independent mutations: there might be no reason to expect certain sorts of mutations under certain circumstances, and to ignore this is to assume away path-dependence when it may be an essential feature of the real-life process. The model here seeks to avoid this problem by providing an *economically justified* model of mutations which generates irreducibility.

4.1 The Model

The basic KMR structure of a finite population of N players repeatedly playing the 2×2 game Λ (equation (1)) is retained, with associated cost-adjusted payoff matrices Λ_1 and Λ_2 (equations (4) and (5)). The first major change to the model is that the switching cost c in any given period is now determined for each player individually as an independent and identically distributed draw from the switching cost random variable C, which is assumed to have a cumulative distribution function (cdf) F with a nonnegative mean $\bar{c} \geq 0$ and variance σ^2. This is a natural, general representation of differing switching costs across players, the positive mean focusing attention on switching *costs* but the infinite support delivering a small probability of switching *benefits*.

The expected payoffs are as they were in section 3 (see equations (2) and (3), page 7). Define $\varpi_1(z) = \pi_1(z) - \pi_2(z)$, and $\varpi_2(z) = \pi_2(z) - \pi_1(z)$, to be the expected payoff gains at stake from strategy switches in state z. Clearly a 1-incumbent will switch to being a 2-strategist if and only if

$$\varpi_2(z) > c$$

whilst a 2-incumbent will switch to being a 1-strategist if and only if

$$\varpi_1(z) > c$$

Defining s_l^t to be the strategy of player l in period t, the switching probabilities conditional on player l having been selected for review are then immediate:

$$\Pr(s_l^{t+1} = 2 \mid s_l^t = 1) = F\left(\varpi_2(z-1)\right)$$
$$\Pr(s_l^{t+1} = 1 \mid s_l^t = 2) = F\left(\varpi_1(z)\right)$$

The probability of a selected 2-incumbent switching to strategy 1 in a given state is illustrated in figure 5 ($f(\cdot)$ representing the switching cost random variable C's probability density function (pdf)).[24]

The second major change to the KMR-style model of section 3 is in the selection dynamic employed. In this more complicated setting, it is desirable for the sake of mathematical convenience to focus attention on the best-response dynamic $B(z)$ - where players play their *best response* to current strategy frequencies - rather than KMR's more general "Darwinian" dynamic $b(z)$. This simplifies the analysis and also serves to isolate the effect of inertia on the system's review rate, without the distraction of having other forms of bounded rationality at work. The effect of the selection dynamic can still be analysed in this setting by comparing the extreme cases

[24] Whilst the pdf depicted in figure 5 is a Normal density, $f(\cdot)$ is a general pdf in the model, and thus can take any form.

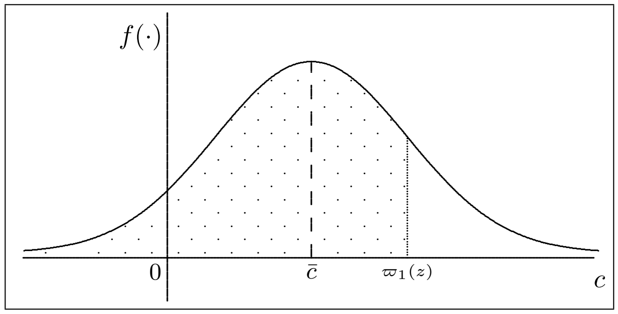


Figure 5: The switching cost C's pdf

of the simultaneous-revisions dynamic of KMR (1993) and others, and the "one-step" single-revisions dynamic favoured by Binmore and Samuelson (1997), Blume (1999), and Myatt and Wallace (1998). Under the "one-step" dynamic, a *single* randomly selected member of the population has the opportunity to revise his strategy at the end of each period. Given his realised value of c, this player observes the strategy distribution among the incumbent population and selects a best response to this frequency.[25] In the simultaneous-revisions model by contrast, *each* player has this opportunity to revise his strategy each period. Evolution can thus proceed far more rapidly under simultaneous revisions, ceteris paribus.[26]

4.2 Analysis

4.2.1 Single revisions

In this subsection, use will be made of the "one-step" best-response dynamic $B^1(z)$, where just one player at a time has the opportunity to revise his strategy. Given that the reviewing player l is randomly selected, the probability that he is a 1-incumbent is simply $\frac{z}{N}$, implying the following transition probabilities between states.

[25]An equivalent scenario is the familiar story of player exit and entry, whereby a randomly selected player leaves at the end of each period, and is replaced by another player with a new draw from C.

[26]Both the single- and simultaneous-revisions dynamics require the implicit assumption that the switching cost of an updating individual will have changed since the last revision, since he must take a fresh draw from C. This is a reasonable assumption when the noise is small, and thus unproblematic for vanishing heterogeneity results obtained as σ^2 tends to zero. More generally, the procedure can be justified by noting that players are more likely to revise their strategy whenever their switching costs change.

Lemma 3 *The transition probabilities p_{ij} of the Markov matrix P satisfy:*

$$p_{ij} = \left\{ \begin{array}{ll} \left(\frac{i}{N}\right) F\left(\varpi_2(i-1)\right) & j = i-1 \\ \left(\frac{i}{N}\right)\left(1 - F\left(\varpi_2(i-1)\right)\right) + \left(\frac{N-i}{N}\right)\left(1 - F\left(\varpi_1(i)\right)\right) & j = i \\ \left(\frac{N-i}{N}\right) F\left(\varpi_1(i)\right) & j = i+1 \end{array} \right\}$$

and are zero elsewhere.

Proof. Given the single-revisions framework, the process cannot move from state i to $j < i-1$ or $j > i+1$. A move to state $i+1$ requires that a 2-incumbent be selected for review, and that having been selected his best response to the current strategy frequency be to switch to strategy 1. The former occurs with probability $(N-i)/N$; the latter with probability $F(\varpi_1(i))$. Similar arguments apply for the cases $j = i$ and $j = i-1$. ■

Proposition 4 *The adaptive response dynamic defined by the transition probabilities p_{ij} in Lemma 3 is an irreducible, aperiodic Markov process on the finite state space Z. Consequently, it has a unique invariant distribution.*

Proof. Since the normal distribution has full support, either strategy may be chosen by any reviewing player. The process can thus move in either direction from any state i (except the extreme states), as formalised in Lemma 3, so that every state is accessible from all others in finite time - i.e., the process is irreducible. Moreover, since in every state there is a positive probability of the system remaining in that state in the next period, the process is aperiodic. The process thus has a unique invariant distribution by Lemma 1. ■

Long-run equilibrium with vanishing heterogeneity In the previous section, long-run equilibrium results were obtained by analysing the limit of the ergodic distribution as the probability of mutation ε tended to zero. The analogue in this model, as in the model of Myatt and Wallace (1998), is vanishing player heterogeneity over switching costs - i.e., taking $\sigma^2 \to 0$.

Given the departure from a uniform mutation rate ε, the simple Young stochastic potential technique of Lemma 2 cannot be applied here. Fortunately, Young's method fits into a wider *graph-theoretic* approach to the analysis of the long-run behaviour of perturbed Markov chains. This graph-theoretic approach is in turn derived from general Markovian theory.[27] If ε were known precisely, it would (in theory) be possible to compute the actual distribution μ^ε by simply solving the stationarity equations $\mu^\varepsilon P^\varepsilon = \mu^\varepsilon$. However, in most applications of interest in economics, the size of the state

[27] A standard reference for Markovian theory is Karlin and Taylor (1975).

space would make this a very cumbersome task. This fact led to the importing of the simplified graph-theoretic techniques of Freidlin and Wentzell (1984) into the economics discipline by Foster and Young (1990), KMR (1993), and Young (1993), and the resulting birth of stochastic adjustment dynamics.

Like the computation of Young's stochastic potential function, the graph-theoretic computation of μ^ε is based on the notion of rooted trees, but this time constructed on the whole state space Z rather than merely on the set of recurrent classes. Specifically, consider a directed graph whose vertex set is the state space Z. The edges of this graph form a *z-tree* (for some particular $z \in Z$) if it consists of $|Z| - 1$ edges and from every vertex $i \neq z$ there is a unique directed path from i to z. A z-trees's edges are weighted with the appropriate Markov transition probabilities p_{ij}. Representing any given directed edge $i \to j$ by the ordered pair of vertices (i, j), a z-tree T can then be represented as a subset of ordered pairs. Let $\mathcal{T}_z$ be the family of all z-trees for a given z. Define the *likelihood* of z-tree $T \in \mathcal{T}_z$ to be

$$p(T) = \prod_{(i,j)\in T} p_{ij}$$

Lemma 4 (Freidlin and Wentzell (1984)) *Let P be an irreducible Markov process on a finite state space Z.*[28] *Its stationary distribution μ has the property that the probability $\mu(z)$ of each state z is proportional to the sum of the likelihoods of its z-trees, that is,*

$$\mu(z) = \frac{v(z)}{\sum_{i\in Z} v(i)}, \textit{ where } v(z) = \sum_{T\in\mathcal{T}_z} p(T) \tag{21}$$

This result allows computation of an *exact* estimate of a system's ergodic distribution μ^ε for each $\varepsilon > 0$;[29] it is not a limiting result as ε tends to zero.

The additional analytical power of the Freidlin-Wentzell method in providing an immediate closed form for the invariant distribution μ comes at the price of the greater complexity inherent in constructing trees on the whole state space (rather than on the set of recurrent classes, as in Young's method[30]). The number of z-trees to be considered soon becomes prohibitively large as the state space grows. This is where the use of the single-revisions dynamic becomes useful, since it means that $v(z)$ in equation (21) takes a very simple form. With only one revision at a time, there is only

[28] Note that Lemma 2's condition that P be a *regular* perturbed process is no longer required.

[29] See Young (1998), section 3.4, for illustrative examples.

[30] Young's technique in fact follows from that of Freidlin and Wentzell, taking advantage of uniform mutation rates, and of the zero resistance of paths along the deterministic dynamic, in order to achieve greater analytical simplicity by simply "counting mutations" between recurrent classes.

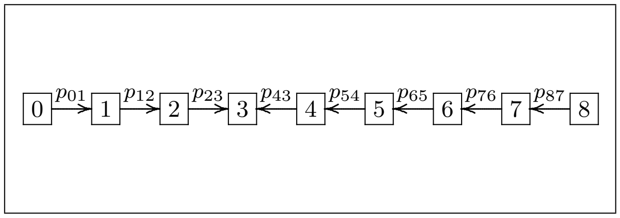


Figure 6: Unique positively weighted 3-tree, $N = 8$

one possible positively weighted z-tree for any given state z - that z-tree involving successive one-step jumps from every state in the direction of z, as illustrated in figure 6.[31] Any other one-step z-tree violates the requirement of a unique path to z from every other state. It follows that $v(z)$ is given by

$$v(z) = \prod_{0 \le i < z} p_{i(i+1)} \prod_{z < i \le N} p_{i(i-1)}$$

which implies, in combination with the transition probabilities in Lemma 3, that

$$v(z) = \frac{1}{N^N} \prod_{0 \le i < z} (N - i) \, F(\varpi_1(i)) \prod_{z < i \le N} i \, F(\varpi_2(i-1)) \tag{22}$$

The unperturbed process P^0 in this model is that where there is no heterogeneity in switching costs, $\sigma^2 = 0$. In this case, the switching cost pdf (figure 5, for example) collapses to a point mass on the mean switching cost $\bar{c}$, and the model becomes that of section 3, with $c = \bar{c}$ (see figure 3, page 18). Because the selection dynamic is the "one-step" single-revisions dynamic, the unperturbed process has a set of recurrent classes consisting only of the mixed absorbing states E_M of Proposition 2 (page 16). As usual, the ergodic distribution will focus all weight on these states as perturbations go to zero, but to select between them it is necessary to consider their *relative weight* in the ergodic distribution. Note that Lemma 4 implies that the relative weight of any two states z and z' in the ergodic distribution μ may be assessed by considering the ratio

$$\frac{\mu(z)}{\mu(z')} = \frac{v(z)}{v(z')} \tag{23}$$

Following Myatt and Wallace, a state z will be said to *dominate* another state z' for vanishing heterogeneity whenever $\lim_{\sigma^2 \to 0} \frac{v(z)}{v(z')} = \infty$. If a state dominates all others in this sense, it is clearly the unique stochastically stable state of the system.[32]

[31] For a proof, see Myatt and Wallace's (1998) Lemma 3.

[32] Refer to the definition of stochastic stability on page 13.

Definition 7 *A distribution F with mean $\bar{y}$ has a* likelihood ratio unbounded in the tails *if*

$$\lim_{\sigma^2 \to 0} \frac{F(y)}{F(y-\epsilon)} \to \infty, \qquad \forall y \leq \bar{y}, \forall \epsilon > 0$$

$$\text{and} \qquad \lim_{\sigma^2 \to 0} \frac{1-F(y)}{1-F(y+\epsilon)} \to \infty, \qquad \forall y > \bar{y}, \forall \epsilon > 0$$

Proposition 5 *Assume that the switching cost distribution F has a likelihood ratio unbounded in the tails. Then the unique stochastically stable state is the integer α just above the mixed-strategy equilibrium z^*.*

Proof. The unique positively weighted α-tree shares all of its branches with that of any given state z, except for those branches lying between α and z. Hence, after cancelling these shared terms,

$$\begin{aligned} \lim_{\sigma^2 \to 0} \frac{v(\beta)}{v(z)} &= \lim_{\sigma^2 \to 0} \frac{\prod_{\alpha < i \leq z} iF(\varpi_2(i-1))}{\prod_{\alpha \leq i < z} (N-i) F(\varpi_1(i))} \\ &= \lim_{\sigma^2 \to 0} \prod_{\alpha \leq i < z} \frac{(i+1)}{(N-i)} \frac{F(\varpi_2(i))}{F(\varpi_1(i))} \end{aligned} \tag{24}$$

assuming without loss of generality that $z > \alpha$, and re-indexing the numerator in the second step. As $\sigma^2 \to 0$ (i.e., the game approaches that of section 3),[33] $F(\varpi_2(i))$ (the probability of a selected 1-incumbent switching to strategy 2) tends to 1 in the limit for $i > z_H$, but to 0 for $i \leq z_H$.[34] Similarly, $F(\varpi_1(i))$ (the probability of a selected 2-incumbent switching to strategy 1) tends to 1 for $i < z_L$, but to 0 for $i \geq z_L$. It is thus clear that both the numerator and the denominator of equation (24) tend to 0 as heterogeneity vanishes. However, consider the likelihood ratio term $(F(\varpi_2(i))/F(\varpi_1(i)))$ for a given $i \in (z_H, z)$. The numerator of this term tends to 1 in the limit for any $i \in (z_H, N]$, whilst the denominator tends to 0, so that the limit of the likelihood ratio term over this range is infinity. Over the range $i \in [\alpha, z_H]$, meanwhile, both the numerator and the denominator of any given likelihood ratio term tend to 0. But since $\varpi_2(i) > \varpi_1(i)$ for $i > z^*$, a sufficient condition for this ratio to tend to infinity as $\sigma^2 \to 0$ is that the switching cost distribution F have a likelihood ratio unbounded in the tails, as defined in Definition 7. For such switching cost distributions then, all of the likelihood ratio terms $(F(\varpi_2(i))/F(\varpi_1(i)))$ for $i \in [\alpha, z)$ tend to infinity as heterogeneity vanishes. Hence, $\lim_{\sigma^2 \to 0}(v(\alpha)/v(z)) = \infty$, and α dominates all states $z \in (\alpha, N]$.

By a similar line of argument, β dominates all states $z \in [0, \beta)$. Now,

$$\lim_{\sigma^2 \to 0} \frac{v(\alpha)}{v(\beta)} = \lim_{\sigma^2 \to 0} \frac{(N-\beta)}{\alpha} \frac{F(\varpi_1(\beta))}{F(\varpi_2(\beta))}$$

[33] Refer to the unperturbed Hawk-Dove game of figure 3 (page 18).

[34] z_L and z_H are as defined in equations (15) and (16) (page 17), with $c = \bar{c}$.

And since $\varpi_1(i) > \varpi_2(i)$ for $i < z^*$, a sufficient condition for this ratio to tend to infinity as $\sigma^2 \to 0$ is that the switching cost distribution F have a likelihood ratio unbounded in the tails. For such switching cost distributions then, α dominates β, and hence all states $z \in [0, \beta]$. It follows that α is the unique stochastically stable state. ■

Remark 1 *It should be noted that α is selected over β here because of the integer effect of players ignoring their own strategy when calculating current strategy frequencies. As $N \to \infty$ - and such integer effects disappear - the limiting ergodic distribution will place equal weight on α and β, which in any case will both converge on the mixed-strategy equilibrium z^* as N gets large.*

Thus, for any switching cost distribution satisfying the property of a likelihood ratio unbounded in the tails, the "attractive" mixed-strategy equilibrium of Hawk-Dove games is selected as the long-run equilibrium of the single-revisions stochastic switching costs model. And a wide variety of reasonable distributions *do* satisfy this property - for example, the Normal, Student's t, exponential, logistic and gamma distributions. Moreover, the likelihood ratio unbounded in the tails property is a sufficient, but not a necessary, condition for the result, so that an even wider class of distributions is admissible.[35]

However, this selection of the mixed-strategy equilibrium should come as no surprise, since using the single-revisions dynamic with Hawk-Dove games is begging the question somewhat; the "one-step" dynamic satisfies both KMR's "contraction" condition (page 15 above) and the "modified contraction" condition of Definition 6 (page 18). In the next section, the stochastic switching cost model is extended to the more general simultaneous-revisions dynamic, and it is shown that in this context convergence to the mixed-strategy equilibrium can be given more satisfactory foundations.

4.2.2 Simultaneous revisions

The selection dynamic employed in this subsection is the simultaneous-revisions best-response dynamic $B^N(z)$. Under this dynamic, all players have the opportunity to revise their strategies each period; each player takes an independent and identically distributed draw from the switching cost random variable C, and then best-responds with respect to the current strategy frequency and his c draw in deciding whether or not to switch strategies.

[35] Note, however, that the uniform mutation rate ε of section 3 does *not* satisfy this property.

Lemma 5 *Under the simultaneous-revisions dynamic, the new transition probabilities* p_{ij} *which constitute the perturbed Markov matrix* P^{σ^2} *are*

$$p_{ij} = \sum_{k=\max\{j-i,0\}}^{\min\{j,N-i\}} \binom{i}{i+k-j} \binom{N-i}{k} \times F(\varpi_2(i-1))^{i+k-j}(1-F(\varpi_2(i-1)))^{j-k} \times F(\varpi_1(i))^k (1-F(\varpi_1(i)))^{N-i-k}$$

Proof. There are $(\min\{j, N-i\} - \max\{j-i, 0\})$ possible combinations of strategy switches in a period that will lead the system from state i to state j, each of which has a number of permutations (essentially relabelling players) given by the product of the two binomial terms. The product of F terms is then the probability of each permutation. ■

Proposition 6 *The adaptive response dynamic defined by the transition probabilities* p_{ij} *in Lemma 5 is an irreducible, aperiodic Markov process on the finite state space* Z*. Consequently, it has a unique invariant distribution.*

Proof. Since F has full support, either strategy may be chosen by any reviewing player. The process can thus move in either direction from any state i (except the extreme states). Moreover, given that all players have the opportunity to revise their strategies in each period, the process can move anything from 0 to N states in one period. Thus, every state is accessible from all others within one period (i.e. every entry p_{ij} in the perturbed Markov matrix P^{σ^2} of Lemma 5 is strictly positive), so that the process is irreducible. Moreover, since in every state there is a positive probability of the system remaining in that state in the next period, the process is aperiodic. By Lemma 1, the process thus has a unique invariant distribution. ■

Long-run equilibrium with vanishing heterogeneity As under single revisions, the unperturbed Markov process P^0 under simultaneous revisions is that where there is no heterogeneity in switching costs, $\sigma^2 = 0$. And similarly, as the switching cost pdf (figure 5) collapses to a point mass on the mean switching cost $\bar{c}$, the model again becomes that of section 3 (see figure 3, page 18). However, because the selection dynamic is now *simultaneous-revisions* best-response, the set of recurrent classes of the unperturbed process is widened to include the extreme state cycle $\{0, N\}$ in addition to the mixed absorbing states E_M (see page 14). All weight will be focused on these classes by the ergodic distribution as heterogeneity vanishes, but as before, limiting relative weights must be considered in order to select between these candidate equilibria.

However, this task is considerably less straightforward than under the single-revisions dynamic. The transition probabilities of Lemma 5 provide a stark illustration of the potential complexity of the Freidlin-Wentzell approach to long-run equilibrium analysis. Every branch of every possible z-tree is weighted by a transition probability p_{ij}, and under simultaneous revisions the number of possible z-trees per state soon becomes prohibitively large as the state space grows. Fortunately, some simplifications are available.

Lemma 6 *The probability p_{ij} of the transition between states i and j is of the order of*

$$\varrho_{ij} = \max_{k \in [\max\{j-i,0\},\min\{j,N-i\}]} \begin{pmatrix} i \\ i+k-j \end{pmatrix} \begin{pmatrix} N-i \\ k \end{pmatrix} \times F\left(\varpi_2(i-1)\right)^{i+k-j} \left(1-F\left(\varpi_2(i-1)\right)\right)^{j-k} \times F\left(\varpi_1(i)\right)^{k} \left(1-F\left(\varpi_1(i)\right)\right)^{N-i-k} \quad (25)$$

as $\sigma^2 \to 0$. Moreover, for switching cost distributions satisfying the condition $F(-x) < (1-F(x))$, and with a likelihood ratio unbounded in the tails, this maximum is achieved at $k = \max\{j-i, 0\}$.

Proof. The first part of the result is immediate from Lemma 5, given that the order of a summation of terms is determined by the highest order term. To see the second part, consider the effect of increasing k by 1 in equation (25): $(1-F(\varpi_2(i-1)))(1-F(\varpi_1(i)))$ is removed from the expression, and replaced by the strictly lower $F(\varpi_2(i-1))F(\varpi_1(i))$ (under the assumption that $F(-x) < (1-F(x))$). A likelihood ratio unbounded in the tails is then sufficient to guarantee that the maximand in equation (25) is decreasing in k as $\sigma^2 \to 0$. ■

Intuitively, the minimum value of k is selected in Lemma 6 because this minimizes the number of strategy switches used to effect a given transition. Two strategy switches in opposite directions merely cancel each other out, and two players remaining inert is more probable than two players switching in opposite directions; hence, anything above the minimum number of strategy switches serves to reduce a transition path's probability. This is a natural feature of a model of evolution under inertia, and indeed a wide class of distributions satisfy the conditions of the Lemma given a positive mean. For example, symmetry about $\bar{c}$ is sufficient (but not necessary) to ensure that $F(-x) < (1-F(x))$ for $\bar{c} > 0$.

Lemma 7 *Under the adaptive response dynamic defined by the transition probabilities p_{ij} in Lemma 5, state z dominates another state z' for vanishing heterogeneity whenever*

$$\lim_{\sigma^2 \to 0} \frac{\varrho(z)}{\varrho(z')} = \infty,$$

where

$$\varrho(z) = \prod_{(i,j)\in T_z^{max}} \varrho_{ij}$$

and $T_z^{max} = \arg\max_{T\in\mathcal{T}_z} p(T)$ *is the most probable* z*-tree for a given state* z.

Proof. Recall that a state z dominates another state z' for vanishing heterogeneity whenever $\lim_{\sigma^2\to 0}(v(z)/v(z')) = \infty$. Recall also from Lemma 4 that $v(z)$ is given by the sum of the likelihoods of all possible z-trees for a given state z. Thus,

$$\lim_{\sigma^2\to 0} \frac{v(z)}{v(z')} = \lim_{\sigma^2\to 0} \frac{\sum_{T\in\mathcal{T}_z} p(T)}{\sum_{T\in\mathcal{T}_{z'}} p(T)},$$

where $p(T) = \prod_{(i,j)\in T} p_{ij}$ is the likelihood of the tree T, which belongs to a family of z-trees $\mathcal{T}_z$ (for a given state z).

Now, $\lim_{\sigma^2\to 0}(v(z)/v(z')) = \infty$ if and only if $v(z)$ is of higher order than $v(z')$ (i.e. $v(z') = o(v(z))$).[36] Hence

$$\lim_{\sigma^2\to 0} \frac{v(z)}{v(z')} = \infty \quad \Leftrightarrow \quad \lim_{\sigma^2\to 0} \frac{\max_{T\in\mathcal{T}_z} p(T)}{\max_{T\in\mathcal{T}_{z'}} p(T)} = \infty$$

Since this will in turn be true if and only if

$$\max_{T\in\mathcal{T}_{z'}} p(T) = o(\max_{T\in\mathcal{T}_z} p(T)),$$

dominance is seen to be determined by a comparison of the order of the likelihood of each state's highest-order z-tree.[37] The likelihoods of the most probable z-trees are themselves products of the transition probabilities p_{ij} of Lemma 5, each of which is of the order of ϱ_{ij} as $\sigma^2 \to 0$ by Lemma 6. ∎

In order to employ Lemma 7 precisely, it is first necessary to identify the most probable z-tree $T_z^{\max}$ for each state z. This task is far from straightforward, and is addressed in Norman (2003b). There it emerges that the most probable way of escaping a given basin of attraction depends on the assumed "noise model" - in this case the density of switching costs - but in general is unlikely to involve either one-step transitions or direct jumps. As a result, fully operationalizing Lemma 7 is a complex task, feasible only for particular noise models. However, the long-run equilibrium can be obtained without knowledge of the precise form of the most probable z-trees.

[36] In general, for two functions $f(x)$ and $g(x)$, if $(f(x)/g(x)) \to 0$ as $x \to \infty$, then f is of smaller order than g, denoted $f(x) = o(g(x))$. If, on the other hand, $\lim_{x\to\infty}(f(x)/g(x)) \leq$ constant, then f is of the same order as g, denoted $f(x) = O(g(x))$.

[37] This is a key result in graph-theoretic Markovian theory which, for example, underlies Young's stochastic potential technique.

Lemma 8 *The most probable z-trees have the following properties:*

1. *For the most probable* 0*- and* N*-trees,* T_N^{max} *and* T_0^{max}*,*
   (a) *all branches rooted outside of the* E_M *region (and hence inside* $D(\{0, N\})$*) are those of the deterministic dynamic (whose probabilities tend to* 1 *as* $\sigma^2 \to 0$*);*
   (b) *the branches rooted in the* E_M *region constitute the (unknown) most probable way of escaping the region of mixed absorbing states* E_M *(and thus of entering* $\{0, N\}$*'s basin of attraction* $D(\{0, N\})$*).*
2. *For the most probable* α*- and* β*-trees,* T_α^{max} *and* T_β^{max}*,*
   (a) *all branches rooted inside the* E_M *region are one-step transitions towards* α *for* T_α^{max}*, and towards* β *for* T_β^{max}*;*
   (b) *the branches rooted outside of the* E_M *region (and hence inside* $D(\{0, N\})$*) constitute the (unknown) most probable way of escaping* $D(\{0, N\})$ *(i.e., of entering the* E_M *region).*

Proof.

1. (a) The transitions of the underlying deterministic dynamic are obviously the most probable transitions available in the limit.
   (b) Given 1(a), any state $z \in D(\{0, N\})$ can be entered from the E_M region, and hence the most probable such escape is selected. Note that, by an "escape" from the E_M region here is meant a sequence of transitions that provides a directed path from *each* mixed absorbing state out of the E_M region.
2. (a) The most probable event within the E_M region is inertia, but a z-tree requires a transition rooted in every state. And the most probable transition from any given mixed absorbing state is the minimum number of strategy switches in the more probable direction - i.e., a one-step transition towards the mixed-strategy equilibrium z^*.
   (b) Given 2(a), any mixed absorbing state can be entered from $D(\{0, N\})$, and hence the most probable such escape is selected.

■

Proposition 7 *Assume that the switching cost distribution* F *has a likelihood ratio unbounded in the tails. Then, as* $\sigma^2 \to 0$*,*

1. *the set* $\{0, N\}$ *dominates all states in the ranges* $(0, z_L)$ *and* (z_H, N)*;*
2. *the set* $\{\alpha, \beta\}$ *dominates all of the other mixed absorbing states.*

Proof.

1. Consider $\varrho(N)/\varrho(0)$, and call this the *order ratio* for convenience. By Lemma 8, the most probable 0- and N-trees share all of their branches, except for those leading from and to each other, $0 \to N$ and $N \to 0$. Thus, cancelling shared terms,

$$\lim_{\sigma^2 \to 0} \frac{\varrho(N)}{\varrho(0)} = \lim_{\sigma^2 \to 0} \frac{\varrho_{0N}}{\varrho_{N0}}$$

   Since both the numerator and the denominator of this ratio tend to 1 as $\sigma^2 \to 0$, it follows that the limit itself is equal to 1, so that neither state 0 nor N dominates the other. This corresponds with the fact that the extreme state cycle $\{0, N\}$ is a recurrent class of the unperturbed process.

   Consider now the relative weight of the extreme states $\{0, N\}$ by comparison with states in the regions $(0, z_L)$ and (z_H, N). Clearly any $z \in \{(0, z_L), (z_H, N)\}$ lies in $\{0, N\}$'s basin of attraction $D(\{0, N\})$. And in $T_N^{\max}$ and $T_0^{\max}$, $D(\{0, N\})$ is entered in the most probable fashion (by Lemma 8). Moreover, $T_N^{\max}$'s (resp., $T_0^{\max}$'s) path from z to N (resp., 0) is clearly more probable than any path from N (resp., 0) to z. F's likelihood ratio unbounded in the tails then guarantees $\{0, N\}$'s dominance over z for vanishing heterogeneity.

2. Consider the order ratio $\varrho(\alpha)/\varrho(\beta)$ for the states α and β that lie immediately above and below the mixed-strategy equilibrium z^*. By Lemma 8, the most probable α- and β-trees share all of their branches, except for those leading from and to each other, $\beta \to \alpha$ and $\alpha \to \beta$. Thus, cancelling shared terms,

$$\lim_{\sigma^2 \to 0} \frac{\varrho(\alpha)}{\varrho(\beta)} = \lim_{\sigma^2 \to 0} \frac{\varrho_{\beta\alpha}}{\varrho_{\alpha\beta}} \tag{26}$$

   Absent integer problems (i.e., as N gets large), $\varrho_{\beta\alpha}$ and $\varrho_{\alpha\beta}$ converge to 0 at the same rate as $\sigma^2 \to 0$ (since $\varpi_1(\beta) \approx \varpi_2(\alpha)$, with equality as $N \to \infty$). Hence, the limit in equation (26) is a positive constant, and neither state α nor β dominates the other for vanishing heterogeneity.

   However, both α and β dominate all of the other mixed absorbing states for switching cost distributions with a likelihood ratio unbounded in the tails. To see this, note first that any directed path from α (resp., β) to any other mixed absorbing state $z_M \in E_M \backslash \{\alpha, \beta\}$ is strictly less probable than $T_\alpha^{\max}$'s (resp., $T_\beta^{\max}$'s) one-step-at-a-time directed path from z_M to α (resp., β). Given that the E_M region is also entered in the most probable manner in $T_\alpha^{\max}$ and $T_\beta^{\max}$ (by Lemma 8), this is sufficient to guarantee $\{\alpha, \beta\}$'s dominance over all $z_M \in E_M \backslash \{\alpha, \beta\}$ for vanishing heterogeneity.

■

Proposition 8 *Assume that the switching cost distribution F has a likelihood ratio unbounded in the tails. Then there exists a threshold level of the mean switching cost, $\hat{c}$, below which the ergodic distribution places all weight equally on the extreme states 0 and N as $\sigma^2 \to 0$, but above which the mixed-strategy equilibrium states α and β are the long-run equilibria.*

Proof. Recall that if a state dominates all others for vanishing heterogeneity, then it is the unique stochastically stable state (or long-run equilibrium) of the system. Lemma 7 thus implies that if $\lim_{\sigma^2 \to 0}(\varrho(z)/\varrho(z')) = \infty, \forall z' \neq z$, then z is the unique long-run equilibrium.

By Proposition 7, there are only two possible candidate equilibria for long-run selection; all weight in the limiting ergodic distribution will be focused *either* on the extreme state cycle $\{0, N\}$ *or* on the mixed-strategy equilibrium states $\{\alpha, \beta\}$. Hence, a comparison of the relative weights of α and 0 will suffice to indicate which set of states is dominant. Lemma 8 implies that, after cancelling shared terms,

$$\lim_{\sigma^2 \to 0} \frac{\varrho(\alpha)}{\varrho(0)} = \lim_{\sigma^2 \to 0} \left(\Upsilon_{0M} \times \frac{\prod_{\alpha_L \leq i \leq \beta} \varrho_{i(i+1)} \prod_{\alpha < i \leq \beta_H} \varrho_{i(i-1)}}{\Upsilon_{M0}} \right) \tag{27}$$

where Υ_{0M} is the (unknown) sequence of ϱ_{ij} terms (one for each branch) leading from 0 into the region of mixed absorbing states E_M, and Υ_{M0} is the (unknown) sequence of ϱ_{ij} terms leading (all of) the mixed absorbing states into 0's basin of attraction. The remaining product of ϱ_{ij} terms represents the one-step transitions towards α within the E_M region of the most probable α-tree. Υ_{0M} clearly tends to 0 as $\sigma^2 \to 0$, but the ratio term tends to infinity. To see this, note that every i-term in the numerator has a corresponding i-term in the denominator Υ_{M0}. And each term in the numerator is the most probable transition available - i.e., towards z^* - whilst the denominator must contain enough transitions *away from* z^* to escape the E_M region, by definition of Υ_{M0}. Hence the numerator of the ratio term is greater than the denominator, and dominates it as $\sigma^2 \to 0$ for switching cost distributions with a likelihood ratio unbounded in the tails.

Now, raising the mean switching cost $\bar{c}$ has a number of effects on equation (27). First, it increases Υ_{0M} by reducing the number of improbable failures to switch strategy required to escape $\{0, N\}$'s basin of attraction (and enter the E_M region), and by increasing the likelihood of these failures. Second, it reduces both the numerator and the denominator of the ratio term by increasing the number of improbable transitions, and by reducing the likelihood of each of these transitions. The net effect at any given point may be either to increase or to decrease the ratio term, but it will never reduce it below 1, and hence does not threaten the numerator's dominance as $\sigma^2 \to 0$. Now, when $\bar{c} = 0$, the set of mixed absorbing states is empty, $\Upsilon_{0M} = 0$, and $\{0, N\}$ is the unique long-run equilibrium. By contrast, when $\bar{c} = (a_{21} - a_{11})$, the region of mixed absorbing states E_M subsumes the extreme state N, and $\Upsilon_{0M} = 1$. At this point, the limit in equation (27) is determined by the limit of the ratio term, and is thus infinity. Hence, $\{\alpha, \beta\}$ is selected as the unique long-run equilibrium. Since Υ_{0M} is strictly increasing in $\bar{c}$, it follows that there exists a threshold level of the mean switching cost, $\hat{c} \in (0, (a_{21} - a_{11}))$, above which $\lim_{\sigma^2 \to 0}(\varrho(\alpha)/\varrho(0)) = \infty$ and $\{\alpha, \beta\}$ is selected, but below which $\lim_{\sigma^2 \to 0}(\varrho(\alpha)/\varrho(0)) = 0$ and $\{0, N\}$ is selected. This threshold $\hat{c}$ is defined implicitly by the equation

$$\frac{\Upsilon_{M0}}{\Upsilon_{0M}} = \prod_{\alpha_L \leq i \leq \beta} \varrho_{i(i+1)} \prod_{\alpha < i \leq \beta_H} \varrho_{i(i-1)}$$

for switching cost distributions with a likelihood ratio unbounded in the tails. ■

Remark 2 *Note that α and β converge to z^* as $N \to \infty$, so that essentially the mixed-strategy equilibrium is selected for $\bar{c} > \hat{c}$.*

Hence, the "attractive" mixed-strategy equilibrium of Hawk-Dove games emerges as a possible long-run equilibrium of the simultaneous-revisions stochastic switching costs model. Its selection does, however, require sufficiently high mean switching costs, otherwise the extreme state cycle $\{0, N\}$ remains the long-run equilibrium. Essentially, increasing $\bar{c}$ has two effects: it expands the region of mixed absorbing states E_M; and it reduces the probability of any given transition. These effects make the E_M region more likely to be entered and harder to escape - properties which Ellison (2000) (and indeed the whole literature on basins of attraction) tells us are likely to deliver stochastic stability.

4.3 Discussion

Thus, in the presence of high enough switching costs, long-run convergence to the mixed-strategy equilibrium of Hawk-Dove games has a coherent theoretical justification. This justification is not the same as KMR's "contraction" condition; indeed, it is more realistic. Whilst the "contraction" condition locks the selection dynamic into an arbitrary pattern that makes the mixed-strategy equilibrium inevitable, the stochastic switching costs model provides sound economic reasons why review is *likely* to proceed in the required fashion. However, it is not certain to do so, and the probabilistic nature of the long-run analysis is retained, with long-run selection depending upon the size of the mean switching cost.

Because of the complexity of the formal analysis, it is difficult to calculate the exact size of the threshold mean switching cost $\hat{c}$. It will, however, certainly be "small" in the sense of being below the question-begging level of the maximum possible payoff gain from a strategy switch, $(a_{12} - a_{22})$[38]. In fact, it is guaranteed to be "smaller" still, since it must be below the lower of the two maximum possible payoff gains from switching strategies, $\hat{c} < (a_{21} - a_{11}) < (a_{12} - a_{22})$. Moreover, depending on the switching cost density and population size, $\hat{c}$ could be considerably lower than this upper bound.

Selection of the mixed-strategy equilibrium is of course more robust if one favours the single-revisions dynamic employed in subsection 4.2.1. This increasingly popular dynamic is appealing for its simplicity, but it can also be given compelling theoretical foundations. Binmore and Samuelson (1997) justify the "one-step" dynamic on the basis of a continuous-time framework in which individual players revise periodically according to an underlying Poisson process. They note that, in this case, at most one revision will be observed with high probability during any small period of time.[39] However, it does seem an unsatisfactory feature of the single-revisions model that the reviewing player is *randomly selected*, particularly given that under consideration is the very phenomenon of players *choosing* not to review under certain conditions - i.e., inertia. The simultaneous-revisions model solves this problem by allowing the review rate to be *endogenised*, with a player reviewing his strategy only when his switching cost draw is exceeded by the expected payoff gain from a strategy switch. This overcomes the weakness in models (like those using the single-revisions dynamic) that assume an exogenous review rate, by instead having it determined as part of the model.

These considerations do not matter for the coordination games of Norman (2003a), where the rate of review is irrelevant for long-run selection results: the deterministic dynamic in a given state only points in the direction of one absorbing state (0 or N), and so however quickly or slowly players review, they are only going to end up at one equilibrium. In Hawk-Dove games, however, the speed of adjustment has been seen in section 3 to be crucial for long-run selection; fast review led play to the extreme recurrent class $\{0, N\}$, whilst slower review (as delivered by the "contraction" conditions) was required for selection of the mixed-strategy equilibrium. As discussed, the "contraction" conditions essentially require that a sufficiently small fraction of those players whose current best-response is to switch strategies actually do so, and that this fraction shrink the closer is the current state to the mixed-strategy equilibrium. It was seen in section 3 that this either implied a requirement of very slow review, or left one with the project of trying to provide a theoretical foundation for a dynamic which

[38] Assuming, without loss of generality, that $z^* > \frac{N}{2}$.

[39] This is also the approach of Myatt and Wallace (1998) and Blume (1999).

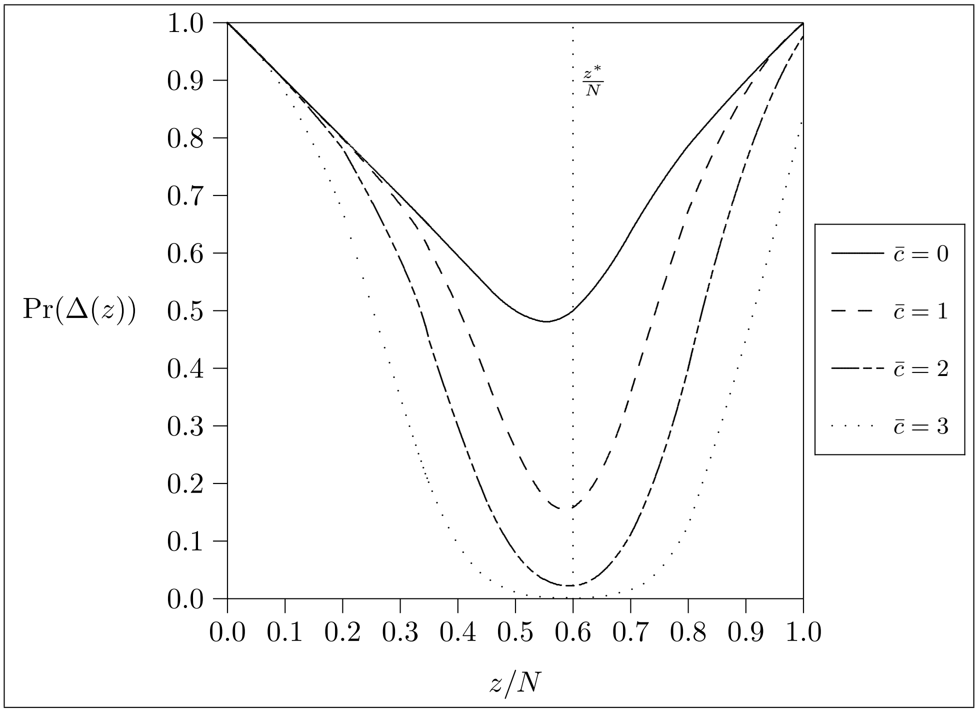


Figure 7: Probability $\Pr(\Delta(z))$ of a randomly selected player switching strategies

sees the fraction of adjusting players fall (sufficiently quickly) as the system approaches the mixed-strategy equilibrium.

But this is precisely what is accomplished by the simultaneous-revisions stochastic switching costs model of subsection 4.2.2. Essentially, Proposition 8's threshold mean switching cost for long-run selection of the mixed-strategy equilibrium captures the requirement that evolution be sufficiently slow; the endogenous review rate is lower the higher is the mean switching cost, as illustrated in figure 7.[40] Moreover, the review rate *falls* as the system approaches the mixed-strategy equilibrium, since lower expected payoff gains are at stake. Under stochastic switching costs and simultaneous revisions then, convergence to the mixed-strategy equilibrium has more satisfactory theoretical foundations.

A closer look at figure 7 yields some intuition for the selection results.

[40]The diagram assumes normally distributed switching costs $C \sim N(\bar{c}, 1)$ for the particular Hawk-Dove game

$$\Lambda_{HD} = \begin{matrix} -4 & 0 \\ 0 & -6 \end{matrix}$$

The probability that a randomly selected player switches strategies,

$$\Pr(\Delta(z)) = \left(\frac{z}{N}\right) F(\varpi_2(z-1)) + \left(\frac{N-z}{N}\right) F(\varpi_1(z))$$

tends to zero at (z^*/N) as $\bar{c} \to \infty$. Meanwhile, it tends to 1 in the extreme states $(z/N) = 0$ and $(z/N) = 1$ as $\bar{c} \to 0$. Since the values $\Pr(\Delta(0)) = \Pr(\Delta(N)) = 1$ would make $\{0, N\}$ a recurrent class (of the perturbed model), whilst $\Pr(\Delta(z^*)) = 0$ would make z^* an absorbing state, one should expect there to exist a threshold level of the mean switching cost $\bar{c}$ below which $\{0, N\}$ is selected in the long run, and above which z^* is selected. This has of course been demonstrated to be the case.

It is also worth noting that the selection results could be strengthened in one obvious way. The best-response dynamic was employed for simplicity and in order to focus attention on the significance of switching costs for the rate of review, but there is no reason why more boundedly rational response mechanisms could not be admitted. Indeed, a truly realistic model would require inertia to be present with various other forms of bounded rationality, such as simple rules of thumb, imitation and so on. Such a retreat from best-response would serve further to slow the pace of evolution, further reducing the attractiveness of the extreme states, with fewer players likely to make the optimal strategy switch in any given period. As a result, a lower threshold $\hat{c}$ would be required for long-run selection of the mixed-strategy equilibrium.

5 Conclusion

This paper is motivated by the belief that player inertia is an important phenomenon in repeated game contexts, and that it is driven in large part by the presence of switching costs to changes in behaviour. Such switching costs are introduced within a stochastically adaptive population repeatedly playing a Hawk-Dove game. The mixed-strategy equilibrium of such games is appealing by virtue of being the unique *symmetric* Nash equilibrium, and yet its long-run selection in stochastic evolutionary models has proved problematic; the required conditions on the deterministic dynamic generally dictate extremely slow evolution, or a review rate that falls as the system approaches the mixed-strategy equilibrium. The presence of switching costs emerges as unhelpful in this regard in a uniform-mutations setting, serving further to illustrate the limitations of this modelling framework. By contrast, the (simultaneous-revisions) state-dependent mutations model of section 4 overcomes the difficulties. By allowing the mutation (or review) rate to be endogenously determined by stochastic switching costs, a mechanism for slower, payoff-sensitive evolution is built into the model. Indeed, since the expected payoff gains at stake fall the closer is the system to the mixed-strategy equilibrium, so does the endogenous review rate, in precisely the manner required for selection of the mixed-strategy equilibrium. This is duly delivered for sufficiently high mean switching costs.

And so, unusually, the mixed-strategy equilibrium has a satisfactory theoretical foundation as the long-run equilibrium of a stochastic evolutionary model. Happily, it is clearly superior in terms of expected payoffs to the alternative equilibrium - the miserable extreme state cycle $\{0, N\}$ where players are constantly coordinating their actions. Nonetheless, this inefficient cycle remains the long-run equilibrium if the mean switching cost is below the threshold required for selection of the mixed-strategy equilibrium. In this case, one might anticipate that players would attempt to condition their behaviour on their roles in the game (i.e., Row player or Column player), which in any case is likely to deliver outcomes superior to the mixed-strategy equilibrium.[41] Relatedly, they might attempt to condition their strategies on some other information (e.g., race, gender, income, eye colour), leading to the decentralised emergence of "discriminatory norms" and even classes, as in Axtell, Epstein, and Young (2001). When such conditioning is not possible, however, the results of this paper suggest an appealing foundation for the selection of the mixed-strategy equilibrium, and the avoidance of the worst-case scenario of the extreme state cycle.

References

Arthur, W. B. (1993): "On Designing Economic Agents that Behave Like Human Agents," *Journal of Evolutionary Economics*, 3, 1–22.

Axtell, R. L., J. M. Epstein, and H. P. Young (2001): "The Emergence of Classes in a Multi-Agent Bargaining Model," in *Social Dynamics*, ed. by S. N. Durlauf, and H. P. Young. Brookings Institution Press, Washington, D.C.

Beggs, A., and P. Klemperer (1992): "Multi-Period Competition with Switching Costs," *Econometrica*, 60, 651–666.

Bergin, J., and B. Lipman (1996): "Evolution with State-Dependent Mutations," *Econometrica*, 64, 943–956.

Binmore, K., and L. Samuelson (1997): "Muddling Through: Noisy Equilibrium Selection," *Journal of Economic Theory*, 74, 235–265.

——— (2001): "Evolution and Mixed Strategies," *Games and Economic Behavior*, 34, 200–226.

Blume, L. (1999): "How Noise Matters," Cornell University Working Paper.

[41] For more on such role-conditioning, see Maynard Smith (1982), Selten (1980), and Binmore and Samuelson (2001).

Börgers, T., and R. Sarin (1995): "Naive Reinforcement Learning with Endogenous Aspirations," Discussion Paper, University College, London.

——— (1997): "Learning through Reinforcement and Replicator Dynamics," *Journal of Economic Theory*, 77, 1–14.

Bush, R., and F. Mosteller (1955): *Stochastic Models for Learning.* Wiley, New York.

Crawford, V. P. (1985): "Learning Behavior and Mixed-Strategy Nash Equilibria," *Journal of Economic Behavior and Organization*, 6, 69–78.

Ellison, G. (2000): "Basins of Attraction, Long-Run Stochastic Stability, and the Speed of Step-by-Step Evolution," *Review of Economic Studies*, 67, 17–45.

Ellison, G., and D. Fudenberg (2000): "Learning Purified Mixed Equilibria," *Journal of Economic Theory*, 90, 84–115.

Eshel, I., and E. Sansone (1995): "Owner-Intruder Conflict, Grafen Effect and Self-Assessment: The Bourgeois Principle Re-examined," *Journal of Theoretical Biology*, 177, 341–356.

Farrell, J. (1987): "Competition with Lock-In," University of California, Berkeley: Mimeo.

Farrell, J., and P. Klemperer (2001): "Coordination and Lock-In: Competition with Switching Costs and Network Effects," Working Paper.

Farrell, J., and C. Shapiro (1988): "Dynamic Competition with Switching Costs," *RAND Journal of Economics*, 19, 123–137.

——— (1989): "Optimal Contracts with Lock-in," *American Economic Review*, 79, 51–68.

Foster, D. P., and H. P. Young (1990): "Stochastic Evolutionary Game Dynamics," *Theoretical Population Biology*, 38, 219–232.

Freidlin, M. I., and A. D. Wentzell (1984): *Random Perturbations of Dynamical Systems*. Springer-Verlag, Berlin.

Fudenberg, D., and D. M. Kreps (1993): "Learning Mixed Equilibria," *Games and Economic Behavior*, 5, 320–367.

Harsanyi, J. (1973): "Games with Randomly Disturbed Payoffs: A New Rationale for Mixed-Strategy Equilibrium Points," *International Journal of Game Theory*, 2, 1–23.

Kandori, M., G. J. Mailath, and R. Rob (1993): "Learning, Mutation and Long-Run Equilibria in Games," *Econometrica*, 61, 29–56.

Karlin, S., and H. M. Taylor (1975): *A First Course in Stochastic Processes*. Academic Press, San Diego, second edn.

Klemperer, P. (1987a): "The Competitiveness of Markets with Switching Costs," *RAND Journal of Economics*, 18, 138–151.

——— (1987b): "Markets with Consumer Switching Costs," *Quarterly Journal of Economics*, 102, 375–394.

——— (1995): "Competition when Consumers have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics, and International Trade," *Review of Economic Studies*, 62, 515–539.

Lipman, B. L., and R. Wang (2000): "Switching Costs in Frequently Repeated Games," *Journal of Economic Theory*, 93, 149–190.

Maynard Smith, J. (1982): *Evolution and the Theory of Games*. Cambridge University Press, Cambridge.

Myatt, D. P., and C. Wallace (1998): "Adaptive Play by Idiosyncratic Agents," Discussion Paper, Nuffield College, Oxford.

Norman, T. (2003a): "The Evolution of Coordination under Inertia," Working Paper.

——— (2003b): "Step-By-Step Evolution with State-Dependent Mutations," Working Paper.

Oechssler, J. (1997): "An Evolutionary Interpretation of Mixed-Strategy Equilibria," *Games and Economic Behavior*, 21, 203–237.

Osborne, M. J., and A. Rubinstein (1994): *A Course in Game Theory*. The MIT Press, Cambridge, Massachusetts.

Padilla, A. J. (1995): "Revisiting Dynamic Duopoly with Consumer Switching Costs," *Journal of Economic Theory*, 67, 520–530.

Radner, R. (1980): "Collusive Behavior in Oligopolies with Long but Finite Lives," *Journal of Economic Theory*, 22, 136–156.

Robson, A., and F. Vega-Redondo (1996): "Efficient Equilibrium Selection in Evolutionary Games with Random Matching," *Journal of Economic Theory*, 70, 65–92.

Roth, A. E., and I. Erev (1995): "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term," *Games and Economic Behavior*, 8, 164–212.

Selten, R. (1980): "A Note on Evolutionarily Stable Strategies in Asymmetric Animal Contests," *Journal of Theoretical Biology*, 84, 93–101.

Sethi, R. (1998): "Strategy-Specific Barriers to Learning and Nonmonotonic Selection Dynamics," *Games and Economic Behavior*, 23, 284–304.

Stokey, N. L., R. E. Lucas, and E. C. Prescott (1989): *Recursive Methods in Economic Dynamics*. Harvard University Press, Cambridge, Massachusetts.

Suppes, P., and R. Atkinson (1960): *Markov Learning Models for Multiperson Interactions*. Stanford University Press, Stanford.

van Damme, E., and J. W. Weibull (1998): "Evolution with Mutations Driven by Control Costs," Tilburg University: Mimeo.

Young, H. P. (1993): "The Evolution of Conventions," *Econometrica*, 61, 57–84.

——— (1998): *Individual Strategy and Social Structure*. Princeton University Press, Princeton, New Jersey.