Equilibrium Selection and the Rate of Convergence in Coordination Games with Simultaneous Play

Hansen, Per Svejstrup; Kaarbøe, Oddvar M.

Publication date:
1998

Document version
Early version, also known as pre-print

Citation for published version (APA):
Hansen, P. S., & Kaarbøe, O. M. (1998). *Equilibrium Selection and the Rate of Convergence in Coordination Games with Simultaneous Play*. Department of Economics, University of Copenhagen.

Discussion Papers

Department of Economics

University of Copenhagen

No. 98-14

Equilibrium Selection and the Rate of Convergence in Coordination Games with Simultaneous Play

Per Svejstrup Hansen

Oddvar M. Kaarbøe

Studiestræde 6, DK-1455 Copenhagen K., Denmark

Tel.: +45 35 32 30 82 – Fax: +45 35 32 30 00

http://www.econ.ku.dk

ISSN: 1601-2461 (online) ISSN 0902-6452 (print)

Equilibrium Selction and the Rate of Convergence in Coordination Games with Simultaneous Play[1]

Per Svejstrup Hansen

University of Copenhagen, Department of Economics,

Studiestraede 6, 1455 Copenhagen K, Denmark

Oddvar M. Kaarbxe[2]

Department of Economics, University of Bergen,

Fosswinckelsgate 6, 5007 Bergen, Norway.

August 31, 1998

[1]Sjur D. Flem read several drafts of this paper, and his suggestions have improved it enormously. Helpful suggestions were also provided by Mogens Jensen, Kjell Erik Lommerud, Larry Samuelson, Birgitte Sloth, Peter Sxrensen and Peyton Young. The paper also bene¿ted from discussions following several seminar presentations. The usual disclaimer applies.

[2] Correspondence: Oddvar M. Kaarbxe, phone + 47 55 58 92 26, fax + 47 55 58 92 10, email: oddvar.kaarboe@econ.uib.no.

Abstract

We apply the dynamic stochastic framework proposed in the recent evolutionary literature to a class of coordination games played simultaneously by the entire population. In these games, payo¾s whence best replies are determined by a summary statistic of the population strategy pro¿le. We demonstrate that with simultaneous play, the equilibrium selection depends crucially on how best responses to the summary statistic remain piece-wise constant. In fact, all the strict Nash equilibria in the underlying stage game can be declared stochastically stable depending on how the best response mapping generates piece-wise constant best responses. Furthermore, we show that if the best response mapping is suÁciently asymmetric, the expected waiting time until the unique stochastically stable state is reached is of the same order as the mutation rate, even in the limit as the population size grows to in¿nity.

Per Svejstrup Hansen

University of Copenhagen, Department of Economics

Studiestraede 6, 1455 Copenhagen K, Denmark

1 Introduction

We apply the dynamic stochastic framework proposed in Kandori, Mailath and Rob (1993) (henceforth KMR) to a class of coordination games played simultaneously by the entire population. This is taken to refer to a context where the interaction between players are weak and di¾use and therefore does not lend themselves to modeling with pairwise interaction, anonymous or otherwise.

We choose to model the weak and di¾use interaction among agents in such away that individual payo¾s depend on the player's own strategy and a summary statistic of the population's strategy pro¿le. Speci¿cally, we assume that each player's strategy space is discrete and consist of M linearly ordered strategies, and, as is often assumed in economic models, the statistic is taken to be the mean of the current strategy distribution.

In the class of games studied in this paper players would try to coordinate since they receive a strictly higher payo¾ from playing a strategy that matches the current population-wide mean, than from playing any other strategy. This implis that there are M strict Nash equilibria in this class of games. In addition we posit that the strategies are totally rankable in the Pareto sense, and that coordinating on a higher value of the statistic gives the player a strictly higher payo¾ than coordination on a lower one.

Note that since there are more average numbers than strategies, the best response mapping cannot be one-to-one. Therefore best responses are piece-wise constant around a given strategy. One of the contributions of this paper is to demonstrate that with simultaneous play, the determination of the stochastically stable states depends crucially on how best responses remain piece-wise constant. In fact all the strict Nash equilibria in the underlying stage game can be declared stochastically stable depending on the way piece-wise constant best responses are specified. This result holds even when we approximate a continuous strategy space, i.e. when we by-pass any artificial considerations that could be associated to the discreteness of the players' strategy space. Furthermore, we show that if piece-wise constant best responses are sufficiently asymmetric, the unique stochastically stable state consists of all players playing one of their extreme strategies (which one depends on the way the asymmetry goes). In this case the expected waiting time until the unique stochastically stable state is reached is of the same order as the mutation rate, even in the limit as the population size grows to infinity. Hence, unlike in many models of random pairing interaction, convergence may in fact be very rapid even though the mutation rate is small.

Our motivation is threefold. First, much research in both traditional and evolutionary game theory has been devoted to discriminate between equilibria in games that exhibit multiple strict Nash equilibria. In coordination games many hold the belief that the Pareto dominant equilibrium stands out as a focal point, and thus should be selected as the equilibrium. Other apply the concept of risk dominance introduced by Harsanyi and Selten (1988), as the refinement criterion. In general, the two concepts, Pareto efficiency and risk dominance, differ. However, in symmetric pure coordination games they coincide. Kandori and Rob (1995) show that for general $n \times n$ pure coordination games the Pareto efficient equilibrium is selected as the unique stochastically stable state, when players are randomly matched in pairs. In a recent article Robles (1997) considers a model which is similar in structure to ours. That is, he studies a simultaneous play coordination game that also applies the evolutionary dynamics of KMR. What Robles (1997) shows is that in coordination games with simultaneous play and payoffs determined by “averaged strategies,” the stochastically stable states are bounded away from the extreme strategies, including the Pareto efficient Nash equilibrium. Apparently, there is a stark contrast between random pairing and simultaneous play. But as we show, the results in Robles (1997) are accounted for by the way he defines the piece-wise constant best response mapping. The Pareto efficient equilibrium may be selected as stochastically stable as may any other strict Nash equilibrium, depending on details of the best response mapping.

Second, one of the criticisms of the relevance of the concept of stochastic stability is that the speed of convergence may be very slow, indeed. The inclusion of a noise term meant to capture for instance mutations or trembles, makes all the strict Nash equilibria occur with positive probability. However, some may be more likely than others. If the long run probability of a (subset of) strict Nash equilibria does not vanish as the noise approaches zero, these states are stochastically stable. The problem is, as pointed out by Ellison (1993) among others, that if the state initially is in a non-stochastically stable state, convergence may be so slow that for all practical purposes, the stochastically stable states are never reached. In fact, Binmore, Samuelson and Vaughan (1995) have estimated that going from the payoff dominant equilibrium to the risk dominant one in the KMR-model, has an expected wait of 1.7×10^{72} periods, when the number of players are 100, the noise-rate is $1/100$, and the payoffs are such that at least 33 of a player's opponents must play the risk dominant equilibrium strategy to induce a switch in the agent in question's best reply. Our model which have features in common with Ellison's (1996) analysis of step-by-step evolution, shows that if piece-wise constant best responses are sufficiently asymmetric, convergence is of the same order as the mutation rate even in the limit as the population size grows to infinity. Thus, another important difference between random pairing and simultaneous play.

Thirdly, apart from Robles (1997), the evolutionary literature has not thoroughly analysed games with simultaneous play, even though Crawford (1991, 1997) forcefully argues for introducing genuine simultaneous interaction into this literature. What seems relevant in many models of economic theory, be it of oligopolies, macroeconomic coordination failure models or models of individual consumers' demand for goods such as popular restaurant seats or theater tickets (Becker, 1990), is an interaction structure characterized by simultaneous play rather than random pairing, anonymous or otherwise. In addition, agents react to some average of other agents' behaviour in these models. Hence, we argue that what is relevant for many economic applications is a simultaneous play interaction pattern with a payoff structure determined in part by the mean of the current strategy distribution. However, from a game theoretical perspective the equilibrium selection mechanism in these games is rather discomforting, since all the strict Nash equilibria of the underlying stage game can be selected as part of the set of stochastically stable states by an appropriate specification of the best response mapping. Unless the best response mapping generates sufficiently asymmetric piece-wise constant best responses in which case our model has strong predictive power, as well as fast convergence to the predicted stochastically stable states.

A natural question that arises is how the best response mapping ought to be defined? Robles postulates without any further argumentation that population averages, which lie between two adjacent discrete strategy choices, should be transformed onto the nearest one of these strategies. That is, if a value of the average is, say, 3.51 the optimal individual strategy is to play 4, whereas it is to play 3 if the population average is 3.49.[1] In pure coordination games, this way of defining the best response mapping does not seem appealing. In this class

[1] Robles (1997) breaks ties such that 3.50 is mapped onto 4.

of games, individual payo¾s are positive if the player's choice of strategy equals the summary statistic, otherwise individual payo¾s are zero. The strategies are also totally rankable in the Pareto sense, such that coordinating on a higher value of the summary statistic give the players a strictly higher payo¾ than coordination on a lower one. We argue that if players look at their payo¾s, the natural way of specifying the best response mapping is such that any population average slightly above a discrete strategy, ought to lead a player to choose the next higher strategy. This gives the player a higher payo¾ and involves no greater risk since both actions are equally secured.[2] These theoretical arguments suggest that a symmetric de¿nition of the best response mapping in pure coordination games is questionable. However, how individuals are likely to perceive what is a best reply to a given statistic is an empirical matter. After all, the perception of best responses is not a choice variable but intrinsic to agents.

The paper is organised as follows. Section 2 serves for motivation and heuristics. It presents the general idea by way of a simple example. Sections 3 and 4 turn the intuition into formal analysis. Section 3 provides the general model, and section 4 states a general possibility theorem saying that, in symmetric coordination games with simultaneous play and an average payo¾ structure, any strict Nash equilibrium can be selected as part of the set of stochastically stable states by an appropriate de¿nition of piece-wise constant best responses. To illustrate the theorem, we calculate numerically the set of stochastically stable states for a given error rate and di¾erent ways of de¿ning piece-wise constant best responses. Section 5 considers the rate of convergence and step-by-step evolution, while section 6 discusses the

[2] A secure action is an action whose lowest payo¾ is at least as large as the lowest payo¾ to any other feasible action. (Van Huyck, Battalio and Beil, 1991).

results and suggests how the best response mapping could be de¿ned for di¾erent classes of coordination games.

2 An Example

Consider a situation where a ¿nite number of players, each having the same ¿nite set of strategies, play a simultaneous coordination game. Individual payo¾s are determined by the player's own action and a population-wide average of the opponent players' strategy choices.[3] For this class of games, we show that any strict Nash equilibrium of the underlying stage game can be stochastically stable depending on how the best response mapping generates piece-wise constant best reponses. This will be derived formally in the following sections but before addressing the theoretical issues, we shall illustrate the point by a simple example.

Consider a symmetric pure coordination game with $N = 9$ players and $M = 5$ linearly ordered strategies for each of them. Let $\pi(m, \overline{\mu})$ be the payo¾ to an individual playing strategy $m \in \{1, 2, 3, 4, 5\}$ when the mean of the population's current strategy pro¿le equals $\overline{\mu}$. Since there are more average numbers $\overline{\mu}$ than strategies m, the best response mapping, $B(\overline{\mu})$, cannot be one-to-one. Speci¿cally, suppose the level-set´$B^{-1}(m) = [m-a, m+1-a)$ is a half-open intervall for some constant $a \in [0, 1)$. In other words, we de¿ne an integer-value function which takes $\overline{\mu} \in [m - a, m + 1 - a) \to m$, such that the best response is piece-wise

[3] The opponents are taken in a wide sense, i.e. the player himself ¿gures among the opponents. If the players knew they could alter the population-wide average by their strategy choices, non of the results in this paper would change in qualitative terms.

constant around a given integer-value of m. Hence,

$$B(\overline{\mu}) = B(m) = m$$

whenever $m - a \leq \overline{\mu} < m + 1 - a$.

Introducing myopic best responses and mutation dynamics as in Kandori and Rob (1995), we follow Young (1993) in de¿ning the stochastically stable states of the game as those states which are the roots of the least resistant paths, where the resistance in this case is the minimum number of players who must mutate in order to move from a state where everyone plays m to a state where everyone plays $m' \neq m$. Denote the minimum resistance of going from m to m' by $r_{mm'}$. It has been shown in Kandori and Rob (1995) that only adjacent states need to be compared for obtaining the stochastically stable states in pure coordination games. Thus, we have to compare $r_{m,m+1}$ and $r_{m+1,m}$ where $m \in \{1, ..., 4\}$. Now assume the integer-function de¿nes pieces symmetrically; that is, $a = \frac{1}{2}$. This is the case in Robles (1997) and as we will demonstrate, the key to understand his equilibrium selection mechanism. Set up the tree as below where the numbers above and below the arrows indicate the resistances of going upward and downward, respectively.

Figure 1, $a = 1/2$

$$1 \;\underset{5}{\overset{2}{\rightleftarrows}}\; 2 \;\underset{3}{\overset{2}{\rightleftarrows}}\; 3 \;\underset{2}{\overset{3}{\rightleftarrows}}\; 4 \;\underset{2}{\overset{5}{\rightleftarrows}}\; 5$$

It is easily seen that state 3 is stochastically stable since it is easier (i.e. requires fewer mutations) to go from 1 to 2 than the opposite way. Similarly for 2 to 3. It also involves fewer mutations to go from 5 to 4 than from 4 to 5. The same applies for 4 to 3. For comparison assume instead that $a = \frac{5}{9}$. This makes more numbers go up to m than for $a = \frac{1}{2}$. Setting up a new tree, we observe that the stochastically stable states are 3 and 4.

Figure 2, $a = 5/9$

$$1 \quad \underset{3}{\overset{1}{\rightleftarrows}} \quad 2 \quad \underset{3}{\overset{2}{\rightleftarrows}} \quad 3 \quad \underset{2}{\overset{2}{\rightleftarrows}} \quad 4 \quad \underset{2}{\overset{4}{\rightleftarrows}} \quad 5$$

This shows that just a small change in how the average is transformed onto a strategy choice, signi¿cantly alters the equilibrium selection. By changing the pieces slightly in favour of going upward, (by increasing $a \in [0, 1)$), the stochastically stable state(s) are biased towards the Pareto eÁcient outcome. The example suggests that by an even higher choice of $a \in [0, 1)$, players would coordinate on the Pareto eÁcient equilibrium. Similar arguments apply for tending towards the least eÁcient equilibrium $\{1\}$. If the same game is played with a random pairing interaction structure, Kandori and Rob (1995) show that the Pareto eÁcient equilibrium is stochastically stable. So clearly there is a di¾erence between random pairing and simultaneous play, but as the above example illustrates, the di¾erence seems to lie in how each player's best response remains piece-wise constant in response to other players' averaged strategy pro¿le, and not so much in the di¾erence in the interaction structure per se. The way of de¿ning piece-wise constant best responses determines the equilibrium selection.

3 The Model

Following Robles (1997), we consider a ¿nite population $\mathcal{N}$ of size N composed of players $n \in \mathcal{N} := \{1, 2, ..., N\}$. At each time $t = 1, 2, ...$ these individuals play simultaneously a symmetric coordination game with linearly ordered strategies $m \in \mathcal{M} := \{1, 2, ..., M\}$.[4] Individual payo¾, $\pi(m, \overline{\mu}(s))$, depends on own action $m \in \mathcal{M}$ and the population-wide mean, $\overline{\mu}(s) := \frac{1}{N} \sum_{m \in \mathcal{M}} m\# \left(\text{players using} m\right)$, which is observable.[5] The (unobservable) state $s = (s_1, ..., s_M)$ is a vector, whose mth element, s_m, represents the number of players using strategy $m \in \mathcal{M}$. Thus, the state space of the system is chosen equal to M^N, where M is the strategy grid introduced above. We assume that $\pi(m, m) < \pi(m', m')$ whenever $m < m'$, and $\pi(m, m') < \pi(m', m')$ whenever $m \neq m'$.

The stage game described above, exhibits M strict Nash equilibria in which all players choose the same strategy. In addition, the Nash equilibria are totally ranked in the Pareto sense; when all players choose strategy 1 the least eÁcient equilibrium is generated, and Pareto optimum results when every player plays her highest strategy M.

Like Kandori and Rob (1995) we assume that strategy adjustment is not instantaneous but is subject to some friction. Speci¿cally, it is assumed that at every $t = 1, 2, ...$ each player takes an independent draw from a Bernoulli trial. With probability $(1 - \zeta) \in (0, 1)$ this draw produces the outcome ˝do not learn´and the player stays with her strategy. With the complementary probability ζ the draw produces the outcome ˝learn´. In this case the

[4] Schelling (1973) introduced the simultaneous play model in economics. In biology the term ˝playing the ¿eld´is used to indicate interaction with a whole population, (Maynard Smith, 1978).

[5] Robles (1997) considers games where individual payo¾s depend on own action and convex combinations of the order statistics of the population's current strategy con¿guration. It should be noted though that the results in our paper generalize, in a qualitatively way, to all convex combinations of order statistics as long as all order statistics have positive weight.

player is able to observe the average of the population's current strategy pro¿le and switches to a best response to the period t average.[6] We assume that she believes her opponents to stay with their strategies and that her choice has a negligible e¾ect on the average. Hence, her myopic best response is to match the current value of the mean.

We assume the existence of a partition of the real interval $[1, M]$ into neighbourhoods (vicinities) $V_1, ..., V_M$ of $1, ..., M$ respectively such that the best response

$$B(\overline{\mu}) := \arg\max \pi(\cdot, \overline{\mu})$$

is constant on each V_m, $m = 1, ..., M$. We shall therefore speak of piece-wise constant best responses. For tractability we assume that

$$V_m = [1, M] \cap [m - a, m + 1 - a)$$

for some $a \in [0, 1)$. In other words,

$$B(\overline{\mu}) = B(m) = m$$

whenever $m - a \leq \overline{\mu} < m + 1 - a$.

In addition to the myopic best-response dynamics, idiosyncratic behaviour is modelled in the following way. For all t, each player $n \in \mathcal{N}$ is subject to some probability $\varepsilon > 0$ of ´mutating˝ in which case the player chooses any strategy $m \in \mathcal{M}$ in a purely arbitrarily

[6] It is without importance that the player observes the average and not the state s since the payo¾ to the player depends on this average and not on how many players who are playing the di¾erent strategies.

manner with positive probability on each $m \in \mathcal{M}$. These events, which occur after the best-response adjustments, are assumed to be independent across players and over time.

The composition of myopic best responses and mutations generates a discrete-time Markov-process over the ¿nite state space S, whose transition matrix is denoted $P(\varepsilon) = (p_{ss'}(\varepsilon))$. An element $p_{ss'}(\varepsilon)$ represents the transition probability of moving to state s' at time $t+1$ conditional on being in state s at time t. The ḿutation-free´dynamics itself corresponds to $P(0)$.

The presence of mutations implies that every transition has positive probability. It is a standard result that such Markov chains have a unique stationary probability distribution. Let $\mu(\varepsilon)$ denote the unique invariant distribution of $P(\varepsilon)$ for each $\varepsilon > 0$. The aim is to characterize the limit

$$\mu^* := \lim_{\varepsilon \to 0} \mu(\varepsilon).$$

Based upon arguments in Freidlin and Wentzell (1984), Foster and Young (1990) have shown that this limit exists and they called it the stochastically stable distribution. Call the support of this limit distribution the set of stochastically stable states and denote it Θ.

As a ¿rst step towards computing the set of stochastically stable states we will identify the set of the recurrent classes under $P(0)$. Denote this set Γ and let e_m be the state where all players play strategy m.

Proposition 1 Using the arithmetic mean as a summary statistic, the set of recurrent classes in the unperturbed game is $\Gamma = \{\{e_1\}, \{e_2\}, ..., \{e_M\}\}$ for any integer-value function $\lceil \cdot \rceil_a : \mathbb{R} \to Z$, de¿ned by $\lceil \cdot \rceil_a (r) := \lceil r \rceil_a = z$ whenever $r \in [z-a, z+1-a)$, z being an integer and $a \in [0, 1)$.

Proof. If $s = e_m$ then $\overline{\mu} = m$. Therefore $B(m) = m$, irrespectively of $\lceil \cdot \rceil_a$, $a \in [0, 1)$. Hence, $e_m \in \Gamma$. If $s' \neq e_m$ but $\lceil \overline{\mu}(s') \rceil_a = m$, then there are individuals who do not play a best response to the current state s'. Let all those players revise their strategy choices. Since they will all change their strategy to m, e_m is reached in one step. Combined with the fact that e_m is an absorbing set this implies that s' is a transient state and therefore $s' \notin \Gamma$. ■

In order to determine μ^*, we need to know the relative size of the transition probabilities, $p_{ss'}(\varepsilon)$, that are converging to zero. Since mutations are independent across players and over time, the elements of $P(\varepsilon)$ are polynomials in ε. In fact, the leading terms of $p_{ss'}(\varepsilon)$ have the form $\varepsilon^{r(s,s')}$, where $r(s, s')$ is the number of mutations needed to move from s to s'. Hence, the number of mutations corresponds to the order (in ε) of the corresponding transition probability. The stochastically stable states are precisely those states which can be reached from any other state with the fewest number of mutations. In addition, since $\mu(0)$ is the limit distribution of $P(0)$, it puts zero probability on every transient state. We may therefore restrict attention to the recurrent states to determine the set of stochastically stable states, Θ.

We now consider moving between two distinct recurrent states e_m and $e_{m'}$, $m \neq m'$, e_m, $e_{m'}$ $\in\Gamma$. For each pair of distinct recurrent states e_m and $e_{m'}$, $m \neq m'$, an mm$'$-path is a sequence of states $\sigma = (s^1, s^2, ..., s^q)$ which begins in e_m and ends in $e_{m'}$ for $m \neq m'$. The resistance of this path, $r(\sigma)$, is the sum of the resistances of its edges, that is $r(\sigma) = \sum_{k=1}^{q-1} r(s^1, s^{k+1})$ where $r(s^k, s^{k+1}) \in N_0 \cup \{\infty\}$ is the number of mutations required to move from state s^k to state s^{k+1}. Let $r_{mm'}$ be the least resistance over all mm'-paths σ.

In fact,

$$r_{mm'} = \min_{\sigma: s^1 = e_m, s^q = e_{m'}} r(\sigma).$$

A tree rooted at vertex m' (an $e_{m'}$-tree), is a set of $M-1$ directed edges, each for one recurrent state, such that from every vertex di¾erent from m', there is a unique directed path in the tree to m'. The weight on the directed edge $m \to m'$ is $r_{mm'}$. The resistance of a rooted tree, T, is the sum of the resistances $r_{mm'}$ on the $M-1$ edges that composes it. Let $T(e_{m'})$ be the set of $e_{m'}-$trees. Following Young (1993), we de¿ne the stochastic potential of the recurrent state $e_{m'}$ by

$$\gamma_{m'} = \min_{T \in T(e_{m'})} \sum_{(e_m, e_{m''}) \in T} r_{mm''}$$

We now state the theorem for determining the stochastically stable states (Young, 1993, Theorem 4).

Theorem 2 The stochastically stable states, $e_m \in \Theta$, are exactly the state(s) with minimum stochastic potential.

4 Equilibrium Selection

In this section we characterize the set of stochastically stable states for the average payo¾ games described in section 3. Since the stochastic potential of $e_m \in \Gamma$ is de¿ned to be the minimum resistance over all trees rooted at m, standard tree constructions determine which e_m has the lowest stochastic potential.

When a player $n \in \mathcal{N}$ learns, her myopic best response is to match the integer-discretised mean of the population's current strategy pro¿le. Hence, to assess the likelihood of a move from the state e_m to $e_{m'}$, we need to ¿nd the minimum number of mutations required to change the average from m to m'. Since large jumps in an individual strategy change the average more then small jumps, having players mutate to extreme strategies is often the ¿rst step along a minimum resistance path. If $1 \leq m < m' < h \leq M$, then evidently there are more strategies above m than m'. This means that one mutation to h has a larger impact on the average when the state is e_m, than when the state is $e_{m'}$. Therefore, the number of mutations needed to destabilize equilibrium e_m upwards must be less than or equal to the number of mutations needed to destabilize equilibrium $e_{m'}$ in the same direction. A similar argument applies to the number of mutations needed to make the transition from a higher to a lower state.[7] Furthermore, a slight modi¿cation of Proposition 3.2 in Robles (1997), allowing for a general speci¿cation of piece-wise constant best responses, implies that to ¿nd the minimum resistance path from e_m we only need to consider adjacent recurrent states, i.e. e_{m-1} and e_{m+1}. Therefore, to ¿nd the resistance for the transition $e_m \rightarrow e_{m+1}$, we need to ¿nd the number of players, de¿ned as $K_{m,m+1}$, who must mutate to M such that the best reply for an m-player, who learns, is to play a strategy $\geq m+1$. Hence, $K_{m,m+1}$ must satisfy $\frac{K_{m,m+1}}{N}M + \frac{(N-K_{m,m+1})}{N}m \geq m+1-a$. Now, de¿ne $\overline{K}_{m,m+1}$ as the minimum number of players who must play M for the above expression to be satis¿ed. Clearly $\overline{K}_{m,m+1}$ depends on a. In fact, $\overline{K}_{m,m+1}(a) := \min\left\{K_{m,m+1} : K_{m,m+1} \geq \frac{(1-a)N}{M-m}\right\}$. Similarly, for the transition $e_{m+1} \rightarrow e_m$, we need to ¿nd the minimum number of players who must mutate to 1 in order for an $(m+1)$-player's best response to be to play a strategy $\leq m$, assuming the $(m+1)$-player

[7] This is what we state formally in Lemma 8.

receives a learning draw. This is de¿ned as $\underline{K}_{m+1,m}(a) := \min\left\{K_{m+1,m} : K_{m+1,m} > \frac{aN}{m}\right\}$. For completeness set $\overline{K}_{M,M+1} = \underline{K}_{1,0} = \infty$.[8]

The following proposition yields a simple characterization of the resistance between two states in Γ.

Proposition 3 $\quad r_{m,m+1}(a) = \overline{K}_{m,m+1}(a)$ and $r_{m+1,m}(a) = \underline{K}_{m+1,m}(a)$.

Proof. Appendix. ■

The next proposition states the conditions for e_m to be supported by the stochastically stable states. It asserts that e_m is a (part of) the stochastically stable states if and only if more mutations are required to move the state from e_m to e_{m-1} and from e_m to e_{m+1} than the other way around. In other words, each inward resistance must be less than the corresponding outward one.

Proposition 4 $\quad e_m \in \Theta$ i¾ $r_{m-1,m}(a) \leq r_{m,m-1}(a)$ and $r_{m+1,m}(a) \leq r_{m,m+1}(a)$.

Proof. The proof follows with a slight modi¿cation from Robles (1997, Proposition 3.2). ■

From the de¿nitions of $\overline{K}_{m,m+1}(a)$ and $\underline{K}_{m+1,m}(a)$ it follows that the resistance between two states in Γ depends on how best responses are piece-wise constant. The next proposition gives necessary and suÁcient conditions for the lowest and highest strategies, respectively, to be stochastically stable.

[8] One potential problem is that it might be possible for $\overline{K}_{m,m+1}$ players who mutate to M, to raise the mean above the new state $m+1$, but not exactly to $m+1$. Lemma 3.1 in Robles (1997) shows that in that case players can mutate to a strategy $h < M$ and reach $m+1$ and that $\overline{K}_{m,m+1} - 1$ players is not suÁcient to increase the mean to $m+1$.

Proposition 5 Let $a \in [0, 1)$. i) If $a \in \left[0, \frac{N+M-1}{NM}\right]$ then $e_1 \in \Theta$, ii) if $a \in \left[1 - \frac{N+M-1}{NM}, 1\right)$ then $e_M \in \Theta$. If in addition $N > M - 1$ and iii) $a \in \left[0, \frac{N-M+1}{NM}\right]$ then $e_1 \in \Theta$ uniquely, or if iv) $a \in \left[1 - \frac{N-M+1}{NM}, 1\right)$ then $e_M \in \Theta$ uniquely.

Proof. A necessary and suÁ cient condition for $e_M \notin \Theta$ is that the number of mutations required to move the state from e_M to e_{M-1} is strictly less than the number of mutations required to move the state the opposite way. This follows from Proposition 4. In fact, using Proposition 3, and the de¿nitions of $\overline{K}_{M-1,M}(a)$ and $\underline{K}_{M,M-1}(a)$, a necessary and suÁ cient condition for $e_M \notin \Theta$ is $\frac{aN}{M-1} + 1 < \frac{(1-a)N}{M-(M-1)}$. From this expression it is easy to establish ii). The corresponding argument concerning i) is essentially identical.

To prove uniqueness, all that is required is that $\overline{K}_{M-1,M} < \underline{K}_{M,M-1}$ for e_M to be unique, and $\underline{K}_{2,1} < \overline{K}_{1,2}$ for e_1 to be unique. Then condition iii) and iv) follows from the de¿nition of $\overline{K}$ and $\underline{K}$. ■

In Robles (1997), where the integer-value function is de¿ned symmetrically, i.e., $a = \frac{1}{2}$, the stochastically stable states are bounded away from the extreme strategies for most parameter con¿gurations. The following corollary gives conditions for this to happen.

Corollary 6 Let $N \geq 5$, $M \geq 3$ and $a = \frac{1}{2}$. Then i) $e_1 \notin \Theta$, ii) $e_M \notin \Theta$.

Proof. The proof follows from Proposition 5. ■

We are now ready to state the main theorem, saying that in symmetric coordination games with simultaneous play and an average payo¾ structure, all the strict Nash equilibria of the underlying stage game can be decleared stochastically stable by an appropriate choice of how the population-wide average is transformed into a discrete strategy choice.

Theorem 7 If $N \geq 5$ and $M \geq 3$, any of the recurrent states $e_m \in \Gamma$ for $m \in \mathcal{M}$ can be selected as stochastically stable by appropriate choice of $a \in [0, 1)$.

Before we prove the theorem, we need to prove that the number of mutations needed to destabilize equilibrium e_m upwards (downwards) is less (larger) than or equal to the number of mutations needed to destabilize equilibrium e_{m+1} in the same direction.

Lemma 8 If $1 \leq m < M$, then $r_{m,m+1}(a) \leq r_{m+1,m+2}(a)$ and $r_{m+2,m+1}(a) \leq r_{m+1,m}(a)$.

Proof. From Proposition 3 we know that the resistances can be expressed in terms of $\overline{K}$'s and $\underline{K}$'s. Then, we have $\overline{K}_{m,m+1}(a) = \min\left\{K_{m,m+1} : K_{m,m+1} \geq \frac{(1-a)N}{M-m}\right\}$ and $\overline{K}_{m+1,m+2}(a) = \min\left\{K_{m+1,m+2} : K_{m+1,m+2} \geq \frac{(1-a)N}{M-(m+1)}\right\}$. Hence $\overline{K}_{m,m+1}(a) \leq \overline{K}_{m+1,m+2}(a)$. Similarly, $\underline{K}_{m+2,m+1}(a) = \min\left\{K_{m+2,m+1} : K_{m+2,m+1} > \frac{aN}{m+1}\right\}$ and $\underline{K}_{m+1,m}(a) = \min\left\{K_{m+1,m} : K_{m+1,m} > \frac{aN}{m}\right\}$. Therefore, $\underline{K}_{m+2,m+1}(a) \leq \underline{K}_{m+1,m}(a)$, and the Lemma is established. ■

Proof. (Theorem 7). From Proposition 5 we know that we can select e_1 as a stochastically stable state by choosing $a \in \left[0, \frac{N+M-1}{NM}\right]$. Call the upper bound of this interval a^1. >From the same proposition, e_M is stochastically stable when $a \in \left[1 - \frac{N+M-1}{NM}, 1\right)$. Denote the lower bound of this interval a^M. It is easy to verify that $a^1 < a^M$ when $N \geq 5$, $M \geq 3$.

Observe that for all $m : 1 \leq m < M$, $r_{m,m+1}(a)$ is non-increasing in $a \in [0, 1)$. This follows directly from the de¿nition of $\overline{K}_{m,m+1}(a)$. In fact, for a given $a = \overline{a}$, a small increase in a implies a change in $r_{m,m+1}(a) \in \{-1, 0\}$, for $1 \leq m < M$. Similarly, a small increase in a implies a change in $r_{m+1,m}(a) \in \{0, 1\}$, for $1 < m \leq M$.

To select any $e_m \in \Theta$ choose $a \in [0, a^1]$. If $m = 1$, then $e_1 \in \Theta$. If $m > 1$, then choose a slightly above a^1. As noted above, this increase in a implies a change in $r_{m,m+1} \in \{-1, 0\}$ and in $r_{m+1,m} \in \{0, 1\}$, and from Lemma 8 it follows that the changes in the resistances $r_{m,m+1}$ ($r_{m+1,m}$) are monotonically non-decreasing (non-increasing). If the increase in a changes the resistance such that the conditions given in Proposition 4 are satis¿ed, then $e_m \in \Theta$. If not, then continue to increase a until they are. If $m = M$, then we have to increase a until $a^M \leq a < 1$. Note that $a^M < 1$ for all $N \geq 5$, $M \geq 3$. In fact, $\lim_{M\to\infty} a^M = \frac{N-1}{N} < 1$ and $\lim_{N\to\infty} a^M = \frac{M-1}{M} < 1$. If $m < M$, it follows from Lemma 8 that the conditions will be satis¿ed for $a < a^M$ and the proof is complete. ■

To illustrate how the probability distribution accumulates on the di¾erent Nash equilibria of the underlying stage game, we can solve for the stationary distribution as a function of the discretisation parameter a and the mutation rate ε directly. The results for $\varepsilon = 0.01$ and selected values of a in a game with $N = 5$ players each having $M = 3$ strategies are summarized in Table 4.1. (We explain how probabilities are calculated in the appendix.)

Table 4.1. Long-run probabilities

	e_1	e_2	e_3
$a = \frac{9}{10}$	–	–	.967
$a = \frac{7}{10}$	–	.488	.479
$a = \frac{5}{10}$	–	.965	–
$a = \frac{3}{10}$	.479	.488	–
$a = \frac{1}{10}$	.976	–	–

A ´–´indicates less than .05 probability.

5 Rate of Convergence

In this section we argue that the way best responses to the summary statistic remain piece-wise constant has important consequences for the expected waiting time required to reach the stochastically stable states. Speci¿cally it is argued that if the best response mapping is suÁciently asymmetric, i,e.a is close to zero or one, then the expected wait to reach the stochastically stable states is relatively short, even if the mutation rate is small. Moreover, in the limit when a approaches zero or one, the expected wait remains of the same order as the mutation rate even when the population size grows to in¿nity. Hence, convergence is fast also in the second sense discussed in Ellison (1993, pp. 1060-1063). This is due to the fact that the system can easily escape the basin of attraction of each Nash equilibrium except the unique stochasticalle stable state e_1 or e_M.

The observation that evolution is more rapid when it may proceed via a series of small steps between intermediate recurrent states is analysed in Ellison (1996). Ellison gives the following biological example to provide intuition: Consider two di¾erent environments in which three major genetic mutations are necessary to produce the more ¿t animal w from animal x. In the ¿rst environment each single genetic mutation on its own, provides an increase in ¿tness that allows the mutants to take over the population. In the second, all three genetic mutations must occur simultaneously to create the animal with a higher ¿tness than x. If mutations are rare phenomena, the expected waiting time to see animal w being created is much larger in the latter case. Hence, the large cumulative change from x to w seems more plausible when gradual changes are possible.

As the analysis in section 4 shows, the minimum resistance paths in coordination games with simultaneous play are constructed between adjacent recurrent classes. Therefore, evolutionary changes occur step-by-step. As a result, the expected wait to reach the stochastically stable state from any given state in Ellison's step-by-step model and in the present one is the same.[9]

To show that convergence is fast when the best response mapping is suÁciently asymmetric we follow Ellison (1996) and de¿ne $\max_{s \in S} W(s, \Theta, \varepsilon)$ as the maximal expected wait until a state belonging to the set Θ is ¿rst reached given that play begins in state $s \in S$ when the mutation rate is $\varepsilon > 0$. If the expected wait is small, convergence is fast and Θ can be regarded as a good prediction of play, even in the medium run.

From the de¿nition of resistance, i.e. from Proposition 4, it follows that e_1 or e_M can be reached via a chain of single mutations when a is close to zero or one. More importantly, this

[9] Kaarbxe (1998) shows that it is easy to construct examples of simultanous play coordination games where Ellison's analysis is not applicable.

result holds also when the population size approaches in¿nity. As a result, the convergence rate is independent of the population size, N, and convergence is fast also in the second sense discussed in Ellison (1993).

Proposition 9 If i) $a \in [0, \frac{1}{N}]$ or ii) $a \in [1 - \frac{1}{N}, 1)$, $\max_{s \in S} W(s, \Theta, \varepsilon)$ is of order ε^{-1} as $\varepsilon \to 0$. Moreover, in the limit when a approaches zero or one, this result holds true when the population size subsequently grows to in¿nity.

Proof. The proof follows from the de¿nition of the resistances. If $\overline{K}_{M-1,M} = 1$ we know that the resistance of going upward from any other state is also one. This follows from the fact that the $\overline{K}$'s are non-decreasing and is proven formally in Lemma 8. Hence $e_M \in \Theta$ is reached with just one mutation. Correspondingly for case i). ■

6 Discussion

Theorem 7 demonstrates that in coordination games with simultaneous play and payo¾s determined by áverage strategies´ any of the strict Nash equilibria of the stage game can be stochastically stable. Which equilibria depend solely on the way the best response mapping transforms the average of other players' strategy pro¿le onto a discrete strategy. This implies that when di¾erent game structures are compared, one should be careful in ascribing di¾erences in the equilibrium selection to the game as such. What is crucial is how best responses remain piece-wise constant. If for instance, piece-wise constant best responses are de¿ned symmetrically the stochastically stable states are bounded away from the extreme strategies. This leads Robles (1997) to conclude that there is a stark contrast in equilibrium selection between coordination games with random pairing and games with a simultaneous play interaction structure. He reaches this conclusion because Kandori and Rob (1995) show that the stochastically stable state is Pareto eÁ cient in pure coordination games where players are randomly matched in pairs. However, this is not in contrast to simultaneous play, but merely a result of the speci¿c way Robles de¿nes piece-wise constant best responses. It should be noted, though, that as the number of players increases, a^M, i.e. the lower bound on $a \in [0, 1)$ that makes $e_M \in \Theta$, goes to one. This indicates that for a given de¿nition of piece-wise constant best responses, it becomes increasingly diÁ cult to coordinate on eÁ cient outcomes when the number of players is large. This result ¿ts intuition as well as much research (see e.g. KMR, Van Huyck, Battalio and Beil (1990, 1991) and Crawford (1995)).

A natural question arises though. Namely, how are individuals most likely to perceive $a \in [0, 1)$, and hence their best responses? Note that this question is not tantamount to asking how an experimenter would de¿ne the best response mapping. He can choose any integer value function to his liking (and hence determine payo¾s), but that does not imply a speci¿c behaviour of players. Their best replies depend on their perception of what is a best response to a given statistic. Unfortunately, we know of no experiments like the ones in Van Huyck, Battalio and Beil (1991) where the payo¾s (hence best replies) are determined by some averages, that could shed light on this issue. Intuitively, however, it is diÁ cult to understand why the best response mapping should be de¿ned and percieved as symmetric in pure coordination games. In this class of games, individual payo¾s are positive if the player match the current average, otherwise individual payo¾s are zero. Hence, all actions are equally secured (see footnote 2). The strategies are Pareto ranked, such that coordinating on a higher value of the summary statistic gives the players a strictly higher payoff than coordinating on a lower one. Thus why should a population-wide average of say, 4.48, induce a player to play strategy 4 instead of 5? In particular since playing 5 gives the player a higher payoff and in addition involves no greater risk than playing 4. Though the experiments in Van Huyck et al (1991) do not cover this case, some indication in favour of this argument can be found in the experiments concerning the median as the payoff relevant summary statistic. In one treatment they considered a case where all disequilibria outcomes give a payoff of zero (the period game Ω). This resembles our pure coordination game with an average payoff structure if the median is interpreted as a proxy for the average. In that experiment, they find that everyone playing their highest strategy is likely to be the equilibrium outcome. Thus agents may perceive a as close to one even though an experimenter has defined it differently.

It is also worth pointing out that pure coordination games are potential games, and Monderer and Shapley (1996) show that for potential games with an average payoff structure, the unique strategy profile that maximises the potential, is the Pareto efficient one. This too, clearly lends support to our claim that for an experimenter an asymmetric way of defining the best response mapping is not something that should be dismissed. In fact, it actually accords with theoretical results as well as empirical equilibrium observations.

For more general coordination games we also expect that if the payoffs the players get when missing the summary statistic differ for different strategies, both defining and perceiving the best response mapping symmetrically is highly unlikely to be a focal point.

References

[1] G. S. Becker, A note on restaurant pricing and other examples of social inÀ uences on price, J. Political Econ. 99, (1990), 1109-1116.

[2] K. Binmore, L. Samuelson and R. Vaughan, Musical chairs: Modelling noisy evolution, Games and Econ. Behavior, 11, (1995), 1-35.

[3] V. P. Crawford, An évolutionary´interpretation of Van Huyck, Battailo, and Beil's experimental results on coordination, Games and Econ. Behavior, 3, (1991) 25-59

[4] V. P. Crawford, Adaptive dynamics in coordination games, Econometrica 63 (1995), 103-144.

[5] V. P. Crawford, Theory and experiment in the analysis of strategic interaction, in Advances in Economics and Econometrics: Theory and Applications, Vol. I (D. M. Kreps and K. F. Wallis, Eds). New York: Cambrigde Univ. Press, 1997.

[6] G. Ellison, Learning, local interaction, and coordination, Econometrica, 61 (1993), 1047-1072.

[7] G. Ellison, Basins of attraction, long run equilibria, and the speed of step-by-step evolution, Mimeo (1996), MIT.

[8] D. Foster and P. Young, Stochastic evolutionary game dynamics, Theor. Population Biol. 38 (1990), 219-232.

[9] M. I. Freidlin and A. D. Wentzell, Ŕandom Perturbations of Dynamical Systems,´ Springer Verlag, New York, 1984.

[10] J.C. Harsanyi and R. Selten, 'A General Theory of Equilibrium Selection in Games,´ MIT Press, Cambrigde, MA, 1988.

[11] O. M. Kaarbxe, Interaction structure, long run equilibria, and the rate of convergence, Mimeo (1998), University of Bergen.

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

[13] M. Kandori and R. Rob, Evolution of equilibria in the long run: A general theory and applications, J. Econ. Theory 65 (1995), 383-414.

[14] J. Maynard Smith, 'Evolution and the Theory of Games,´Cambridge University Press, Cambridge, 1982.

[15] D. Monderer and L. S. Shapley, Potential games, Games Econ. Behav. 14 (1996), 124-143.

[16] J. Robles, Evolution and long run equilibria in coordination games with summary statistic payo¾ technologies, J. Econ. Theory 75 (1997), 180-193.

[17] T. C. Schelling, Hockey helmets, concealed weapons, and daylight saving: A study of binary choices with externalities, J. ConÀict Resolution 17 (1973), 381-428.

[18] W. S. Stewart, 'Introduction to the Numerical Solution of Markov Chains,´Princeton University Press, Princeton, New Jersey, 1994.

[19] J. Van Huyck, R. C. Battalio, and R. O. Beil, Tacit coordination games, strategic uncertainty, and coordination failure, American Econ. Review 80 (1990), 234-248.

[20] J. Van Huyck, R. C. Battalio, and R. O. Beil, Strategic uncertainty, equilibrium selection, and coordination failure in average opinion games, Quart. J. Econ. 106 (1991), 885-910.

[21] P. Young, The evolution of conventions, Econometrica 61 (1993), 57-84.

Appendix

Proof of Proposition 3. We shall show that the transition $m \to m+1$ for $1 \leq m < m+1 \leq M$ can happen after $\overline{K}_{m,m+1}(a)$ mutations and not fewer. The corresponding argument concerning any transition $m+1 \to m$ for $1 \leq m < m+1 \leq M$ is essentially identical and omitted.

First, note that after $\overline{K}_{m,m+1}(a)$ mutations to M, the best response for an m-player, who learns, is to play strategy $m+1$.(This follows trivially from the de¿nition of $\overline{K}_{m,m+1}(a)$). Call the state that results if, starting at e_m, $\overline{K}_{m,m+1}(a)$ players mutate to M, for s^1. Suppose $\lceil \overline{\mu}(s^1) \rceil_a = m+1$. Since it is assumed that at every $t = 1, 2, ...$ each player enjoys a strictly positive probability of learning, let all players revise their strategy choices. From the best-response dynamics it follows that all players adjust to strategy $m+1$ and e_{m+1} is reached with $\overline{K}_{m,m+1}(a)$ mutations. If $\lceil \overline{\mu}(s^1) \rceil > m+1$, then there exists mutations to $h < M$ such that $\lceil \overline{\mu}(s') \rceil = m+1$. (See footnote 9). Here s' is the state that results if, starting at e_m, $\overline{K}_{m,m+1}(a)$ players mutate to h. Again let all players learn. Since $B(\overline{\mu}(s')) = m+1$, they all adjust to e_{m+1}. Hence, e_m is reached with $\overline{K}_{m,m+1}(a)$ mutations.

We now show that $\overline{K}_{m,m+1}(a)-1$ mutations are not suÁcient to reach e_{m+1}. Let the state which results after $\overline{K}_{m,m+1}(a)-1$ mutations be s^2. From the de¿nition of $\overline{K}_{m,m+1}(a)$, it follows that the best response for an m-player, who learns, is to play m. Now, let an M-player receive the learning draw. Her best response is by de¿nition to play m as well. Call the resulting state after the M-player has played her best response for s^3. Since $\overline{\mu}(s^3) < \overline{\mu}(s^2)$, e_{m+1} is not reachable from e_m with $\overline{K}_{m,m+1}(a)-1$ mutations. ■

Calculation of the Stationary Distribution. The composition of myopic best responses and mutations generates an irreducible and aperiodic Markov chain over the ¿nite state space S. We now show how to compute the unique invariant distribution, $\mu(\varepsilon)$, $\varepsilon > 0$, for the games described in this paper.

To simplify the computation burden, we assume that each player $n \in \mathcal{N}$ enjoys the probability of revising her strategy choice with probability one, i.e. $\zeta = 1$. We refer to this as the deterministic best-response dynamics. It is called deterministic since every player switches to a best reply in every period.[10] Therefore, from any initial state s, the deterministic best-response dynamics implies a transition to the state e_m, where $m \in \mathcal{M}$ is the best reply to $\overline{\mu}(s)$. This transition happens before the mutation dynamics. The probability of the one-period transition $s = (i_1, ..., i_m, ..., i_M) \to s' = (i'_1, ..., i'_m, ..., i'_M)$, is then the probability of the transition $e_m \to s'$ via the mutation dynamics, where $B(\overline{\mu}(s)) = m$.

When a player mutates, we simply assume she chooses any strategy $m \in \mathcal{M}$ with a time-invariant positive probability which is distributed uniformly over all possible choices. Hence,

$$p_{ss'}(\varepsilon) = \sum_{\substack{0 \le i_m(0) \le i_m \\ \overline{i}_1(r_m)+...+\overline{i}_M(r_m)=r(m)=i_m-i_m(0) \\ i_m(0)+\sum\limits_{m'} \overline{i}_m(r_{m'})=i'_m}} \binom{i_m}{i_m(0)} \binom{r(m)}{\overline{i}_1(r(m)), ..., \overline{i}_M(r(m))} (1-\varepsilon)^{i_m(0)} \left(\frac{\varepsilon}{M}\right)^{r(m)}$$

where, $i_m(0)$ is the number of players playing strategy i who do not mutate, $r(m)$ is the

[10] The assumption is not crucial for the point emphasized in Table 4.1. First, each player's probability of revising her strategy choice can be chosen arbitrarily close to 1. Secondly, the least resistance paths are always constructed with transitions between adjacent recurrent sets. Hence, assuming all players learn every period does not change the number of mutations in the least resistance paths.

number of players who play m and mutate, $\binom{i_m}{i_m(0)}$ is the binomial coeÁ cient $\left(\frac{i_m!}{(i_m - i_m(0))! i_m!}\right)$ and $\binom{r(m)}{\bar{i}_1(r(m)),\ldots,\bar{i}_M(r(m))}$ is the multinomial coeÁ cient $\left(\frac{r(m)!}{\bar{i}_1(r(m))! \ldots \bar{i}_M(r(m))!}\right)$.

To illustrate the above formula, let $M = 3$ and $N = 5$. In this game there are 21 states. Now assume $a = \frac{1}{2}$, $s = (1, 4, 0)$ and $s' = (1, 2, 2)$. In state s, $\overline{\mu} = \frac{9}{5}$ and $\lceil \overline{\mu}(s) \rceil_{\frac{1}{2}} = 2$. Hence, $B(\overline{\mu}) = 2$, and e_2 is reached via the deterministic best-response dynamics. For a given value of ε, the probability of the transition $s \to s'$ is then the probability of the transition $e_2 \to s'$ by the mutation dynamics. Hence,

$$
\begin{aligned}
p_{ss'}(\varepsilon) &= \binom{5}{2}\binom{3}{1,0,2}(1-\varepsilon)^2\left(\frac{\varepsilon}{3}\right)^3 + \binom{5}{1}\binom{4}{1,1,2}(1-\varepsilon)\left(\frac{\varepsilon}{3}\right)^4 + \binom{5}{1,2,2}\left(\frac{\varepsilon}{3}\right)^5 \\
&= 30(1-\varepsilon)^2\left(\frac{\varepsilon}{3}\right)^3 + 60(1-\varepsilon)\left(\frac{\varepsilon}{3}\right)^4 + 30\left(\frac{\varepsilon}{3}\right)^5.
\end{aligned}
$$

When the transition matrix, $P(\varepsilon)$, is calculated, the stationary distribution $\mu(\varepsilon)$ is found by power iteration on $P(\varepsilon)$ until it converges. (See e.g. Stewart (1994) for a systematic and detailed treatment of the numerical solution of Markov chains.)