TI 2001-041/1
Tinbergen Institute Discussion Paper

Idiosyncratic and Aggregate Time-Varying Mutation Rates in Coordination Games

Harold Houba[1,2]
Alexander F. Tieman[1,3]

[1] Faculty of Economics and Business Administration, Vrije Universiteit Amsterdam, [2] Tinbergen Institute, [3] Econometric Research and Special Studies Department, De Nederlandsche Bank

Idiosyncratic and Aggregate Time-Varying Mutation Rates in Coordination Games*

Harold Houba† and Alexander F. Tieman‡

March 22, 2001

Abstract

We extend the standard evolutionary model of Kandori, Mailath, and Rob (1993) to incorporate time-varying aggregate and idiosyncratic shocks separately in coordination games. We show that both types of shocks have a different effect on the invariant distribution over the different equilibria of the game. While idiosyncratic shocks are shown to be neutral, the aggregate shocks introduce a systematic bias against the risk dominant equilibrium. Different from Kandori, Mailath, and Rob (1993) we derive a sufficient condition under which this bias prevents equilibrium selection with probability one.

*The authors thank Gerard van der Laan for comments and stimulating discussions.

†Free University and Tinbergen Institute. E-mail: hhouba@econ.vu.nl.

‡Corresponding author. Free University and De Nederlandsche Bank, Econometric Research and Special Studies Department, P.O. Box 98, 1000 AB Amsterdam, The Netherlands. E-mail: a.f.tieman@dnb.nl, Phone: +31-20-5242916, Fax: +31-20-5242916.

1. Introduction

Coordination failure on the macro-economic level in an economy, see e.g. Cooper (1999), or coordination on for instance the silver versus golden standard in a monetary economy, see e.g. Young (1998), can both be modelled as a coordination game. Coordination games typically have several Nash equilibria and standard game theory cannot discriminate between them. The evolutionary game theoretic models of Kandori, Mailath, and Rob (1993) (henceforth KMR) and Young (1993) provide a framework that does discriminate in coordination games and that singles out the so-called risk-dominant Nash equilibrium. However, contrary to the macro-economic literature KMR do not distinguish between aggregate and idiosyncratic shocks. Aggregate shocks capture the idea that some shocks directly affect populations as a whole, e.g. hard times such as a flooding, earthquake, or a recession. Idiosyncratic shocks capture individual differences among players that are too small to model explicitly. Here we introduce both types of shocks into the evolutionary framework of KMR. The aim of this paper is to show the different consequences of these shocks in coordination games.

In the evolutionary model of KMR each individual optimizes in a myopic way against the last round aggregate distribution of actions in the economy, which is observed by everyone, and each individual has a probability of making a mistake in implementing his optimum, which is called the mutation rate. In KMR, this rate is assumed to be the *same* for everyone and it is *constant* over time. The model reduces to an ergodic Markov process and, as the mutation rate vanishes, the unique distribution converges to the degenerate distribution in which everyone behaves according to the risk-dominant Nash equilibrium.

Below we assume time-varying individual mutation rates in 2×2 coordination games. To be precise, in each time period each agent has a mutation rate consisting of a common component for all agents, i.e. the aggregate shock, and an individual component reflecting the idiosyncratic shock. Below we characterize the invariant (probability) distribution over the basins of attraction associated with the two pure Nash equilibria. This invariant distribution is compared with the invariant distribution in the standard KMR model, where the fixed mutation rate is equal to the mean of the volatile process. Then both types of shocks have a different effect on the invariant distribution in the standard KMR model. While idiosyncratic shocks are shown to be neutral, the aggregate shocks introduce a systematic bias against the risk dominant equilibrium, i.e. the equilibrium with the largest basin. As a consequence, the expected fraction of time the system spends in the risk dominated equilibrium is lower than in the standard model of KMR. Furthermore, the introduction of individual time-varying mutation rates increases the volatility between the two basins of attraction in the dynamics and hence the coordination failure aggravates.

The limit of the invariant distribution is also investigated for vanishing mutation rates. First, it is shown that the limit probability of the risk dominant equilibrium is bounded away from one if the smallest upper bound upon the support of the probability distribution does not go to zero in taking this limit. This limit result differs from the limit result in KMR. Furthermore, in Bergin and Lipman (1996) a sufficient condition for equilibrium selection of any arbitrary strict Nash equilibrium is derived when mutation rates are state dependent. Our results differ in this respect, because the limit invariant distribution puts positive probability upon both pure Nash equilibria. Moreover, the limit probability of the risk dominant equilibrium is larger than the limit probability of the risk dominated equilibrium. Next, it is shown that letting the smallest upper bound of the support go to zero is an almost sufficient condition for equilibrium selection of the risk dominant equilibrium. So, taking the limit in case of time varying mutation rates matters. Finally, our results are valid for populations of all sizes, contrary to the large population results in KMR.

This paper is organized as follows. The next section introduces the model. The bias of aggregate shocks in the absence of idiosyncratic shocks is derived in section 3, whereas the model with both types of shocks is analyzed in section 4. The last section deals with equilibrium selection. The appendix contains the proofs of several lemmas used in the main theorems.

2. The Model

We consider a population labelled $\mathcal{N} = \{1, 2, \ldots, N\}$ of $N \geq 2$ players, with N even, who are all randomly matched in every time period in $\frac{1}{2}N$ pairs to play $\frac{1}{2}N$ identical (fixed) coordination games, each of which is given by

1\2	s_1	s_2
s_1	a, a	$0, 0$
s_2	$0, 0$	$1, 1$

with $a > 1$. The off-diagonal elements are normalized to 0. The pure Nash equilibrium (s_1, s_1) is the risk-dominant equilibrium, see e.g. Harsanyi and Selten (1988). Time is discrete and denoted by $t \in \mathbb{N}$. By $z_t \in Z$, $Z = \{0, 1, \ldots, N\}$, we denote the number of players adopting action s_1 at time t. Thus $z_t = N$ means that the risk dominant equilibrium (s_1, s_1) is played in every one of the $\frac{1}{2}N$ pairs at time t, while $z_t = 0$ yields the risk dominated equilibrium (s_2, s_2) in every pair. We denote the states $z_t = N$ and $z_t = 0$ by E_1 and E_2 respectively. By $z^* = \frac{N+a-1}{a+1} < \frac{1}{2}N$ we denote the critical level of the population for which an arbitrary player is indifferent between s_1 and s_2 (see also section 5 of KMR). Generically z^* will not be integer and we denote $[z^*]$ as the largest integer smaller or equal to z^*, i.e. the Entier function of z^*. We assume that, at the beginning of period $t+1$, all players observe z_t and play a best response to z_t which induces the best-reply deterministic dynamics given by

$$z_{t+1} = BR(z_t) = \begin{cases} 0, & \text{if } z_t \leq [z^*], \\ N, & \text{if } z_t \geq [z^*] + 1. \end{cases}$$

The basin of attraction of E_1 is equal to $B_1 = \{[z^*] + 1, \ldots, N\}$, whereas $B_2 = Z \backslash B_1 = \{0, \ldots, [z^*]\}$ is the basin of attraction of E_2.

We add a stochastic component to the dynamics representing random mutation by the players. A player who mutates plays an action selected at random from $\{s_1, s_2\}$ with probability $\frac{1}{2}$ on either action. Thus with a probability half as large as the mutation probability, a mutant player plays the action s_i which is not a best response to z_t. Similar as in KMR we assume that if $z_t \in B_i$, $i = 1, 2$, then in period $t+1$ the best reply dynamics first take the system from the state z_t to the state E_i and, subsequently, the stochastic process determines the number of mutants. Thus, if $z_t \in B_2$, then the number of mutants from E_2 at period $t+1$ is reflected by z_{t+1}. Similarly, if $z_t \in B_1$, then $N - z_{t+1}$ states the number of mutants from state E_1 at period $t+1$. In case the number of mutants exceeds a certain threshold the state moves from the current basin of attraction into the other. Since in order to switch equilibrium, mutations have to offset either E_1 or E_2, it is without loss of generality to reduce the N-state Markov chain to a two-state Markov chain on the level of the basins of attraction B_1 and B_2. The invariant distribution on Z can be easily derived from the invariant distribution of the reduced Markov chain.

The innovative feature of our model lies in the introduction of an individual mutation rate $2\zeta_{i,t}$ for player i at time t that is stochastic or volatile over time. The mutation rate $\zeta_{i,t}$ is thus the individual probability of not playing a best reply at time t.[1] At every time t we assume that $\zeta_{i,t} = \varepsilon_t + \varepsilon_{i,t}$, where the common term ε_t represents an aggregate shock on the population as a whole and $\varepsilon_{i,t}$ denotes an idiosyncratic shock due to small individual differences between players that are not explicitly modelled. Define the $(N+1)$-dimensional stochastic variable θ_t as $(\varepsilon_{1,t}, \ldots, \varepsilon_{N,t}, \varepsilon_t) \in \mathbb{R}^{N+1}$. We assume that *i)* θ_t is i.i.d. over time, *ii)* at each time t, all $\varepsilon_{i,t}$ are i.i.d. and ε_t is independent of all $\varepsilon_{i,t}$, *iii)* $\zeta_{i,t} \in (0, \delta)$ for some $\delta \in (0,1)$ and *iv)* $\mathbb{E}\varepsilon_{i,t} = 0$. Note that the latter corresponds to $\varepsilon_{i,t}$ being drawn from a (possibly asymmetric) fair probability distribution that includes the special case of white noise (i.e. a symmetric fair distribution). So, $\mathbb{E}\zeta_{i,t} = \mathbb{E}\varepsilon_t \in (0, \delta)$ and all $\zeta_{i,t}$ can be regarded as (asymmetric) noise around a volatile aggregate mutation rate ε_t. Two interesting special cases involving degenerate probability distributions are *i)* $\varepsilon_t = \varepsilon > 0$ for all t, i.e. a constant aggregate mutation rate with idiosyncratic variation around ε, and *ii)* $\varepsilon_{i,t} = 0$ for all i and t, i.e. no agent-specific shocks. The model of KMR corresponds to $\zeta_{i,t} = \varepsilon > 0$ for all i and t, i.e. in KMR both special cases hold simultaneously.

For explanatory reasons let us first assume that $\zeta_{i,t}$ denotes a realisation. Then the proba-

[1] Note that from this point on we use the term mutation probability or rate to mean the probability or rate at which a player plays a non-best reponse. As explained above, the actual mutation probability or rate is twice as large.

bility that all players in a subset $\mathcal{M} \subset \mathcal{N}$ mutate and all other players in the population don't is given by

$$\prod_{i\in\mathcal{M}} \zeta_{i,t} \cdot \prod_{i\in\mathcal{N}\setminus\mathcal{M}} \left(1-\zeta_{i,t}\right).$$

Given $z_t \in B_1$ the system first moves from the state z_t to the state E_1 and, subsequently, at least $N - [z^*]$ mutations are needed to take it away from B_1 into B_2. Then to obtain the probability of going from B_1 to B_2 conditional on $z_t \in B_1$ we have to sum up over all groups M with at least $N - [z^*]$ players. This probability is given by

$$p\left(\theta_t\right) = \sum_{M:|M|\geq N-[z^*]} \prod_{i\in M} \zeta_{i,t} \cdot \prod_{i\in\{1,\dots,N\}\setminus M} \left(1-\zeta_{i,t}\right),$$

where $|M|$ denotes the cardinality of M. Similar, the probability of going from B_2 to B_1, i.e. $z_{t+1} > z^*$, is given by

$$q\left(\theta_t\right) = \sum_{M:|M|\geq [z^*]+1} \prod_{i\in M} \zeta_{i,t} \cdot \prod_{i\in\{1,\dots,N\}\setminus M} \left(1-\zeta_{i,t}\right),$$

because the transition from E_2 to E_1 takes at least $[z^*]+1$ players that mutate. Note that taking $\theta_t = (0,\dots,0,\varepsilon)$ yields the transition probabilities in KMR, i.e.

$$p\left(\theta_t\right) = \sum_{j=N-[z^*]}^{N} \binom{N}{j} \varepsilon^j \left(1-\varepsilon\right)^{N-j} \quad \text{and} \quad q\left(\theta_t\right) = \sum_{j=[z^*]+1}^{N} \binom{N}{j} \varepsilon^j \left(1-\varepsilon\right)^{N-j}. \tag{2.1}$$

The transition matrix $P\left(\theta_t\right)$ of the two-state Markov chain on B_1 and B_2 is given by

$$P\left(\theta_t\right) = \begin{pmatrix} 1-p\left(\theta_t\right) & p\left(\theta_t\right) \\ q\left(\theta_t\right) & 1-q\left(\theta_t\right) \end{pmatrix},$$

where the (i,j)-th entry denotes the probability of the transition $B_i \to B_j$. The associated unique invariant distribution, denoted by $\mu\left(P(\theta_t)\right) = \left(\mu_1\left(P\left(\theta_t\right)\right), \mu_2\left(P\left(\theta_t\right)\right)\right)$, is given by

$$\mu_1\left(P(\theta_t)\right) = \frac{q\left(\theta_t\right)}{p\left(\theta_t\right)+q\left(\theta_t\right)} \text{ and } \mu_2\left(P(\theta_t)\right) = 1-\mu_1\left(P(\theta_t)\right) = \frac{p\left(\theta_t\right)}{p\left(\theta_t\right)+q\left(\theta_t\right)},$$

where μ_i, $i = 1, 2$, is the probability on the state $z_t \in B_i$. The expected length of time spent in B_1 (B_2) conditional on $z_t \in B_1$ (B_2) is given by $\frac{1}{p(\theta_t)}$ ($\frac{1}{q(\theta_t)}$). KMR show that $\lim_{\varepsilon\to 0} \mu_1\left(P(\varepsilon)\right) = 1$ for the case $\theta_t = (0,\dots,0,\varepsilon)$.

The case of volatile individual mutation rates can be handled as follows, see e.g. Bellman (1954). In every period, the dynamics consist of a *compound* probability distribution. At each period t, first θ_t is drawn to determine each player's individual mutation rate $\zeta_{i,t}$ for period t, and next, these N mutation rates $\zeta_{i,t}$ determine the transition probabilities of the Markov chain. Doing so yields a homogeneous Markov chain in which the conditional probabilities $\Pr(B_i \to B_j | B_i)$, $i, j = 1, 2$, of going from B_i into B_j conditional on $z_t \in B_i$ are given by

$$\Pr(B_1 \to B_2 | B_1) = \mathbb{E}\ p(\theta_t) \quad \text{and} \quad \Pr(B_2 \to B_1 | B_2) = \mathbb{E}\ q(\theta_t).$$

Thus, the two stochastic events can be compounded into a single transition matrix $\mathbb{E}P(\theta_t)$, $\theta_t \sim \Theta$, where

$$\mathbb{E}P(\theta_t) = \begin{pmatrix} 1 - \mathbb{E}p(\theta_t) & \mathbb{E}p(\theta_t) \\ \mathbb{E}q(\theta_t) & 1 - \mathbb{E}q(\theta_t) \end{pmatrix},$$

and $\mathbb{E}(.)$ is the expectation operator. The associated unique invariant distribution is denoted by $\mu(\mathbb{E}P(\theta_t)) = (\mu_1(\mathbb{E}P(\theta_t)), \mu_2(\mathbb{E}P(\theta_t)))$. Finally, the aim of the analysis below is to compare the invariant distribution $\mu(\mathbb{E}P(\theta_t))$ of the volatile process with the invariant distribution $\mu(P(\mathbb{E}\theta_t))$, where we interpret the latter as the invariant distribution in the standard KMR model under the assumption $\theta_t = (0, \ldots, 0, \mathbb{E}\varepsilon_t)$, i.e. the KMR model with $\varepsilon = \mathbb{E}\varepsilon_t$. To put it differently, we investigate the effect on the invariant distribution of introducing a volatile θ_t around $(0, \ldots, 0, \varepsilon)$ in the standard KMR model.

3. Aggregate shocks

In this section we restrict the analysis to the special case of volatile aggregate shocks and abstract from idiosyncratic shocks. Formally, we take $\theta_t = (0, \ldots, 0, \varepsilon_t)$, where ε_t is i.i.d. over time. For notational convenience we treat the vector θ_t as a real number and drop its time index t.

The derivation of the systematic bias in the invariant distribution is derived in two steps. First, we show that aggregate shocks have a systematic bias upon the invariant distribution if we compare $\mu(\mathbb{E}P(\theta))$ with $\mu(P(\mathbb{E}\theta))$. This result holds locally, meaning that it is derived for all probability distributions with the property that the support of θ lies within close distance of its mean. Then, second, we relax the first result by showing that close distance means a support restricted to $\left[0,\overline{\theta}\right]$ for some upperbound $\overline{\theta}>0$, for which we derive explicit expressions.

Our first result states a sufficient condition under which $\mu_2(\mathbb{E}P(\theta))$ is bounded from below by $\mu_2(P(\mathbb{E}\theta))$.

Theorem 3.1. *Let* $\overline{\theta}=\min\left\{\frac{1}{2}\frac{N}{N-1},\frac{N-[z^*]}{N}-\frac{\sqrt{[z^*](N-[z^*])(N-1)}}{N(N-1)}\right\}$. *If* $\mathbb{E}\theta<\overline{\theta}$, *then there exists a* $\gamma>0$ *such that for every probability distribution with support in the interval* $(\mathbb{E}\theta-\gamma,\mathbb{E}\theta+\gamma)$ *it holds that* $\mu_2(P(\mathbb{E}\theta))<\mu_2(\mathbb{E}P(\theta))$. *Moreover,* $\overline{\theta}>0$.

Proof

We have to show that $\mu_2\left(\mathbb{E}P\left(\theta\right)\right)-\mu_2\left(P\left(\mathbb{E}\theta\right)\right)>0$, which is equivalent to

$$\frac{\mathbb{E}p\left(\theta\right)}{\mathbb{E}p\left(\theta\right)+\mathbb{E}q\left(\theta\right)}-\frac{p\left(\mathbb{E}\theta\right)}{p\left(\mathbb{E}\theta\right)+q\left(\mathbb{E}\theta\right)}>0\Leftrightarrow\mathbb{E}\left[p\left(\theta\right)q\left(\mathbb{E}\theta\right)-p\left(\mathbb{E}\theta\right)q\left(\theta\right)\right]>0.$$

Define the *asymmetric* $N\times N$ matrix $A\left(\theta\right)$ by its (ℓ,m)-th element $a_{\ell,m}\left(\theta\right)$

$$a_{\ell,m}\left(\theta\right)=\binom{N}{\ell}\binom{N}{m}\left(\mathbb{E}\theta\right)^{\ell}\left(1-\mathbb{E}\theta\right)^{N-\ell}\left(\theta\right)^{m}\left(1-\theta\right)^{N-m},\ l,m=1,\ldots N.$$

Then we have that $p\left(\theta\right)q\left(\mathbb{E}\theta\right)$ is equal to

$$\left[\sum_{k=N-[z^*]}^{N}\binom{N}{k}\left(\theta\right)^{k}\left(1-\theta\right)^{N-k}\right]\left[\sum_{j=[z^*]+1}^{N}\binom{N}{j}\left(\mathbb{E}\theta\right)^{j}\left(1-\mathbb{E}\theta\right)^{N-j}\right]=\sum_{j=[z^*]+1}^{N}\sum_{k=N-[z^*]}^{N}a_{j,k}\left(\theta\right)$$

and $p\left(\mathbb{E}\theta\right)q\left(\theta\right)$ is equal to

$$\left[\sum_{k=N-[z^*]}^{N}\binom{N}{k}\left(\mathbb{E}\theta\right)^{k}\left(1-\mathbb{E}\theta\right)^{N-k}\right]\left[\sum_{j=[z^*]+1}^{N}\binom{N}{j}\left(\theta\right)^{j}\left(1-\theta\right)^{N-j}\right]=\sum_{j=[z^*]+1}^{N}\sum_{k=N-[z^*]}^{N}a_{k,j}\left(\theta\right).$$

This means that $p\left(\theta\right)q\left(\mathbb{E}\theta\right)$ sums over all the elements in the $([z^*]+1)\times(N-[z^*])$ submatrix in the lower-right corner of $A\left(\theta\right)$ while $p\left(\mathbb{E}\theta\right)q\left(\theta\right)$ sums over all the elements in the $(N-[z^*])\times([z^*]+1)$ submatrix in the lower-right corner of $A\left(\theta\right)$. So, $p\left(\theta\right)q\left(\mathbb{E}\theta\right)$ and $p\left(\mathbb{E}\theta\right)q\left(\theta\right)$ share the terms in the $([z^*]+1)\times([z^*]+1)$ submatrix in the lower-right corner of $A\left(\theta\right)$, which cancel. Formally,

$$p\left(\theta\right)q\left(\mathbb{E}\theta\right)-p\left(\mathbb{E}\theta\right)q\left(\theta\right)=\sum_{j=[z^*]+1}^{N-[z^*]}\sum_{k=N-[z^*]}^{N}\left[a_{j,k}\left(\theta\right)-a_{k,j}\left(\theta\right)\right].\tag{3.1}$$

In Figure 3.1 we provide a picture which shows the general form of expression (3.1). Note that $a_{j,k}(\theta) - a_{k,j}(\theta) = 0$ for all $\theta \in [0,1]$ if $k = j = N - [z^*]$ and that for all other combinations of j and k we have $j < k$. In order to determine the second derivative of expression (3.1), we first determine the second derivative of the individual terms of the double sum. With respect to this derivative, note that

$$\frac{\partial^2}{\partial\theta^2}(\theta)^m(1-\theta)^{N-m} = \frac{m(m-1)+\theta(\theta N-2m)(N-1)}{\theta^2(1-\theta)^2}(\theta)^m(1-\theta)^{N-m}, \quad m = 1,\ldots,N,$$

and, therefore,

$$\frac{\partial^2}{\partial\theta^2}a_{\ell,m}(\theta) = \frac{m(m-1)+\theta(\theta N-2m)(N-1)}{\theta^2(1-\theta)^2}a_{\ell,m}(\theta), \quad m = 1,\ldots,N.$$

Denote $c_m(\theta) = m(m-1)+\theta(\theta N-2m)(N-1)$. Substitution of this result yields the second derivative of (3.1)

$$\frac{\partial^2}{\partial\theta^2}[p(\theta)q(\mathbb{E}\theta) - p(\mathbb{E}\theta)q(\theta)] = \sum_{j=[z^*]+1}^{N-[z^*]}\sum_{k=N-[z^*]}^{N}\frac{[c_k(\theta)a_{j,k}(\theta) - c_j(\theta)a_{k,j}(\theta)]}{\theta^2(1-\theta)^2}. \quad (3.2)$$

From Lemma A.2 it follows that $c_k(\theta)a_{j,k}(\theta) - c_j(\theta)a_{k,j}(\theta) > 0$ in $\theta = \mathbb{E}\theta$ for $j = [z^*]+1,\ldots,N-[z^*]$, $k = N-[z^*],\ldots,N$ and $j < k$ provided $\mathbb{E}\theta < \overline{\theta}$. So, for all $j < k$ the functions $a_{j,k}(\theta) - a_{k,j}(\theta)$ are strictly convex around $\theta = \mathbb{E}\theta$ for all $j = [z^*]+1,\ldots,N-[z^*]$, $k = N-[z^*],\ldots,N$ and $j < k$. Thus (3.2) is positive. Continuity of the second derivative implies that it is also positive locally around $\theta = \mathbb{E}\theta$. Finally, applying Jensen's inequality locally around $\theta = \mathbb{E}\theta$ to the convex function $f(\theta) = p(\theta)q(\mathbb{E}\theta) - p(\mathbb{E}\theta)q(\theta)$ results in $\mathbb{E}[f(\theta)] > f(\mathbb{E}\theta) = 0$, which concludes the proof of Theorem 3.1. □

Theorem 3.1 states that fluctuations of θ around the mean $\mathbb{E}\theta$ introduce a systematic bias in the invariant distribution in favor of the risk dominated equilibrium E_2 and against the risk dominant equilibrium E_1. This can be explained as follows. First of all, the fluctuations in the

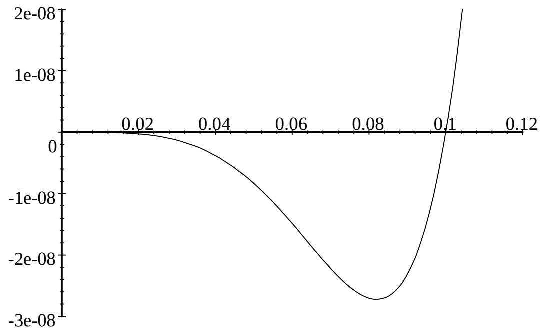

Figure 3.1: Expression (3.1) for $\theta \in [0, 0.12]$ with parameters $N = 10$, $z^* = 3$ and $\mathbb{E}\theta = 0.1$.

mutation rates increase the probability of a switch to the other basin of attraction, because[2]

$$\mathbb{E}p(\theta) > p(\mathbb{E}\theta) \quad \text{and} \quad \mathbb{E}q(\theta) > q(\mathbb{E}\theta).$$

So, random aggregate shocks in the mutation rates increase the volatility in the dynamical system and the coordination failure aggravates. Second, the result of Theorem 3.1 implies that the probability $\mathbb{E}p(\theta)$ of a switch from E_1 to E_2 increases relatively faster than the probability $\mathbb{E}q(\theta)$ of a switch from E_2 to E_1. So, it becomes relatively more easier to leave E_1 than E_2 and, therefore, the fraction of time the stochastic process spends in E_1 decreases due to these fluctuations.

The sufficient condition of Theorem 3.1 allows every probability distribution with its mean $\mathbb{E}\theta$ smaller than the positive threshold $\overline{\theta}$. Since $[z^*] \approx \frac{N+a-1}{a+1}$ this threshold is a function of the parameters N and a. Then it is easy to verify that for $\frac{1}{2}\frac{N}{N-1}$ to be the minimum in the

[2]This follows immediately from the expression $\frac{\partial^2}{\partial\theta^2}\theta^m (1-\theta)^{N-m}$ in the proof of theorem 3.1, because

$$\frac{\partial^2}{\partial\theta^2}\theta^m (1-\theta)^{N-m} > 0 \Leftrightarrow \theta \notin \left[\frac{m}{N} - \frac{\sqrt{m(N-m)(N-1)}}{N(N-1)}, \frac{m}{N} + \frac{\sqrt{m(N-m)(N-1)}}{N(N-1)}\right],$$

and the lower bound is positive if $m \geq 2$. Since $N - [z^*] > [z^*] + 1 \geq 2$ this holds for all the individual terms in $p(\theta)$ and $q(\theta)$, provided θ sufficiently close to 0. Then application of Jensen's inequality yields the stated results.

definition of $\overline{\theta}$ corresponds to

$$a \geq \frac{2N^2 - 4N + 4 + 4\sqrt{(N-1)\left(N - \frac{1}{2} + \frac{1}{2}\sqrt{5}\right)\left(N - \frac{1}{2}\sqrt{5} - \frac{1}{2}\right)}}{2\left(N - 3 + \sqrt{5}\right)\left(N - 3 - \sqrt{5}\right)}, \tag{3.3}$$

provided $N \geq \max\left\{\frac{1}{2}\sqrt{5} + \frac{1}{2}, 3 + \sqrt{5}\right\} \approx 5.2$. Figure 3.2 represents the latter area in the (N, a)-plane for $N \geq 6$ and the curve is asymptotic to $a = 1$ as N goes to infinity. So, for every $a > 1$ there exists a $N(a) < \infty$ such that $\overline{\theta} = \frac{1}{2}\frac{N}{N-1} > \frac{1}{2}$ for every $N \geq N(a)$. In that case $\overline{\theta} > \frac{1}{2}$ is quite large. Nevertheless, Theorem 3.1 also holds for small population size N, which is at the expense of $\overline{\theta}$.

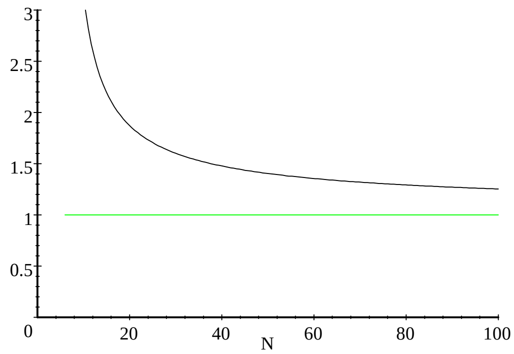

Figure 3.2: The area of inequality (3.2) represented in the (N, a)-plane, $a > 1$. The area on and above the downward sloping curve features $\overline{\theta} = \frac{1}{2}\frac{N}{N-1}$.

Theorem 3.1 states that adding small volatility of the mutation rates to the KMR model is sufficient to obtain a systematic bias. The next theorem states a sufficient condition under which the entire interval $\left[0, \overline{\theta}\right]$ can be regarded as an upper bound upon the support of the volatile mutations. This sufficient condition is more restrictive than the one in Theorem 3.1.

Theorem 3.2. *Let* $\widetilde{\theta} = \min\left\{\frac{1}{2}\frac{N}{N-1}, \frac{[z^*]+1}{N} - \frac{\sqrt{([z^*]+1)(N-[z^*]-1)(N-1)}}{N(N-1)}\right\}$. *If* $\mathbb{E}\theta \leq \widetilde{\theta}$, *then for each probability distribution with support in* $\left[0, \overline{\theta}\right]$ *it holds that* $\mu_2(P(\mathbb{E}\theta)) < \mu_2(\mathbb{E}P(\theta))$. *Moreover,* $0 < \widetilde{\theta} \leq \overline{\theta}$.

Proof

Lemma A.4 states that there exists a unique inflection point $\widehat{\theta} \in (0, \mathbb{E}\theta)$ if $\mathbb{E}\theta \leq \widetilde{\theta}$. Therefore, each function $a_{j,k}(\theta) - a_{k,j}(\theta)$ in (3.2) is strictly concave on $\left(0, \widehat{\theta}\right)$ and strictly convex on $\theta \in \left[\mathbb{E}\theta, \overline{\theta}\right)$ (recall also Lemma A.2). We introduce the following convex approximation for $a_{j,k}(\theta) - a_{k,j}(\theta)$. First, denote the function $r(\theta)$ as the ray (or line) that starts in the origin $(0,0)$ and that is tangent to the function $a_{j,k}(\theta) - a_{k,j}(\theta)$ in some point $\theta^* \in \left(\widehat{\theta}, \mathbb{E}\theta\right)$. The uniqueness of the inflection point $\widehat{\theta}$ guarantees that θ^* is uniquely determined. Moreover, $r(\theta) \leq a_{j,k}(\theta) - a_{k,j}(\theta)$ for all $\theta \in [0, \theta^*]$. Now, define the convex function $f^*_{j,k} : \left[0, \widetilde{\theta}\right] \to \mathbb{R}$ as

$$f^*_{j,k}(\theta) = \begin{cases} r(\theta), & \text{if } \theta \in [0, \theta^*), \\ a_{j,k}(\theta) - a_{k,j}(\theta), & \text{if } \theta \in \left[\theta^*, \widetilde{\theta}\right]. \end{cases}$$

Then $f^*_{j,k}$ is a convex approximation of $a_{j,k}(\theta) - a_{k,j}(\theta)$ on $\left[0, \widetilde{\theta}\right]$ and $f^*_{j,k}(\theta) \leq a_{j,k}(\theta) - a_{k,j}(\theta)$ for all $\theta \in \left[0, \widetilde{\theta}\right]$. Furthermore, $f^*_{j,k}(\mathbb{E}\theta) = a_{j,k}(\mathbb{E}\theta) - a_{k,j}(\mathbb{E}\theta) = 0$. But then

$$\begin{aligned} \mathbb{E} \sum_{j=[z^*]+1}^{N-[z^*]} \sum_{k=N-[z^*]}^{N} \left[a_{j,k}(\theta) - a_{k,j}(\theta)\right] &\geq \sum_{j=[z^*]+1}^{N-[z^*]} \sum_{k=N-[z^*]}^{N} \mathbb{E} f^*_{j,k}(\theta) \\ &> \sum_{j=[z^*]+1}^{N-[z^*]} \sum_{k=N-[z^*]}^{N} f^*_{j,k}(\mathbb{E}\theta) = 0. \end{aligned}$$

where the last inequality is Jensen's inequality. □

Similar as before, the sufficient condition of Theorem 3.2 allows every probability distribution on $\left[0, \overline{\theta}\right]$ provided $\mathbb{E}\theta < \widetilde{\theta}$. The threshold $\widetilde{\theta}$ is also a function of the parameters N and a through $[z^*] \approx \frac{N+a-1}{a+1}$. Somewhat unexpected we obtain that $\widetilde{\theta} \geq \frac{1}{2}\frac{N}{N-1}$ in the (N, a)-plane coincides with the condition stated in (3.3). This may seem strange, but recall that $\widetilde{\theta} = \frac{1}{2}\frac{N}{N-1}$ implies $\overline{\theta} = \frac{1}{2}\frac{N}{N-1}$. As before, the mild restriction $\widetilde{\theta} = \overline{\theta} = \frac{1}{2}\frac{N}{N-1} > \frac{1}{2}$ can be achieved for every $a > 1$ provided its population size N exceeds a certain minimum but finite threshold.

Finally, Theorem 3.1 and Theorem 3.2, respectively, are derived by requiring that *each* individual term under the double summation sign of (3.2) is locally convex and has a unique inflection point on $(0, \mathbb{E}\theta)$. This suggests that the upper bounds $\overline{\theta}$ and $\widetilde{\theta}$ upon the mean can be relaxed by considering the double sum of (3.2) as one function.

4. Aggregate and idiosyncratic shocks

In this section both types of shocks will be analyzed, i.e. $\theta_t = (\varepsilon_{1,t}, \ldots, \varepsilon_{N,t}, \varepsilon_t)$. The main result in the next theorem states that the invariant distribution of the general model coincides with the invariant distribution of the model with aggregate shocks only.

Theorem 4.1. *For each probability distributions of θ_t it holds that $\mu(\mathbb{E}P(\theta_t)) = \mu(\mathbb{E}P((0, \ldots, 0, \varepsilon_t)))$.*

Proof

We will first show that $\mathbb{E}\mathrm{p}(\theta_t) = \mathbb{E}\sum_{j=N-[z^*]}^{N} \binom{N}{j} \varepsilon_t^j (1-\varepsilon_t)^{N-j}$. The assumptions on $\zeta_{i,t}$ imply that $\mathbb{E}\zeta_{i,t} = \mathbb{E}\varepsilon_t$. Furthermore, $\prod_{i=1}^{k} \zeta_{i,t} = \prod_{i=1}^{k} (\varepsilon_t + \varepsilon_{i,t})$ is a polynomial with cross terms containing powers of ε_t and products of $\varepsilon_{i,t}$, $i = 1, \ldots, k$. Since all $\varepsilon_{i,t}$ are i.i.d. with $\mathbb{E}\varepsilon_{i,t} = 0$, also $\mathbb{E}(\varepsilon_{i_1,t} \cdot \varepsilon_{i_2,t} \cdot \ldots \cdot \varepsilon_{i_l,t}) = 0$, $i_k \in \{1, \ldots, N\}$, $k \in \{1, \ldots, N\}$. Combined with the independence between all $\varepsilon_{i,t}$ and ε_t, we have that $\mathbb{E}\left(\varepsilon_t^j \cdot \varepsilon_{i_1,t} \cdot \varepsilon_{i_2,t} \cdot \ldots \cdot \varepsilon_{i_{k-j},t}\right) = 0$, with $i_l \in \{1, \ldots, N\}$, $l = 1, \ldots, k-j$. Thus, $\mathbb{E}\left(\prod_{i=1}^{k} \zeta_{i,t}\right) = \varepsilon_t^k$, $k = 1, \ldots, N$. This gives us $\mathbb{E}\left\{\prod_{i \in M} \zeta_{i,t} \cdot \prod_{i \in \{1,\ldots,N\} \setminus M} (1-\zeta_{i,t})\right\} = \mathbb{E}\varepsilon_t^{|M|}(1-\varepsilon_t)^{N-|M|}$. Using this yields

$$\begin{aligned}\mathbb{E} \sum_{M:|M| \geq N-[z^*]} \left\{ \prod_{i \in M} \zeta_{i,t} \cdot \prod_{i \in \{1,\ldots,N\} \setminus M} (1-\zeta_{i,t}) \right\} &= \sum_{M:|M| \geq N-[z^*]} \mathbb{E}\left\{\varepsilon_t^{|M|} (1-\varepsilon_t)^{N-|M|}\right\} \\ &= \mathbb{E} \sum_{j=N-[z^*]}^{N} \binom{N}{j} \varepsilon_t^j (1-\varepsilon_t)^{N-j}.\end{aligned}$$

Similar, $\mathbb{E}q(\theta_t) = \mathbb{E}\sum_{j=[z^*]+1}^{N} \binom{N}{j} \varepsilon_t^j (1-\varepsilon_t)^{N-j}$. Then the stated result immediately follows.□

Theorem 4.1 implies that the idiosyncratic shocks can be neglected when studying the effects of varying mutation rates. So, the effect of idiosyncratic shocks on the invariant distribution is neutral. This theorem also allows for a reinterpretation of the standard model in KMR, namely idiosyncratic shocks with expectation 0 around ε, i.e. formally $\zeta_{i,t} = \varepsilon_{i,t}$ with $E[\varepsilon_{i,t}] = \varepsilon$ for all $i = 1, \ldots, N$.

5. Limit results

The main result in KMR translated to time-varying mutation rates can be formulated as $\mu_2 (P (\mathbb{E}\theta_t)) \to 0$ as $\mathbb{E}\theta_t \to 0$. In this section it is shown that the limit of $\mu_2 (\mathbb{E}P (\theta_t))$ as $\mathbb{E}\theta_t \to 0$ depends upon how we treat the support and the distribution of probabilities over its support in the limit as the mean vanishes. To show this we treat $\theta_t = (0, \dots , 0, \varepsilon_t)$ as a scalar and drop the time index t.

Before we state the main results, we consider the following example, which nicely illustrates the results derived below.

Example 5.1. *Suppose* $\theta \in \{0, \gamma\}$ *for some* $\gamma > 0$ *and* $\Pr\{\theta = \gamma\} = \rho > 0$. *Then* $\mathbb{E}\theta = \rho\gamma$ *and*

$$\mu_2 (\mathbb{E}P (\theta)) = \frac{p (\gamma)}{p (\gamma) + q (\gamma)} > 0$$

is independent of $\rho > 0$. *Obviously,*

$$\lim_{\rho \to 0} \mu_2 (\mathbb{E}P (\theta)) > 0 \tag{5.1}$$

and, by standard arguments from KMR,

$$\lim_{\gamma \to 0} \mu_2 (\mathbb{E}P (\theta)) = 0,$$

whereas both $\lim_{\rho \to 0} \mathbb{E}\theta = \lim_{\gamma \to 0} \mathbb{E}\theta = 0$. *Moreover, the limit in (5.1) depends upon* γ. *Furthermore, suppose that the parameters* ρ *and* γ *of the probability distribution of* θ *can be represented as functions of some single parameter* $\lambda \in [0, 1]$. *Formally, let both* $\rho = \rho(\lambda) : [0, 1] \to [0, 1]$ *and* $\gamma = \gamma(\lambda) : [0, 1] \to [0, 1]$ *be continuous functions of* λ. *Then for all functions* $\rho(\lambda)$ *and* $\gamma(\lambda)$ *with* $\gamma(0) = 0$ *it holds that* $\lim_{\lambda \to 0} \mu_2 (\mathbb{E}P (\theta(\lambda))) = 0$, *whereas for all functions* $\rho(\lambda)$ *and* $\gamma(\lambda)$ *with* $\rho(0) = 0$ *and* $\gamma(0) > 0$ *it holds that* $\lim_{\lambda \to 0} \mu_2 (\mathbb{E}P (\theta(\lambda))) > 0$. *Note that* $\lim_{\lambda \to 0} \mu_2 (\mathbb{E}P (\theta(\lambda))) < \frac{1}{2}$, *because (2.1) implies that* $q(\gamma(\lambda)) > p(\gamma(\lambda))$.

This example shows that *the way in which the limit is taken matters for the convergence results.* Note that $\mathbb{E}\theta \to 0$ can be accomplished in two different ways. Either the support is fixed and the probability distribution converges to a degenerate probability distribution over the support that is degenerated in the point 0, i.e. $\{0, \gamma(\lambda)\}$ fixed and $\rho(\lambda) \to 0$ in the example, or the support collapses into a single point in the limit 0 without restrictions upon the distributions over the support, i.e. $\{0, \gamma(\lambda)\} \to \{0\}$ and no restrictions upon $\rho(\lambda)$ in the example. (Or a combination of both.) Moreover, the limit $\lim_{\lambda \to 0} \mu_2 (\mathbb{E}P(\theta(\lambda)))$ depends upon $\gamma(0)$ but is always bounded from above by $\frac{1}{2}$.

In order to make things precise we introduce the following notation. Let $\lambda \in [0,1]$. Since every probability distribution can be approximated arbitrarily close by some discrete probability distribution we restrict attention to such distributions. Denote $\theta(\lambda)$ as the stochastic variable for $\lambda \in [0,1]$. For some $n \geq 1$ the functions $\gamma_s : [0,1] \to [0,1]$, $s = 0, \ldots, n$, determine the finite support $\{\gamma_0(\lambda), \gamma_1(\lambda), \ldots, \gamma_n(\lambda)\}$ of $\theta(\lambda)$ with $0 = \gamma_0(\lambda) < \gamma_1(\lambda) < \ldots < \gamma_n(\lambda)$. Furthermore, the functions $\rho_s : [0,1] \to [0,1]$, $s = 0, \ldots, n$, determine the distribution of the probability mass over the support of $\theta(\lambda)$ where $\Pr\{\theta = \gamma_s(\lambda)\} = \rho_s(\lambda) > 0$ for some $\lambda > 0$ and $\sum_{s=0}^{n} \rho_s(\lambda) = 1$. So, $\mathbb{E}\theta(\lambda) = \sum_{s=0}^{n} \rho_s(\lambda)\gamma_s(\lambda)$ and $\lim_{\lambda \to 0} \mathbb{E}\theta(\lambda) = 0$ if and only if $\lim_{\lambda \to 0} \rho_s(\lambda)\gamma_s(\lambda) = 0$ for all $s \geq 1$.

The next theorem states a sufficient condition such that generically[3] $\mu_2(\mathbb{E}P(\theta(\lambda)))$ does not converge to 0 if λ goes to 0.

Theorem 5.2. *Suppose ρ_s and γ_s are continuously differentiable for all $s = 0, \ldots, n$. If there is some s^*, $1 \leq s^* \leq n$ such that $\lim_{\lambda \to 0} \gamma_s(\lambda) = \gamma_s > 0$ and $\lim_{\lambda \to 0} \rho_s(\lambda) = 0$ for $s \geq s^*$, then generically*

$$\lim_{\lambda \to 0} \mu_2 (\mathbb{E}P(\theta(\lambda))) = \frac{\sum_{s=s^*}^{n} \rho_s'(0)\, p(\gamma_s)}{\sum_{s=s^*}^{n} \rho_s'(0) \left[p(\gamma_s) + q(\gamma_s)\right]} \neq 0.$$

[3]By generic we mean that the measure of the area in the parameter space for which our result holds is equal to one. We cannot do better than generic, because the proof of theorem 5.2 shows that a nongeneric counter example can be constructed.

Proof

We will derive necessary and sufficient condition for $\lim_{\lambda\to0}\mu_2(\mathbb{E}P(\theta(\lambda)))=0$, which will be nongeneric. Since $\lim_{\lambda\to0}\mathbb{E}p(\theta(\lambda))=\lim_{\lambda\to0}\mathbb{E}q(\theta(\lambda))=0$ both the denominator and numerator are equal to 0 in the limit. Application of De L'Hôpital's rule yields that $\lim_{\lambda\to0}\mu_2(\mathbb{E}P(\theta(\lambda)))$ is equal to

$$\frac{\lim_{\lambda\to0}\sum_{s=s^*}^{n}\left[\rho_s'(\lambda)\,p(\gamma_s(\lambda))+\rho_s(\lambda)\,p'(\gamma_s(\lambda))\,\gamma_s'(\lambda)\right]}{\lim_{\lambda\to0}\sum_{s=s^*}^{n}\left[\rho_s'(\lambda)\left[p(\gamma_s(\lambda))+q(\gamma_s(\lambda))\right]+\rho_s(\lambda)\left[p'(\gamma_s(\lambda))+q'(\gamma_s(\lambda))\right]\gamma_s'(\lambda)\right]}. \quad (5.2)$$

Since $p(0)=p'(0)=0$, $q(0)=q'(0)=0$ and $\rho_s(0)=0$ for all $s\geq s^*$ it follows that the numerator is equal to $\sum_{s=s^*}^{n}\rho_s'(0)\,p(\gamma_s)$. So, the latter can only be 0 if and only if $\sum_{s=s^*}^{n}\rho_s'(0)\,p(\gamma_s)=0$, which is a one-dimensional restriction upon the $n-s^*$-dimensional space of first derivatives in $\lambda=0$ of the vector function $(\rho_{s^*}(\lambda),\ldots,\rho_n(\lambda))$. This restriction is nongeneric. Therefore, the numerator is generically not equal to 0. Similarly, the denominator $\sum_{s=s^*}^{n}\rho_s'(0)\left[p(\gamma_s)+q(\gamma_s)\right]$ is generically not equal to 0. So, generically $\lim_{\lambda\to0}\mu_2(\mathbb{E}P(\theta(\lambda)))\neq0$. □

The condition of Theorem 5.2 translates as follows. If the support $\{\gamma_0(\lambda),\gamma_1(\lambda),\ldots,\gamma_n(\lambda)\}$ converges to the set $\{0,\gamma_{s^*},\ldots,\gamma_n\}$, then it is generically impossible that $\mu_2(\mathbb{E}P(\theta(\lambda)))$ converges to 0 as λ goes to 0. The limit depends upon the $n-s^*$ first-derivatives of ρ_s and the $n-s^*$ points γ_s in the limit support.

Theorem 5.2 implicitly imposes the necessary condition $\lim_{\lambda\to0}\{\gamma_0(\lambda),\gamma_1(\lambda),\ldots,\gamma_n(\lambda)\}=\{0\}$ in order to have convergence of $\mu_2(\mathbb{E}P(\theta(\lambda)))$ to 0. This means that the limit of the support of $\theta(\lambda)$ has to vanish or collapse into the singleton $\{0\}$. The following theorem states that this necessary condition is generically also a sufficient condition for the subclass of functions $\rho_s(\lambda)$ and $\gamma_s(\lambda)$ that are all $[z^*]+1$ times continuously differentiable. This means that for this subclass no further restrictions have to be made with respect to the functions ρ_s in order to have convergence of $\mu_2(\mathbb{E}P(\theta(\lambda)))$ to 0 as λ goes to 0.

Theorem 5.3. *Suppose $\rho_s(\lambda)$ and $\gamma_s(\lambda)$ are $[z^*]+1$ times continuously differentiable for all*

$s = 0, \ldots, n$. *If* $\lim_{\lambda \to 0} \gamma_s(\lambda) = 0$ *for all* s, *then generically* $\lim_{\lambda \to 1} \mu_2(\mathbb{E}P(\theta(\lambda))) = 0$.

Proof

Equation (5.2) in the proof of Theorem 5.2 is also valid here. However, since $\gamma_s(0) = 0$ for all s, $p(0) = p'(0) = 0$ and $q(0) = q'(0) = 0$ both the numerator and denominator are 0. Therefore, we proceed in an iterative manner in applying De L'Hôpital's rule. Within finite iterations this rule will yield a numerator and a denominator that are not equal to 0, because the $([z^*]+1)$-th derivative of $\gamma_s^{[z^*]+1}(1-\gamma_s)^{N-[z^*]-1}$ in $q(\gamma_s)$ is positive and all other terms are equal to 0. This positive term is multiplied by a weighted sum of the limit probabilities $\rho_s(0)$, $s \geq 1$. Generically, this weighted sum is not equal to zero So, the numerator is equal to 0 while the denominator is not at the $[z^*]+1$ step of the iterative process. □

References

Bellman, R. (1954). Limit theorems for non-commutative operations. I. *Duke Mathematical Journal 21*, 491–500.

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

Cooper, R.W. (1999). *Coordination Games: Complementarities and Macroeconomics.* Cambridge, U.K.: Cambridge University Press.

Harsanyi, J.C. and R. Selten (1988). *A General Theory of Equilibrium Selection in Games.* Cambridge, Massachusetts: The MIT Press.

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

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

Young, H.P. (1998). *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions.* Princeton, New Jersey: Princeton University Press.

A. Appendix

This appendix provides four lemma's, two of which (Lemmata A.2 and A.4) are used in the proofs of Theorem 3.1 and 3.2. The other two Lemmata, namely Lemmata A.1 and A.3, are used in the proofs of Lemmata A.2 and A.4 respectively.

Lemma A.1 focuses on the inequality $c_\ell(\theta) = \ell(\ell-1) + \theta(\theta N - 2\ell)(N-1) > 0$ for sufficiently small $\theta > 0$, which appears several times in the analysis. The lemma states several properties of the upper bound upon θ. Before we proceed by stating the lemma, note that

$$0 \leq \theta < \frac{\ell}{N} - \frac{\sqrt{\ell(N-\ell)(N-1)}}{N(N-1)} \Rightarrow \ell(\ell-1) + \theta(\theta N - 2\ell)(N-1) > 0.$$

Lemma A.1. *Let* $f(\ell) = \frac{\ell}{N} - \frac{\sqrt{\ell(N-\ell)(N-1)}}{N(N-1)}$. *Then*

$$0 < \min_{\ell=[z^*]+1,\dots,N-[z^*]} f(\ell) = f([z^*]+1) < f(N-[z^*]) = \min_{\ell=N-[z^*],\dots,N} f(\ell) < 1.$$

Proof

Note that $f(0) = f(1) = 0$, $f(N) = 1$and $0 < f(\ell) < \frac{\ell}{N} < 1$ for all other ℓ. Next, we show that $f(\ell)$ is increasing on $\ell \in (1, N-1)$. The first derivative of $f(\ell)$ with respect to $\ell \in (1, N)$ is given by

$$\frac{\partial}{\partial \ell} f(\ell) = \frac{2\sqrt{\ell(N-\ell)(N-1)} - (N-2\ell)}{2N\sqrt{\ell(N-\ell)(N-1)}}.$$

This derivative is larger than 0 if and only if

$$N\left(-4\ell^2 + (4(N-1)+4)\ell - N\right) > 0.$$

This inequality holds at $\ell = 1$ and $\ell = N-1$ (because $N \geq 2$) and, therefore, it holds for all $\ell = 1, \dots, N-1$. Since $f(N-1) < 1 = f(N)$ the function f also increases in going from $\ell = N-1$ to $\ell = N$. Then,

$$\min_{\ell=[z^*]+1,\dots,N-[z^*]} f(\ell) = f([z^*]+1), \qquad \min_{\ell=N-[z^*],\dots,N} f(\ell) = f(N-[z^*])$$

and $f([z^*]+1) < f(N-[z^*])$. □

The second lemma is used in the proof of Theorem 3.1.

Lemma A.2. *If* $\mathbb{E}\theta < \overline{\theta}$*, then* $c_k(\theta) a_{j,k}(\theta) - c_j(\theta) a_{k,j}(\theta) > 0$ *for* $\theta \in [\mathbb{E}\theta, \overline{\theta})$*,* $j = [z^*] + 1, \ldots, N - [z^*]$*,* $k = N - [z^*], \ldots, N$ *and* $j < k$*. Moreover,* $\overline{\theta} > 0$*.*

Proof

First, $a_{j,k}(\theta) \geq a_{k,j}(\theta) > 0$ if and only if $\theta \geq \mathbb{E}\theta$. Second, if $0 < \theta < \frac{k}{N} - \frac{\sqrt{k(N-1)(N-k)}}{N(N-1)}$, then $c_k(\theta) > 0$ and, thus, $c_k(\theta) a_{j,k}(\theta) > 0$. From Lemma A.1 it follows that $c_k(\theta) > 0$ for all $k = N - [z^*], \ldots, N$ if θ is smaller than

$$\min_{k=N-[z^*],\ldots,N} \frac{k}{N} - \frac{\sqrt{k(N-1)(N-k)}}{N(N-1)} = \frac{N-[z^*]}{N} - \frac{\sqrt{[z^*](N-[z^*])(N-1)}}{N(N-1)} > 0.$$

Third,

$$c_k(\theta) - c_j(\theta) = (k-j)\left[(k+j-1) - 2(N-1)\theta\right] > 0 \Leftrightarrow \theta < \frac{1}{2}\frac{j+k-1}{N-1}.$$

The latter bound upon θ is larger than $\frac{1}{2}\frac{N}{N-1} > \frac{1}{2}$, because $j \geq [z^*] + 1$ and $k \geq N - [z^*]$. To conclude, for $\theta \in [\mathbb{E}\theta, \overline{\theta})$ (recall $\overline{\theta}$ is the minimum of the two bounds upon θ just derived) it holds that $c_k(\theta) a_{j,k}(\theta) > c_j(\theta) a_{k,j}(\theta)$ for all $k > j$. □

The next lemma is used in the proof of Lemma A.4.

Lemma A.3. *Let* $g(j,k) = \frac{k+j-1}{2(N-1)} - \frac{\sqrt{N^2(k+j-1)^2 - 4N(N-1)kj}}{2N(N-1)}$*. Then*

$$0 < \min_{\substack{k=N-[z^*],\ldots,N \\ j=[z^*]+1,\ldots,N-[z^*]}} g(j,k) = g([z^*]+1, N) = \frac{N-[z^*]}{N-1}.$$

Proof

First, $g(j,k) > 0$ for all j,k. The derivative with respect to k is given by

$$\frac{\partial}{\partial k} g(j,k) = \frac{N\sqrt{N^2(k+j-1)^2 - 4N(N-1)kj} - N^2(k+j-1) + 2N(N-1)j}{2(N-1)N\sqrt{N^2(k+j-1)^2 - 4N(N-1)kj}}.$$

Then $\frac{\partial}{\partial k} g(j,k) < 0$ is equivalent to

$$\left(N\sqrt{N^2(k+j-1)^2 - 4N(N-1)kj}\right)^2 - \left(N^2(k+j-1) + 2N(N-1)j\right)^2$$
$$= -4N^2(N-1)j\left(2N(k+j-1) + (N-j)\right) < 0.$$

Note that $\frac{\partial}{\partial k} g(j,k) < 0$ for all j. By symmetry $\frac{\partial}{\partial j} g(j,k) < 0$ for all k. So,

$$\min_{\substack{k=N-[z^*],\dots,N \\ j=[z^*]+1,\dots,N-[z^*]}} g(j,k) = g(N-[z^*],N) = \frac{N-[z^*]}{N-1}.$$

$\square$

The last lemma is applied in the proof of Theorem 3.2.

Lemma A.4. *If $\mathbb{E}\theta < \widetilde{\theta}$, then for all $j = [z^*]+1,\dots,N-[z^*]$, $k = N-[z^*],\dots,N$ and $j<k$ the function $a_{j,k}(\theta) - a_{k,j}(\theta)$ admits a unique inflection point $\widehat{\theta} \in (0,\mathbb{E}\theta)$ and it is strictly convex on $\theta \in \left(\widehat{\theta},\overline{\theta}\right)$. Moreover, $0 < \widetilde{\theta} < \overline{\theta}$.*

Proof

Assume $\mathbb{E}\theta < \overline{\theta}$. Lemma A.2 states that each function $a_{j,k}(\theta) - a_{k,j}(\theta)$ is convex for $\theta \in \left[\mathbb{E}\theta,\overline{\theta}\right)$. Furthermore, from (3.1) and (3.2) it follows that each $a_{j,k}(\theta) - a_{k,j}(\theta)$ is locally concave for sufficiently small $\theta > 0$. So, there exists an inflection point $\widehat{\theta} \in (0,\mathbb{E}\theta)$. The remainder of this proof consist of showing that the third derivative of $a_{j,k}(\theta) - a_{k,j}(\theta)$ is increasing in every $\widehat{\theta} \in (0,\mathbb{E}\theta)$. We proceed as follows. Since

$$\frac{\partial}{\partial\theta}(\theta)^{\ell-2}(1-\theta)^{N-\ell-2} = \frac{\ell-2-\theta(N-4)}{\theta(1-\theta)}(\theta)^{\ell-2}(1-\theta)^{N-2-\ell}$$

we obtain

$$\frac{\partial}{\partial\theta}\frac{a_{m,\ell}(\theta)}{\theta^2(1-\theta)^2} = \frac{\ell-2-\theta(N-4)}{\theta^3(1-\theta)^3}a_{m,\ell}(\theta).$$

Differentiation of the second derivative $c_k(\theta)a_{j,k}(\theta) - c_j(\theta)a_{k,j}(\theta)$ with respect to θ yields

$$\frac{c_k'(\theta)a_{j,k}(\theta) - c_j'(\theta)a_{k,j}(\theta)}{\theta^2(1-\theta)^2} + \frac{(k-2-\theta(N-4))c_k(\theta)a_{j,k}(\theta) - (j-2-\theta(N-4))c_j(\theta)a_{k,j}(\theta)}{\theta^3(1-\theta)^3},$$

where $c_\ell'(\theta)$ denotes $\frac{\partial}{\partial\theta}c_\ell(\theta)$, $\ell = j,k$. By definition of $\widehat{\theta}$ it must hold that $c_k\left(\widehat{\theta}\right)a_{j,k}\left(\widehat{\theta}\right) = c_j\left(\widehat{\theta}\right)a_{k,j}\left(\widehat{\theta}\right)$, which yields

$$\left((k-j) + \widehat{\theta}\left(1-\widehat{\theta}\right)\left(\frac{c_k'\left(\widehat{\theta}\right)}{c_k\left(\widehat{\theta}\right)} - \frac{c_j'\left(\widehat{\theta}\right)}{c_j\left(\widehat{\theta}\right)}\right)\right)\frac{c_k\left(\widehat{\theta}\right)a_{j,k}\left(\widehat{\theta}\right)}{\widehat{\theta}^3\left(1-\widehat{\theta}\right)^3}.$$

From Lemma A.1 it follows that both $c_k(\theta)$ and $c_j(\theta)$ are positive for $\theta < f([z^*]+1) = \frac{[z^*]+1}{N} - \frac{\sqrt{([z^*]+1)(N-[z^*]-1)(N-1)}}{N(N-1)} < f(N-[z^*])$. So, $f([z^*]+1) \leq \overline{\theta}$. Since $k-j \geq 0$, $\widehat{\theta}\left(1-\widehat{\theta}\right) > 0$ and the term outside the large round brackets are all positive for $\widehat{\theta} < \overline{\theta}$ it suffices to prove that $c_k'\left(\widehat{\theta}\right)/c_k\left(\widehat{\theta}\right) - c_j'\left(\widehat{\theta}\right)/c_j\left(\widehat{\theta}\right)$ is positive in $\widehat{\theta} < f([z^*]+1)$. So,

$$\frac{c_k'\left(\widehat{\theta}\right)}{c_k\left(\widehat{\theta}\right)} - \frac{c_j'\left(\widehat{\theta}\right)}{c_j\left(\widehat{\theta}\right)} = \frac{2(N-1)(k-j)\left(N(N-1)\widehat{\theta}^2 - N(k+j-1)\widehat{\theta} + kj\right)}{c_k\left(\widehat{\theta}\right)c_j\left(\widehat{\theta}\right)}.$$

The numerator is positive for sufficiently small $\widehat{\theta} > 0$ (which can be established by choosing $\mathbb{E}\theta$ small enough). Then, necessarily, the numerator is positive for

$$\widehat{\theta} < \frac{k+j-1}{2(N-1)} - \frac{\sqrt{N^2(k+j-1)^2 - 4N(N-1)kj}}{2N(N-1)}.$$

Lemma A.3 implies that the latter bound upon $\widehat{\theta}$ is minimal at $\frac{N-[z^*]}{N-1} > \frac{1}{2}\frac{N}{N-1} \geq \overline{\theta}$, because $[z^*] < \frac{1}{2}N$. So, if $\mathbb{E}\theta \leq \widetilde{\theta} = \min\left\{\overline{\theta}, f([z^*]+1), \frac{N-[z^*]}{N-1}\right\}$, then the third derivative in $\widehat{\theta} \in (0, \mathbb{E}\theta)$ is positive and $a_{j,k}(\theta) - a_{k,j}(\theta)$ admits a unique inflection point on $(0, \mathbb{E}\theta)$. The rest trivially follows. $\square$