Stochastic Stability and Equilibrium Selection in Games

Alexandr Matros

AKADEMISK AVHANDLING

Som för avläggande av ekonomie doktorsexamen
vid Handelshögskolan i Stockholm
framläggs för offentlig granskning
fredag den 8 juni 2001, kl 13.15 i sal 750
Handelshögskolan, Sveavägen 65

Stochastic Stability and Equilibrium Selection in Games

STOCKHOLM SCHOOL OF ECONOMICS
EFI, THE ECONOMIC RESEARCH INSTITUTE

EFI Mission
EFI, the Economic Research Institute at the Stockholm School of Economics, is a scientific institution which works independently of economic, political and sectional interests. It conducts theoretical and empirical research in management and economic sciences, including selected related disciplines. The Institute encourages and assists in the publication and distribution of its research findings and is also involved in the doctoral education at the Stockholm School of Economics.
EFI selects its projects based on the need for theoretical or practical development of a research domain, on methodological interests, and on the generality of a problem.

Research Organization
The research activities are organized in nineteen Research Centers within eight Research Areas. Center Directors are professors at the Stockholm School of Economics.

ORGANIZATION AND MANAGEMENT	
Management and Organisation; (A)	Prof Sven-Erik Sjöstrand
Center for Ethics and Economics; (CEE)	Adj Prof Hans de Geer
Public Management; (F)	Prof Nils Brunsson
Information Management; (I)	Prof Mats Lundeberg
Center for People and Organization (PMO)	Acting Prof Jan Löwstedt
Center for Innovation and Operations Management; (T)	Prof Christer Karlsson
ECONOMIC PSYCHOLOGY	
Center for Risk Research; (CFR)	Prof Lennart Sjöberg
Economic Psychology; (P)	Prof Lennart Sjöberg
MARKETING	
Center for Information and Communication Research; (CIC)	Adj Prof Bertil Thorngren
Center for Consumer Marketing; (CCM)	Associate Prof Magnus Söderlund
Marketing, Distribution and Industrial Dynamics; (D)	Prof Lars-Gunnar Mattsson
ACCOUNTING, CONTROL AND CORPORATE FINANCE	
Accounting and Managerial Finance; (B)	Prof Lars Östman
Managerial Economics; (C)	Prof Peter Jennergren
FINANCE	
Finance; (FI)	Prof Clas Bergström
ECONOMICS	
Center for Health Economics; (CHE)	Prof Bengt Jönsson
International Economics and Geography; (IEG)	Prof Mats Lundahl
Economics; (S)	Prof Lars Bergman
ECONOMICS STATISTICS	
Economic Statistics; (ES)	Prof Anders Westlund
LAW	
Law; (RV)	Prof Erik Nerep

Chairman of the Board: Prof Sven-Erik Sjöstrand. *Director:* Associate Prof Bo Sellstedt,

Adress
EFI, Box 6501, S-113 83 Stockholm, Sweden • Internet: www.hhs.se/efi/
Telephone: +46(0)8-736 90 00 • Fax: +46(0)8-31 62 70 • E-mail efi@hhs.se

Stochastic Stability and Equilibrium Selection in Games

Alexandr Matros

STOCKHOLM SCHOOL OF ECONOMICS
EFI, THE ECONOMIC RESEARCH INSTITUTE

Dissertation for the Degree of Doctor of Philosophy, Ph.D.
Stockholm School of Economics, 2001

ISBN NR 91-7258-571-4

Keywords:
Game Theory, Evolutionary Game Theory, Industrial Organization, Markov chain, Stochastic stability, Imitation, Better Replies, Tournaments, Contest.

Printed by:
Elanders Gotab, Stockholm 2001

Distributed by:
EFI, The Economic Research Institute
Stockholm School of Economics,
P.O. Box 6501, S-113 83 Stockholm, Sweden
www.hhs.se/efi

To my parents and grandparents

Contents

Preface

This thesis consists of five papers, presented as separate chapters within three parts: Industrial Organization, Evolutionary Game Theory and Game Theory. The common basis of these parts is research in the field of game theory and more specifically, equilibrium selection in different frameworks. As each chapter is independent their introductory sections in Part II contain overlaps.

The first part, Industrial Organization, consists of one paper co-authored with Prajit Dutta and Jörgen Weibull. It is the result of my first research project as a Ph.D. student in Economics. The first draft, written together with Jörgen Weibull, appeared on January 20, 1998. Later on, Prajit Dutta joined the project and, three years after the first draft, we produced this paper. We analyze forward-looking consumers in a Bertrand framework and assume that if firms can anticipate a price war and act accordingly, so can consumers.

The second part, Evolutionary Game Theory, contains three chapters. All models in these papers are based on Young's (1993, 1998) approach. In Chapter 2, I generalize the Saez Marti and Weibull's (1999) model from the Nash Demand Game to generic two-player games. In Chapter 3, co-authored with Jens Josephson, we introduce a special set of stochastically stable states, minimal construction, which is the long-run prediction under imitation behavior in normal form games. In Chapter 4, I consider best reply and imitation rules on extensive form games with perfect information.

The last part, Game Theory, consists of my most recent paper. It asks how players must distribute their endowments during elimination cup tournaments, such as in tennis tournaments.

I warmly welcome any comments, remarks or suggestions the reader may have with regard to this research.

Acknowledgements

During the writing of this thesis, I have received useful comments and advice from many people. My greatest debt is to my scientific advisor, Jörgen Weibull, who taught me game theory and showed me how its tools could fruitfully be applied to various economic questions. He has provided me with invaluable guidance, profound insights, and general encouragement. His comments on all my papers were always precise and greatly improved them: he made me reconsider many things once more and reformulate imprecision in the text.

Larry Blume and Martin Dufwenberg were also very generous with their advice and encouragement.

It is a pleasure to acknowledge the support I receive from the Economics Department of the Stockholm School of Economics and, especially, from Karl Wärneryd, Tore Ellingsen, Jens Josephson and Maria Saez Marti.

I have been fortunate to receive moral, intellectual and financial support from the Stockholm Institute of Transition Economics (SITE). The director, Erik Berglöf, not only invited me to the Ph.D. program at the Stockholm School of Economics, but also constantly helped me in many respects. I have benefited greatly from long conversations with Guido Friebel. Gun Malmquist has given me more good advice and help than I thought was humanly possible. I am also indebted to Jody Lanfrey and Tommaso Milani, who were always ready to answer my "short questions" and correct my English.

I am indebted to Ritva Kiviharju, Pirjo Furtenbach, Kerstin Niklasson and Ingrid Nilsson for providing me with assistance in administrative matters. I would like to thank Christina Loennblad for her fine editing. I am grateful for financial support from the Jan Wallander and Tom Hedelius Foundation and the Stockholm School of Economics.

Without the encouragement of my family and friends, the completion of this thesis would have been impossible. In particular, I would like to express my love and gratitude to my parents and grandparents, who have always supported me in everything. I thank my wife, Zhanna, for her forbearance and support on those nights when writing and revising papers temporarily took priority over spending time with our family.

Thank you all.

Stockholm, April 2001.

Alexandr Matros

I

Part I: Industrial Organization

Chapter 1: Dynamic Bertrand Competition with Forward-Looking Consumers

1

Bertrand Competition with Forward-Looking Consumers

PRAJIT DUTTA, ALEXANDER MATROS AND JÖRGEN W. WEIBULL*
COLUMBIA UNIVERSITY AND THE STOCKHOLM SCHOOL OF ECONOMICS

April 27, 2001

ABSTRACT. In the standard model of dynamic Bertrand competition, competing firms meet the same demand function every time period. This is not a satisfactory model of the demand side if consumers can make intertemporal substitution between periods, and if they have some foresight. Consumers who observe price undercutting may (correctly) anticipate a subsequent price war, and may therefore postpone purchases. This effect may drastically reduce the profits to a deviation from collusive pricing. Hence, consumers' intertemporal substitution possibilities and foresight may facilitate collusion against them. However, such a richer model of the demand side complicates the analysis, since the interaction between the firms no longer constitutes a repeated game - and hence falls outside the domain of the usual Folk theorems. We formally analyze collusive pricing in such a setting, and identify cases both when collusion is facilitated and when it is made more difficult when consumers have perfect foresight. We also consider cases of imperfect foresight.

1. INTRODUCTION

The Coase conjecture, Coase (1972), stipulates that a monopolist selling a new durable good cannot credibly commit to the monopoly price, because once this price has been announced, the monopolist will have an incentive to reduce his price in order to capture residual demand from consumers who value the good below the monopoly price. This in turn, Coase claims, would be foreseen also by consumers with valuations above the monopoly price, and therefore some of these (depending on their time preference) will chose to postpone their purchase in anticipation of a price fall. Coase's argument is not relevant only for a monopoly firm in a transient market for an new durable good, but also for oligopolistic firms in a perpetually ongoing market for durable and non-durable goods. If such firms maintain a price above the competitive price, based on a threat of a price war or other severe punishments in case of defection, as the literature on repeated games suggests, then consumers might foresee such price wars in the wake of a defection, and hence not buy from a firm that slightly undercuts the others but instead postpone purchase to the anticipated

*Matros and Weibull thank the Hedelius and Wallander Foundation for financial support.

subsequent price war. Such dynamic aspects of the demand side runs against the spirit of the standard text-book model of dynamic competition viewed as a repeated game.[1] Indeed, the interaction is no longer a dynamic game, since the market demand faced by the firms today in general depends on history. Consequently, the model falls outside the domain of the standard Folk Theorems. Moreover, unlike in the case of a monopoly for a new durable product, this application of Coase's argument leads to a very different conclusion: under many circumstances such intertemporal substitution and foresight on behalf of the consumers in a recurrent market setting facilitates, rather than undermines collusion against them.[2]

There is a literature on the Coase conjecture, building on models of consumers who have the possibility of intertemporal substitution and are endowed with foresight, see e.g. Gul, Sonnenschein and Wilson (1986), Gul (1987) and Ausubel and Deneckere (1987). We here model consumers very much in the same vein. However, while the demand structure in those models is transient, we here develop a stationary demand structure, more precisely an infinite sequence of overlapping cohorts of new consumers entering and leaving the market. All consumers have a fixed life span of one time unit, while each market period, during which firms' prices are held constant, has length $1/m$. A new cohort of consumers enters the market in each market period, the size of each cohort is $1/m$. Hence, except for the first $m-1$ market periods, the population size of consumers is constantly equal to one in every market period, and each consumer lives during m consecutive market periods. Consumers differ as to their individual valuation of the good. In each cohort, the individual valuation v is distributed according to some cumulative distribution function. The good in question is assumed to be sold in indivisible units, and each consumer wants to acquire at most one unit of the good in her life time. There is no resell. Following the above-mentioned analyses, we treat firms as players in the game-theoretic sense but model consumers as price-taking and expectation forming economic agents with no strategic incentive or power. Their aggregate behavior will constitute a state variable in a dynamic game played by the firms. Most of our analysis is focused on the case of consumers with perfect foresight, but we also consider a case when consumers have imperfect foresight.

This paper is not a plea that analysts should always assume all economic agents to have perfect foresight. We believe that consumers and firms may more realistically be modelled as having more or less imperfect foresight. Our position is rather that the contrast in current models of dynamic oligopolistic competition between, on the

[1] See e.g. Tirole (1988) for repeated-games models of dynamic oligopoly, and Fudenberg and Tirole (1991) for various versions of the folk theorem.

[2] However, we show that in some special cases the effect may go in the same direction as in the Coase conjecture: collusion may be more difficult if consumers have foresight.

one hand, the great sophistication and expectations coordination ascribed to firms and, on the other hand, the complete lack of sophistication ascribed to consumers, should be replaced by a milder contrast.[3] Even taking a small step in this direction requires the analyst to go outside the familiar class of repeated games to the less familiar class of stochastic games - containing repeated games as a subclass. We here outline in the simplest possible context how such a generalization can be made, and what are its most direct implications.

The paper is organized as follows. The model is defined in Section 2. Section 3 considers briefly the special case $m = 1$, corresponding to the standard model of Bertrand competition as a repeated game. Section 4 treats in some detail the case $m = 2$, i.e., where each consumer lives in two periods. Most of the action comes out already in this case. Section 5, finally, considers briefly the cases $m > 2$.

2. The model

Suppose that there are n firms in the market for a homogenous indivisible good. The market operates over an infinite sequence of periods $k \in \mathbb{N} = \{0, 1, 2, ...\}$. The firms simultaneously announce their *ask prices* every period. Let $x_{ik} \geq 0$ be the price that firm i asks in period k, and let $x_k = (x_{1k}, ..., x_{nk})$ be the vector of ask prices in that period. All consumers are assumed to observe all ask prices in each period, and they buy only from the firms with the lowest ask price. The lowest ask price in any period k,

$$p_k = \min\{x_{1k}, ..., x_{nk}\},$$

will accordingly be called the *market price* in that period. If more than one firm ask the lowest price, then we assume that sales are split equally between them.[4] Each *market period* is of length $\Delta = 1/m$, for some positive integer m. Thus Δ is the duration of the commitment that firms make to their ask prices. The firms face no capacity constraint on production, and production costs are normalized to zero. Hence, each firm's profit in a market period is simply its sales multiplied by its ask price. They all discount future profits at the same time (interest) rate $r > 0$, resulting in the common *discount factor*

$$\delta = \exp(-r\Delta) = \exp(-r/m) \tag{1}$$

[3]It can be argued, on evolutionary grounds, that firms should be more sophisticated than consumers, since the selection pressure on firms' behaviors might be stronger. However, in our opinion, this does not justify the stark difference assumed in the standard models.

[4]This may seem to contradict the assumed indivisibility of the good. However, we assume that the number of consumers is very large, and model the consumer population as a continuum.

between successive market periods.

There is a continuum of consumers, all with a fixed *life span* of one time unit - which thus amounts to m market periods. We keep the life span of consumers fixed in this study but vary the length of market periods. A new cohort of consumers enters the market in each market period. The size of each cohort is $1/m$. Hence, except for the first $m-1$ market periods, the population size of (living) consumers is constantly equal to one in every market period. Our analysis is focused on periods $k = m$, $m+1$, $m+2, \ldots$, when each market period contains all m consumer cohorts, and the population size is 1.

Consumers differ as to their individual valuation of the good. In each cohort, the individual valuation v is distributed according to some cumulative distribution function $F : \mathbb{R}_+ \to [0,1]$. Each consumer wants to acquire at most one unit of the (indivisible) good in her life time. A consumer with valuation v derives utility $(v-p)\,\beta_h$ from buying one unit of the good at price p in market period $h = 0, 1, 2, \ldots, m-1$ of her life, where all $\beta_h \in [0,1]$ and $\beta_0 = 1$. We assume dynamically consistent temporal preferences in the sense that each consumer attaches the same relative weights to any pair of future market periods in her life, independently of her current age. Hence, β_h is the consumers' subjective *discount factor* between market periods $h-1$ and h in her life time (for $h = 1, 2, ..$). We will pay special attention to (a) maximally patient consumers, with $\beta_h = 1$ for all h, (b) maximally impatient consumers, with $\beta_h = 0$ for all $h > 0$, and (c) consumers with the same exponential time preference as the firms, $\beta_h = \delta^h$ for all h. Moreover, we assume that

[$A1$] Consumers hold identical expectations about future prices,

[$A2$] There is no resell.

3. Bertrand competition when a market period is the life span of each consumer

The case $m = 1$ corresponds to the standard Bertrand model in the literature on oligopolistic competition: The firms face one and the same demand function in each market period. Hence, in this case we have a repeated game, and the Folk Theorem applies. In the present case, with a valuation distribution F, the firms with the lowest ask price, p, face the period demand $D(p) = 1 - F(p)$; the population share of consumers with valuations exceeding p. In particular, a trigger strategy profile according to which all firms quote the same price p^*, in all periods, until a price deviation occurs, and thereafter quote zero, constitutes a subgame perfect equilibrium in this infinitely repeated game if and only if

$$\frac{1}{n}p^*\left[1 - F(p^*)\right] \geq \left(1 - e^{-r}\right) p \left[1 - F(p)\right] \tag{2}$$

for all $p < p^*$. In particular, if the revenue function $R(p) = p[1 - F(p)]$ is single-peaked, then the right-hand side of equation (2) is maximized at $p = p^*$, granted p^* does not exceed the monopoly price. Hence, the deviating firm wants to undercut the going price only slightly, and p^* is a sustainable equilibrium price if and only if

$$e^{-r} = \delta \geq 1 - 1/n. \tag{3}$$

4. Bertrand competition when a market period is half the life span of each consumer

In the case $m = 2$, the market period is half as long. In each market period, firms now face two consumer groups, each half the size of the consumer group in the case $m = 1$. Aggregate demand in any market period (after the very first) can accordingly be decomposed into two components, one arising from all the young, and another, residual demand, arising from those old individuals who did not buy while young. The demand of the young depends not only on the current market price; it also depends on their expectations about the market price in the subsequent period. If a young individual with valuation v faces a current market price p and expects the market price p^e in the next period, then she should buy in the present market period if and only if her consumer surplus from buying now, $v - p$, exceeds (or equals) the present expected surplus from buying in the next period, $\beta(v - p^e)$.

In particular, if $\beta < 1$ and the expected price next period equals the current price, $p^e = p$, then all young consumers with valuation $v \geq p$ will buy in the present period. By contrast, if the expected price next period were zero, then only those young consumers who have valuation $v \geq p/(1 - \beta)$ will buy in the present period. All young individuals with lower valuations will wait until the next period. In general, the cut-off valuation level, when the current price is p and the price expected for the next period is p^e, is

$$v^+ = \frac{p - \beta p^e}{1 - \beta}. \tag{4}$$

This cohort's demand in the next period, when they are old, stems from those currently young individuals who have valuations below v^+. Since this residual demand depends on previous prices - directly via p and indirectly via p^e - and this residual demand affects the profit function of the firms in that period, we no longer have a repeated game. In fact, what we have is a so-called stochastic or Markovian game,[5] with state variable v^+ which we will call the *residual valuation.* The accompanying residual demand is $D^+(p) = [1 - F(p)]/2$ for all $p \leq v^+$ and $D^+(p) = 0$ for all $p > v^+$.

[5]For a discussion of Markovian games, see Fudenberg and Tirole (1995), chapter 12. For a discussion of equilibrium characterization in such games, see Dutta (1995).

We assume that firms have complete information about the market interaction, and that they observe all past prices. In particular, all firms know the residual valuation inherited from the previous period. A *strategy* for a firm is thus a rule that specifies its ask price in each market period, given any history of prices (and thus also of residual valuations) up to that period. Firms' price strategies constitute a subgame perfect equilibrium in this game provided each firm maximizes its discounted future stream of profits after any price history, given all other firms' strategies and the consumers' expectations. In particular, a *trigger strategy* can be defined in much the same way as in the standard repeated-games model: All firms ask the same price p^* in the first period, and continue to post that price as long as all firms quote that price in all periods. In the event of any deviation from that price, all firms ask the price zero in all subsequent periods.[6]

It is easily verified (see equation (4)) that if consumers have perfect foresight, in the sense of correctly anticipating the market price to fall to zero after any price deviation, and to otherwise to remain constant, then such a trigger strategy profile constitutes a subgame perfect equilibrium if and only if

$$\frac{1}{2n}\frac{p^*\left[1-F(p^*)\right]}{1-\delta} \geq \frac{1}{2}p\left[F(p^*)-F(p)+1-F\left(\frac{p}{1-\beta}\right)\right] \tag{5}$$

for all $p < p^*$. The factor $\frac{1}{2}$ on both sides of this inequality reflects the fact that half the population in any period is young while the other half is old. The left hand side is the present value of the stream of future profits to each firm, from the present period onwards, when the price p^* is asked by all firms in all periods. The expression on the right-hand side is the present value of the stream of profits, from the present period onwards, to a firm that unilaterally deviates in the present period. In all periods after the current, such a deviator earns zero profit. The expression in large square brackets is the demand faced by the deviating firm if its present ask price is $p < p^*$. This demand is composed of two components: the residual demand from the old who did not buy when they were young, $F(p^*) - F(p)$, and the demand from the young, $1 - F\left[p/(1-\beta)\right]$, where $p/(1-\beta) = v^+$, the residual valuation when $p^e = 0$ (see equation (4)). Rearranging the terms, and using equation (1), inequality (5) can be re-written as

$$\frac{1}{n}p^*\left[1-F(p^*)\right] \geq \left(1-e^{-r/2}\right)p\left[1-F(p)+F(p^*)-F\left(\frac{p}{1-\beta}\right)\right] \tag{6}$$

for all $p < p^*$. The left-hand side is identical with the left-hand side in the incentive

[6] Recall that the marginal cost is zero, so to quote the price zero is a Nash equilibrium of the stage game in any given period, irrespective of the inherited residual valuation.

constraint (2) in the standard case, $m = 1$, while the right hand sides in these two cases differ.

To analyze the conditions under which equation (6) holds - and to contrast those conditions with the standard Bertrand model - we will focus on two countervailing forces at work: one working for the firms and against the consumers, to be called the *anti-competitive* force and another force, working against the firms and for the consumers, to be called the *pro-competitive* force.

4.1. The anti-competitive force - the young may not bite. When a young consumer, with perfect foresight, sees a price below p^* he knows that an even lower price is coming in the next period. Provided he is patient he will leave the market and wait with gleeful anticipation for the price war. That will diminish the fruits to the firm that undercuts the going price. Such lowering of deviation profits may then convince a prospective under-cutter to not drop his price - giving more room for anti-competitive or collusive outcomes.[7]

To make the point most stark, consider the case of consumers who are maximally patient, $\beta = 1$.[8] Inequality (6) then boils down to

$$\frac{1}{n}p^*\left[1 - F(p^*)\right] \geq \left(1 - e^{-r/2}\right) p\left[F(p^*) - F(p)\right] \tag{7}$$

for all $p < p^*$. Note that a deviating firm only sells to old consumers because every young consumer, regardless of his valuation, prefers to wait for the zero price in the next period. In comparing the two incentive constraints, we note that $F(p^*) < 1$ and $e^{-r/2} > e^{-r}$, so the right-hand side of (7) is less than the right-hand side in (2), at any price p (while the left-hand sides are identical). Hence,

Proposition 1. *If consumers have perfect foresight and are maximally patient, then any price that can be sustained in subgame perfect equilibrium in the standard Bertrand model, $m = 1$, can also be sustained in subgame perfect equilibrium when $m = 2$.*

Put differently, patient consumers with perfect foresight are more likely to meet collusive prices when the market period is shorter - beyond the mere effect of increasing the discount factor between periods.

[7]The logic is similar to the anti-competitive implication of a polcy that promises to "match the competitor's price." In this case a firm that undercuts the competition gets no additional sales because consumers stay with their original seller - they simply get their seller to match the undercut price. In our model a firm that undercuts may not see much by way of additional sales because consumers wait for a price war.

[8]In order to avoid complete temporal indiffernce, we actually mean the case when β is slightly below 1. Alternatively, one could here think of a lexicographic preference for early timing.

One way to evaluate the effect of this anti-competitive force is to determine the cut-off discount factor for firms at which any given price p^* becomes a sustainable equilibrium price. Denote this cut-off discount factor $\delta(p^*)$. What the previous proposition has established is that whenever consumers are maximally patient, then $\delta(p^*)$ is less than $1 - 1/n$, the cut-off discount factor in the standard Bertrand model. An interesting question to ask is whether collusion is comparatively easier to sustain when the price p^* is close to the monopoly price or when it is far away from the monopoly price. We provide a regularity condition on the demand function that implies that $\delta(p^*)$ is increasing in p^*, i.e., the monopoly price is hardest to sustain while prices close to the (zero) marginal cost are the easiest to sustain. Contrast this with the standard Bertrand model, in which no price is sustainable below the cut-off discount factor $1 - 1/n$, and every price is sustainable at and above that discount factor. The regularity condition we provide is that both the demand function, $D(p) = 1 - F(p)$, and the revenue function $R(p) = p[1 - F(p)]$ be concave. This regularity condition holds when F is generated by a uniform distribution and hence D linearly decreasing - the standard text-book example in the industrial organization literature.

Proposition 2. *Suppose that F is convex and that the revenue function $p[1 - F(p)]$ is concave. For every price p^* there exists a discount factor $\delta(p^*) < 1$ such that $p^*$is a subgame perfect equilibrium price if and only if $e^{-r/2} \geq \delta(p^*)$. Moreover, $\delta(p^*)$ is strictly increasing at all prices below the monopoly price.*

Proof: That $\delta(p^*)$ is uniquely defined for every p^* follows straighforwardly from equation (7). Consider the maximum deviation payoff: $M(p^*) = \sup_{p<p^*} p\,[F(p^*) - F(p)]$. From an application of the Envelope Theorem, it follows that M is a convex function if F is. The incentive condition (7) can evidently be rewritten as $\pi(p^*) - n\left(1 - e^{-r/2}\right) M(p^*) \geq 0$. Since the profit function π is concave and M is convex, it follows that their difference is a concave function. Moreover this difference is zero at $p^* = 0$ and has a strictly positive slope at that price. Hence there is a lowest positive price at which the difference becomes zero again; let that price be denoted $p(\delta)$. It follows that $p(\cdot)$ is an increasing function. However, its inverse function is $\delta(\cdot)$. Hence, $\delta(\cdot)$ is strictly increasing. **End of proof.**

4.2. The pro-competitive force - the old will always bite. When market periods are shorter, there is however also an advantage to dropping the price, an advantage that is absent in the standard Bertrand model. This arises from the fact that older consumers are still present, who have not yet bought, and who might now be tempted to do so (in the twilight of their lives). As long as the undercutting firm does not lose too many younger consumers - who may wait for the price war - this additional residual demand from the old (absent in the standard model) might make price cuts even more profitable than in the standard model. Consequently collusion may in fact be harder to sustain when market periods are shorter.

Again to make the point most stark, consider the case where young consumers buy from an undercutting firm, even if they expect an even lower market to follow. This is the case if consumers are maximally impatient, $\beta = 0$. In this case, inequality (6) boils down to

$$\frac{1}{n}p^* \left[1 - F(p^*)\right] \geq \left(1 - e^{-r/2}\right) p \left[1 - 2F(p) + F(p^*)\right] \tag{8}$$

for all $p < p^*$. In comparing inequalities (2) and (8), we note that the best deviation payoff is higher in the present case if and only if[9]

$$\left(1 + e^{-r/2}\right) p^* \left[1 - F(p^*)\right] < \max_{0<p<p^*} p \left[1 - 2F(p) + F(p^*)\right] \ . \tag{9}$$

What makes the analysis somewhat messy is that the best deviation price is in general difficult to characterize in closed form without some further assumptions about the valuation distribution F. Instead of analyzing (9) for a general c.d.f. F, we therefore illustrate this possibility by means of a simple example.

Example. Suppose the c.d.f. F has valuations concentrated on two values, v_1 and v_2, where $v_1 < v_2$. Let θ denote the population share at the high valuation, v_2, and suppose without loss of generality that v_2 is the monopoly price, i.e., $\theta v_2 > v_1$. From inequality (8) we see that the monopoly price, $\hat{p} = v_2$, is *not* sustainable in subgame perfect equilibrium in the case $m = 2$, i.e., with short market periods, if

$$\frac{\theta v_2}{n} < \left(1 - e^{-r/2}\right) v_1 \left(2 - \theta\right) \tag{10}$$

By contrast, in the standard Bertrand model, i.e., the case $m = 1$, the monopoly price *is* sustainable in subgame perfect equilibrium if

$$\frac{\theta v_2}{n} \geq \left(1 - e^{-r}\right) \theta v_2 \ , \tag{11}$$

or, equivalently, if $1 - e^{-r} \leq 1/n$. Hence, both conditions hold if

$$1 - e^{-r} \leq \frac{1}{n} < \left(1 - e^{-r/2}\right) \frac{v_1 \left(2 - \theta\right)}{\theta v_2} \tag{12}$$

[9]We have here used the identity $\left(1 - e^{-r}\right) = \left(1 - e^{-r/2}\right)\left(1 + e^{-r/2}\right)$.

Proposition 3. *If consumers have prefect foresight, are maximally impatient ($\beta = 0$), their valuation distribution is concentrated at two distinct values, and condition (12) is satisfied, then the monopoly price is sustainable in subgame perfect equilibrium in the standard Bertrand model ($m = 1$) but not when $m = 2$.*

Put differently, impatient consumers with perfect foresight may see less collusive prices when the market period is short than when it is long. That condition (12) is non-vacuous is easily checked. For example, all parameter combinations in a neighborhood of $n = 2$, $e^{-r/2} = 0.72$, $\theta \leq 0.2$ and any $v_1/\theta v_2 = 1$ satisfy the condition.

4.3. A more general case. We here generalize the above results in two directions: first we briefly present the analysis for trigger strategies, and any level of consumer patience β. Recall the incentive condition (6) for consumers with perfect foresight and arbitrary time preferences β. From the right-hand side of this inequality it immediately follows that the more patient are consumers, i.e., the higher is β, the lower is the payoff to any one firm undercutting a given price p^*. Consequently, firms find it easier to collude when consumers are patient:

Proposition 4. *For any price p^* (possibly the monopoly price) and any consumer discount factor $\beta \geq 0$ there exists a critical discount factor $\delta(\beta) \in (0,1)$ for the firms, such that p^* is a subgame perfect equilibrium price if and only if $e^{-r/2} \geq \delta(\beta)$. Moreover, $\delta(\beta)$ is strictly decreasing in β.*

In particular, we already know that $\delta(1)$ is less than $1 - 1/n$, the cut-off discount factor in the standard Bertrand model. If $\delta(0)$ exceeds $1 - 1/n$ - as for instance in the example discussed in the previous subsection - it follows by continuity that there is some intermediate level of consumer patience, say β^* at which $\delta(\beta^*) = 1 - 1/n$. At that degree of consumer impatience, the cut-off discount factor for firms is the same as in the standard Bertrand model.[10]

So far, we have been concentrating on the case when consumer correctly anticipate firms to use the grim trigger strategy. Of course, firms may actually choose to initiate less severe price wars, or temporary severe price wars ("forgiving trigger strategies"). As equation (4) makes clear, the less severe the price war the larger is the share of the young consumers who will buy in the current market period. Hence, a less severe price war is more profitable for a firm that undercuts a collusive price, making collusion less sustainable. It follows from this observation that whenever a price can

[10]Of course, the exact value of β^* depends on the price p^* that is being sustained in equilibrium. If F is convex it can be shown, by a similar argument as in the proof of proposition 2, that β^* is an increasing function of p^*.

be sustained by a less severe price war, and consumer correctly anticipate this, it can also be sustained by the most severe price war - zero price reversion. Moreover, a severe but temporary price war, zero price for a finite number of market periods, may be sufficient to deter firms' deviations but will induce consumers to behave exactly as in the case of zero price forever.

4.4. Consumers with imperfect foresight. Another comparison that may be relevant is between consumers with perfect foresight, as modelled in the preceding sections, and consumers who always expect the current market price to prevail also in the future - "martingale expectations."[11] When faced with a price cut, all young consumers with valuations above the new market price will then buy, just as the above case of maximally impatient consumers. Hence, for any time preference $\beta \in (0,1)$, the profit to a deviating firm with ask price $p < p^*$ is

$$\frac{1}{2}\left(1 - e^{-r/2}\right) p \left[1 - 2F(p) + F(p^*)\right] \tag{13}$$

if all consumers hold non-rational expectations in this sense, while the profit to the same deviation would have been

$$\frac{1}{2}\left(1 - e^{-r/2}\right) p \left[1 - F(p) + F(p^*) - F\left(\frac{p}{1-\beta}\right)\right] . \tag{14}$$

if all consumers had perfect foresight (see right-hand side of (5)). The latter quantity never exceeds the first, so the range of collusive prices against consumers with perfect foresight always contains the range of collusive prices against consumers with martingale expectations, for any time preference $\beta > 0$ that the consumers may have. In this precise sense, consumers are worse off if they have perfect foresight.

The standard model can be identified with this particular case of imperfect foresight. In this perspective, the comparison with the standard model is clear: consumers' foresight unambiguously facilitate collusion against them. We believe this qualitative conclusion to hold also in intermediate cases, when, say, some consumers have perfect foresight and others have martingale expectations, or when all consumers hold expectations between these two extreme cases.

5. Bertrand competition when market periods are shorter

In the cases $m > 2$, firms face more than two consumer cohorts in each market period. None of the currently youngest, and some of the individuals in the intermediate cohorts, have not yet bought the good. Unlike the oldest, all of these have the option

[11] More precisely, we here assume $p^e_{k+1} \equiv p_k$.

of postponing their purchased another market period. Thus, all but the oldest base their buying decisions in part on their expectations about future market prices. Do the above observations concerning the case $m = 2$ carry over to these more complex cases? The answer turns out to be in the affirmative: the analysis is readily generalized, and though quantitatively different, the results are qualitatively the same.

In each market period, firms now face $m > 2$ consumer cohorts, each cohort being of size $1/m$ as compared with the standard case $m = 1$. Aggregate demand in any market period can conveniently be decomposed into three components, one arising from all the young, another arising from the old individuals who did not yet buy, and a third, new component arising from individuals in intermediate cohorts, who did not yet buy. If a young or intermediate individual with valuation v faces a current market price p and expects the market prices $\left(p_1^e, p_2^e, ..., p_{m-1}^e\right)$ in the following $m-1$ periods, then she should buy in the present period if and only if her consumer surplus from buying now, $v - p$, exceeds (or equals) the expected discounted surplus, $\beta_h\left(v - p_h^e\right)$, from buying in each of the remaining market periods during her life span, where $\beta_h \in [0, 1]$ is the consumer's discount factor in the h-th market period of her life.

Assuming exponential discounting and expectations such that the prices in future periods are expected not to be lower than the price expected in the next period (which is consistent with perfect foresight under trigger strategies), then the cut-off valuation level, when the current price is p and the price expected for the next period is p^e, is

$$v^+ = \frac{p - \beta p^e}{1 - \beta}, \tag{15}$$

precisely as in the case $m = 2$, see equation (4). The accompanying residual demand in the next period is $D^+(p) = [1 - F(p)]/m$ for all $p \leq v^+$ and $D^+(p) = 0$ for all $p > v^+$.

Just as in the case $m = 2$, we focus on trigger strategies. More exactly, all firms ask the same price p^* in the first m periods, and continue to post that price as long as all firms quote that price. In the event of any deviation from that price, all firms ask the price zero in all subsequent periods. It is easily verified (see equation (15)) that if consumers correctly anticipate the market price to fall to zero after any price deviation, and to otherwise to remain constant, then such a trigger strategy profile constitutes a subgame perfect equilibrium if and only if

$$\frac{p^*\left[1 - F(p^*)\right]}{mn\left(1 - \delta\right)} \geq \frac{p}{m}\left[F(p^*) - F(p) + M + 1 - F\left(\frac{p}{1 - \beta}\right)\right], \tag{16}$$

for all $p < p^*$, where

$$M = (m - 2)\max\left\{0, F(p^*) - F\left(\frac{p}{1 - \beta}\right)\right\}.$$

The factor $1/m$ represents the size of each cohort, where $1/m$ is the population share of young individuals (in their first market period), and of old individuals (in their last market period), while $(m-2)/m$ is the population share of individuals in intermediate ages (in an intermediate market period of their lives). The left-hand side is the present value of the stream of future profits to each firm, from the present period onwards, when the price p^* is asked by all firms in all periods. The right-hand side is the present value of the stream of profits from the present period onwards to a firm that unilaterally deviates in the current period - the revenue in that period if its present ask price is $p < p^*$ (the revenue in all subsequent periods being zero). The deviating firm faces demand that can be decomposed into three components: the residual demand from the old who did not yet buy, $F(p^*) - F(p)$, the residual demand from those individuals in intermediate cohorts who did not yet buy, M, and the demand from the young, $1 - F[p/(1-\beta)]$ (see equation (15)). Rearranging the terms, and using equation (1), inequality (16) can be written as

$$\frac{p^*}{n}[1-F(p^*)] \geq \left(1-e^{-r/m}\right)p\left[F(p^*)-F(p)+M+1-F\left(\frac{p}{1-\beta}\right)\right] \tag{17}$$

The left-hand side is identical with the left-hand side in the incentive constraint (2) in the standard case, $m = 1$, while the right hand sides in these two cases differ.

5.1. The anti-competitive force - the young may not bite. In the case of consumers who are maximally patient, $\beta = 1$, inequality (17) boils down to

$$\frac{1}{n}p^*[1-F(p^*)] \geq \left(1-e^{-r/m}\right)p\left[F(p^*)-F(p)\right], \tag{18}$$

c.f. (7). Note that a deviating firm only sells to old consumers because every young or intermediate consumer, regardless of his valuation, prefers to wait for the zero price in the next period. In comparing with the case $m = 1$, we note that $F(p^*) < 1$ and $e^{-r/m} > e^{-r}$, so the right-hand side of (18) is less than the right-hand side in (2), at any price p, while the left-hand sides are identical. Moreover, the right-hand side of (18) is decreasing function of m. Hence,

Proposition 5. *If consumers have perfect foresight and are maximally patient ($\beta =$ 1), then any price that can be sustained in subgame perfect equilibrium in the standard Bertrand model, $m = 1$, can also be sustained in subgame perfect equilibrium when $m \geq 2$. Moreover, any price that can be sustained in subgame perfect equilibrium for some m can be sustained for all $m' > m$.*

5.2. The pro-competitive force - the old will always bite. Again to make the point most stark, consider the case where consumers are maximally impatient, $\beta = 0$. In this case, inequality (17) boils down to

$$\frac{1}{n}p^*\left[1 - F(p^*)\right] \geq \left(1 - e^{-r/m}\right) p\left[1 - mF(p) + (m-1)\, F(p^*)\right], \tag{19}$$

c.f. (8). In comparing with the case $m = 1$ and $m = 2$, we note that $e^{-r/m} > e^{-r/2} > e^{-r}$ and $[1 - F(p)] < [1 - 2F(p) + F(p^*)] < [1 - mF(p) + (m-1)\, F(p^*)]$, so the right-hand side in (19), might be less or more than the right-hand side in (2) and in (8), while the left-hand sides are identical. Hence, we cannot say more until $F(p)$, r and m are specified.

REFERENCES

[1] Ausubel L. and R. Deneckere (1987): "One is almost enough for monopoly", *Rand Journal of Economics* 18, 255-274.

[2] Coase R.H. (1972): "Durability and monopoly", *Journal of Law and Economics* 15, 143-149.

[3] Dutta, P. K. (1995): "A Folk Theorem for Stochastic Games", *Journal of Economic Theory* 66, 1-32.

[4] Fudenberg, D. and J. Tirole (1991): *Game Theory*, MIT Press, Cambridge, MA.

[5] Gul F., H. Sonnenschein and R. Wilson (1986): "Foundations of dynamic monopoly and the Coase conjecture", *Journal of Economic Theory* 39, 155-190.

[6] Gul F. (1987): "Noncooperative collusion in durable goods oligopoly", *Rand Journal of Economics 18*, 248-254.

[7] Tirole J. (1988): *The Theory of Industrial Organization*, MIT Press, Cambridge, MA.

Part II: Evolutionary Game Theory

II

Chapter 2: Clever agents in adaptive learning

Clever agents in adaptive learning

ALEXANDR MATROS*
DEPARTMENT OF ECONOMICS
STOCKHOLM SCHOOL OF ECONOMICS

October 3, 2000

2

ABSTRACT. Saez-Marti and Weibull [4] investigate the consequences of letting some agents play a myopic best reply to the myopic best reply in Young's [8] bargaining model, which is how they introduce "cleverness" of players. I analyze such clever agents in *general* finite two-player games and show Young's [9] prediction to be robust: *adaptive learning with clever agents* does select the same minimal curb set as in the absence of clever agents, if their population share is less than one. However, the long-run strategy distribution in such a curb set may vary with the share of clever agents.

Key words: Evolution; game theory; bounded rationality; Markov chain; stochastic stability.

Journal of Economic Literature Classification: C72, C73.

1. INTRODUCTION

While bounded rationality and learning models have been studied extensively in the last few years, game theory has been unsuccessful in explaining where the bounds on rationality should be placed in a model of boundedly rational agents.

Recently, Young [7] – [9] suggested an evolutionary model which explains how agents can make their choices on basis of their own preferences and a sample of occurrences in the recent past only. A two-player game is played repeatedly by the members of two large populations and agents in the same population have the same preferences. In every round, two agents, one from each population, are randomly selected to play the game. Each of the two agents simultaneously chooses a strategy in the game, and each agent has access to a random sample, drawn from the recent

*Previous version - licentiate paper - February 15, 2000. I would like to thank Martin Dufwenberg, Jens Josephson, Martin Meier, Vladimir Rodionov, Maria Saez-Marti, Jana Vyrastekova, Peyton Young, Karl Wärneryd, conference participants at the Stockholm School of Economics and the Tenth International Conference at Stony Brook for helpful comments, and Tatiana Damjanovic for assistance on computer issues. Special thanks to Kerstin Niklasson and Roland Nilsson for their fine editing. I am indebted to Jörgen Weibull, who has provided personal encouragement for this paper. Financial support from the Jan Wallander and Tom Hedelius Foundation is gratefully acknowledged. Remaining errors are solely due to the author.

history of play. They use their sample as a predictor of the behavior of the agent they face, and almost always play a best reply to the empirical strategy distribution of the opponent population in the sample. Occasionally, agents "mutate", however, and instead choose a strategy that is not a best reply to any possible sample from the recent history of play.

Using Young's [8] model, Saez-Marti and Weibull [4] consider agents that are "clever" in a certain sense. They study the effect of letting a share of one of the populations know the preferences of the opponent population, denoting these agents as "clever", in the Nash Demand Game. Saez-Marti and Weibull [4] first assume, that the population without clever agents plays its best reply to a sample of past strategies played by the other population and, second, that clever agents play a best reply to the opponent population's best reply to the clever agent's sample. In other words, clever agents try to anticipate their opponent's choice on the basis of the sample of strategies played by their own population. They show that Young's prediction is robust to the introduction of any share of clever agents less than one.

The purpose of the present paper is to demonstrate that this robustness holds for generic finite two-player games. Young [9] proves that the *adaptive learning* process in a generic class of finite games settles down in a minimal curb configuration which minimizes the stochastic potential in such games when the mutation rate goes to zero. My main result is that the *adaptive learning with* any share of *clever agents* less than one converges to the same minimal curb configuration as Young's adaptive learning process, when the mutation rate goes to zero. However, I show that the presence of clever agents may influence the long-run strategy distribution inside the minimal curb configuration.

I also analyze the question, which was asked in Saez-Marti and Weibull [4], how well clever agents fare among non-clever agents given that there are fixed population shares of clever and non-clever agents. More specifically, we consider the Matching-Pennies Game and find that the gain of clever agents depends on their share in the population. Moreover, clever agents do not only outsmart the agents in the other population but, indirectly, also the non-clever agents in their own population. As a result, non-clever agents in both populations earn expected negative payoffs while the "clever subpopulation" on average earns expected positive payoffs in this zero-sum game. The larger the share of clever agents, the larger the gain for the population. On the margin, an additional clever agent gains less as the share of clever agents increases, here called "decreasing returns to cleverness".

Saez-Marti and Weibull [4] show that the "clever" population gets the whole pie when playing with the "non-clever" population in the *Nash Demand Game*. However, I demonstrate that "cleverness" does not guarantee an advantage in asymmetric coordination games, even if all agents in one population are clever. In the *Strict Demand*

Game, where two players must coordinate to get exactly the size of the pie, otherwise they get nothing, the population without clever agents obtains the whole pie.

Young [9] shows that for 2×2 coordination games, in the limiting case when the mutation rate goes to zero, adaptive play converges to the *risk-dominant* convention - a repetition of the *risk-dominant* equilibrium. He also analyzes the case with two different sample sizes for two populations. I demonstrate that clever agents matter: if the sample sizes are different in the two populations, then the stochastically stable convention may differ from Young's [9] prediction. In the extreme case, when *all agents* in one population are clever, the outcome is *the risk-dominant convention* for the other population.[1] The intuition for this result is as follows. The clever agent in population 1 predicts the strategy of her opponent and also chooses the predicted strategy, because of the specific structure of 2×2 coordination games.

The paper is organized as follows. In Section two, I describe the unperturbed and perturbed versions of the adaptive play with clever agents. In Section three, I derive general results for the stationary distribution of this process. In Section four, the specific nature of the limiting distribution inside a minimal curb set is studied in detail. In Section five, I investigate the properties of 2×2 coordination games when agents have asymmetric information and Section six concludes. Proofs are given in the Appendix.

2. Adaptive play with clever agents

In the evolutionary model described below, I consider clever agents, introduced in Saez-Marti and Weibull [4] for the Nash bargaining game, in *two-player games.* The basic setting without clever agents is Young's [9] model.

Let Γ be a two-player game with finite strategy space, $X_1 \times X_2$, and payoff functions $u_i : X_1 \times X_2 \to \mathbf{R}$, $i = 1, 2$. I assume that there exist two finite populations of agents. In each discrete time period, $t = 1, 2, \ldots$, one agent is drawn at random from each of the populations to play the game. Agents in population 1 (2) can only play role 1 (2) in the game. Population 1 consists of clever and non-clever agents, in fixed population shares λ and $1 - \lambda$, respectively, while the agents in population 2 are only non-clever. All agents are equally likely to be drawn to play. An agent in role i chooses a strategy x_i^t from the set X_i at time t according to a rule defined below. The *play* at time t is the vector $x^t = (x_1^t, x_2^t)$. The *history of play* up to time t is the sequence $h^t = (x^{t-m+1}, \ldots, x^t)$.

Strategies are chosen as follows. Fix integers s and m, where $1 \leq s \leq m$. At time $t + 1$, each agent drawn to play the game inspects a sample of size s, taken without replacement from the history of play up to time t. The draws of samples

[1] To find the risk-dominant convention, we only consider payoffs of the non-clever population. The formal definition is found on pp.16.

are statistically independent across agents and time. A non-clever agent chooses a best reply to the opponent population's empirical strategy distribution in her sample. Clever agents are assumed to know the preferences of the other population and they use this knowledge to choose a best reply to the anticipated choice by their opponent. More precisely, a clever agent - these always play in role 1 - inspects her own population's play in her sample, and calculates player 2's best reply to this sample. Then, the clever agent chooses a best reply to this predicted strategy. If there are more than one best reply, an agent chooses each of these with positive probability. These probabilities will be specified later on.

Consider the sampling process to begin in period $t = m+1$ from some arbitrary initial sequence of m plays h^m. We then obtain a finite Markov chain on the state space $(X_1 \times X_2)^m = H$ of sequences of length m drawn from strategy space $X_1 \times X_2$, with an arbitrary initial state h^m. As we will see below, the resulting process is *ergodic*; thus, in the long-run, the initial state is irrelevant. Given a history $h^t = (x^{t-m+1}, ..., x^t)$ at time t, the process moves to a state of the form $h^{t+1} = (x^{t-m+2}, ..., x^t, x^{t+1})$, in the next period, a state called a *successor* of h^t.

The process moves from the current state h to a successor state h' in each period, according to the following transition rule. For each $x_i \in X_i$, let $p_i(x_i \mid h)$ be the conditional probability that agent i chooses x_i, given that the current state is h. We assume that $p_i(x_i \mid h)$ is independent of t and $p_i(x_i \mid h) > 0$ if and only if there exists a sample s such that x_i is a best reply to this sample for a non-clever agent or x_i is a best reply to the opponent's best reply to this sample for a clever agent in population 1. If $x = (x_1, x_2)$ is the rightmost element of h', the probability of moving from h to h' is $R^{m,s,\lambda,0}_{hh'} = p_1(x_1 \mid h)p_2(x_2 \mid h)$ if h' is a successor of h and $R^{m,s,\lambda,0}_{hh'} = 0$ if h' is not a successor of h. Following Young [10], we call the process $R^{m,s,\lambda,0}$ unperturbed *adaptive play with clever agents* with memory m, sample size s, and share λ of clever agents in population 1.

The perturbed process can be described as follows. In each period, there is a small probability $\varepsilon > 0$ that any drawn agent in role i experiments by choosing a random strategy from X_i instead of applying the best reply rule. The event that i experiments is assumed to be independent from the event that the other agent playing this game in the opponent role j, experiments. For every i, let $q_i(x_i \mid h)$ be the conditional probability that i chooses $x_i \in X_i$, given that i experiments and the perturbed process is in state h. Assume that $q_i(x_i \mid h)$ is independent of t and $q_i(x_i \mid h) > 0$ for all $x_i \in X_i$ and all h. Suppose that the perturbed process is in state h at time t. The probability is $\varepsilon(1-\varepsilon)$ that exactly one of the agents playing the game experiments and that the other does not. Conditional on this event, the transition probability of moving from h to h' is $Q^1_{hh'} = q_i(x_i \mid h)p_j(x_j \mid h)$, where $i \neq j$, if h' is a successor of h and x is the rightmost element of h' and $Q^1_{hh'} = 0$, if h' is not a successor of h. Similarly, ε^2 is the probability that both drawn agents experiment. Conditional on this event, the transition probability of moving from h to h' is $Q^2_{hh'} = q_1(x_1 \mid h)q_2(x_2 \mid h)$, if h' is a successor of h and x is the rightmost element of h' and $Q^2_{hh'} = 0$, if h' is not a successor of h. This gives the following transition probability of the perturbed Markov process:

$$R^{m,s,\lambda,\varepsilon}_{hh'} = (1-\varepsilon)^2 R^{m,s,\lambda,0}_{hh'} + 2\varepsilon(1-\varepsilon)Q^1_{hh'} + \varepsilon^2 Q^2_{hh'}. \tag{1}$$

The process $R^{m,s,\lambda,\varepsilon}$ is a denoted (perturbed) adaptive play with clever agents with memory m, sample size s, share λ of clever agents in population 1 and error rate ε.

As is usual in evolutionary models, two forces drive the perturbed Markov process. The first - $R^{m,s,\lambda,0}_{hh'}$ - is the *selection* rule. The second - $Q^1_{hh'}$ and $Q^2_{hh'}$ - is the *mutation*. Note that 1) if $\lambda = 0$, then $R^{m,s,0,\varepsilon}_{hh'}$ is Young's [9] adaptive learning, 2) if Γ is the Nash bargaining game, we are in the framework of Saez-Marti and Weibull [4].

3. How cleverness does not matter

In this section, I discover when the introduction of a positive share λ of clever agents does not change the long-run prediction of the model without clever agents, starting with useful definitions.

A *product set* of strategies is a set of form $C = C_1 \times C_2$, where each C_i is a non-empty subset of X_i, $i = 1, 2$. Let ΔC_i denote the set of probability distributions over C_i, and let $\Delta C_1 \times \Delta C_2$ denote the product set of such distributions. Let $BR_i(C_j)$ denote the set of strategies in X_i that are player i's best replies to some distribution $p_j \in \Delta C_j$, $i \neq j$. Denote $BR(C) = BR_1(C_2) \times BR_2(C_1)$.

Definition 1. *(Basu and Weibull [1]) A non-empty Cartesian product set $C = C_1 \times C_2 \subset X$ is* **closed under best replies** *(or C is a* **curb set***) if $BR(\Delta C_1 \times \Delta C_2) \subset C$. Such a set is a* **minimal curb set** *if it does not properly contain a curb set.*

It is straightforward to show that $BR(\Delta C_1 \times \Delta C_2) = C$ for any minimal curb set C. Following Young [9], a *span* of a subset $H' \subseteq H = (X_1 \times X_2)^m$, denoted by $S(H')$, is the product set of all pure strategies that appear in some history in H'. H' is a *minimal curb configuration* if its span is a minimal curb set.

We say that a *recurrent class* of the process $R^{m,s,\lambda,0}$ is a set of states such that there is zero probability of moving from any state in the class to any outside state and there is a positive probability of moving from any state in the class to any other state in the class.

I will work with generic games and need to introduce a generic condition, which is a common condition in economics models. For discussion of this issue see, for example, Samuelson [5, *pp.*30].

Given a two-player game Γ on the finite strategy space $X_1 \times X_2$, let $BR_i^{-1}(x_i)$ denote the set of all probability mixtures $p_j \in \Delta_j = \Delta X_j$, where $j \neq i$, such that x_i is a best reply to p_j. We will work with Young's [9] generic condition.

Definition 2. *Γ is a* **nondegenerate in best replies** *if for every player i and every $x_i \in X_i$, either $BR_i^{-1}(x_i)$ is empty or contains a non-empty subset open in the relative topology of Δ_j, where $j \neq i$.*

The following result shows that a prediction in generic games for the unperturbed process $R^{m,s,\lambda,0}$ with $\lambda \in (0,1)$ is the same as in Young [9] for the unperturbed process $R^{m,s,0,0}$. In other words, recurrent classes for process $R^{m,s,\lambda,0}$ are independent of λ.

Theorem 1. *Let Γ be a nondegenerate in best replies two-player game on the finite strategy space $X_1 \times X_2$. If s/m is sufficiently small, the unperturbed process $R^{m,s,\lambda,0}$ converges to a minimal curb configuration with probability one.*

Proof: See the Appendix.

If $\lambda = 1$, then Theorem 1 can fail. Consider the game in Figure 1.

	a	b
A	$12,1$	$0,0$
B	$4,0$	$4,1$
C	$8,1$	$3,0$

FIGURE 1

In this game, A is a best reply to a and B a best reply to b. For any history h, an agent in player position 2 has three opportunities. She can have a or b as the only best reply to a sample of player 1 from history or she can be indifferent between a and b. In the last case, we assume that an agent in player position 2 will randomize $50:50$ between two choices. A clever agent in player position 1 will also calculate these three opportunities and play A against a, B against b, and A in the third case. Hence, there are only strategies A and B in any sample of player 1 in the long-run if $\lambda = 1$. However, strategy C also belongs to a minimal curb set, for it is a best reply to, for example, mixed strategy $\frac{2}{7}a + \frac{5}{7}b$.

A process is said to be *irreducible* if and only if there is a positive probability of moving from any state to any other state in a finite number of periods. We will need the following definitions:

Definition 3. *(Young [10]) $R(\varepsilon)$ is a* **regular perturbed Markov process** *if $R(\varepsilon)$ is irreducible for every $\varepsilon \in (0, \varepsilon^*]$, and for every state h, $h' \in H$, $R_{hh'}(\varepsilon)$ approaches $R_{hh'}(0)$ at an exponential rate, i.e. $\lim_{\varepsilon \to 0} R_{hh'}(\varepsilon) = R_{hh'}(0)$ and if $R_{hh'}(\varepsilon) > 0$ for some $\varepsilon > 0$, then $0 < \lim_{\varepsilon \to 0} \frac{R_{hh'}(\varepsilon)}{\varepsilon^{r_{h \to h'}}} < \infty$ for some $r_{h \to h'} \geq 0$. The real number $r_{h \to h'}$ is the* **resistance** *of the transition $h \longrightarrow h'$.*

Lemma 1. *An adaptive play with clever agents is a regular perturbed Markov process.*

Proof: $R^{m,s,s',\lambda,\varepsilon}$ is a regular perturbed Markov process for the same reason as shown by Young [9] when he considers adaptive play. **End of proof.**

Definition 4. *(Young [7]) Let $\mu(\varepsilon)$ be the unique stationary distribution of an irreducible process $R(\varepsilon)$. A state h is* **stochastically stable** *if $\lim_{\varepsilon \to 0} \mu_h(\varepsilon) > 0$.*

Let process $R^{m,s,\lambda,0}$ have recurrent classes $E_1, ..., E_K$. For each pair of distinct recurrent classes, a *pq-path* is a sequence of states $\zeta = (h_p, ..., h_q)$ beginning in E_p and ending in E_q. The *resistance* of this path is the sum of the resistances on the edges composing it. Let r_{pq} be the least resistance over all pq-paths. Construct a complete directed graph with K vertices, one for each recurrent class. The weights on the directed edge $E_p \longrightarrow E_q$ is r_{pq}. A tree rooted at E_l is a set of $K - 1$ directed edges such that, from every vertex different from E_l, there is a unique directed path in the tree to E_l. The *resistance of such a rooted tree* $\Im(E_l)$ is the sum of resistances r_{pq} on its $K - 1$ edges. The *stochastic potential* $\rho(E_l)$ of a recurrent class E_l is the minimum resistance over all trees rooted at E_l. I will use the following result in the main theorem.

Theorem 2. *(Young [9]) Let $R(\varepsilon)$ be a regular perturbed Markov process and let $\mu(\varepsilon)$ be the unique stationary distribution of $R(\varepsilon)$ for $\varepsilon > 0$. Then, $\lim_{\varepsilon \to 0} \mu(\varepsilon) = \mu(0)$ exists and is a stationary distribution of $R(0)$. The stochastically stable states are precisely the states contained in the recurrent classes of $R(\varepsilon)$, having minimum stochastic potential.*

We are now in a position to state the main result.

Theorem 3. *Let Γ be a nondegenerate in best replies two-player game on the finite strategy space $X_1 \times X_2$. If s/m and ε are sufficiently small, s and m are sufficiently large and $\lambda \in (0, 1)$, the perturbed process $R^{m,s,\lambda,\varepsilon}$ puts arbitrarily high probability on the minimal curb configuration(s) that minimize the stochastic potential of the perturbed process $R^{m,s,0,\varepsilon}$.*

Proof: See the Appendix.

This theorem shows that strategies taken by agents in the two populations are the same for the perturbed process $R^{m,s,\lambda,\varepsilon}$ with $\lambda \in (0,1)$ and the perturbed process $R^{m,s,0,\varepsilon}$, without clever agents. In other words, the same recurrent classes will be chosen in the long-run by the perturbed process $R^{m,s,\lambda,\varepsilon}$ for all $\lambda \in [0,1)$. However, in the next section, it will be shown by means of an example that the distribution of strategies taken by agents in the two populations is different for different values of λ.

4. How cleverness matter

4.1. Intra-curb effects with $\lambda \in [0,1)$. Although the presence of clever agents does not influence the choice of the limiting curb set if $\lambda \in [0,1)$, as we saw in the previous section, I show here that it can influence the distribution of strategies *inside* the limiting curb configuration. This is clarified by means of the following example. Consider the Matching-Pennies game with the payoff matrix in Figure 2.

	a	b
A	$1,-1$	$-1,1$
B	$-1,1$	$1,-1$

Figure 2

Choose the parameters in adaptive learning with clever agents as follows: memory, $m=2$; sample size, $s=1$; proportion $\lambda \in [0,1]$ of clever agents in population 1.[2] Each state can be represented by a 1×4 block of A's, B's, a's, and b's, where the first two squares represent the previous two strategies of the agent in population 1 and the last two represent the previous two strategies of the agent in population 2. For example, state $ABab$ means that the agent in population 1 chose A two periods ago, and B one period ago, while the agent in population 2 chose a two periods ago, and b one period ago. There are 16 possible states for the process.

The asymptotic properties of the finite Markov process $R^{m,s,\lambda,\varepsilon}$ can be studied algebraically as follows. Let $z_1 = AAaa, ..., z_{16} = BBbb$ be an enumeration of the states, let R be a transition matrix of the Markov process $R^{m,s,\lambda,\varepsilon}$ on the finite state space $(\{A,B\}\times\{a,b\})^2$, and let $\mu(\varepsilon,\lambda) = \begin{pmatrix} \mu_{z_1}(\varepsilon,\lambda) \\ \vdots \\ \mu_{z_{16}}(\varepsilon,\lambda) \end{pmatrix}$ be a column vector of probability distribution on the finite state space $(\{A,B\}\times\{a,b\})^2$. Consider a

[2]For our purpose, we only need to consider case $\lambda \in [0,1)$.

system of linear equations

$$R \cdot \mu(\varepsilon, \lambda) = \mu(\varepsilon, \lambda), \text{ where } \mu(\varepsilon, \lambda) \geq 0 \text{ and } \sum_{l=1}^{16} \mu_{z_l}(\varepsilon, \lambda) = 1. \tag{2}$$

It is well known that this system (for the irreducible process $R^{m,s,\lambda,\varepsilon}$) always has exactly one solution $\mu(\varepsilon, \lambda)$, called a *stationary distribution* of the process $R^{m,s,\lambda,\varepsilon}$. From Theorem 2, it follows that

$$\lim_{\varepsilon \to 0} \mu(\varepsilon, \lambda) = \mu(0, \lambda), \tag{3}$$

where $\mu(0, \lambda)$ is the stationary distribution of $R^{m,s,\lambda,0}$. Note that the process $R^{m,s,\lambda,\varepsilon}$ has only one recurrent class - the whole state space $(\{A, B\} \times \{a, b\})^2$ - and without loss of generality, we can only analyze the unperturbed process $R^{m,s,\lambda,0}$. Solving the system of linear equations:

$$R \cdot \mu(0, \lambda) = \mu(0, \lambda), \tag{4}$$

gives the stationary distribution $\mu(0, \lambda)$ for different values of λ, where R is the matrix described in Figure 3. Empty squares in the matrix correspond to zeros. If an agent is indifferent between two pure strategies, then she is assumed to play either of them with equal probability.

Now, we can calculate the expected payoffs for both populations. It is enough only to consider the expected payoffs for the agent drawn from population 1 to play the game, since this is a zero-sum game.

A clever agent correctly predicts the only best reply of the other agent from population 2 if the agents in two last periods in population 1 played the same strategy. Hence, a clever agent always receives payoff $u_1^c(\lambda, z_j) = 1$ in the following 8 states: $AAaa$, $AAab$, $AAba$, $AAbb$, $BBaa$, $BBab$, $BBba$, $BBbb$. In the remaining 8 states, the expected payoff to a clever agent is zero. Therefore, the total expected payoff to a clever agent is

$$u_1^c(\lambda) = \sum_{j=1}^{16} \mu_{z_j}(0, \lambda) \, u_1^c(\lambda, z_j) = \sum_{j=1}^{4} \mu_{z_j}(0, \lambda) + \sum_{j=13}^{16} \mu_{z_j}(0, \lambda) > 0, \tag{5}$$

where $z_1 = AAaa, ..., z_4 = AAbb$, $z_{13} = BBaa, ..., z_{16} = BBbb$.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	AA aa		1-λ				λ										
2	ab				.5(1-λ)				.5(1+λ)								
3	ba		.5(1-λ)				.5(1+λ)										
4	bb								1								
5	AB aa									.5(1-.5λ)	.5(1-.5λ)			.25λ	.25λ		
6	ab											.25	.25			.25	.25
7	ba									.25	.25			.25	.25		
8	bb											.25λ	.25λ			.5(1-.5λ)	.5(1-.5λ)
9	BA aa	.5(1-.5λ)	.5(1-.5λ)			.25λ	.25λ										
10	ab			.25	.25			.25	.25								
11	ba	.25	.25			.25	.25										
12	bb			.25λ	.25λ			.5(1-.5λ)	.5(1-.5λ)								
13	BB aa									1							
14	ab											.5(1+λ)				.5(1-λ)	
15	ba									.5(1+λ)				.5(1-λ)			
16	bb											λ				1-λ	

FIGURE 3.

A non-clever agent in population 1 plays a best reply to an opponent's probability distribution in the sample. As a result of this behavior, she always receives payoff -1 in the following two states: $AAaa$ and $BBbb$, and payoff 1 in the following two states: $AAbb$ and $BBaa$. In the remaining 12 states, the expected payoff to a non-clever agent is zero. Therefore, the total expected payoff to a non-clever agent in population 1 is

$$u_1^n(\lambda) = \sum_{j=1}^{16} \mu_{z_j}(0,\lambda)\, u_1^n(\lambda, z_j) = \mu_{z_1}(0,\lambda) - \mu_{z_4}(0,\lambda) - \mu_{z_{13}}(0,\lambda) + \mu_{z_{16}}(0,\lambda). \tag{6}$$

The expected payoff to a clever agent is positive if at least one of the states $z_1, ..., z_4, z_{13}, ..., z_{16}$ shows up in the stationary distribution $\mu(0,\lambda)$, which is the case for any value of $\lambda \in (0,1]$. Hence, "cleverness" is an advantage in the Matching-Pennies game, because the expected payoff to a clever agent is higher than the expected payoff to a non-clever agent from the same population.[3] What is the expected average payoff

[3]In my view, it should be a general result that the expected payoff of the clever agent is equal to or higher than the expected payoff of the non-clever agent from the same population, inside any minimal curb set.

of population 1? Figure 4 shows how the expected payoffs to a clever agent, a non-clever agent and population 1 (on average) depend on the share λ of clever agents in population 1.

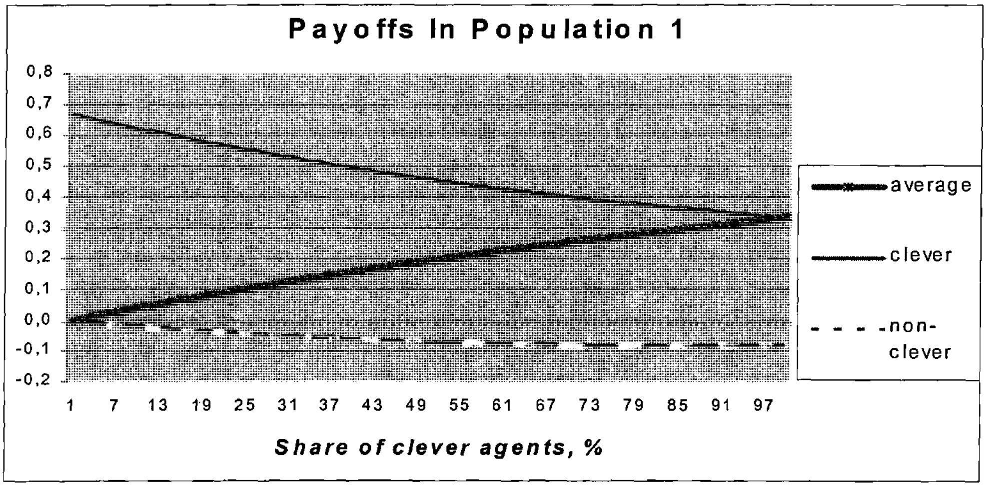

FIGURE 4.

There are different stationary distributions for different values of λ. The expected payoff to a non-clever agent in population 1 is negative and falling as the share of clever agents increases. The expected average payoff to population 1 is positive for $\lambda > 0$ and depends positively on the share of clever agents. Accordingly, all agents in population 2 earn a negative expected payoff, which is decreasing in λ. Clever agents outsmart agents from the other population. The larger is the share of clever agents in population 1, the smaller is the expected payoff to each clever agent. In this sense, the marginal return to cleverness is decreasing. Moreover, the presence of clever agents in population 1 imposes a negative externality on the non-clever agents in the same population.

4.2. Clever population against non-clever population, $\lambda = 1$. Consider an extreme case, where all agents in population 1 are clever. In this case, strategies that agents choose in the two populations may differ for the perturbed process $R^{m,s,1,\varepsilon}$ with all clever agents in population 1 and the perturbed process $R^{m,s,0,\varepsilon}$ without clever agents. The question now arises whether the clever agent who has more information about the opponent is better off than the non-clever agent. The answer depends on a game. The following examples illustrate this point. We start from the *Strict Demand* game and then compare that game with the *Nash Demand* game, studied in Saez-Marti and Weibull [4].

The Strict Demand Game. Consider two finite populations, 1 and 2, who periodically bargain pair-wise over their shares of a common pie. Let x denote the share of player 1, and let y denote the share of player 2. Suppose that all agents in population 1 have the same concave, increasing, and differentiable utility function, which is a function of the share x

$$u : [0,1] \to \mathbb{R}, \tag{7}$$

and all agents in population 2 have the same concave, increasing, and differentiable utility function as a function of the share y

$$v : [0,1] \to \mathbb{R}. \tag{8}$$

Without loss of generality, we can normalize u and v so that $u(0) = v(0) = 0$.

In each period $t = 1, 2, ...,$ one agent is drawn at random from each population. They play the *Strict Demand Game*, later *SDG*: player 1 demands some number $x \in (0,1]$, and simultaneously, player 2 demands some number $y \in (0,1]$. The outcomes and payoffs are as in Figure 5.

Demands	Outcomes	Payoffs
$x+y=1$	x, y	$u(x), v(y)$
$x+y \neq 1$	$0, 0$	$0, 0$

FIGURE 5

To keep the state space finite, we shall discretize demands. Let a finite set $D(\delta) = \{\delta, 2\delta, ..., 1-\delta\}$ be the space of demands. Furthermore, let $R^{m,s,1,\varepsilon}$ be an adaptive play with all clever agents in population 1. Let (x^t, y^t) denote the amounts demanded by the agents in population 1 and population 2 respectively in period t. At the end of period t, the state is

$$h^t = \left(\left(x^{t-m+1}, y^{t-m+1}\right), ..., (x^t, y^t)\right). \tag{9}$$

At the beginning of period $t+1$, the current clever agent, playing the game, draws a sample of size s from the $x-$values in h^t. Simultaneously and independently, the agent in population 2 also draws a sample of size s from the $x-$values in h^t.

A *conventional division* is a state of the form

$$h_x = ((x, 1-x), ..., (x, 1-x)), \tag{10}$$

where $0 < x < 1$. We say that a division $(x, 1-x)$ is *stochastically stable* for a given precision δ, if the corresponding convention h_x is stochastically stable.

Definition 5. *The* **basin of attraction** *of state h is the set of states h' such that there is a positive probability of moving from h' to h in a finite number of periods under the unperturbed process $R^{m,s,1,0}$.*

For every real number r, let $\lceil r \rceil$ denote the least integer greater than or equal to r.

Lemma 2. *For every $x \in D(\delta)$, the minimum resistance of moving from convention h_x to a state in some other basin of attraction is $\lceil sr_\delta(x) \rceil$, where*

$$r_\delta(x) = \frac{1}{1 + \frac{v(1-\delta)}{v(1-x)}}.$$

Proof: See the Appendix.

Lemma 3. *A division $(x, 1-x)$ is stochastically stable if and only if x maximizes the function $r_\delta(x)$ on $D(\delta)$.*

Proof. It follows from Theorem 2.

Corollary 1. *The division $(\delta, 1-\delta)$ is stochastically stable.*

Proof. x maximizes the function $r_\delta(x)$ on $D(\delta)$ at $x = \delta$. **End of the proof.**

The main result of this part follows immediately from the corollary.

Proposition 1. *Assume all agents in population 1 to be clever. Then, for every $\delta > 0$, there exists one stable division, which converges to $(x, y) = (0, 1)$ as $\delta \rightarrow 0$.*

The Nash Demand Game. In this subsubsection, I refresh an example analyzed by Saez-Marti and Weibull [4]: the *Nash Demand Game.* The set up is the same as in the case of the Strict Demand Game, but now players will also get a demanding part of the whole pie, if their common demand is less than 1. For the case $\lambda = 1$, Saez-Marti and Weibull [4] show that *for any precision $\delta > 0$, there exist at least one and at most two stable divisions, and as $\delta \to 0$, they converge to $(x, y) = (1, 0)$.*

5. 2×2 Coordination games and asymmetric sampling

So far, agents have been assumed to have the same amount of information in the sense of having the same sample size. What happens if the two populations differ in sample sizes? In this section, I analyze this question for 2×2 *Coordination Games* by assuming that the agents in population 1 have sample size s and the agents in population 2 have sample size s'. Let the error rate, ε, be the same for all agents, and let the memory, m, be the same for both populations and satisfy $m \geq \max\{2s, 2s'\}$. Altogether, these factors define a Markov process $R^{m,s,s',\lambda,\varepsilon}$ with memory m, sample sizes s and s', the share λ of clever agents in population 1, and the error rate ε. We denote any game in this class of 2×2 games by G.

Several research studies have analyzed the question of which equilibrium would be observed in 2×2 coordination games in the long-run. Two papers, Kandori, Mailath, and Rob [3] and Young [7], pioneer this topic. I will compare my results with predictions from the existing literature. Young [7] shows that the stochastically stable states of the perturbed process $R^{m,s,s,0,\varepsilon}$ correspond one to one with the risk-dominant conventions.

Consider a two-player coordination game G with payoff matrix in Figure 6.

	a	b
A	u_{Aa}, v_{Aa}	u_{Ab}, v_{Ab}
B	u_{Ba}, v_{Ba}	u_{Bb}, v_{Bb}

FIGURE 6

G is a *Coordination Game* with pure strategy Nash equilibria (A, a) and (B, b) if the following inequalities hold:

$$u_{Aa} > u_{Ba}, \; u_{Bb} > u_{Ab}, \; v_{Aa} > v_{Ab}, \; v_{Bb} > v_{Ba}. \tag{11}$$

The following concept from Young [7] will play a crucial role in our study of 2×2 coordination games.

Definition 6. *A* **convention** *is a state of the form $h_x = (x, x, ..., x)$, where x is a strict Nash equilibrium of G.*

Definition 7. *(Harsanyi and Selten [2]) Equilibrium (A, a) is* **risk-dominant** *if*

$$(u_{Aa} - u_{Ba})(v_{Aa} - v_{Ab}) \geq (u_{Bb} - u_{Ab})(v_{Bb} - v_{Ba}).$$

A convention h_x is said to be *risk-dominant* if the strict Nash equilibrium x is risk-dominant. It turns out that the risk-dominant convention is stochastically stable in Young's [9] framework if $s = s'$. The following proposition is a corollary of Theorems 1 and 3.

Proposition 2. *If $max\{s, s'\} \leq m/2$, then from any initial state, the unperturbed process $R^{m,s,s',\lambda,0}$ converges with probability one to a convention and locks in.*

If $\lambda \in (0,1)$ and $s = s'$, then Young's [9] result holds. More precisely, if $s/m \leq 1/2$, and s and m are sufficiently large, the stochastically stable states of the perturbed process $R^{m,s,s',\lambda,\varepsilon}$, as $\varepsilon \to 0$, correspond one to one with risk-dominant conventions.

Let us examine what happens if the first and the second populations have different sample sizes.

Proposition 3. *Suppose that $\lambda \in (0,1)$, $s < s'$, $s'/m \leq 1/2$, and s, s' and m are sufficiently large. Then, the stochastically stable states of the perturbed process $R^{m,s,s',\lambda,\varepsilon}$, as $\varepsilon \to 0$, correspond one to one with risk-dominant conventions.*

Proof: See the Appendix.

The intuition behind this result is that if the sample size of population 1 is less than the sample size of population 2, then agents in population 1 need to make less mistakes to switch to another strategy. This means that the role of the first population becomes crucial for finding the minimum stochastic potential. In population 1, all agents have the same sample size and clever agents, as before, play a best reply to a best reply of an agent from population 2. In a 2×2 coordination game, a clever agent in population 1 chooses the same strategy as an agent in population 2, with positive probability. A clever agent thus needs the same number of mistakes to switch to another strategy as a non-clever agent in population 2, if their respective sample size is the same. Hence, we have returned to Young's [7] framework.

Define

$$\alpha = (u_{Aa} - u_{Ba}) \,/\, (u_{Aa} - u_{Ab} - u_{Ba} + u_{Bb}) \tag{12}$$

and

$$\beta = (v_{Aa} - v_{Ab}) \,/\, (v_{Aa} - v_{Ab} - v_{Ba} + v_{Bb}). \tag{13}$$

Proposition 4. *Suppose that* $\lambda \in (0,1)$, $s > s'$, $s/m \leq 1/2$, *and* s, s' *and* m *are sufficiently large, and* ε *sufficiently small. Then, the state* $h_{(A,a)}$ *of the perturbed process* $R^{m,s,s',\lambda,\varepsilon}$ *is stochastically stable if and only if*

$$\min\{\lceil \alpha s \rceil, \lceil \beta s' \rceil\} \geq \min\{\lceil (1-\alpha)s \rceil, \lceil (1-\beta)s' \rceil\}. \tag{14}$$

Proof: See the Appendix.

What does proposition 3 say if the sample size $s \gg s'$? According to the proposition, the state $h_{(A,a)}$ is stochastically stable if and only if inequality (14) holds. For sufficiently large s', the inequality becomes $\beta \geq 1-\beta$. This means that only the payoffs of the agents in population 2 matter. If the sample size of the first population is "much larger" than the sample size of the second population, then the risk-dominant convention for population 2 will be stochastically stable.

Definition 8. *The convention* $h_{(A,a)}$ *is* **risk-dominant for population 2** *if:*

$$v_{Aa} - v_{Ab} \geq v_{Bb} - v_{Ba}. \tag{15}$$

If inequality (15) *is reversed, then convention* $h_{(B,b)}$ *is risk-dominant for population 2.*

Observe that (15) is equivalent to

$$\beta \geq 1-\beta. \tag{16}$$

The following proposition considers the extreme case when population 1 consists entirely of clever agents. It turns out that there is a discontinuity at this end of the spectrum: the whole population will move to the risk-dominant equilibrium of the population without clever agents.

Proposition 5. *Assume that every agent in population 1 is clever. If max* $\{s, s'\} \leq m/2$, *and* s, s' *and* m *are sufficiently large, the stochastically stable states of the perturbed process* $R^{m,s,s',1,\varepsilon}$, *as* $\varepsilon \to 0$, *correspond one to one with the risk-dominant convention(s) of the population without clever agents.*

Proof: See the Appendix.

Corollary 2. *Consider a two-player symmetric coordination game, G^S, with the following payoff matrix in Figure 7*

$$\begin{array}{c|c|c|} & a & b \\ \hline A & u,u & c,d \\ \hline B & d,c & v,v \\ \hline \end{array} \tag{17}$$

FIGURE 7

where $u > d$, $v > c$. Find α and β from (12) and (13):

$$\alpha = \frac{u-d}{u-d-c+v} = \beta. \tag{18}$$

This means that the stochastically stable state(s) of the perturbed process $R^{m,s,s',\lambda,\varepsilon}$ correspond(s) one to one with the risk-dominant convention(s), for any share of clever agents in population 1.

Corollary 3. *If $d = c$ in (17), then a two-player symmetric coordination game becomes a two-player doubly symmetric coordination game.*[4] *It follows straightforward that the stochastically stable state(s) of the perturbed process $R^{m,s,s',\lambda,\varepsilon}$ correspond(s) one to one with the Pareto dominant convention(s), for any share of clever agents in population 1.*[5]

The results for two-player 2×2 coordination games are summarized in Figure 8, showing the stochastically stable states for different λ, s, and s'.

	$\lambda = 0$	$\lambda \in (0,1)$	$\lambda = 1$
$s = s'$	*risk-dominant*	*risk-dominant*	*risk-dominant(2)*
$s > s'$	*	*	*risk-dominant(2)*
$s < s'$	*	*risk-dominant*	*risk-dominant(2)*

FIGURE 8

Here

[4]A symmetric two-player game is doubly symmetric if $A^T = A$, where A is the symmetric payoff matrix. See also Weibull [6].

[5]We say that convention $h_{(A,a)}$ is Pareto dominant if $u \geq v$, and convention $h_{(B,b)}$ is Pareto dominant if the reversed inequality holds.

1. *risk-dominant* means that the outcome is *risk-dominant* convention.
2. * means that the convention $h_{(A,a)}$ is stochastically stable if inequality (14) holds. If inequality (14) is reversed, then the convention $h_{(B,b)}$ is stochastically stable.
3. *risk-dominant(2)* means that the outcome is a *risk-dominant* convention for population 2.

The question now arises whether the population with more information about the opponent, population 1, is better off. The answer depends on the structure of the game.

Consider the following games in Figure 9, and the process $R^{m,s,s',\lambda,\varepsilon}$ with $s = s'$.

	a	b
A	3, 2	0, 0
B	0, 0	1, 4

Game 1

	a	b
A	3, 4	4, 3
B	0, 0	5, 2

Game 2

FIGURE 9

Then (A, a) is a risk-dominant equilibrium and $h_{(B,b)}$ is a risk-dominant convention for population 2 in both games. Hence, if $\lambda \in [0, 1)$, then convention $h_{(A,a)}$ is stochastically stable (Proposition 2), otherwise if $\lambda = 1$, then convention $h_{(B,b)}$ is stochastically stable (Proposition 5). Cleverness (or $\lambda = 1$) is a disadvantage in Game 1 (the clever population switches from 3 in $h_{(A,a)}$ to 1 in $h_{(B,b)}$) and an advantage in Game 2 (the clever population switches from 3 in $h_{(A,a)}$ to 5 in $h_{(B,b)}$).

6. CONCLUDING REMARKS

In this paper, I answer some questions asked in Saez-Marti and Weibull [4]. They study the consequences of letting some agents play a myopic best reply to the myopic best reply in Young's [8] bargaining model, which is how they introduce "cleverness" of players. Saez-Marti and Weibull [4] ask whether their results can be generalized. I use the "cleverness" approach from their paper to analyze *generic* two-player games in Young's [9] set-up. The resulting Markov process is denoted as an adaptive play with clever agents.

Saez-Marti and Weibull [4] prove that an introduction of any share of clever agents less than one in the special case will not change the long-run behavior for the Nash Demand Game. I have shown this result to be robust in *generic* two-player games: *adaptive learning with clever agents* will settle down in a minimal curb configuration, which minimizes the stochastic potential for adaptive learning without clever agents. However, the share of clever agents does matter inside the minimal curb configuration, as shown in the Matching-Pennies Game, where the gain of clever agents depends on the share of these agents in the population.

Furthermore, in the extreme case, if the share of clever agents equals one, then we have a discontinuity in the following sense. In the case of all clever agents in one of the populations, the stochastically stable states might differ from previous ones. Saez-Marti and Weibull [4] find this discontinuity for the Nash Demand Game.

I also study whether it is advantageous to be a member of the population consisting of clever agents only, and show the answer to be ambiguous even in coordination games. On the one hand, Saez-Marti and Weibull [4] show that the "clever" population gets the whole pie in the *Nash Demand Game.* On the other hand, in the *Strict Demand Game,* where two players must coordinate to get exactly the size of the pie, otherwise they both get nothing, the population without clever agents obtains the whole pie. Hence "cleverness" does not guarantee an advantage in general coordination games, even if all agents in one of the populations are clever.

I analyze the properties of 2×2 coordination games when agents have asymmetric information. It is a classical result (see Young [7] and Kandori, Mailath, and Rob [3] - the first, among others) for the symmetric setting that in the limiting case, when the mutation rate goes to zero, the *risk-dominant* equilibrium will be observed as the long-run outcome in such games. I prove that if the sample sizes are different in the two populations, the long-run outcome may differ from this prediction.

Appendix.

Proof of Theorem 1: This proof is similar to the proof in [9, *Theorem* 7.2]. We need only show that clever agents will change nothing.

Given a positive integer s, we say that the probability distribution $p_i \in \Delta_i$ has *precision* s if sp_i is an integer for all $x_i \in X_i$. We shall denote the set of all such distributions by Δ_i^s. For each subset $Y_i \subset X_i$, let $\Delta_i^s(Y_i)$ denote the set of distributions $p_i \in \Delta_i^s$ such that $p_i(x_i) > 0$ implies $x_i \in Y_i$. For each positive integer s, let $BR_i^s(X_{-i})$ be the set of pure-strategy best replies by a non-clever agent in population i to some product distribution $p_{-i} \in \Delta_{-i}^s(X_{-i}) = \Delta_j^s(X_j)$, where $j \neq i$. Similarly, $BR_i^s(Y_{-i})$ denotes the set of all best replies by a non-clever agent in population i to some product distribution $p_{-i} \in \Delta_{-i}^s(Y_{-i})$. Let $BR_1^s(X_1)$ be the set of pure-strategy best replies by a clever agent in population 1 to some product distribution $p_1 \in \Delta_1^s(X_1)$.

For each product set Y and an agent in population i, define the mappings

$$\beta_i(Y) = Y_i \cup BR_i(Y_{-i}) \text{ and } \beta_1'(Y) = Y_1 \cup BR_1(Y_1), \tag{19}$$

and let $\beta(Y) = [\beta_1(Y) \cup \beta_1'(Y)] \times \beta_2(Y)$, where $\beta_1'(Y)$ is the mapping for the clever agents from population 1 and $\beta_i(Y)$ is the mapping for the non-clever agent i. Note that $\beta(Y) = [\beta_1(Y) \cup BR_1(Y_1)] \times \beta_2(Y)$. Similarly, for each integer $s \geq 1$ let

$$\beta_i^s(Y) = Y_i \cup BR_i^s(Y_{-i}),\ \beta_1'^s(Y) = Y_1 \cup BR_1^s(Y_1) \tag{20}$$

and

$$\beta^s(Y) = [\beta_1^s(Y) \cup \beta_1'^s(Y)] \times \beta_2^s(Y). \tag{21}$$

In the same way, as it appears in the proof in [9, *Theorem* 7.2], we can show that $\beta^s(Y) = \beta(Y)$ for all sufficiently large s.

Consider the process $R^{m,s,\lambda,0}$. We show that if s is large enough and s/m is small enough, the spans of the recurrent classes correspond one to one with the minimal curb sets of game Γ.

Fix a recurrent class E_k of $R^{m,s,\lambda,0}$, and choose any $h^0 \in E_k$ as the initial state. We shall show that the span of E_k, $S(E_k)$, is a minimal curb set. As shown in the proof in [9, *Theorem* 7.2], there is a positive probability of reaching a state h^1 where the most recent s entries involve a repetition of some fixed $x^* \in X$, because there is a positive probability that a non-clever agent will be chosen from population 1 in every period. Note that $h^1 \in E_k$, because E_k is a recurrent class. Let $\beta^{(j)}$ denote the j-fold iteration of β and consider the nested sequence

$$\{x^*\} \subseteq \beta(\{x^*\}) \subseteq \beta^{(2)}(\{x^*\}) \subseteq ... \subseteq \beta^{(j)}(\{x^*\}) \subseteq ... \tag{22}$$

Since X is finite, there exists some point at which this sequence becomes constant, say

$$\beta^{(j)}(\{x^*\}) = \beta^{(j+1)}(\{x^*\}) = Y^*. \tag{23}$$

By construction, Y^* is a curb set.

The proof that Y^* is, in fact, a minimal curb set is the same as in the proof in [9, *Theorem* 7.2]. **End of proof.**

Proof of Theorem 3: It follows immediately from Theorems 1 that minimal curb configurations are recurrent classes of the regular perturbed Markov process $R^{m,s,\lambda,\varepsilon}$. By Theorem 2, one (or some) of this minimal curb configuration(s) is (are) stochastically stable. We must show that this minimal curb configuration is the same as in the absence of the clever agents.

Take any two recurrent classes, two minimal curb configurations, E_p and E_q. Note that every mistake made in population 2 can only influence the behavior of the non-clever agents in population 1. It means that all mistakes made in population 2 have the same effect for both processes $R^{m,s,\lambda,\varepsilon}$ and $R^{m,s,0,\varepsilon}$.

Suppose that l mistakes in a row in population 1 are necessary to move process $R^{m,s,0,\varepsilon}$ from recurrent class E_p to recurrent class E_q. The clever agents in process $R^{m,s,\lambda,\varepsilon}$ anticipate this. Hence, if there were less than l mistakes in population 1, then the clever agent in role 1 expects an agent in role 2 to play as if they are in recurrent class E_p. There must be at least l mistakes in population 1 to change these expectations of the clever agent. It means that there must be at least l mistakes in population 1 to move process $R^{m,s,\lambda,\varepsilon}$ from recurrent class E_p to recurrent class E_q. At the same time, there is a positive probability that only non-clever agents will be chosen from population 1 in every period. Therefore, it is enough to make exactly l mistakes in population 1 to move process $R^{m,s,\lambda,\varepsilon}$ from recurrent class E_p to recurrent class E_q. **End of proof.**

Proof of Lemma 2: Suppose that the process is in the convention h_x, where $x \in D(\delta)$. Let π be a path of the least resistance from h_x to a state that is in some other basin of attraction. Clearly, π must pass through some state $\mathbf{w}$ such that some best reply of agent in population 2 to a sample from $\mathbf{w}$ is different from $1-x$. Let $\mathbf{w}$ be the first such a state.

To compute the least number of mistakes necessary to exit from convention h_x, it suffices to consider, for every $x' \neq x$, the least number of initial mistakes x' by the agents in population 1 that will cause an agent in population 2 to reply with $1-x'$. The number of mistakes in population 2 does not matter in this setting, because both agents only look at the $x-$values in h^t.

Choose an arbitrary $x' \neq x$. Suppose that the agents in population 1 make j successive demands of x' that cause some agent's best reply in population 2 to switch to $1-x'$ instead of $1-x$. We can assume that $j \leq s$. When the agent in population 2 samples these j mistaken demands x', together with $s-j$ of the previous "conventional" demands x, she switches to $1-x'$ provided that

$$\frac{j}{s}v(1-x') \geq \frac{s-j}{s}v(1-x),$$

that is

$$j \geq \frac{v(1-x)}{v(1-x')+v(1-x)}s.$$

Over all feasible $x' \neq x$, the minimum value of j occurs when $x' = \delta$ and

$$j = \frac{1}{1+\frac{v(1-\delta)}{v(1-x)}}s.$$

Hence, the lowest number of mistakes to exit from the h_x–basin of attraction is $\lceil sr_\delta(x) \rceil$, where

$$r_\delta(x) = \frac{1}{1 + \frac{v(1-\delta)}{v(1-x)}}.$$

This completes the proof of Lemma 2.

Proof of Propositions 3 and 4: It is straightforward to show that the resistances to transiting between the absorbing states $h_{(A,a)}$ and $h_{(B,b)}$ are

$$r^{s,s'}_{Aa\to Bb} = \min\{\lceil \alpha s \rceil, \lceil \beta s' \rceil, \lceil \beta s \rceil\} \tag{24}$$

and

$$r^{s,s'}_{Bb\to Aa} = \min\{\lceil (1-\alpha)s \rceil, \lceil (1-\beta)s' \rceil, \lceil (1-\beta)s \rceil\}, \tag{25}$$

where α and β are from (12) and (13).

The last terms $\lceil \beta s \rceil$ and $\lceil (1-\beta)s \rceil$ appear because of the clever agents in population 1, with sample size s, who imitate the behavior of an agent in population 2 in a 2×2 coordination game.

By Theorem 2, $h_{(A,a)}$ is stochastically stable if and only if

$$r^{s,s'}_{Aa\to Bb} \geq r^{s,s'}_{Bb\to Aa}. \tag{26}$$

Let $s > s'$, then we have

$$r^{s,s'}_{Aa\to Bb} = \min\{\lceil \alpha s \rceil, \lceil \beta s' \rceil, \lceil \beta s \rceil\} = \min\{\lceil \alpha s \rceil, \lceil \beta s' \rceil\} \tag{27}$$

and

$$r^{s,s'}_{Bb\to Aa} = \min\{\lceil (1-\alpha)s \rceil, \lceil (1-\beta)s' \rceil, \lceil (1-\beta)s \rceil\} = \min\{\lceil (1-\alpha)s \rceil, \lceil (1-\beta)s' \rceil\}, \tag{28}$$

which is exactly the statement of proposition 4.

If $s < s'$, then for sufficiently large s

$$r^{s,s'}_{Aa\to Bb} = \min\{\lceil \alpha s \rceil, \lceil \beta s' \rceil, \lceil \beta s \rceil\} = \min\{\lceil \alpha s \rceil, \lceil \beta s \rceil\} = s\min\{\alpha, \beta\} \tag{29}$$

and

$$r^{s,s'}_{Bb \to Aa} = \min\left\{\lceil(1-\alpha)s\rceil, \lceil(1-\beta)s'\rceil, \lceil(1-\beta)s\rceil\right\} = s\min\left\{(1-\alpha),(1-\beta)\right\}. \tag{30}$$

What does it mean that

$$\min\{\alpha,\beta\} \geq \min\{(1-\alpha),(1-\beta)\} \tag{31}$$

from a "risk-dominance" point of view? Suppose that

$$\min\{\alpha,\beta\} = \alpha. \tag{32}$$

Hence,

$$\min\{(1-\alpha),(1-\beta)\} = 1-\beta. \tag{33}$$

Find the domain where (31), (32), (33) and $0 \leq \alpha,\beta \leq 1$ hold. We have the following inequalities:

$$\alpha+\beta \geq 1 \tag{34}$$

and

$$0 \leq \alpha \leq \beta \leq 1. \tag{35}$$

Suppose now that

$$\min\{\alpha,\beta\} = \beta. \tag{36}$$

Hence,

$$\min\{(1-\alpha),(1-\beta)\} = 1-\alpha. \tag{37}$$

Find the domain where (31), (36), (37) and $0 \leq \alpha,\beta \leq 1$ hold. We get the following inequalities:

$$\alpha+\beta \geq 1 \tag{38}$$

and

$$0 \leq \beta \leq \alpha \leq 1. \tag{39}$$

Hence, the inequality

$$\min\{\alpha,\beta\} \geq \min\{(1-\alpha),(1-\beta)\} \tag{40}$$

is equivalent to inequalities

$$\alpha+\beta \geq 1,\ 0 \leq \alpha,\beta \leq 1. \tag{41}$$

By definition, equilibrium (A,a) is risk-dominant if

$$\alpha\beta \geq (1-\alpha)(1-\beta),\ 0 \leq \alpha,\beta \leq 1, \tag{42}$$

or

$$\alpha+\beta \geq 1,\ 0 \leq \alpha,\beta \leq 1, \tag{43}$$

which is exactly inequalities (41), and the statement of proposition 3 follows immediately. **End of proof.**

Proof of proposition 5: As in the proof of the previous proposition, it is straightforward to show that the resistances of transitions between the absorbing states $h_{(A,a)}$ and $h_{(B,b)}$ for sufficiently large s and s' are

$$r^{s,s'}_{Aa\to Bb} = \min\{\lceil\beta s'\rceil, \lceil\beta s\rceil\} = \beta\min\{s',s\} \tag{44}$$

and

$$r^{s,s'}_{Bb\to Aa} = \min\{\lceil(1-\beta)s'\rceil, \lceil(1-\beta)s\rceil\} = (1-\beta)\min\{s',s\}. \tag{45}$$

Note that there are only clever agents in population 1 and therefore, the sample of population 2 is irrelevant.

The convention $h_{(A,a)}$ is stochastically stable if and only if

$$r^{s,s'}_{Aa\to Bb} \geq r^{s,s'}_{Bb\to Aa}, \tag{46}$$

or

$$\beta \geq 1-\beta. \tag{47}$$

This means, see (16), that the convention $h_{(A,a)}$ is risk-dominant for population 2.

The same logic can be used if the convention $h_{(B,b)}$ is stochastically stable. Hence, the stochastically stable state(s) of the perturbed process $R^{m,s,s',1,\varepsilon}$ correspond one to one with the risk-dominant convention(s) of population 2. **End of proof.**

REFERENCES

[1] K. Basu and J. Weibull, Strategy subsets closed under rational behavior, *Economics Letters* **36** (1991), 141-146.

[2] J. Harsanyi and R. Selten, "*A General Theory of Equilibrium in Games*", Cambridge, MIT Press, 1988.

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

[4] M. Saez-Marti and J. Weibull, Clever agents in Young's evolutionary bargaining model, *J. Econ. Theory* **86** (1999), 268-279.

[5] L. Samuelson, "*Evolutionary Games and Equilibrium Selection*", Massachusetts, MIT Press, 1997.

[6] J. Weibull, "*Evolutionary Game Theory*", Massachusetts, MIT Press, 1995.

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

[8] P. Young, An evolutionary model of bargaining, *J. Econ. Theory* **59** (1993), 145-168.

[9] P. Young, "*Individual Strategy and Social Structure*", Princeton University Press, 1998.

Stochastic Imitation in Finite Games

JENS JOSEPHSON AND ALEXANDR MATROS*
DEPARTMENT OF ECONOMICS
STOCKHOLM SCHOOL OF ECONOMICS
P.O. BOX 6501, S-113 83 STOCKHOLM, SWEDEN

January 21, 2001

3

ABSTRACT. In this paper we model an evolutionary process with perpetual random shocks, where individuals sample population-specific strategy and payoff realizations and imitate the most successful behavior. For finite n-player games we prove that in the limit, as the perturbations tend to zero, only strategy-tuples in minimal sets closed under the better-reply graph will be played with positive probability. If the strategy-tuples in one such minimal set have strictly higher payoffs than all outside strategy-tuples, then the strategy-tuples in this set will be played with probability one in the limit, provided the minimal set is a product set.

Keywords: Evolutionary game theory, bounded rationality, imitation, Markov chain, stochastic stability, better replies, Pareto dominance.

JEL classification: C72, C73.

1. INTRODUCTION

In most game-theoretical models of learning, the agents are assumed to know a great deal about the structure of the game, such as their own payoff function and all players' available strategies. However, for many applications this assumption is neither reasonable nor necessary. In many cases agents may not even be aware that they are playing a game. Moreover, equilibrium play may be achieved even with agents who have very little knowledge of the game. This observation was made already in 1950 by John F. Nash. In his unpublished Ph.D. thesis (1950), he referred to it as "the 'mass-action' interpretation of equilibrium points." Under this interpretation "it is

*The authors are grateful for helpful comments and suggestions from Jörgen Weibull. We have also benefited from comments by Bo Becker, Ken Binmore, Maria Sáez-Marti, Philippe Solal, and seminar participants at the Stockholm School of Economics 1999, the Young Economist Meeting 2000, and the First World Congress of the Game Theory Society 2000. We gratefully acknowledge financial support from the Jan Wallander and Tom Hedelius Foundation.

unnecessary that the participants have full knowledge of the total structure of the game, or the ability and inclination to go through any complex reasoning processes. But the participants are supposed to accumulate empirical information on the relative advantages of the various pure strategies at their disposal."

In the present paper we develop a model in this spirit, where individuals are only required to know their own available strategies and a sample of the payoffs that a subset of these strategies have earned in the past. In spite of this weak assumption, our model predicts equilibrium play in many games. We use an evolutionary framework with perpetual random shocks similar to Young (1993), but our assumption of individual behavior is different from his. Whereas the individuals in his model play a myopic best reply to a sample distribution of their opponents' strategies, the individuals in our model imitate other agents in their own population. Imitational behavior has both experimental, empirical, and theoretical support. For experimental support of imitation see for example Huck *et al.* (1999, 2000) and Duffy and Feltovich (1999), for empirical support see Graham (1999), Wermers (1999), and Griffiths *et al.* (1998), and for theoretical support see Björnerstedt and Weibull (1996) and Schlag (1998, 1999).

More specifically we assume that in every period individuals are drawn at random from each of n arbitrary-size populations to play a finite n-player game. Each of these individuals observes a sample from a finite history of her population's past strategy and payoff realizations. She thereafter imitates by choosing the most attractive strategy in her sample. This could for instance be the strategy with the highest average payoff, or the strategy with the highest maximum payoff.[1,2] With some small probability the individuals also make errors or experiment, and instead choose any strategy at random from their set of strategies.[3] This results in an ergodic Markov process, which we denote *imitation play*, on the space of histories. We study the stationary distribution of this process as the experimentation probability tends to zero.

Imitation in a stochastic setting has previously been studied by Robson and Vega-Redondo (1996), who modify the framework of Kandori *et al.* (1993) to allow for random matching. More precisely, they assume that agents each period are randomly matched for a finite number of rounds and tend to adopt the strategy with the highest

[1]In the special case when each population consists of only one individual, this behavior can be interpreted as a special kind of reinforcement learning.

[2]This behavior is related to one of the interpretations of individual behavior in Osborne and Rubinstein (1998), where each individual first samples each of her available strategies once and then chooses the strategy with the highest payoff realization.

[3]An alternative interpretation, which provides a plausible rationale for experimentation and is consistent with the knowledge of agents in the model, is that if and only if the sample does not contain all available strategies, then with a small probability the individual instead picks a strategy not included in the sample at random.

average payoff across the population. Robson and Vega-Redondo (1996) assume either single- or two-population structures and obtain results for symmetric 2×2 games and 2-player games of common interest.

However, our model differs considerably from this and other stochastic learning models, and has several advantages. First, we are able to prove general results, applicable to any finite n-player game, about the limiting distribution of imitation play. We are thus not restricted to the two classes of games in Robson and Vega-Redondo (1996), or even to a generic class of games, as in Young (1998). Second, we find that this distribution has some interesting properties. For instance, it puts probability one on an efficient set of outcomes in a large class of n-player games. Third, the speed of convergence of our process is relatively high. We show that in 2×2 Coordination Games, for small experimentation probabilities the expected first passage time may be considerably shorter than in Young (1993), Kandori *et al.* (1993), and Robson and Vega-Redondo (1996).

The perturbed version of imitation play is a regular perturbed Markov process. This implies that the methods employed by Young (1993) can be used to calculate the states that will be played with positive probability by the stationary distribution of the process as the experimentation probability tends to zero, the *stochastically stable states.* We prove three results which facilitate this calculation and enable us to characterize the set of such states. These results hold in finite n-player games provided the information available to the agents is sufficiently incomplete and their sample size is sufficiently large.

First, we show that from any initial state the unperturbed version of imitation play converges to a state which is a repetition of a single strategy-tuple. For this reason we denote such states *conventions.* Hence, the stochastically stable states of the process belong to the set of conventions. Note that our definition of a convention differs from Young's (1993) in that it does not require the repeated strategy-tuple to be a strict Nash equilibrium.

Second, we prove that in the limit, as the experimentation probability tends to zero in the perturbed process, only strategy-tuples in particular subsets of the strategy-space are played with positive probability. These sets, which we denote *minimal constructions,* are defined in the following way. By drawing directed edges from each strategy-tuple to strategy-tuples which gives exactly one of the players at least as high payoff, a *better-reply graph* is obtained. We define a *construction* as a set of strategy-tuples from which there are no outgoing such edges. A minimal construction is a minimal set with this property. Minimal constructions are similar to Sobel's (1993) definition of *non-equilibrium evolutionary stable* (*NES*) sets for 2-player games and to what Nöldeke and Samuelson (1993) call *locally stable components* in their analysis of extensive form games. They are also closely related to *minimal sets closed under better* *replies* (Ritzberger and Weibull, 1995). We show that every minimal set closed under better replies contains a minimal construction and that if a minimal construction is a product set, then it is also a minimal set closed under better replies. The relationship between minimal constructions and the limiting distribution of imitation play should be contrasted with Young's (1998) finding that adaptive play for generic games selects strategy-tuples in minimal sets closed under best replies.

Finally, we show that in a certain class of games, imitation play selects efficient outcomes. If the strategy-tuples in a minimal construction have strictly higher payoffs than all outside strategy-tuples, then the strategy-tuples in this set will be played with probability one in the limit, provided the minimal construction is a product set. This is a generalization of Robson and Vega-Redondo's (1996) result that a Pareto-dominant strategy-tuple under certain conditions corresponds to a unique stochastically stable state in 2-player games of common interest.

Applied to 2×2 games these three results give clear predictions. In Coordination Games where one equilibrium strictly Pareto dominates the other, imitation play selects the strictly Pareto-superior Nash-equilibrium. This result differs sharply from the predictions in Young's (1993) and Kandori *et al.* (1993) models, where the stochastically stable states correspond to the risk-dominant equilibria, but it is consistent with the predictions of Robson and Vega-Redondo's (1996) model for symmetric Coordination Games. However, if neither equilibria Pareto dominates the other, the latter model may select the risk-dominant equilibrium, whereas both of the equilibria are played with positive probability on in our model. In games without pure Nash-equilibria all of the four conventions are stochastically stable.

The paper is organized as follows. In Section 2 we define the unperturbed and perturbed versions of imitation play. In Section 3 we derive general results for the limiting distribution of the process. In Section 4 we investigate some of the properties of our solution concept, the minimal construction. In Section 5 we apply our results to 2×2 games and compare our findings to those in previous literature. In Section 6 we conclude.

2. The Model

The model we describe below is similar to Young (1993), but the sampling procedure is modified and the agents in our model employ a different decision-rule. Let Γ be a n-player game in strategic form and let X_i be the finite set of pure strategies x_i available to player $i \in \{1, ..., n\} = N$. Define the product sets $X = \prod_i X_i$ and $\Delta = \prod_i \Delta_i$, where $\Delta_i = \{p \in R_+^{|X_i|} \mid p'\mathbf{1} = 1\}$ is the simplex of individual strategies. Let $C_1, ..., C_n$ be n finite and non-empty populations of agents. These populations need not be of the same size, nor need they be large. Each member of population C_i is a candidate to play role i in the game Γ. All agents in population C_i have payoffs represented by the utility function $\pi_i : X \to \Pi_i$, where $\Pi_i \subset R$.[4] Expected payoffs are represented by the function $u_i : \Delta \to R$. Note that we write "players" when referring to the game Γ and "agents" or "individuals" when referring to the members of the populations.

Play proceeds as follows. Let $t = 1, 2, ...$ denote successive time periods. The game Γ is played once every period. In period t, one agent is drawn at random from each of the n populations and assigned to play the corresponding role. An individual in role i chooses a pure strategy x_i^t from a subset of her strategy space X_i according to a rule that will be defined below. The strategy-tuple $x^t = (x_1^t, ..., x_n^t)$ is recorded and referred to as *play* at time t. The *history* of plays up to time t is the sequence $h^t = (x^{t-m+1}, ..., x^t)$, where m is a given positive integer, the *memory size* of all individuals.

Let h be an arbitrary history. Denote by $w_i = (x_i^{k_1}, ..., x_i^{k_s}) \in X_i^s$ a sample of s, $1 \leq s \leq m$, elements from the m most recent strategy choices by individuals in population C_i and by $v_i = (\pi_i^{k_1}, ..., \pi_i^{k_s}) \in \Pi_i^s$ the corresponding payoff realizations. For any history h, the *maximum average correspondence*, $\alpha_i : X_i^s \times \Pi_i^s \twoheadrightarrow X_i$, maps each pair of strategy sample w_i and payoff sample v_i to the strategy (or the set of strategies) with the highest average payoff in the sample. Following Young (1993), we can think of the sampling process as beginning in period $t = m+1$ from some arbitrary initial sequence of m plays h^m. In this period and every period thereafter, each agent in player position i inspects a pair (w_i, v_i) and plays a strategy $x_i \in \alpha_i(w_i, v_i)$. This defines a finite Markov process on the finite state space $H = X^m$ of histories. Given a history $h^t = (x^{t-m+1}, ..., x^t)$ at time t, the process moves in the next period to a state of the form $h^{t+1} = (x^{t-m+2}, ..., x^t, x^{t+1})$. Such a state is called a *successor* of h^t. The process moves from the current state h to a successor state h' in each period according to the following transition rule. For each $x_i \in X_i$, let $p_i(x_i \mid h)$ be the probability that agent i chooses strategy x_i. We assume $p_i(x_i \mid h) > 0$ if and only if there exists a sample of population-specific strategy choices w_i and payoff outcomes v_i, such that $x_i \in \alpha_i(w_i, v_i)$. If x is the rightmost element of h', the probability of moving from h to h' is $B_{hh'}^{m,s,0} = \prod_{i=1}^n p_i(x_i \mid h)$ if h' is a successor of h and $B_{hh'}^{m,s,0} = 0$ if h' is not a successor of h. We call the process $B^{m,s,0}$ *imitation play with memory m and sample size s.*

As an example consider imitation play with memory $m = 6$ and sample size $s = 3$

[4]Actually, utility functions need not be identical within each population for any of the results in this paper. It is sufficient if each agent's utility function is a positive affine transformation of a population-specific utility function.

in the 2×3 game in Figure 1.

	a	b	c
A	2,2	0,0	3,3
B	0,0	1,1	0,0

FIGURE 1

Let $h = ((A,a),(B,a),(A,b),(B,b),(A,c),(B,c))$ be the initial history. Assume that the individual in the row player position (player 1) draws the last three plays from this history, such that $w_1 = (B,A,B)$ and $v_1 = (1,3,0)$. This gives an average payoff of 3 to strategy A and 1/2 to strategy B. Hence, the individual in the row player position will choose strategy A in the next period. Further assume the individual in the column player position (player 2) draws the first three plays, such that $w_2 = (a,a,b)$ and $v_2 = (2,0,0)$. This gives an average payoff of 1 to strategy a and 0 to strategy b. Since strategy c is not included in the sample it cannot be chosen. Hence, the column player will choose strategy a in the next period. Altogether this implies that the unperturbed process will move to the state $h' = ((B,a),(A,b),(B,b),(A,c),(B,c),(A,a))$ in the next period.

The perturbed process can be described as follows. In each period there is some small probability $\varepsilon > 0$ that each individual i drawn to play chooses a strategy at random from X_i instead of according to the imitation rule. The event that i experiments is assumed to be independent of the event that j experiments for every $j \neq i$ and across time periods. For every i let $q_i(x_i \mid h)$ be the conditional probability that i chooses $x_i \in X_i$, given that i experiments and the process is in state h. We assume that $q_i(x_i \mid h) > 0$ for all $x_i \in X_i$. Suppose that the process is in state h at time t. Let $D \subset N$ be any subset of d player positions. The probability is $\varepsilon^d(1-\varepsilon)^{n-d}$ that exactly the individuals drawn to play the positions in D experiment and the others do not. Conditional on this event the transition probability of moving from h to h' is $Q^D_{hh'} = \prod_{i \in D} q_i(x_i \mid h) \prod_{i \notin D} p_i(x_i \mid h)$ if h' is a successor of h and x is the rightmost element of h' and $Q^D_{hh'} = 0$ if h' is not a successor of h. This gives the following transition probability of the perturbed Markov process, similar to equation (2) in Young (1993, p. 67):

$$B^{m,s,\varepsilon}_{hh'} = (1-\varepsilon)^n B^{m,s,0}_{hh'} + \sum_{D \subset N, D \neq \emptyset} \varepsilon^{|D|}(1-\varepsilon)^{|N \setminus D|} Q^D_{hh'}. \tag{1}$$

The process $B^{m,s,\varepsilon}$ is denoted *imitation play with memory m, sample size s, and experimentation probabilities ε.*

3. Stochastic Stability in Finite Games

In this section we turn our attention to the limiting distribution of imitation play as the experimentation probability tends to zero. We first show that we can apply some of the tools in Young (1993) to calculate this distribution. We thereafter prove that it puts positive probability only on the strategy-tuples in particular subsets of the strategy-space.

3.1. Preliminaries. In what follows we will make use of the following definitions. A *recurrent class* of the process $B^{m,s,0}$ is a set of states such that there is zero probability of moving from any state in the class to any state outside, and there is a positive probability of moving from any state in the class to any other state in the class. We call a state h' *absorbing* if it constitutes a singleton recurrent class. We refer to a state $h_x = (x, x, ..., x)$, where x is any strategy-tuple from X, as a *convention.*[5] In other words, a convention is a state where the individuals in each player position played the same strategy during last m periods. If each player i has $|X_i| \geq 1$ strategies in the game Γ, then there are $|X| = \prod_i |X_i|$ conventions in this game. The following result shows that conventions correspond one to one with the recurrent classes of imitation play.

Theorem 1. *All conventions are absorbing states of the unperturbed process $B^{m,s,0}$. If s/m is sufficiently small, the process converges with probability one to a convention.*

Proof: It is evident that for s such that $1 \leq s \leq m$, any convention is an absorbing state, since any sample from a convention will contain only one strategy. We shall prove that if $s/m \leq 1/2$, then the conventions are the only recurrent classes of the unperturbed process. Consider an arbitrary initial state $h^t = (x^{t-m+1}, ..., x^t)$. If $s/m \leq 1/2$ there is a positive probability that all agents drawn to play sample from $x^{t-s+1}, ..., x^t$ (i.e. i, for $i \in N$, sample $x_i^{t-s+1}, ..., x_i^t$ and $\pi_i^{t-s+1}, ..., \pi_i^t$) in every period from $t+1$ to $t+s$ inclusive. All of them play the strategy with the highest average payoff in their sample. Assume without loss of generality that this is a unique strategy x_i^* for each of the player positions (if there are more than one strategy all of them have positive probability according to the assumptions). With positive probability all the agents drawn to play thereafter sample only from plays more recent than x^t in every period from $t+s+1$ to $t+m$ inclusive. Since all of these samples have the form $w_i^* = (x_i^*, ..., x_i^*)$ and $v_i^* = (\pi_i^*, ..., \pi_i^*)$, the unique strategy with the highest payoff in the sample is x_i^*. Hence, there is a positive probability of at time $t+m$

[5]Note that this definition differs from Young's (1993) definition of a convention, in that it does not require x to be a strict Nash equilibrium. Our definition corresponds to what Robson and Vega-Redondo (1996) denote a *monomorphic* or *pure* state.

obtaining a history $h^{t+m} = (x^*, ..., x^*)$, a convention. It follows that for $s/m \leq 1/2$ the only recurrent classes of the unperturbed process are the conventions. *Q.E.D.*

Lemma 1 below implies that Theorem 3.1 in Young (1998, p. 56) applies to imitation play. This theorem, with a slightly modified notation, is referred to as Theorem 2 in this paper.

Definition 1. *(Young, 1998) $B^{m,s,\varepsilon}$ is a* **regular perturbed Markov process** *if $B^{m,s,\varepsilon}$ is irreducible for every $\varepsilon \in (0, \varepsilon^*]$, and for every states h, h', $B^{m,s,\varepsilon}_{hh'}$ approaches $B^{m,s,0}_{hh'}$ at an exponential rate, i.e. $\lim_{\varepsilon \to 0} B^{m,s,\varepsilon}_{hh'} = B^{m,s,0}_{hh'}$ and if $B^{m,s,\varepsilon}_{hh'} > 0$ for some $\varepsilon > 0$, then $0 < \lim_{\varepsilon \to 0} \frac{B^{m,s,\varepsilon}_{hh'}}{\varepsilon^{r_{hh'}}} < \infty$ for some $r_{hh'} \geq 0$. The real number $r_{hh'}$ is called the* **resistance** *of the transition from h to h'.*

Lemma 1. *Imitation play is a regular perturbed Markov process.*

PROOF: See the Appendix.

Definition 2. *(Young, 1993) Let μ^ε be the unique stationary distribution of an irreducible process $B^{m,s,\varepsilon}$. A state h is* **stochastically stable** *if $\lim_{\varepsilon \to 0} \mu^\varepsilon(h) > 0$.*

Let an unperturbed Markov process $B^{m,s,0}$ with $s/m \leq 1/2$ have recurrent classes/conventions $h_{x^1}, ..., h_{x^{|X|}}$ (according to Theorem 1 there is a one to one correspondence between the recurrent classes and the conventions for $s/m \leq 1/2$). For each pair of distinct recurrent classes, an *xy-path* is a sequence of states $\zeta = (h_x, ..., h_y)$ that begins in h_x and ends in h_y. The *resistance* of this path is the sum of the resistances on the edges that compose it. Let r_{xy} be the least resistance over all xy-paths. Construct a complete directed graph with $|X|$ vertices, one for each recurrent class. The weights on the directed edge $h_x \to h_y$ is r_{xy}. A tree rooted at h_x is a set of $|X| - 1$ directed edges such that, from every vertex different from h_x, there is a unique directed path in the tree to h_x. The *resistance of such a rooted tree* $\Im(x)$ is the sum of the resistances $r_{x'x''}$ on the $|X|-1$ edges that compose it. The *stochastic potential* $\rho(x)$ of a recurrent class h_x is the minimum resistance over all trees rooted at h_x.

Theorem 2. *(Young, 1998) Let $B^{m,s,\varepsilon}$ be a regular perturbed Markov process and let μ^ε be the unique stationary distribution of $B^{m,s,\varepsilon}$ for $\varepsilon > 0$. Then $\lim_{\varepsilon \to 0} \mu^\varepsilon = \mu^0$ exists, and is a stationary distribution of $B^{m,s,0}$. The stochastically stable states are precisely those states that are contained in the recurrent classes of $B^{m,s,0}$ having minimum stochastic potential.*

In order to illustrate how to calculate the stochastic potential under imitation play, we present an example of a 2-player game. In the game in Figure 2, every player has three strategies, labeled A, B and C for the first player and a, b and c for the second player. The game has one strict Nash equilibrium (A, a), where both players gain less than in a mixed equilibrium with the probability mixture 1/2 on $B\,(b)$ and 1/2 on $C\,(c)$ for the first (second) player.[6]

	a	b	c
A	$1,1$	$0,0$	$0,0$
B	$0,0$	$3,2$	$0,3$
C	$0,0$	$0,3$	$3,2$

FIGURE 2

The conventions $h_{(B,b)}$ and $h_{(C,c)}$ are Pareto superior to convention $h_{(A,a)}$. Denote by $x_1 \in \{A, B, C\}$ some strategy choice by player 1 and $x_2 \in \{a, b, c\}$ some strategy choice by player 2. To find the stochastically stable conventions, construct directed graphs with nine vertices, one for each convention. In Figure 3 we illustrate two such trees. The numbers in the squares correspond to the resistances of the directed edges and the numbers in the circles represent the payoffs associated with the conventions. It is easy to check that for $s > 2$, $\rho(A, a) = 8$ and all other conventions have a stochastic potential of 9. Hence convention $h_{(A,a)}$ is *stochastically stable.*

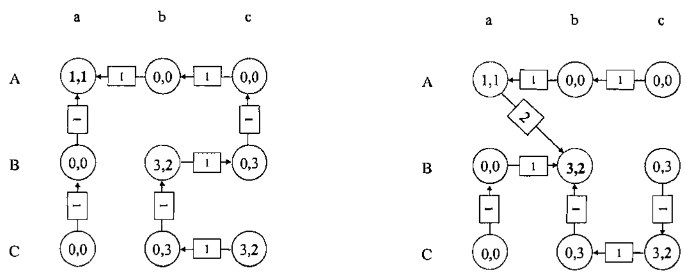

FIGURE 3–Minimum-resistance trees rooted at $h_{(A,a)}$ and $h_{(B,b)}$ respectively.

3.2. Sets of Stochastically Stable States. In this subsection we will show that the stochastically stable states of imitation play correspond to strategy-tuples in particular subsets of the strategy-space. In order to do this we need to introduce

[6]There is also a third equilibrium, $((\frac{5}{7}, \frac{1}{7}, \frac{1}{7}), (\frac{3}{5}, \frac{1}{5}, \frac{1}{5}))$.

some new concepts. Given n-player game Γ with finite strategy space $X = \prod X_i$, associate each strategy-tuple $x \in X$ with the vertex of a graph. Draw a directed edge from vertex x to vertex x' if and only if:
i) there exists exactly one player i such that $x_i \neq x_i'$, $u_i\left(x_i', x_{-i}\right) \geq u_i\left(x\right)$ and
ii) there does not exist $x_i'' \neq x_i$ and $x_i'' \neq x_i'$ such that
$u_i\left(x_i', x_{-i}\right) > u_i\left(x_i'', x_{-i}\right) > u_i\left(x\right)$.
Call the graph that is obtained in this manner the *better-reply graph* of game Γ.[7] A *better-reply path* is a sequence of the form $x^1, x^2, ..., x^l$ such that each pair $\left(x^j, x^{j+1}\right)$ corresponds to a directed edge in the better-reply graph. A *sink* is a vertex with no outgoing edges. Clearly, x is a sink if and only if it is a strict Nash equilibrium. A better-reply graph can also contain directed cycles.

Figure 4 shows the better-reply graph for a game with two sinks - (C, a) and (B, b). The *basin* of a sink is the set of all vertices from which there exists a directed path ending at that sink. Note that a vertex may be in several basins simultaneously. For example, (B, a) is in the basin of (C, a) and also in the basin of (B, b).

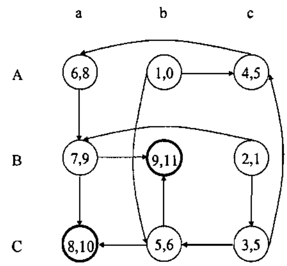

FIGURE 4–Better-reply graph.

Lemma 2. *If x and x' are vertices of the better-reply graph of game Γ and there exists a directed edge from vertex x to vertex x', then the following inequality holds for the stochastic potentials $\rho\left(x\right)$ and $\rho\left(x'\right)$ of the conventions h_x and $h_{x'}$ respectively:*

$$\rho\left(x\right) \geq \rho\left(x'\right).$$

[7]Condition ii) is included only for ease of exposition. The omission of this condition would complicate notation considerably without changing our results.

PROOF: Suppose the claim is false, such that $\rho(x) < \rho(x')$. Note that $\rho(x)$ is the minimum resistance over all trees rooted at the state h_x. Construct a new tree rooted at $h_{x'}$ by taking (one of) the tree(s) with minimum resistance rooted at h_x, adding the directed edge from vertex x to vertex x' and deleting the directed edge from x'. The resistance of the added edge is exactly one and of the deleted edge at least one, so the total resistance of the new tree is at most $\rho(x)$, which implies a contradiction. *Q.E.D.*

Definition 3. *A non-empty set of vertices V is said to be closed under the better-reply graph, or V is a* **construction,** *if there are no directed edges from any of the vertices in V to any vertex outside of V. Such a set is called a* **minimal construction** *if it does not properly contain another construction.*

From the definition follows that every game contains a minimal construction. Moreover, any sink is a minimal construction. The game in Figure 5 has two minimal constructions: the sink (A, a), and the set $V = \{(B,b), (B,c), (C,c), (C,b)\}$, which can be considered as a single directed cycle $(B,b) \rightarrow (B,c) \rightarrow (C,c) \rightarrow (C,b) \rightarrow (B,b)$.

	a	b	c
A	1,1	0,0	0,0
B	0,0	3,2	2,3
C	0,0	2,3	3,2

FIGURE 5

Note that our definition of minimal constructions in normal form games is related to Sobel's (1993) definition of *non-equilibrium evolutionary stable (NES)* sets in 2-player games. Our concept is also similar to Nöldeke and Samuelson's (1993) definition of *locally stable components* in extensive form games.

Lemma 3. *Let V be a minimal construction of an n-player game Γ with finite strategy space $X = \prod_{i=1}^{n} X_i$. Then, for any two vertices $x, x' \in V$, there exist better-reply paths $x, ..., x'$ and $x', ..., x$, which connect these vertices.*

PROOF: Suppose that the claim is false and there exist two vertices $x, x' \in V$ such that there is no better-reply path from x to x'. Consider all better-reply paths, which start at the vertex x. Since X is finite, there exist finitely many vertices along all these paths. Collect all these vertices. By construction this set of vertices has only ingoing edges and by assumption it does not contain the vertex x'. Hence, the constructed set of vertices is a minimal construction and a proper subset of the minimal construction V. This is a contradiction, so the claim must be true. *Q.E.D.*

To every vertex of a minimal construction corresponds a convention, which is a repetition of the associated strategy-tuple $x \in X$. Call the set of these conventions *conventions of a minimal construction.* We are now in a position to state the following main theorem.

Theorem 3. *If $s/m \leq 1/2$, s is sufficiently large and ε is sufficiently small, the perturbed process $B^{m,s,\varepsilon}$ puts arbitrarily high probability on the conventions of minimal construction(s) that minimize stochastic potential.*

PROOF: See the Appendix.

In order to prove this claim, we use Lemma 2 and Lemma 3 to show that: A) all conventions in a minimal construction have equal stochastic potential and B) for every convention which does not belong to any minimal construction, there exists a convention with lower stochastic potential. The theorem establishes a relation between the stochastically stable states of imitation play and minimal constructions, which is similar to the relationship proved between the stochastically stable states of *adaptive play* and minimal curb sets in Theorem 7.2 of Young (1998, p. 111).

We say that a finite set Y of strategy-tuples *strictly Pareto dominates* a strategy-tuple x if for any strategy-tuple $y \in Y$, $\pi_i(y) > \pi_i(x)$, for all i. The following theorem shows that imitation play selects sets of efficient outcomes in a large class of games.

Theorem 4. *Suppose that there exists a minimal construction V, which is a product set and which strictly Pareto dominates all strategy-tuples outside of V. If $s/m \leq 1/2$, s is sufficiently large and ε is sufficiently small, the perturbed process $B^{m,s,\varepsilon}$ puts arbitrarily high probability on the conventions of V.*

PROOF: See the Appendix.

The intuition behind this result is that for sufficiently large sample size, the transition from a state inside of V to any state outside of V requires more mistakes than the number of player positions, while the opposite transition requires at most one mistake per player position. The following corollary follows immediately from Theorem 4.

Corollary 1. *If x is a strict Nash Equilibrium, which strictly Pareto dominates all other strategy-tuples, then for $s/m \leq 1/2$ and s sufficiently large, the convention h_x is a unique stochastically stable state.*

The requirement that V be a product set is necessary, as shown by the game in Figure 6.

	a	b	c
A	1,1	0,0	0,0
B	0,0	0,0	5,3
C	0,0	2,4	5,4

FIGURE 6

In this game, the minimal construction $V = \{(C,b), (C,c), (B,c)\}$ strictly Pareto dominates all strategy-tuples outside of V. It is evident that two mistakes are enough to move from convention $h_{(A,a)}$ to any convention in the minimal construction V. We will show that two mistakes are also enough to move from the minimal construction V to an outside state. Suppose the process is in the state $h_{(C,c)}$ at time t. Further suppose that the agent in player position 1 plays B instead of C at time $t+1$ by mistake. This results in play (B,c) at time $t+1$. Assume the agent in player position 2 makes a mistake and plays b instead of c, and that the agent in player position 1 plays C in period $t+2$. Hence, play at time $t+2$ is (C,b). Assume the agents in both player positions sample from period $t-s+2$ to period $t+2$ for the next s periods. This means that the agents in player position 1 chooses to play B and the agents in player position 2 chooses to play b from period $t+3$ to period $t+s+2$. There is a positive probability that from period $t+s+3$ through period $t+m+2$ the agents in both player positions will sample from periods later than $t+2$. Hence, by the end of period $t+m+2$, there is a positive probability that the process will have reached the convention $h_{(B,b)}$ outside of the minimal construction V. It is now straightforward to show that all of the conventions $h_{(A,a)}$, $h_{(C,b)}$, $h_{(C,c)}$ and $h_{(B,c)}$ have equal stochastic potential.

4. Properties of Minimal Constructions

In this section we investigate the relationship between minimal constructions and other set-wise concepts. We also analyze whether strictly dominated strategies can be included in minimal constructions and possible invariance under payoff transformations.

4.1. Relation to Minimal Sets Closed Under better replies. We start by the relation between minimal constructions and so called minimal sets closed under better replies. It is maybe not surprising that there is a quite strong connection between these two concepts.

Definition 4. *(Ritzberger and Weibull, 1995) Let the* **better-reply correspondence** $\gamma = \times_{i \in N} \gamma_i : \Delta \twoheadrightarrow X$ *be defined as follows*

$$\gamma_i(p) = \{x_i \in X_i \mid u_i(x_i, p_{-i}) \geq u_i(p)\}, \forall i \in N.$$

A product set of strategies is a set of form $Y = \Pi_i Y_i$, where each Y_i is a non-empty subset of X_i. Let $\Delta(Y_i)$ denote the set of probability distributions over Y_i, and let $\Delta(Y) = \Pi_i \Delta(Y_i)$ denote the product set of such distributions. Let $\gamma_i(Y)$ denote the set of strategies in X_i that are better replies by i to some distribution $p \in \Delta(Y)$ and let $\gamma(Y) = \Pi_i \gamma_i(Y)$.

Definition 5. *A non-empty product set $Y \subset X$ is said to be* **closed under better replies** *(or Y is a* **cubr** *set) if $\gamma(Y) \subset Y$. A set is called a* **minimal cubr** *set if it is cubr and contains no proper subset which is cubr.*

Proposition 1. *Every minimal cubr set contains a minimal construction. If a minimal construction is a product set, then it coincides with a minimal cubr set.*

PROOF: We start from the first claim. Let $Y = \Pi_i Y_i$ be a minimal cubr set. Associate each strategy-tuple $x \in X$ with the vertex of a graph and draw a better-reply graph for game Γ on the finite strategy space X. By the definition of a minimal cubr set, there are no directed edges from the set of vertices in Y to any vertices outside of Y in the better-reply graph. This means that the set of vertices in Y is a minimal construction or a construction containing a minimal construction.

The last claim follows immediately from the definitions of a minimal cubr set and a minimal construction. *Q.E.D.*

Combined with Theorem 3 this implies that if all the minimal constructions of a game are product sets, then imitation play converges to the conventions of minimal cubr sets. Note that every minimal construction not necessarily is included in a minimal cubr set. Consider the game in Figure 7.

	a	b	c
A	2,2	1,1	0,0
B	1,1	0,0	3,3
C	0,0	3,3	3,3

FIGURE 7

In this game, there is a unique minimal cubr set, $\{(A, a)\} = Y$, and two minimal constructions, $V_1 = \{(A, a)\}$ and $V_2 = \{(B, c), (C, b), (C, c)\}$. Hence V_2 is not contained in any minimal cubr set.

4.2. Relation to Curb Sets. The minimal curb set is another set-wise concept. Young (1998) show that there is a strong relation between minimal curb sets and stochastically stable states in his model. Let $Z = \Pi_{i=1}^{n} Z_i$ be a product set of strategies, and let $\beta_i\left(Z_{-i}\right)$ denote the set of strategies in X_i that are best replies by i to some distribution $p_{-i} \in \Delta\left(Z_{-i}\right)$. Define $\beta(Z) = \Pi_{i=1}^{n} \beta_i(Z_{-i})$.

Definition 6. *(Basu and Weibull, 1991) A non-empty product set $Z \subset X$ is said to be* **closed under rational behavior** *(or Z is a* **curb** *set) if $\beta(Z) \subset Z$. Such a set is called a* **minimal curb** *set if it does not properly contain a curb set.*

It is straightforward to show that $\beta(Z) = Z$ for any minimal curb set Z. Let the *span* of a non-empty set of vertices V, denoted by $S\left(V\right)$, be the product set of all strategies that appear in some strategy-tuple of V. By the following two examples, we will show that a minimal curb set and (the span of) a minimal construction are different set-wise concepts. In the game in Figure 8, the minimal cubr set, the minimal construction and the minimal curb set are all different.

	a	b	c
A	1,3	3,2	3,1
B	3,2	2,3	2,1
C	2,3	1,1	1,2

FIGURE 8

In this game, $\{A, B, C\} \times \{a, b, c\} = Y$ is a minimal cubr set. There is a unique minimal construction: $V = \{(A, a), (C, a), (B, a), (B, b), (A, b)\}$ with span $S\left(V\right) = \{A, B, C\} \times \{a, b\}$. It is obvious that $\{A, B\} \times \{a, b\} = Z$ is the minimal curb set. Hence $S\left(V\right) \cap Y = S\left(V\right)$ and $S\left(V\right) \cap Z = Z$ in this game.

One may ask whether the minimal curb set always is included in the span of a minimal construction. In other words, is the set-wise relationship $S\left(V\right) \cap Z = Z$ true in general? The answer is no. Consider the game in Figure 9.

	a	b	c
A	3,1	−10,0	1,3
B	−10,0	1,3	3,1
C	1,3	3,1	−10,0
D	0,0	0,0	0,0

FIGURE 9

In this game, A is a best reply to a, a is a best reply to C, C is a best reply to b, b is a best reply to B, B is a best reply to c, c is a best reply to A. Hence any curb set that involves one or more of $\{A, a, C, b, B, c\}$ must involve all of them. Hence it must involve D, because D is a best reply to the probability mixture 1/3 on a, 1/3 on b,and 1/3 on c. We conclude that $\{A, B, C, D\} \times \{a, b, c\} = Z$ is the unique minimal curb set. However, the unique minimal construction is $V = \{(A, a), (C, a), (C, b), (B, b), (B, c), (A, c)\}$ with span $S(V) = \{A, B, C\} \times \{a, b, c\}$, such that $S(V) \cap Z = S(V) \neq Z$.

4.3. Relation to Strictly Dominated Strategies. Our model assumes that agents are boundedly rational. In this subsection we ask whether any dominated strategies will be played in the stochastically stable states of imitation play. Consider a 2×2 game where strategy A strictly dominates strategy B for the row player. In such a game, the minimal construction will consist either of (A, a) or (A, b) or both of them. From Theorem 3 follows that in 2×2 games, a strictly dominated strategy will never be used with positive probability in the long run. However, as illustrated by the game in Figure 10, a weakly dominated strategy may be used with positive probability. In this game, the unique minimal construction is $V = \{(A, b), (B, b), (B, a)\}$.

	a	b
A	0, 0	3, 3
B	3, 3	3, 3

FIGURE 10

Moreover, in games with more than two strategies, strictly dominated strategies may be played with positive probability. Consider the game in Figure 11.

	a	b
A	0, 3	2, 0
B	2, 0	1, 3
C	1, 1	0, 1

FIGURE 11

In this game, B strictly dominates C, but the unique minimal construction is $V = \{(A, a), (C, a), (C, b), (B, b), (A, b), (B, a)\}$ and by Theorem 3, imitation play will put positive probability on all strategy-tuples in this minimal construction in the long run. Play of strictly dominated strategies may appear an undesirable or counter intuitive feature, but it is merely a consequence of the limited information available to the agents in the model. Since they do not know the true payoffs to every strategy profile, but rely on information provided by a finite sample of previous play, they cannot deduce that a certain strategy is dominated.

4.4. Payoff Transformations. Minimal constructions are invariant under affine payoff transformations and local payoff shifts. This follows since the better-reply graph is unaffected by such changes. The stochastically stable states of imitation play are also invariant under affine payoff transformations, but not under local payoff shifts. The two games in Figure 12 illustrate the latter claim.

	a	b
A	3,3	0,0
B	0,0	2,2

	a	b
A	3,3	2,0
B	0,2	4,4

FIGURE 12

In the left game, (A, a) is stochastically stable according to Theorem 4. By adding 2 units to the each of the row player's payoffs in column b and to each of the column player's payoff in row B in this game, the right game is obtained. In the right game (B, b) is stochastically stable by Theorem 4.

5. APPLICATIONS TO 2×2 GAMES

In this section we apply the results from Section 3 to 2×2 games. First, we find the stochastically stable states in three special classes of games and second, we study the speed of convergence in general 2×2 games. In all of the following games we denote player 1's strategies A and B, and player 2's strategies a and b respectively.

5.1. Stochastically Stable States in Special Classes of Games. In this subsection we analyze the stochastic stability in games with two strict Nash equilibria, games with one strict Nash equilibrium and games without Nash equilibria in pure strategies. We start with the class of games that have a unique strict Nash equilibrium. This class includes for example *Prisoners' Dilemma Games.*

Proposition 2. *In 2×2 games with a unique strict Nash equilibrium, if $s/m \leq 1/2$ and s is sufficiently large, the stochastically stable state correspond one to one with the convention of the unique strict Nash equilibrium.*

PROOF: Games in this class contain exactly one minimal construction, consisting of the unique strict Nash equilibrium. By Theorem 3 the corresponding convention must be the unique stochastically stable state. *Q.E.D.*

We now proceed with the class of games with two strict Nash equilibria. *Coordination Games* and *Hawk-Dove Games* are two examples of games in this class.

Proposition 3. *In 2×2 games with two strict Nash equilibria, where one Nash Equilibrium strictly Pareto dominates the other, if $s/m \leq 1/2$ and s is sufficiently large, then the unique stochastically stable state corresponds one to one with to the Pareto dominant equilibrium.*

PROOF: Games in this class contain two minimal constructions, either $\{(A,a)\}$ and $\{(B,b)\}$ or $\{(A,b)\}$ and $\{(B,a)\}$. Assume without loss of generality that $\{(A,a)\}$ and $\{(B,b)\}$ are the minimal constructions. If (A,a) strictly Pareto dominates (B,b), it must also strictly Pareto dominate the two other strategy-tuples (A,b) and (B,a). Hence, by Corollary 1 the convention $h_{(A,a)}$ is the unique stochastically stable state for $s/m \leq 1/2$ and s is sufficiently large. *Q.E.D.*

This implies that, unlike Young's (1993) process of adaptive play, imitation play does not generally converge to the risk-dominant equilibrium in Coordination Games. In the right game in Figure 12, (A,a) is a risk dominant equilibrium whereas (B,b) is a strictly Pareto superior equilibrium and subsequently the unique stochastically stable state of imitation play. Our result is consistent with Robson and Vega-Redondo's (1996) result for generic symmetric Coordination Games. However, for the non-generic case when $\pi_i(A,a) = \pi_i(B,b)$, the stochastically stable states in their model depend on the details of the adjustment process, whereas imitation play always selects both equilibria. Consider the following proposition.

Proposition 4. *In 2×2 games with two strict Nash equilibria, where neither strictly Pareto dominates the other, if $s/m \leq 1/2$ and s is sufficiently large, then the stochastically stable states correspond one to one with the two equilibria.*

PROOF: See the Appendix.

Finally, we consider games which do not have any Nash equilibrium in pure strategies. One of the games in this class is the *Matching Pennies Game.*

Proposition 5. *In 2×2 games without Nash equilibria in pure strategies, if $s/m \leq 1/2$ and s is sufficiently large, the stochastically stable states correspond one to one with the four conventions of the game.*

PROOF: Games in this class contain exactly one minimal construction $\{(A,a),(A,b),(B,a),(B,b)\}$. By Theorem 3 the four corresponding conventions must all be stochastically stable. *Q.E.D.*

5.2. Speed of Convergence. In this subsection we analyze the speed of convergence of imitation play.

Proposition 6. *In 2×2 games, if $s/m \leq 1/2$ and s is sufficiently large, the maximum expected first passage time for the perturbed process $B^{m,s,\varepsilon}$ from any state to a stochastically stable state is at most $\delta\varepsilon^{-2}$ units of time, for some positive constant δ.*

PROOF: The claim follows from the observation that in any 2×2 game, the transition from an arbitrary state to the basin of attraction of a stochastically stable convention requires at most two mistakes. *Q.E.D.*

This result should be contrasted with the speed of convergence in Young (1993), Kandori *et al.* (1993), and Robson and Vega-Redondo (1996). In Young's (1993) model the maximum expected first passage for a 2×2 Coordination Game is at most $\delta_Y \varepsilon^{-v}$ where v depends on the sample size and both players' payoffs. In Kandori *et al.* (1993) the maximum expected first passage time is of order $\delta_{KMR}\varepsilon^{-Nu}$, where N is the size of the population and u is determined by the game's payoff structure. In Robson and Vega-Redondo (1996) the corresponding figure is $\delta_{RV}\varepsilon^{-q}$, where q is a positive integer independent of the payoffs and the current state. Thus, if v, Nu and q all are greater than two and ε is sufficiently small, then imitation play converges considerably faster than the processes in these three models.

6. EXTENSIONS

All the results in this paper hold for a more general class of imitation dynamics. Let the *maximum correspondence* be a correspondence which maps a strategy sample w_i and the associated payoff sample v_i to a strategy with the highest payoff in the sample. This correspondence defines a new Markov process on the space of histories with the same set of absorbing states and of stochastically stable states as imitation play. Moreover, if each population consists of arbitrary shares of individuals make choices based on the maximum correspondence and the maximum average correspondence respectively, then the results of this paper still hold. Hence, the model allows for a certain kind of population heterogeneity, where individuals make their choices based on different rules.

7. CONCLUSION

In this paper we develop an evolutionary model with perpetual random shocks where individuals every period choose the strategy with the highest average payoff in a finite sample of past play. We denote the resulting Markov process imitation play. We prove that, provided information is sufficiently incomplete and the sample size sufficiently large, the stochastically stable states of imitation play are repetitions of the strategy-tuples in minimal sets closed under the better-reply graph. We call such sets minimal constructions. These sets are related to minimal sets closed under better replies and other set-wise concepts. We also prove that if the strategy-tuples in a minimal construction have strictly higher payoffs than all outside strategy-tuples, then, provided the minimal construction is a product set, the strategy-tuples in this set will be played with probability one in the limit as the experimentation probability tends to zero. Our results give clear predictions in 2×2 games. In Coordination Games where one equilibrium strictly Pareto dominates the other, imitation play selects the strictly Pareto superior Nash equilibrium. If neither equilibria strictly Pareto dominates the other, then both are stochastically stable. Finally, we show that the speed of convergence for imitation play in many cases is higher than in other known models.

The objective of this paper is to derive predictions for general finite games in a world of truly boundedly rational agents. The assumption underlying the model, that individuals do not make decisions based on the predictions of their opponents' future strategies, but rather based on which strategies have been successful in the past, is maybe most appealing in the class of games where it is costly to obtain information about the opponents. A high cost may be due to the size or the complexity of the game or to institutional factors that prevent the release of information about the opponents. It would be particularly interesting to test the implications of our model against empirical or experimental evidence in this class of games.

APPENDIX

PROOF OF LEMMA 1: This proof follows Young (1998). Given a history $h^t = (x^{t-m+1}, ..., x^t)$ at time t, the process moves in the next period to a state of the form $h^{t+1} = (x^{t-m+2}, ..., x^t, x^{t+1})$. Remember that such a state is called a *successor* of h^t. The strategy x_i^{t+1} is an *idiosyncratic choice* or *error* if and only if there exist no sample of strategy choices w_i and payoff outcomes v_i, such that $x_i^{t+1} \in \alpha_i(w_i, v_i)$. For each successor h^{t+1} of h^t, let $r_{h^t h^{t+1}}$ denote the total number of errors in the transition from h^t to h^{t+1}. Evidently $0 \leq r_{h^t h^{t+1}} \leq n$ and the probability of the transition $h^t \to h^{t+1}$ is on the order of $\varepsilon^{r_{h^t h^{t+1}}}(1-\varepsilon)^{n-r_{h^t h^{t+1}}}$. If h^{t+1} is not a successor of h^t, the probability of the transition is zero. Hence, the process $B^{m,s,\varepsilon}$ approaches $B^{m,s,0}$ at an exponential rate in ε. Since the process is also irreducible whenever $\varepsilon > 0$, it follows that $B^{m,s,\varepsilon}$ is a regular perturbed Markov process. *Q.E.D.*

PROOF OF THEOREM 3: In order to prove the claim, we will show that:
A) all conventions in a minimal construction have equal stochastic potential;
B) for every convention which does not belong to any minimal construction, there

exists a convention with lower stochastic potential.
The theorem follows immediately from A) and B), by applying Lemma 1 and Theorem 2.

Let us start from A). On the one hand, by Lemma 3, there exist a better-reply path from an arbitrarily convention h_x in a minimal construction to any other convention $h_{x'}$ in the same minimal construction. Let the sequence $h_x, ..., h_{x'}$ be such a path. By Lemma 2, the following inequalities hold:

$$\rho(x) \geq ... \geq \rho(x'). \tag{2}$$

On the other hand, by applying Lemma 3 once again, there exist a better-reply path from the convention $h_{x'}$ to the convention h_x. Using Lemma 2, gives:

$$\rho(x') \geq ... \geq \rho(x). \tag{3}$$

From the inequalities in (2) and (3) follow that $\rho(x) = \rho(x')$ for any arbitrarily conventions h_x and $h_{x'}$ in a minimal construction.

Part B). Note that for every convention not included in any minimal construction, there exists a finite better-reply path which ends in some minimal construction. Let this path be $h_{x^1}, h_{x^2}, ..., h_{x^{T-1}}, h_{x^T}$, where h_{x^1} is an arbitrarily convention that does not belong to any minimal construction and h_{x^T} is the first convention on the path that belongs to some minimal construction V. By Lemma 2, it follows that

$$\rho(x^1) \geq \rho(x^2) \geq ... \geq \rho(x^{T-1}) \geq \rho(x^T). \tag{4}$$

We will show that in fact $\rho(x^{T-1}) > \rho(x^T)$. Note that $\rho(x^{T-1})$ is the minimum resistance over all trees rooted at the state $h_{x^{T-1}}$. Denote (one of) the tree(s) that minimizes resistance by $\Im(x^{T-1})$. Find in the tree $\Im(x^{T-1})$ a directed edge from some vertex $y \in X$ in the minimal construction V to some other vertex $y' \in X$ outside of this minimal construction. It will be shown later that there is only one such directed edge in the minimal resistance tree $\Im(x^{T-1})$. Delete in the tree $\Im(x^{T-1})$ the directed edge $y \to y'$ and add the directed edge $x^{T-1} \to x^T$. As a result we obtain a tree $\Im(y)$ rooted at the state h_y. By construction, the total resistance of the new tree $\Im(y)$ is less than the stochastic potential $\rho(x^{T-1})$. Moreover, by part A), the convention h_{x^T} of the minimal construction has the same stochastic potential as the convention h_y. Hence $\rho(x^{T-1}) > \rho(x^T)$.

We will now consider the tree $\Im(x^{T-1})$ and show that there is only one directed edge from every minimal construction to a state outside of the construction. Suppose there are finitely many such directed edges $y^j \to z^j$, $j = 1, 2, ..., l$ from some minimal construction, where $y^1, ..., y^l$ are vertices in the minimal construction and $z^1, ..., z^l$

are vertices outside of the construction. It is clear that there cannot be an infinite number of outgoing edges since the game Γ is finite. Recall that a rooted at vertex y^j tree is a set of $|X|-1$ directed edges such that, from every vertex different from y^j, there is a unique directed path in the tree to y^j. The resistance of any directed edge $y^j \to z^j$, $j = 1, 2, ..., l$ is at least two. By Lemma 3, there exists a finite better-reply path from vertex y^1 to vertex y^2 in the minimal construction. Let $y^1, f^1, ..., f^k, y^2$ be such a path.

Consider vertex f^1. There are two mutually exclusive cases:
1.a) there exists a directed path from f^1 to one of the vertices $y^2, ..., y^l$ in the initial tree $\Im\left(x^{T-1}\right)$, or
1.b) there exists a directed path from f^1 to y^1.

In case 1.a) by deleting the directed edge $y^1 \to z^1$ and adding the directed edge $y^1 \to f^1$ to the tree $\Im\left(x^{T-1}\right)$, we obtain a new tree $\Im^1\left(x^{T-1}\right)$ with lower stochastic potential than $\Im\left(x^{T-1}\right)$, because the resistance of the directed edge $y^1 \to f^1$ is one. This means that we are done, since it contradicts the assumption of $\Im\left(x^{T-1}\right)$ being a minimal resistance tree.

In case 1.b) we will use the following procedure for the vertex f^1: delete the initial directed edge from f^1 and add the directed edge $f^1 \to f^2$. As above, there are two cases:
2.a) there exists a directed path from f^2 to one of the vertices $y^2, ..., y^l$ in the initial tree $\Im\left(x^{T-1}\right)$, or
2.b) there exists a directed path from f^2 to y^1.

In case 2.a) we obtain a new tree $\Im^2\left(x^{T-1}\right)$ with lower stochastic potential than $\Im\left(x^{T-1}\right)$, because the resistance of the directed edge $f^1 \to f^2$ is one. This means that we are done, since it contradicts the assumption of $\Im\left(x^{T-1}\right)$ being a minimal resistance tree.

In case 2.b) we repeat the procedure for the vertex $f^2, f^3, ...$ The better-reply path $y^1, f^1, ..., f^k, y^2$ from vertex y^1 to vertex y^2 is finite. Hence, after at most $k+1$ steps we have constructed a tree $\Im''\left(x^{T-1}\right)$ rooted at the state $h_{x^{T-1}}$ with lower stochastic potential than $\Im\left(x^{T-1}\right)$, which is impossible. *Q.E.D.*

PROOF OF THEOREM 4: Define the *basin of attraction* of a state h' as the set of states h such that there is a positive probability of moving in a finite number of periods from h to h' under the unperturbed process $B^{m,s,0}$. Let V be a minimal construction which strictly Pareto dominates all strategy-tuples outside of V. Let h_y be a convention such that $y \notin V$. Consider a minimal resistance tree $\Im(y)$ rooted at h_y. Let $x \in V$ be a vertex such that there is a directed edge from x to a vertex $z \notin V$ in the tree $\Im(y)$. Assume that the stochastic potential of h_x is at least as high as the stochastic potential of h_y.

We claim that for sufficiently large s, the resistance of the edge from $x \in V$ to $z \notin V$ is greater than n. Assume without loss of generality that $s > n$ (otherwise s consecutive mistakes by agents in one population could be sufficient to transfer from h_x to the basin of attraction of h_z). Let $\pi_i(x)$ and $\pi_i(z)$ be payoffs for player i corresponding to the strategy-tuples x and z respectively. The individual in player position i will play z_i instead of x_i only if the average payoff for z_i in her sample is greater than or equal to the average payoff for x_i. With at most n mistakes this is only possible if the following inequality holds

$$\frac{(s-n)\pi_i(x) + (n-1)\pi_i(x_i, x'_{-i})}{s-1} \leq \pi_i(z), \tag{5}$$

where $\pi_i(x_i, x'_{-i})$ is the payoff to player i when she plays x_i and her opponents play any $x'_{-i} \neq x_{-i}$. It is clear that this inequality does not hold for any individual in player position i if s is sufficiently large. Hence, the resistance from h_x to h_z must be larger than n for sufficiently large s.

Create a tree $\Im(x)$ rooted at h_x by adding a directed edge from h_y to h_x and deleting the directed edge from x to z in the tree $\Im(y)$. Provided that s is sufficiently large, the deleted edge has a resistance greater than n and the added edge has a resistance of at most n (n simultaneous mistakes are sufficient to move from any convention to the basin of attraction of h_x). Hence the total resistance of the new tree $\Im(x)$ is less than $\rho(y)$, contradicting the assumption of $\rho(y) \leq \rho(x)$. This establishes that h_x has minimal stochastic potential. Theorem 4 now follows from part A) in the proof of Theorem 3 where it is proved that all the conventions of a minimal construction have equal stochastic potential. *Q.E.D.*

PROOF OF PROPOSITION 4: Like in the proof of Proposition 2, assume without loss of generality that $\{(A, a)\}$ and $\{(B, b)\}$ are the minimal constructions of the game. By Theorem 3 it follows that the conventions $h_{(A,a)}$ and $h_{(B,b)}$ are the only two candidates for the stochastically stable states. Suppose that only one of these conventions is stochastically stable, say $h_{(A,a)}$. Let $\Im(A, a)$ be a minimum resistance tree with resistance $\rho(A, a)$ rooted at (A, a). In this tree there is an outgoing edge from the convention $h_{(B,b)}$.

First note that the resistance of this edge is at least two, such that at least two mistakes are needed to move from the convention $h_{(B,b)}$. This follows since $\pi_1(B, b) > \pi_1(A, b)$ and $\pi_2(B, b) > \pi_2(A, a)$.

Second, note that provided that $s/m \leq 1/2$ and s is sufficiently large, two mistakes are sufficient to move the process from the convention $h_{(A,a)}$ to the convention $h_{(B,b)}$. Suppose the process is in state $h_{(A,a)}$. Since neither of the Nash equilibria is strictly Pareto superior, either $\pi_1(B,b) \geq \pi_1(A,a)$ and/or $\pi_2(B,b) \geq \pi_2(A,a)$. Assume without loss of generality that the first of these inequalities holds. Then there is a positive probability that the agents in both player positions simultaneously make mistakes at time t. There is also a positive probability that the agents in player position 1 draw the fixed sample $(x_1^{t-s+1},...,x_1^t)$ with corresponding payoffs $(\pi_1^{t-s+1},...,\pi_1^t)$ and that the agents in player position 2 sample from plays earlier than x_2^t from period $t+1$ to and inclusive period $t+s-1$. With positive probability the agents in player position 1 play B and the agents in player position 2 play a in all of these periods. This implies that if the agents in both player positions sample from plays more recent than x_i^{t-1} from period $t+s$ to and inclusive period $t+s-1+m$, 1's sample will only contain strategy B and 2's sample will always contain strategy b and possibly strategy a. Furthermore, the average payoff of strategy b will be $\pi_2(B,b)$ as compared to an average payoff of $\pi_2(B,a)$ for strategy a when the latter strategy is included in the sample. Hence, with positive probability the process will be in state $h_{(B,b)}$ at time $t+s-1+m$.

Finally, create a new tree rooted at $h_{(B,b)}$ by deleting the outgoing edge from the convention $h_{(B,b)}$ in the tree $\Im(A,a)$ and adding an edge from $h_{(A,a)}$ to $h_{(B,b)}$. The resistance of the deleted edge is at least two and of the added edge two. Hence the total resistance of the new is at most $\rho(A,a)$, contradicting the assumption that only $h_{(A,a)}$ is stochastically stable. *Q.E.D.*

References

[1] Basu, K., and Weibull J. (1991). "Strategy Subsets Closed under Rational Behavior," *Economics Letters* **36**, 141-146.

[2] Björnerstedt, J., and Weibull J. (1996). "Nash equilibrium and evolution by imitation," in *The Rational Foundations of Economic Behaviour* (K. Arrow et al. Eds.), pp. 155-171. MacMillan, London.

[3] Duffy, J., and Feltovich N. (1999). "Does observation of others affect learning in strategic environments? An experimental study," *International Journal of Game Theory* **28**, 131-152.

[4] Graham, J. R. (1999). "Herding among Investment Newsletters: Theory and Evidence," *Journal of Finance* **54**, 237-268.

[5] Griffiths, M. D., Smith, B. F., Turnbull, D., Alasdair, S., and White R. W. (1998). "Information flows and open outcry: evidence of imitation trading," *Journal Of International Markets, Institutions And Money* **8**, 101-116.

[6] Huck, S., Normann, H.-T., and Oechssler J. (1999). "Learning in Cournout Oligopoly: An experiment," *Economic Journal* **109**, 80-95.

[7] Huck, S., Normann, H.-T., and Oechssler J. (2000). "Does information about competitors' actions increase or decrease competition in experimental oligopoly markets?," *International Journal of Industrial Organization* **18**, 39-58.

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

[9] Nash, J. (1950). "Non Cooperative Games," Ph.D. thesis, Princeton University.

[10] Nöldeke, G., and Samuelson L. (1993). "An Evolutionary Analysis of Backward and Forward Induction", *Games and Econ. Behavior* **5**, 425-454.

[11] Osborne, M. J., and Rubinstein A. (1998). "Procedurally Rational Players," *American Economic Review* **88**, 834-847.

[12] Ritzberger, K., and Weibull J. (1995). "Evolutionary Selection in Normal-Form Games" *Econometrica* **63**, 1371-1400.

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

[14] Schlag, K., (1998). "Why Imitate, and If So, How? A Boundedly Rational Approach to Multi-armed Bandits," *Journal of Economic Theory* **78**, 130-56.

[15] Schlag, K., (1999). "Which One Should I Imitate?," *Journal of Mathematical Economics* **31**, 493-522.

[16] Sobel, J. (1993). "Evolutionary Stability and Efficiency," *Economics Letters* **42**, 301-312.

[17] Wermers, R. (1999). "Mutual Fund Herding and the Impact on Stock Prices,", *Journal of Finance* **54**, 237-268.

[18] Young, P. H. (1993). "The Evolution of Conventions," *Econometrica* **61**, 57-84.

[19] Young, P. H. (1998). *Individual Strategy and Social Structure.* Princeton University Press.

Evolutionary Dynamics on Extensive Form Games with Perfect Information

ALEXANDR MATROS*
DEPARTMENT OF ECONOMICS
STOCKHOLM SCHOOL OF ECONOMICS

January 20, 2000

ABSTRACT. In this paper, I examine dynamic evolutionary processes driven by perpetual random shocks on extensive form games with perfect information. Every period, n individuals are randomly chosen from n finite populations to play an extensive form game. Each individual observes a sample from the memory of past plays. Then she either plays a best reply to that sample or imitates by choosing the action with the highest or the highest average payoff in the sample on the reached nodes. Occasionally, individuals also experiment or make mistakes and choose a pure local strategy at random on the reached nodes. For finite n-player games, I prove that in the limit, as the probability of experimentation tends to zero, the backward induction outcomes occur with positive probability in the best reply and the imitation cases. Moreover, for a special class of games, the backward induction outcome is the unique prediction.

Keywords: Evolutionary game theory, games in extensive form, games of perfect information, backward induction equilibrium, Markov chain, stochastic stability.

1. INTRODUCTION

The key concept in Game Theory is a Nash equilibrium (introduced by J. Nash in 1950). However, most games have several Nash equilibria and the question of which is the most natural one then arises. A large number of studies have tried to answer this question. Evolutionary Game Theory first suggested *evolutionary stable strategy* (introduced by J. Maynard Smith and G. Price in 1973) as an equilibrium

*Previous versions: November 2000, September 2000; May 2000. Financial support is from the Jan Wallander and Tom Hedelius Foundation.
I am grateful to Jörgen Weibull and Lawrence Blume for helpful discussions and suggestions. I thank Carlos Alos-Ferrer, Joerg Oechssler, Klaus Ritzberger, Birgitte Sloth, seminar participants at the Stockholm School of Economics, participants the Eleventh International Conference at Stony Brook, and Evolutionary Game Theory 2000 conference in Odense for helpful comments.

concept and then the sharper *stochastic stability* (introduced by D. Foster and P. Young in 1990). The classical papers Kandori, Mailath, and Rob (1993) and Young (1993) have made the *stochastic stability* concept very well known. These papers show how the technique developed by Freidlin and Wentzell (1984) can be used to select among a number of equilibria. Although both Kandori, Mailath, and Rob (1993) and Young (1993) analyze 2×2 two-player games, they use different approaches. Kandori, Mailath, and Rob (1993) consider one population which plays the field. Young (1993) suggests two populations and only one pair of individuals, one from each population, plays the game every period. Both approaches show the same *risk-dominant* equilibrium (introduced by J. Harsanyi and R. Selten in 1988) to be the prediction for 2×2 symmetric coordination games, when the probability of mutation goes to zero. The start was promising and generalizations for finite n-player games soon followed. Samuelson (1997) develops a straightforward generalization of the model of Kandori, Mailath, and Rob (1993). Almost simultaneously, Young (1998) generalizes his own model. His memory framework is rich enough and also allows to consider imitation rules. Josephson and Matros (2000) generalize Young's (1993) model for finite n-player games, but they analyze an imitation rule instead of his best reply rule.

The concept introduced by Foster and Young (1990) was very fruitful and was also applied to extensive form games. Nöldeke and Samuelson (1993) examine extensive form games based on the "population approach" of Kandori, Mailath, and Rob (1993). Jacobsen, Jensen, and Sloth (1999) study equilibrium selection in monotone two-type signalling games. They have Young's memory framework but restrict the attention to games where individuals have a separate memory on every information set in the game.

In this paper, I examine dynamic evolutionary processes driven by perpetual random shocks on extensive form games with perfect information. Games with a special condition are considered, each player has one node at most along each play in the game. Using the technique developed by Foster and Young (1991), I analyze the limit distribution, when the probability of mutation goes to zero. Every period, n individuals are randomly chosen from each of n finite populations to play an extensive form game. Each drawn individual observes a sample from the memory of the last m plays. As in the existing literature in this field (see, for example, Samuelson (1997)), I assume that players can observe the play through the game but not the complete strategies. In the first part of the paper, an individual knows the extensive game form[1], her own payoff function and expects the individuals in the opponent populations to play according to the probability distribution in her sample on reached

[1]For a formal definition, see Section 2.

nodes in the next period. Given these expectations, the individual plays a local pure best reply to the sample on the reached nodes and *any* local pure strategy on unreached nodes. However, the best reply must be specified, because the individuals only inspect previous plays, not the actual strategies. For that purpose, the individual's expectations about unused moves in the reached nodes in the sample are introduced. *Any* expectations/belief specifications are allowed, which seems more natural than the special beliefs specification in Nöldeke and Samuelson (1993) and Samuelson (1997)[2], and separate memories for different nodes in Jacobsen, Jensen, and Sloth (1999). Nöldeke and Samuelson (1993) show their locally stable outcomes to be subgame perfect. However, even without a particular belief specification, the backward induction outcome occurs with positive probability in the long-run also in my framework. This is the main result for the best reply process on extensive form games with perfect information. It might seem obvious, but traditional game theory predicts the backward induction outcome only if players know the extensive form of the game, every player knows that every other player knows that she knows that and so on. In my model, the individuals deal with very limited information. They know only what they can get in every terminal node, but they know nothing about the payoffs of the others. Thus there is no common knowledge. The individuals use a sample from the past plays to choose a best reply on the reached nodes, expecting that all other individuals will play according to the probability distribution in the sample.

In the second part of the paper, I follow Josephson and Matros (2000): The individuals observe a sample from the memory of the last m plays and the realized payoffs of these plays in the past. Then, every drawn individual imitates by choosing the local pure strategy which gave the highest (or the highest average) payoff in her sample on reached nodes. Sobel (1993) defines non-equilibrium evolutionary stable (NES) sets for two-player games in the normal form. His NES sets are "ideologically" similar to my sets, which will be observed in the long-run. In the limit, as the experimentation probability tends to zero, only a particular subset of all plays is played with positive probability. The main result in this part is that every such subset contains a backward induction outcome. I obtain that result in the situation where individuals only know payoffs on reached terminal nodes. An individual has limited memory and if a terminal node is not included in the sample, then she does not know of its possible payoff.

My analysis demonstrates that in both cases, for the best reply dynamics and for the imitation dynamics, there is a positive probability for the backward induction

[2]Although Nöldeke and Samuelson (1993) and Samuelson (1997) restrict their attention to specific beliefs, they analyze a broader class of games, for example, games with imperfect information.

outcome to occur in the long-run in extensive form games with perfect information. Thus two approaches, the best reply and the imitation approach, support the same outcome, which is not the case for games in normal form, not even in 2×2 two-player games. Robson and Vega-Redondo (1996), who modify the framework of Kandori, Mailath and Rob (1993), show that imitation leads to a Pareto efficient equilibrium, which may differ from the Kandori, Mailath, and Rob (1993) prediction: the risk-dominant equilibrium. Moreover, several studies show that evolutionary models do not select subgame perfect outcomes see, for example, Nöldeke and Samuelson (1993), Gale, Binmore and Samuelson (1995), Cressman and Schlag (1998), Samuelson (1997), and Fudenberg and Levine (1998). I prove that both the best reply and the imitation dynamics will always select a backward induction outcome, although it may not be the unique prediction. Moreover, there are some special cases where the backward induction outcome is the only long-run prediction. This result is different from Hart's (2000) main result, i.e. that the backward induction outcome is the unique evolutionarily stable outcome for dynamic models consisting of selection and mutation, when the mutation rate is low and populations are large. However, if individuals have an infinite memory, then we get the Hart's effect. After finitely many periods, all nodes will be reached with probability one and every individual will play a local pure best reply on *every* node. This leads to the unique prediction, the backward induction outcome. To obtain the same result, it is sufficient to have separate memories on every node, as in Jacobsen, Jensen, and Sloth (1999).

The paper is organized as follows. Section 2 describes the model. The best reply play is studied in Section 3. Section 4 defines the unperturbed and perturbed versions of the imitation play and derives general results for the limiting distribution of the process. Section 5 concludes.

2. The Model

A formulation of an extensive form is required for the analysis. The extensive form captures: (1) the set of players; (2) who moves when - the tree of the game and how players are partitioned; (3) what choices each player can make whenever it is her turn to move; (4) what players know when they move; (5) moves of nature; (6) payoffs to players as functions of the moves they select. This formulation is a la Kuhn (1950, 1953); see also Kreps and Wilson (1982), Weibull (2000), and any text-book in game theory. Call a collection of (1)-(5) as an *extensive game form*; see also Mas-Colell, Whinston, and Green (1995), pp. 226. Defines *play* according to Kuhn (1950): "...every particular instance in which a game is played from beginning to end is called a *play* of that game".

I will only consider games with perfect information, which meet a special condition; every player has one node at most along each play in the game. In such games, a *behavior strategy* for player i is a function that assigns a probability distribution over the finite set of moves available at each player i's node. A behavior strategy is *pure* if it assigns probabilities zero or one. A *local strategy* for player i, at her node θ, is a probability distribution, $s_{i\theta}$, over the finite set of moves available at θ.[3]

Let Γ be a finite extensive game form with perfect information and $N = \{1, ..., n\}$ denote the set of players, and each player $i \in N$ has a finite set of pure behavior strategies, S_i. There are n finite and non-empty populations of individuals. These need not be of the same size, nor need they be large. Each member of population i is a candidate to play position i in the game Γ. All individuals in population i have the payoff function $\pi_i : S_1 \times ... \times S_n \to \mathbb{R}$. Game Γ is played once every period $t \in \{0, 1, 2, ...\}$. In period t, n individuals, one from each population, are drawn at random to play the game. Every drawn individual $i \in N$ chooses a pure behavior strategy $s_i \in S_i$ for position i in the game, according to the behavior rules defined below. *Best reply* and *imitation* behavior rules are considered. A realized vector of the individuals' behavior strategies at period t, $s^t = (s_1^t, ..., s_n^t) \in S_1 \times ... \times S_n$, is called a *behavior profile* in period t. It corresponds to play τ^t of the game and the realized payoffs π^t in period t. Denote the set of all plays by Ω. The *history* of plays up to time t is the sequence $h^t = (\tau^{t-m+1}, ..., \tau^t)$, where m is a given positive integer, the *memory size*.

Let $h \in \Omega^m$ be an arbitrary history. Denote by $v_i = (\pi_i^{t-m+1}, ..., \pi_i^t) \in \mathbb{R}^m$ the corresponding payoff realizations for player i. After an individual i has been drawn to play game Γ, she chooses her behavior strategy s_i according to the behavioral rules based on a sample of size $k \in \mathbb{N}$, where $k \leq m$, from the current state h, the realized payoffs in this sample and some stochastic element. Formally, a behavioral rule for a drawn individual is a correspondence $B : \Omega^k \times \mathbb{R}^k \rightrightarrows \Delta(S_i)$, where $\Delta(S_i)$ represents the probability distribution on S_i. Note that the individuals in the players' positions observe k previous plays but cannot observe previous strategies. Now, the best reply and the imitation rules will be formally specified.

2.1. Best reply rule. The individual drawn to play position i in the game knows her payoff function $\pi_i : S_1 \times ... \times S_n \to \mathbb{R}$, the extensive game form, and observes a sample of size k from the history of size m from previous plays. She expects the individuals in the other player positions to play according to the probability distribution in her sample on the reached nodes. An individual in a player position cannot observe the complete strategies and must form beliefs about how play proceeds after unused moves on her reached nodes. Two types of individuals: (o) *optimists* and (p) *pessimists* are considered. The individual is an optimist if she believes that after her deviation, an unused move on her reached node in the sample, *the highest possible*

[3]See also Weibull (1995).

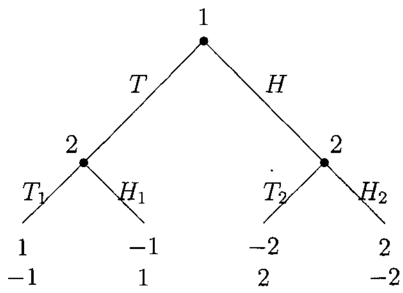

Figure 1: Sequential Matching Pennies game.

payoff will arise. The individual is a pessimist if she believes that after her deviation from the plays in the sample, *the lowest possible payoff* will arise. All possible beliefs give expected payoffs in the range between these two extreme types.

Behavior strategies are chosen as follows. At time $t+1$, each individual drawn to play the game inspects a sample of size k from the history of size m, taken from the plays up to time t. An individual chooses a pure local best reply to the opponent population's empirical play distribution in the sample on her reached nodes, given her beliefs about expected payoffs after unused moves on her reached nodes, and plays *any* pure local strategy on unreached nodes. If there is more than one pure local best reply, an individual chooses each of these with positive probability.

The sampling process begins in period $t = m+1$ from an arbitrary initial sequence of m plays, h^m. Then, a finite Markov chain is obtained on the state space Ω^m of sequences of length m, drawn from the play space Ω, with an arbitrary initial state, h^m. The resulting process is *ergodic*; thus, in the long-run, the initial state is irrelevant. Given a history $h^t = (\tau^{t-m+1}, ..., \tau^t)$ at time t, in the next period the process moves to a state of the form $h^{t+1} = (\tau^{t-m+2}, ..., \tau^t, \tau^{t+1})$. Such a state is called a *successor* of h^t. Call the process $BR^{m,k,0}$ *best-replay play with the memory of size* m *and the sample of size* k.

Example 1. Consider the best-replay play with the memory of size $m = 3$ and the sample of size $k = 3$ in the game in Figure 1. Let

$$h = ((T, H_1), (T, H_1), (T, H_1))$$

be the initial history. Note that the play space $\Omega = \{(T, H_1), (T, T_1), (H, H_2), (H, T_2)\}$. There are two cases.

(o) *Optimist* in player position 1 has one reached node. She expects that after her action T, player 2 will play H_1. This follows from the plays in the sample, because player 2 always played H_1. After move T, the expected payoff of player 1 is -1. Now, the individual in position 1 must form beliefs about her expected payoff after the unused move H on the reached node: she believes that her opponent will play H_2, which gives the highest possible payoff equal to 2 for player 1. This logic defines the strategy choice - H - which is the local best reply for the individual in player position 1. The individual in player position 2 has two nodes; one is reached in the sample, the other is not. The local pure strategies on the unreached node are T_2 and H_2. The local pure best reply on the reached node is H_1. This means that the individual in player position 2 either plays behavior strategy (H_1, T_2) or behavior strategy (H_1, H_2). Altogether, this implies that the unperturbed process will move either to state $h' = ((T, H_1), (T, H_1), (H, T_2))$ or state $h'' = ((T, H_1), (T, H_1), (H, H_2))$ in the next period.

(p) *Pessimist* in player position 1 must form beliefs about her expected payoff after the unused move H on the reached node: she believes that her opponent will play T_2, which gives the lowest possible payoff equal to -2 to player 1. As before, after move T, the probability distribution of the local strategies of player 2 is H_1 with probability one and T_1 with probability zero and the expected payoff of player 1 is -1. This defines the local pure best reply - T - for the individual in player position 1. The individual in player position 2 has two nodes; one is reached in the sample, the other is not. The local pure strategies on the unreached node are T_2 and H_2. The local pure best reply on the reached node is H_1. Therefore, the individual in player position 2 either plays behavior strategy (H_1, T_2) or behavior strategy (H_1, H_2). Altogether, this implies that the unperturbed process will move to state $h''' = ((T, H_1), (T, H_1), (T, H_1))$ in the next period.

The perturbed process can be described as follows. In each period, there is a small probability $\varepsilon > 0$ that any drawn individual in position i experiments by randomly choosing a local pure strategy on reached nodes instead of applying the best reply rule. The event that i experiments is assumed to be independent of the event that another individual playing this game in the opponent's position, j, experiments, which gives the perturbed Markov process $BR^{m,k,\varepsilon}$.

2.2. Imitation. As usual, an imitation behavior rule means that more prevalent actions/plays are more likely to be adopted in the near future. An imitation rule on the reached node might be simple, when an individual imitates the most frequently

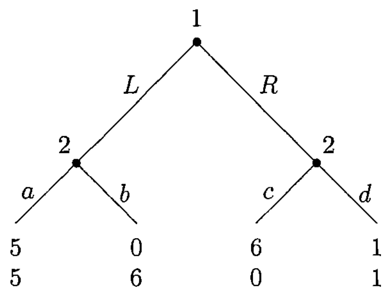

Figure 2: Sequential Prisoners' Dilemma game.

used pure local strategy in her sample from the memory, or more sophisticated, where the pure local strategy the individual imitates does not only depend on its frequency in the sample, but also on the associated realized payoffs. Note that if a player makes a choice on the reached node, then she can no longer influence the play, because every player has one node at most along any play. If there exists a node where the individual must make a choice and she cannot observe any actions on this node in her sample - the previous plays in the sample never reached this node - then the individual chooses each local pure strategy with a positive probability, on such an unreached node.

In every period t, each drawn individual i inspects a pair: her sample of size k from the memory of size m and the associated payoffs, and imitates a local strategy on the reached nodes in the sample. This defines a finite Markov chain on the finite state space Ω^m of plays. Given a history $h^t = (\tau^{t-m+1}, ..., \tau^t)$ at time t, the process moves to a state of the form $h^{t+1} = (\tau^{t-m+2}, ..., \tau^t, \tau^{t+1})$ in the next period. Such a state is called a *successor* of h^t. Call the process $IM^{m,k,0}$ *imitation play with the memory of size m and the sample of size k.*

Example 2. Consider the imitation play with the memory of size $m = 3$ and the sample of size $k = 3$ in the sequential Prisoners' Dilemma game in Figure 2. Let

$$h = ((L, a), (R, d), (L, b))$$

be the initial history. Given this history, the individuals in player positions observe the game tree and three terminal notes, as shown on Figure 3. Assume that all

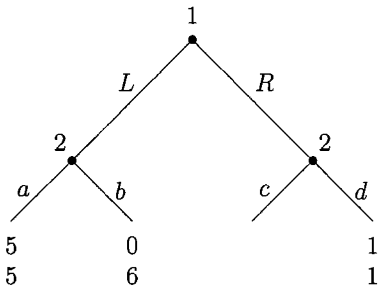

Figure 3: Imitation play on the Sequential Prisoners' Dilemma game.

individuals simply imitate the most frequent actions from the past on the reached nodes. Note that $\Omega = \{(L,a), (L,b), (R,c), (R,d)\}$. Then, the drawn individual in player position 1 chooses to play L (she observes L two times in the sample, but R only once). The drawn individual in player position 2 chooses $s_2 = (a,d)$ or $s_2' = (b,d)$, because she observes actions a and b the same number of times in the sample. Hence, the unperturbed process either will move to state $h' = ((R,d),(L,b),(L,a))$ or state $h'' = ((R,d),(L,b),(L,b))$ in the next period.

The perturbed process can be described as follows. In each period, there is a small probability $\varepsilon > 0$ that each drawn individual in player position i chooses a random local strategy on the reached nodes instead of playing according to the imitation rule. The event that the drawn individual assigned to play position i experiments is assumed to be independent of the event that another drawn individual assigned to play position j experiments for every $j \neq i$ and across time periods, which gives the perturbed Markov process $IM^{m,k,\varepsilon}$. Now, we are in a position to start the analysis.

2.3. Preliminaries. In this subsection, some definitions and results for the limiting distribution of the Markov process B^{ε} as the experimentation probability tends to zero are recalled from Young (1993) and Young (1998). Process B^{ε} can either be the best reply $BR^{m,k,\varepsilon}$ or the imitation $IM^{m,k,\varepsilon}$ play.

Definition 1. *(Young, 1998) B^{ε} is a* **regular perturbed Markov process** *if B^{ε} is irreducible for every $\varepsilon \in (0,\varepsilon^*]$, and for every state h, h', $B^{\varepsilon}_{hh'}$ approaches $B^{0}_{hh'}$*

at an exponential rate, i.e. $\lim_{\varepsilon\to 0} B^{\varepsilon}_{hh'} = B^{0}_{hh'}$ *and if* $B^{\varepsilon}_{hh'} > 0$*, for some* $\varepsilon > 0$*, then* $0 < \lim_{\varepsilon\to 0} \frac{B^{\varepsilon}_{hh'}}{\varepsilon^{r_{hh'}}} < \infty$*, for some* $r_{hh'} \geq 0$*. The real number* $r_{hh'}$ *is called the* **resistance** *of the transition from* h *to* h'*.*

A *recurrent class* of the process B^0 is a set of states such that there is zero probability of moving from any state in the class to any state outside, and there is a positive probability of moving from any state in the class to any other state in the class. A state h' is *absorbing* if it constitutes a singleton recurrent class.

Let an unperturbed Markov process B^0 have recurrent classes $h_1, ..., h_L$. For each pair of distinct recurrent classes, an *xy-path* is a sequence of states $\zeta = (h_x, ..., h_y)$ beginning in h_x and ending in h_y. The *resistance* of this path is the sum of resistances on its edges. Let r_{xy} be the least resistance over all xy-paths. Construct a complete directed graph with L vertices, one for each recurrent class. The weight on the directed edge $h_x \longrightarrow h_y$ is r_{xy}. A tree rooted at h_x is a set of $L-1$ directed edges such that, from every vertex different from h_x, there is a unique directed path to h_x in the tree. The *resistance of such a rooted tree* $\Im(x)$ is the sum of the resistances $r_{x'x''}$ on its $L-1$ edges. The *stochastic potential* $\rho(x)$ of a recurrent class h_x is the minimum resistance over all trees rooted at h_x.

Definition 2. *(Young, 1993) Let* μ^{ε} *be the unique stationary distribution of an irreducible process* B^{ε}*. A state* h *is* **stochastically stable,** *if* $\lim_{\varepsilon\to 0} \mu^{\varepsilon}(h) > 0$*.*

Theorem 1. *(Young, 1998) Let* B^{ε} *be a regular perturbed Markov process and* μ^{ε} *be the unique stationary distribution of* B^{ε} *for* $\varepsilon > 0$*. Then,* $\lim_{\varepsilon\to 0} \mu^{\varepsilon} = \mu^0$ *exists and is a stationary distribution of* B^0*. The stochastically stable states are exactly those contained in the recurrent classes of* B^0*, with a minimum stochastic potential.*

Definition 3. *A* **basin of attraction** *of state* $h \in \Omega^m$ *is a non-empty set of states* $B(h)$ *from which there is a positive probability to reach state* h *in a finite number of periods without any mistakes.*

Define a repetition of the same play τ as a state $h_\tau = (\tau, ..., \tau)$. The best-reply rule is first analysed and then imitation rule.

3. Best Reply

Now, we turn the attention to the limiting distribution of the best-reply play, as the experimentation probability tends to zero in this section. The stochastically stable states are characterized when both populations consist entirely of pessimist individuals and optimist individuals, respectively. Then, I prove that the backward induction outcome occurs with positive probability in the long-run, independent of a belief specification.

3.1. *(p) Pessimists.* Two results are proved in this subsection. First, every Nash equilibrium outcome generates an absorbing state, and second, the backward induction outcome is a stochastically stable state.

Proposition 1. *Let Γ be an extensive form game with perfect information such that*

(i) each player moves at most once along each play in the game,

(ii) all payoffs for each player are different,

(iii) k is large enough and $k/m \leq 1/2$,

then every recurrent class is an absorbing state. Moreover, state h_τ is an absorbing state of the unperturbed process $BR^{m,k,0}$ if and only if there exists a Nash equilibrium, which generates play τ.

Proof: First, show that a recurrent class must be an absorbing state. Suppose that there exists an arbitrary state $h^t = (\tau^{t-m+1}, ..., \tau^t)$ in some recurrent class which is not in the form h_τ. If $k/m \leq 1/2$, there is a positive probability that all individuals are drawn to play sample $\left(\tau^{t-k+1}, ..., \tau^t\right)$ in every period from $t+1$ to $t+k$ inclusive. Consider the last players along every play in the sample, who have at least two choices. They can be indifferent between choices only if these choices give the same expected payoff, which is impossible because of assumptions (ii) and (iii). Hence, the individual in those player positions will play only one action thereafter at those nodes. Now, roll back and consider the players who have at least two choices along the plays in the sample before the players previously considered. Using the same arguments, we conclude that the individual in that player position will also play one particular action on the reached nodes and so on, which defines a unique play τ. With positive probability, all individuals drawn to play thereafter will only sample from plays more recent than τ^t in every period from $t+k+1$ to $t+2k$ inclusive. Since all these samples have the form $(\tau, ..., \tau)$, there is a positive probability that an individual in every player position will choose the same strategy and the same play τ' will be played in every period from $t+2k+1$ to $t+3k$ inclusive. With positive probability, all individuals drawn to play thereafter will only sample from plays more recent than τ^{t+2k} in every period from $t+2k+1$ to $t+3k$ inclusive. Since all these samples have the form $(\tau', ..., \tau')$, a unique play τ'' will be played in every period from $t+3k+1$ to $t+4k$ inclusive. Continuing in the same manner, after a finite number of periods, the unperturbed process settles down to a state h_{τ^*}. State h_{τ^*} is such that if all players playing local best reply to sample $(\tau^*, ..., \tau^*)$, then play τ^* is repeated. Hence, there is a positive probability of obtaining a history $h_{\tau^*} = (\tau^*, ..., \tau^*)$, an absorbing state, in a finite number of periods.

Part *if.* Suppose that $(s_1, ..., s_n)$ is a Nash equilibrium and play τ is the corresponding play to this Nash equilibrium. In state h_τ, an individual in any player position has no incentives to deviate from play τ, because she cannot gain a higher payoff than in the Nash equilibrium, only a lower one, due to assumption (ii). The pessimist is supposed to believe in the lowest possible payoff realization after any deviation, it might be even lower than the payoffs after a deviation in the Nash equilibrium, which means that she will not deviate from play τ either. This means that state h_τ is absorbing.

Part *only if.* Suppose state h_τ to be an absorbing state and that play τ cannot be generated by any Nash equilibrium. Then, there exists a player j, who may play action a_j instead of following play τ, and action a_j is part of her "better strategy" s_j, which gives her a higher payoff. The individual in player position j does not play a_j in the absorbing state h_τ, because she believes such a deviation to lead to the lowest possible payoff, which must be less than her payoff when play τ is played. Fix these "lowest possible" actions as parts of the opponents' strategies on their unreached nodes. It will not influence play τ, but the individual in player position j has no incentives to deviate in the current state. This is true for any player position, and means that there exists a Nash equilibrium, which generates play τ. **End of proof.**

Now, show that the backward induction outcome is the long-run prediction.

Proposition 2. *Let Γ be an extensive form game with perfect information such that*
(i) play τ^* leads to the backward induction outcome,
(ii) each player moves at most once along each play in the game,
(iii) all payoffs for each player are different,
(iv) $k \geq 2$,
then state h_{τ^*} is a stochastically stable state of the perturbed process $BR^{m,k,\varepsilon}$, when $\varepsilon \longrightarrow 0$.

Proof: State h_{τ^*} is absorbing from proposition 1. Show that h_{τ^*} is a stochastically stable state. Unperturbed Markov process $BR^{m,k,0}$ has L absorbing states, $h_{x^1}, ..., h_{x^L}$, each of which corresponds to the play associate with Nash equilibrium. Consider an arbitrarily absorbing state, h_τ. In play τ, find the last player i, who does not play according to backward induction. Let the individual in this player position make a mistake and switch to the choice corresponding to backward induction, and all the following players also play according to backward induction thereafter. This happens with positive probability. The deviating player gets a higher payoff and the individual in this player position will play according to backward induction also in the next period. It means that one mistake leads to a basin of attraction of a new absorbing state $h_{\tau'}$. Once more, find the last player j in play τ', who does not play according to backward induction. This player j must be the predecessor of player

i. "Backward induction" deviation of player j from her new play τ' gives her higher payoff, if all the following players play according to backward induction. Continue in this way and remember that game Γ is finite. The process $BR^{m,k,\varepsilon}$ settles down in the absorbing state h_{τ^*} in a finite number of periods. The resistance of the rooted tree $\Im(h_{\tau^*})$ is equal to $L-1$ and means that the absorbing state h_τ is stochastically stable. **End of proof.**

3.2. (*o*) *Optimists.* I first show that there exists a unique recurrent class, and second, that the backward induction outcome is found in this class of stochastically stable states.

Proposition 3. *Let Γ be an extensive form game with perfect information such that*
(*i*) *each player moves at most once along each play in the game,*
(*ii*) *all payoffs for each player are different,*
(*iii*) *k is large enough,*
then there exists a unique recurrent class of the perturbed process $BR^{m,k,\varepsilon}$, which contains the backward induction outcome, when $\varepsilon \rightarrow 0$.

Proof: Suppose that the perturbed process $BR^{m,k,\varepsilon}$ is in state h_{τ^*}, where play τ^* leads to the backward induction outcome. An individual in any player position can deviate from this play only if she has another choice, which may lead to a higher payoff for her. In the following period, after the deviating move has been made, the individuals in the other player positions play local pure best replies on the new reached nodes. The new play is not the one associated with the backward induction outcome, which means that some players can play optimally (better for them). This means that there is a finite number of such deviations, but with a large enough sample size (assumption (*iii*)) all these plays will be present in the memory and players play local best replies, given the play distribution and their optimistic beliefs about expected payoffs after unused moves on the reached nodes. There are only plays in the memory, which give higher payoffs than what players have in the backward induction equilibrium, at least for one player. There is a finite number of plays different from τ^* and, for k large enough, individuals in all player positions start to play optimally - a local pure best reply, in all reached nodes or according to backward induction, and the play associated with the backward induction is repeated. Take an arbitrary state from a recurrent class. This state must contain play τ^*, because of assumption (*iii*). Hence, the unique recurrent class exists and the backward induction play τ^* is found in the states of this class. **End of proof.**

Main Result. The previous two subsections lead to the main result: the backward induction outcome is the long-run prediction, independent of the belief specification.

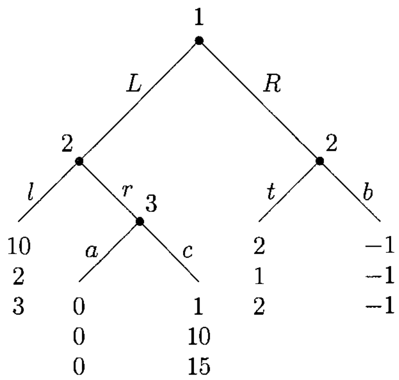

Figure 4: Multuplicity of stochastically stable states.

Theorem 2. *Let* Γ *be an extensive form game with perfect information, such that*
(i) *every player has one node at most along each play in the game,*
(ii) *all payoffs for each player are different,*
(iii) k *is large enough,*
then for any players' beliefs, the perturbed process $BR^{m,k,\varepsilon}$ *puts positive probability on the backward induction outcome when* $\varepsilon \longrightarrow 0$.

Proof: It has been shown that the perturbed process $BR^{m,k,\varepsilon}$ puts a positive probability on the backward induction outcome, either if individuals are optimists or pessimists, when $\varepsilon \longrightarrow 0$. Any individual's beliefs about expected payoffs after unused moves on reached nodes are in the range between those of the optimist and the pessimist. This means that the backward induction outcome will be observed in the long-run for any belief specification, when $\varepsilon \longrightarrow 0$. **End of proof.**

However, the backward induction outcome may not be the unique long-run prediction, which is illustrated by the following example.

Example 3. Consider the game in Figure 4. $(R,(r,t),c)$ is the backward induction equilibrium and it is associated with play $\tau^* = (R,t)$. State h_{τ^*} is a stochastically stable state of the perturbed process $BR^{m,k,\varepsilon}$, according to the previous proposition. Consider two cases. If individuals are pessimists, there are two other absorbing states h_τ and $h_{\tau'}$, where plays $\tau = (L,l)$ and $\tau' = (L,r,c)$ are associated with Nash equilibria $(L,(l,b),a)$ and $(L,(r,b),c)$. One mistake is enough to switch from state h_{τ^*}

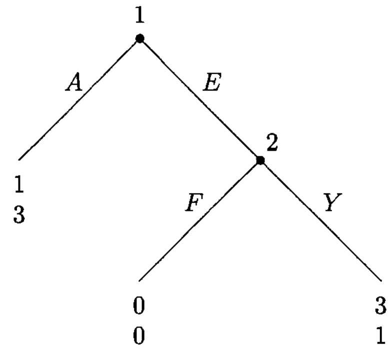

Figure 5: Entry Deterrence game

to the basin of attraction of absorbing state h_τ. Moreover, one mistake is enough to switch from state h_τ to the basin of attraction of absorbing state $h_{\tau'}$. At the same time, one mistake is enough to switch from state $h_{\tau'}$ to the basin of attraction of state h_{τ^*}. Hence, there are three absorbing states which are all stochastically stable. If individuals are optimists, there is one recurrent class of the perturbed process $BR^{m,k,\varepsilon}$, which includes states h_{τ^*}, h_τ and $h_{\tau'}$. This means that plays τ^*, τ, and τ' will be the long-run predictions, when $\varepsilon \to 0$.

3.3. Backward Induction Outcome as a unique prediction. Consider the following example:

Example 4. The Entry Deterrence Game in Figure 5 has two Nash equilibria in pure strategies (E, Y) and (A, F). Only the backward induction equilibrium (E, Y) generates the stochastically stable state $h_{(E,Y)}$, however, which is explained by the following: All nodes are reached in play (E, Y) and individuals in the player positions play local pure best replies. Given that the history was $h_{(E,Y)}$, the individual in player position 1 observes the game tree and her payoffs as shown in Figure 6 and expects that player 2 will play Y. Strategy E is a unique local best reply in this situation. The individual in player position 2 observes the game tree and her payoffs as shown in Figure 7 and expects that player 1 will play E. Strategy Y is a unique local best reply for player 2. There is an evolutionary pressure on every node in the game and individuals in the player positions play (E, Y). If one of the players deviates from play (E, Y), by mistake, then strategy E will be a local best reply for player 1 and

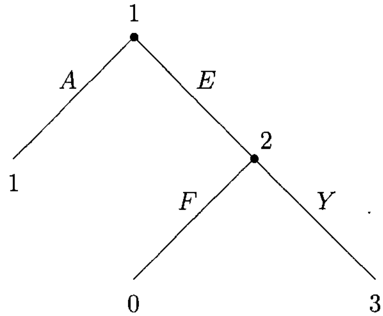

Figure 6: Entry Deterrence game

strategy Y will be a local best reply for player 2 in the next period. This is the driving force in the selection of the backward induction outcome. What about state $h_{(A)}$? The individual in player position 1 observes the game tree and her payoffs as shown in Figure 6 and must form expectations about what player 2 will play on her node. The individual in player position 2 observes the game tree and her payoffs as shown in Figure 7 and expects player 1 to play A. This means that both strategy F and strategy Y are local best replies. Suppose that player 2 chooses strategy Y. If player 1, by mistake, plays strategy E at that time, then play (E, Y) appears. Given that such events happen, the individual in player position 1 expects player 2 to play strategy Y and then strategy E is a unique local best reply. The same applies to the individual in player position 2. She plays a unique local best reply - strategy Y - in the next period. Therefore, the process moves to the basin of attraction of the stochastically stable state $h_{(E,Y)}$.

This example provides an intuition for the backward induction outcome to be a unique outcome of the limit distribution of the best reply play. If no player has a higher payoff after her deviation from the backward induction play, then the backward induction outcome is the unique prediction.

Definition 4. *(Balkenborg, 1993) The backward induction equilibrium outcome, identified by the terminal node ξ^*, is a* **strict outcome** *if no player can force entry into a subgame which has a terminal node with a higher payoff than the backward induction equilibrium.*

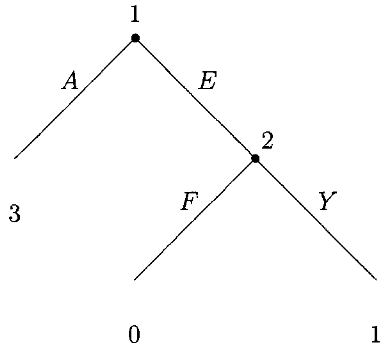

Figure 7: Entry Deterrence game

In other words, suppose that player i can force entry into the subgame $G(a_i)$. Let $\Xi(a_i)$ denote the set of terminal nodes contained in $G(a_i)$. If ξ^* is a strict outcome, then for all $\xi \in \Xi(a_i)$, $\pi_i(\xi^*) > \pi_i(\xi)$ must hold.

Theorem 3. *Let Γ be an extensive form game with perfect information such that*
(i) the backward induction equilibrium outcome is strict,
(ii) every player has one node at most along each play in the game,
(iii) all payoffs for each player are different,
(iv) k is large enough,
then state h_{τ^*}, which is associated with the backward induction play τ^*, is the unique stochastically stable state of the perturbed process $BR^{m,k,\varepsilon}$.

Proof: State h_{τ^*} is stochastically stable from Theorem 2. Show that it is a unique stochastically stable state. Consider two cases. If individuals are optimists, then there exists a unique recurrent class of the perturbed process $BR^{m,k,\varepsilon}$, which contains the backward induction outcome from Proposition 3. However, there is no player who wants to deviate from play τ^*, because the backward induction equilibrium outcome is strict. This means that the unique recurrent class of the perturbed process $BR^{m,k,\varepsilon}$ coincides with state h_{τ^*}, or that state h_{τ^*} is the unique stochastically stable state of the perturbed process, $BR^{m,k,\varepsilon}$.

If individuals are pessimists, then every recurrent class is an absorbing state, from Proposition 1. Using the same logic as in Proposition 2, it can be shown that the minimal resistance tree, rooted at state h_{τ^*}, has resistances on all edges equal to 1.

However, the other trees, rooted at another absorbing state, must have at least one resistance - from state h_{τ^*} - at least equal to 2 which means that state h_{τ^*} has a minimum stochastic potential. **End of proof.**

There are some special classes of games, where the backward induction outcome is a unique prediction, as a corollary of the theorem.

Corollary 1. *If game Γ is such that the backward induction play τ^* reaches every node in the game, then state h_{τ^*} is a unique stochastically stable state of the perturbed process $BR^{m,k,\varepsilon}$.*

The following corollary shows that in games with common interests, the backward induction outcome is a unique long-run prediction.

Corollary 2. *If game Γ is such that every player strictly prefers the terminal node corresponding to play τ^* to all other terminal nodes, then state h_{τ^*} is a unique stochastically stable state of the perturbed process $BR^{m,k,\varepsilon}$.*

4. Imitation

Refer to a state $h_\tau = (\tau, \tau, ..., \tau)$, where τ is any play in the tree, as a *convention.* In other words, a convention is a state where the drawn individuals played the same play in the last m periods. There are $|\Omega|$ conventions in game Γ. The following result shows that conventions correspond one to one with the recurrent classes of imitation plays.

Proposition 4. *Every convention is an absorbing state of the unperturbed process $IM^{m,k,0}$, if $k \leq m/2$.*

Proof: It is evident that for any $m > 0$, any convention is an absorbing state, since any sample from a convention will contain only one choice for each player along the play. We shall prove that the conventions are the only recurrent classes of the unperturbed process. Consider an arbitrary initial state, $h^t = (\tau^{t-m+1}, ..., \tau^t)$. There is a positive probability that all individuals drawn to play pick up the same sample $(\tau^{t-k+1}, ..., \tau^t)$ in every period from $t+1$ to $t+k$ inclusive. There is also a positive probability that all individuals drawn imitate the same play τ^* in every period from $t+1$ to $t+k$. All individuals drawn to play thereafter sample $(\tau^*, ..., \tau^*)$ and imitate behavior strategies which result in play τ^*. Thus, there is a positive probability of obtaining a history $h^{t+m} = (\tau^*, ..., \tau^*)$ at time $t+m$, a convention. It follows that the only recurrent classes of the unperturbed process are the conventions. **End of proof.**

This result is natural for the imitation dynamics, because individuals simply do not have any choice: they must imitate the unique play from their memory. See, for example, Björnerstedt and Weibull (1996). The following lemma is also standard (see, for example, Young (p.55, 1998)) and the proof is skipped.

Lemma 1. *Imitation play $IM^{m,k,\varepsilon}$ is a regular perturbed Markov process.*

I will develop a necessary condition for a convention to be in the support of the limiting distribution. Let us start from the definitions.

Definition 5. *A non-empty set of states $A(h)$ is a single-mutation* **neighborhood** *of state $h \in \Omega^m$, if it is a minimal set with the property that any single mistake (mutation) in state h leads to some state in $A(h)$.*

Now, define a set which will play a crucial role in finding the stochastically stable states of the imitation play.

Definition 6. *A non-empty set of conventions R is a* **construction** *in extensive form, if $A(h_\tau) \cap B(h_{\tau'}) = \emptyset$, for any convention $h_\tau \in R$ and $h_{\tau'} \notin R$. Such a set is a* **minimal construction,** *if it does not properly contain another construction.*

It is obvious that there exists at least one minimal construction in every game with perfect information in extensive form. Moreover, if there is more than one minimal construction, this must entail more than one mistake to move from one minimal construction to another by the previous definition. In this sense, all minimal constructions are disjointed. The following proposition tells us how to choose among minimal constructions in the long-run, which follows immediately from Lemma 1 and Theorem 1.

Proposition 5. *The perturbed process $IM^{m,k,\varepsilon}$ puts an arbitrarily high probability on the minimal construction(s) that minimize the stochastic potential, when $\varepsilon \to 0$.*

If there exists a unique minimal construction, the previous proposition guarantees that the support of the limiting distribution coincides with the conventions in that minimal construction. What is the structure of a minimal construction? To answer this question, a particular imitation rule must be specified. So far, a drawn individual knows the extensive game form. How the imitation rule is related to payoffs has not yet been specified. If individuals choose the most frequent action on every reached node, then *every* convention is a minimal construction for $k \geq 3$, and *all* minimal constructions have the same stochastic potential. We then turn the attention to imitation rules that are sensitive to payoffs.

4.1. Sophisticated Imitation. Assume that every individual does not only know the extensive game form, but also the realized payoffs for every player. A drawn individual i inspects the pair: a sample of size k from the history of m plays and the corresponding payoff realizations in the sample $v = (\pi^{t-m+1}, ..., \pi^{t}) \in \mathbb{R}^k$ for all players. She imitates in the sophisticated or in the successful way on all reached nodes. By the sophisticated imitation, an individual puts herself in every node starting from the end and imitates "the most successful" actions - actions, which give the highest (or the highest average) payoff to the player at that node, among all actions at this node from the sample - given the choices in all following nodes. This procedure is similar to Kuhn's (1953) algorithm. To write this formally, some notation is needed, however.

Let $W_1, ..., W_L$ be a sequence of sets of nodes constructed in the following way. W_1 consists of all nodes followed only by terminal nodes. W_l for $l > 1$ consists of all nodes not included in the union $W_1, ..., W_{l-1}$ and followed only by nodes in the union $W_1, ..., W_{l-1}$ and terminal nodes.

The sophisticated imitation for a drawn individual begins by determining "the most successful" actions for every player position on all reached nodes from set W_1. Formally, for each node ϖ^i of player i from W_1, this is the action maximizing player i's payoff (or the average player i's payoff) conditional on ϖ^i being reached. Note that if ϖ^i was never reached in the sample from the memory, any pure local strategy might be played. The only requirement is that every possible pure local strategy has a positive probability in this case.

Then, the individual works back to set W_2, and determines "the most successful" action among used actions in the sample for every player position, given that the individual in player position i on set W_1 will play the action previously determined. The algorithm proceeds to "roll back", just as in usual backward induction for games with perfect information, until the last set W_L is reached. Call this rule a sophisticated imitation rule. Note once more that on every *unreached node*, a drawn individual puts a positive probability for each pure local strategy.

Example 5. Define a sequence W_1, W_2, W_3 for the game from Example 3, Figure 4. Set W_1 contains two nodes: the node, where player 3 must make her choice and the node, where player 2 has two local strategies, t and b. Set W_2 contains one node, the one, where player 2 has two local strategies, l and r. Set W_3 contains one node, the one, where player 1 must make her choice.

Proposition 6. *Suppose that convention h_τ belongs to a minimal construction, and that player i has node θ on play τ and move a_i on this node, such that a_i is not on play τ. If there exists a terminal node in subgame $G(a_i)$ where player i has the same* *or a higher payoff than when playing play τ, then convention $h_{\tau'}$ is also in the same minimal construction, if play τ' generates that payoff.*

Proof: Note that the drawn individual in player position i needs to make one mistake only to switch from play τ to τ'. By the definition of the minimal construction, convention $h_{\tau'}$ must also be in the same minimal construction. **End of proof.**

That proposition indicates that if a player can get a higher payoff in some subgame where she can force entry, play τ corresponding to this higher payoff generates convention h_τ, which is also in the same minimal construction as the initial convention. The intuition is straightforward. One mistake is always sufficient to reach any terminal node from any convention. The deviating player must make a mistake and all other players can choose all corresponding moves with positive probability. It turns out that if the deviating player receives higher payoff in the new play, the individual in this player position might repeat that new move next time. This means that the process moves to the basin of attraction of the new convention.

Will the backward induction outcomes be observed in the long-run? The following proposition answers this question.

Proposition 7. *Every minimal construction contains a backward induction outcome.*

Proof: Take any convention h_τ from a minimal construction. Consider the player who makes the last move along play τ. Though one mistake, she can get the highest possible payoff among all terminal nodes after her node, with positive probability. It means that the new play τ' might be repeated and another convention $h_{\tau'}$ must be in the same minimal construction. Moreover, play τ' is the backward induction equilibrium play in the last subgame. Consider player who makes the move before last along the play. With one mistake, she can play the backward induction equilibrium play on this subgame and so on. After a finite number of periods, continuing in the same manner, convention h_{τ^*}, where τ^* is the backward induction play, appears in the same minimal construction. **End of proof.**

If a game has a unique backward induction equilibrium, there exists a unique minimal construction.

Proposition 8. *Let Γ be a generic extensive form game with perfect information. Then, there exists a unique minimal construction. Moreover, the minimal construction contains the convention, which generates the backward induction equilibrium outcome.*

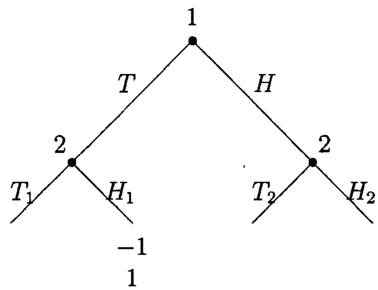

Figure 8: Sequential Matching Pennies game.

Proof: If Γ is a generic extensive form game of perfect information, there exists a unique backward induction equilibrium. The claim follows immediately from the previous proposition. **End of proof.**

The long-run prediction is not so sharp even in "simple" games, however.

Example 6. Consider the sequential Matching-Pennies game in Figure 1. The crucial assumptions is that the individual in player position 2 chooses every pure local strategy, with a positive probability on unreached nodes. This assumption drives results in this example. The unique minimal construction contains conventions $h_{(T,H_1)}$, $h_{(H,H_2)}$, $h_{(H,T_2)}$, $h_{(T,T_1)}$. Why is that? Start from the backward induction play (T, H_1) and the corresponding convention, $h_{(T,H_1)}$. In state $h_{(T,H_1)}$, the individuals in player positions 1 and 2 observe the game tree and the payoffs as shown in Figure 8. If the individual in player position 1, by mistake, plays H, there is a positive probability that play (H, H_2) might appear. In this new state, the individuals in player positions 1 and 2 observe the game tree and the payoffs as shown in Figure 9 with positive probability. The individual in player position 1 expects player 2 to imitate behavior strategy (H_1, H_2) and plays "the most successful" action H against it. There is a positive probability that state $h_{(H,H_2)}$ arises. Now, one mistake from an individual in player 2 position leads to the basin of attraction of state $h_{(H,T_2)}$ and so on. Thus, all plays will be observed in the long-run.

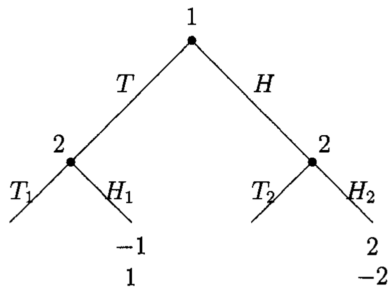

Figure 9: Sequential Matching Pennies game.

4.2. Backward Induction Outcome as a unique prediction. The backward induction outcome might be a unique long-run prediction. Interestingly, the same class of games, which was received for the best-reply play, has this property also.

Proposition 9. *If the backward induction equilibrium outcome is strict, then state h_{τ^*}, the minimal construction, associated with the backward induction play τ^*, is the unique stochastically stable state of the perturbed process $IM^{m,k,\varepsilon}$, when $\varepsilon \longrightarrow 0$.*

Proof: There is a unique minimal construction in the game, from Proposition 8, which contains state h_{τ^*}, from Proposition 7. There must be more than one mistake to move from state h_{τ^*} to the basin of attraction of another convention h_τ, since the backward induction equilibrium outcome is strict. By the definition of the minimal construction, convention h_τ is the only state in this construction. **End of proof.**

There are some special classes of games, where the backward induction outcome is a unique prediction, as a corollary of the proposition.

Corollary 3. *If game Γ is such that the backward induction play τ^* reaches every node in the game, then state h_{τ^*} is a unique stochastically stable state of the perturbed process, $IM^{m,k,\varepsilon}$.*

The following corollary shows that in games with common interests, the backward induction outcome is a unique long-run prediction.

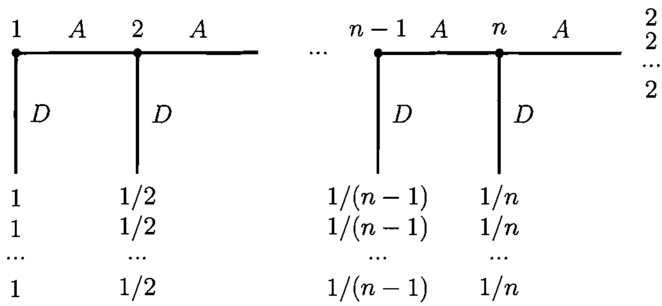

Figure 10: Fudenberg and Tirole's game

Corollary 4. *If game Γ is such that every player strictly prefers the terminal node corresponding to play τ^* to all other terminal nodes, then state h_{τ^*} is a unique stochastically stable state of the perturbed process, $IM^{m,k,\varepsilon}$.*

Example 7. The example from Fudenberg and Tirole (1991) in Figure 10 illustrates the previous corollary. Consider the n-player game in Figure 10, where each player $i < n$ can either end the game by playing "D" or play "A" and give the move to player $i + 1$. If player i plays "D", each player gets $1/i$; if all players play "A", each gets 2.

From the corollaries, it follows that convention $h_{(A,A,...,A)}$ is the unique minimal construction and the stochastically stable state. Note that play $(A, A, ..., A)$ reaches every node in the game and reaches the terminal node, which every player strictly prefers to all other terminal nodes.

5. Conclusion

I study best reply and imitation dynamics on games in extensive form with perfect information, when every player has one node along each play at most. My main result is that the backward induction outcome occurs with positive probability in the long-run in both cases. Cressman and Schlag (1998) who, among others, analyse deterministic dynamics point out that the subgame perfect equilibrium outcome may not be the unique outcome supported by evolutionary dynamics. I provide the conditions when the backward induction outcome is the unique long-run prediction in both cases. Samuelson (1997) receives the same class of games with the unique backward induction outcome, but he imposes restrictions on beliefs.

In the second part of the paper, I consider stochastic imitation on extensive form games. To the best of my knowledge, only deterministic models of imitation have previously been analysed. The stochastic best reply dynamics were studied in the literature, where the common feature was to find the self-confirming equilibria (introduced in D. Fudenberg and D. Levine (1993)) and then choose among these, see, for example, Nöldeke and Samuelson (1993), Samuelson (1997), and Fudenberg and Levine (1998). I make a special assumption about the individuals' imitation behavior on the nodes without a previous past in the sample. These nodes - nodes outside the play - cannot be observed in a sample from the memory and individuals might choose any pure local strategy on these nodes, but every such local strategy must have a positive probability. Individuals have nothing to imitate and their choices are random on such nodes. The first result is that absorbing states - repetitions of the same play - are the same for any specification of imitation rules, which is standard for all imitation dynamics see, for example, Björnerstedt and Weibull (1996) for games in normal form.

I specify imitation rules and analyse sophisticated imitation rules, where the drawn individuals imitate the actions giving the highest (or the highest average) payoff for their player positions in the game. The set of stochastically stable states is characterized and shows that if the backward induction outcome is strict, then it is the unique outcome of the game.

References

[1] Balkenborg, D. (1993): "Strictness, Evolutionary Stability and Repeated Games with Common Interests", *CARESS working paper*, University of Pennsylvania.

[2] Björnerstedt, J. and J. Weibull (1996): "Nash Equilibrium and Evolution by Imitation", in *The Rational Foundations of Economic Behavior.* K. Arrow et al. (Eds.), MacMillan, London, 155-171.

[3] Cressman, R., and K. Schlag (1998): "The Dynamic (In)Stability of Backwards Induction", *Journal of Economic Theory*, 83, 260-285.

[4] Foster, D. and P. Young (1990): "Stochastic Evolutionary Game Dynamics", *J. Theor. Biol.* 38, 219-232.

[5] Fudenberg, D. and D. Levine (1993): "Self-Confirming Equilibrium," *Econometrica*, 61, 523-545.

[6] Fudenberg, D. and D. Levine (1998): *The Theory of Learning in Games.* MIT Press.

[7] Fudenberg, D. and J. Tirole (1991): *Game Theory.* MIT Press.

[8] Freidlin, M. and A. Wentzell (1984): *Random perturbations of dynamical systems.* Springer, New York.

[9] Gale, J., Binmore, K., and L. Samuelson (1995): "Learning to be Imperfect: The Ultimatum Game," *Games and Economic Behavior*, 8, 56–90.

[10] Groes, E., Jacobsen, H., and B. Sloth (1999): "Adaptive learning in extensive form games and sequential equilibrium," *Economic Theory*, 13, 125-142.

[11] Harsanyi J. and R. Selten (1988): *A General Theory of Equilibrium in Games*, Cambridge, MIT Press.

[12] Hart, S. (2000): "Evolutionary Dynamics and Backward Induction," *mimeo*, Center for Rationality and Interactive Decision Theory; Department of Economics; and Department of Mathematics; The Hebrew University of Jerusalem.

[13] Jacobsen, H., M. Jensen, and B. Sloth (1999): "Evolutionary Learning in Signalling Games," *mimeo*, University of Copenhagen.

[14] Josephson, J. and A. Matros (2000): "Stochastic imitation in finite games", *Working Paper*, 363, Stockholm School of Economics.

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

[16] Kreps, D., and R. Wilson (1982): "Sequential Equilibria", *Econometrica*, 50, 863-894.

[17] Kuhn, H. (1950): "Extensive Games," *Proceedings of the National Academy of Sciences*, 36, 570–576.

[18] Kuhn, H. (1953): "Extensive Games and the Problem of Information," in *Contributions to the Theory of Games II*, H. W. Kuhn and A. W. Tucker (eds.), *Annals of Mathematics Studies*, 28, Princeton University Press, 193–216.

[19] Mas-Colell, A., M. Whinston, and J. Green (1995): *Microeconomic Theory*, Oxford University Press.

[20] Maynard Smith, J. and G. Price (1973): "The Logic of Animal Conflict", *Nature*, 246, 15-18.

[21] Nash, J. (1950): "Non Cooperative Games", Ph.D. thesis, Princeton University.

[22] Nöldeke, G., and L. Samuelson (1993): "An Evolutionary Analysis of Backward and Forward Induction", *Games and Economic Behavior*, 5, 425-454.

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

[24] Samuelson, L. (1997): *Evolutionary Games and Equilibrium Selection.* Massachusetts: MIT Press.

[25] Sobel, J. (1993): "Evolutionary stability and efficiency", *Economics Letters*, 42, 301-312.

[26] Weibull, J. (1995): *Lecture Notes on Game Theory*, Stockholm School of Economics.

[27] Weibull, J. (2000): "Testing Game Theory", *Working Paper*, 382, Stockholm School of Economics.

[28] Young, P. (1993): "The evolution of conventions", *Econometrica*, 61, 57-84.

[29] Young, P. (1998): *Individual Strategy and Social Structure.* Princeton University Press.

Part III: Game Theory

Chapter 5: Elimination Cup Tournaments with Player Endowment

II

Elimination Cup Tournaments with Player Endowment

Alexandr Matros*
Department of Economics
Stockholm School of Economics

April 15, 2001

Abstract. This paper models T-round elimination cup tournaments where players have an endowment. I prove that players *always* spend a higher share of their endowments in early than in later rounds in symmetric equilibrium. Equal distribution among T rounds takes place only in the winner takes it all case.

Keywords: Tournaments, contest, symmetric equilibrium, sports.

JEL classification: C72.

1. Introduction

Tournaments are usually associated with sports: tennis, football and chess, for example. There is an economic literature, however, where authors analyze applications of the tournament structure for labor markets, for example, Lazear and Rosen (1981) and Rosen (1981, 1986). So far, the question of incentives has been considered when tournaments and their applications for sports and economics have been analyzed in the literature. The authors insist that high differences in prizes in the last round(s) must provide enough incentives for players to insert an optimal effort in these rounds; see, for example, Lazear and Rosen (1981) and Rosen (1981).

The focus of this paper is on elimination cup tournaments. I assume that players have fixed endowments (energy in tennis and football tournaments, novelties in chess tournaments, budgets for election campaigns, etc.) and have to maximize their expected payoffs by distributing this endowment among all rounds in the tournament. Every round, the winners of the previous round are matched in pairs and invest a part of their endowments to increase the probability of winning the contest. Their behavior must be strategic: Every player must distribute her endowment for the whole

*I would like to thank Guido Friebel and Jonas Björnerstedt for very useful discussions. I have also benefited from comments by Tatiana Damjanovic, Elena Palzeva, Karl Wärneryd, and seminar participants at the Stockholm School of Economics. Special thanks to Christina Loennblad for her fine editing. Financial support from the Jan Wallander and Tom Hedelius Foundation is gratefully acknowledged. Remaining errors are solely due to the author.

tournament, given that if she loses one round, then she is out. This strategic problem is different from the problem analyzed in the contest literature, where players must decide how much effort to spend to win the prize in one contest; see, for example, Dixit (1987, 1999), Baik and Shogren (1992), and Baye and Shin (1999). I consider a chain of contests.

I show that there does not exist a prize scheme where the players spend most of their endowments in the very last rounds. In the symmetric equilibrium, all players actually invested more in the previous round than in the current one, and more in the current round than in the following, if they are still in the tournament. The intuition is straightforward: if you keep your endowment until the very last rounds to get higher prizes, then you will be eliminated long before these very last rounds. The extreme case, equal endowment distribution, arises in the winner takes it all case.

Section 2 introduces the elimination tournament and brings results. Section 3 provides a discussion.

2. The Model

Consider a T-round elimination tournament with 2^T players fighting for prizes (payoffs). In round 1, all players are matched in pairs for fights, where only the winners of the current round continue to fight for higher payoffs in the following rounds. All losers get payoff Z_T and are out of the tournament. In round 2, the winners of the first round, 2^{T-1} players, are matched in pairs for new fights. Every winner proceeds to the next round and all losers receive Z_{T-1} and are out of the tournament, and so on. Finally, in round T, only two players remain. The winner of the final gets Z_0 and the loser receives payoff Z_1. In other words, in every round $k \in \{1, ..., T\}$, there are 2^{T-k+1} players who are matched in pairs for new fights. All losers in round k receive payoff Z_{T-k+1} and are out of the tournament and all winners continue. I make the standard assumption that prizes increase from round to round

$$A1. \ Z_0 \geq Z_1 \geq ... \geq Z_T \geq 0. \tag{1}$$

Each player i has an initial fixed endowment E, and must decide how to distribute this endowment in all T rounds. Denote a distributed part of player i's endowment in round k by x^i_{T-k}, where $T-k$ is exactly the number of remaining rounds in the tournament. If player i chooses to use a part $x^i_{T-k} \in [0, E]$ of her initial endowment in round k, when her opponent in round k, player j, chooses a part $x^j_{T-k} \in [0, E]$, then player i wins this fight with probability

$$\frac{f\left(x^i_{T-k}E\right)}{f\left(x^i_{T-k}E\right) + f\left(x^j_{T-k}E\right)}, \tag{2}$$

where $f(x)$ is a positive and increasing function:

$$A2.\ f(x) > 0,\ f'(x) > 0 \text{ on interval } [0, E]. \tag{3}$$

A pure strategy for player i is a rule $(x_1^i, ..., x_T^i)$, which assigns a part of her endowment for every round in the tournament such that $\sum_{k=1}^{T} x_k^i = E$, $x_k^i \geq 0$ for any $i \in \left\{1, ..., 2^T\right\}$ and $k \in \{1, ..., T\}$. The T-round elimination tournament is a symmetric game with 2^T players and at least one symmetric equilibrium, which follows from an application of Kakutani's Fixed-Point Theorem.

Proposition 1. *Suppose that assumptions* (1), (3) *hold, then the T-round elimination tournament has at least one symmetric equilibrium.*

Proof: The proof is a generalization of the result for two-player symmetric games in Weibull (Proposition 1.5, 1995). The set of all strategies for player i is a T dimensional simplex Δ, where vertex k of the simplex is a strategy where the whole endowment E will be spent in round k and nothing in all the other rounds. Simplex Δ is non-empty, convex, and compact. Fix all players but player i, and denote these players as $-i$. Suppose that players $-i$ can only choose the same strategy $x \in \Delta$, which is the diagonal in simplex Δ^{2^T-1}. This diagonal is exactly simplex Δ. The best reply correspondence $\beta^i(x, ..., x) = \beta^i(x)$ of player i to the same strategies for players $-i$ is upper hemi-continuous. Moreover, $\beta^i(x) \subset \Delta$ is convex and closed. By the Kakutani's theorem, there exists at least one fixed point: $x^* \in \beta(x^*)$, $x^* \in \Delta$. This is true for any player i and leads to the statement of the proposition. **End of proof.**

Proposition 1 guarantees the existence of a symmetric equilibrium. Let $(x_1, ..., x_T)$ be a symmetric equilibrium, where x_k is a part of the endowment every player spends in round $T - k$. It will be shown that a symmetric equilibrium in pure strategies is unique if function $f(x)$ is "not very convex":

$$A3.\ f(x) f''(x) - [f'(x)]^2 \leq 0 \text{ on interval } [0, E]. \tag{4}$$

Note that concave and linear functions belong to this class. I will call the following prize structure

$$Z_0 > Z_1 = Z_2 = ... = Z_T \geq 0,$$

the winner takes it all.

Proposition 2. *Suppose that assumptions* (1), (3) *and* (4) *hold. Then, in symmetric equilibrium $(x_1, ..., x_T)$, it must be $x_1 \leq x_2 \leq ... \leq x_T$, for any prize structure $(Z_0, Z_1, ..., Z_T)$. Equal distribution among all rounds $x_1 = x_2 = ... = x_T$ takes place only in the winner takes it all cases.*

Proof: Given the opponent's endowment distribution $(y_1, ..., y_T)$, the player's distribution decision x_k in the $k-th$ round of the tournament is determined by the solution of:

$$V_k = \max_{x_k} \left[\frac{f(x_k)}{f(x_k) + f(y_k)} V_{k-1} + \frac{f(y_k)}{f(x_k) + f(y_k)} Z_k \right], \text{ for } k = 2, ..., T, \tag{5}$$

and in the final round by:

$$V_1 = \frac{f(E - x_T - ... - x_2)}{f(E - x_T - ... - x_2) + f(E - y_T - ... - y_2)} Z_0 +$$
$$\frac{f(E - y_N - ... - y_2)}{f(E - x_T - ... - x_2) + f(E - y_T - ... - y_2)} Z_1. \tag{6}$$

The first order condition for problem $(5) - (6)$ is

$$\frac{f'(x_k) f(y_k)}{[f(x_k) + f(y_k)]^2} [V_{k-1} - Z_k] +$$
$$\frac{f(x_k)}{f(x_k) + f(y_k)} \times \frac{f(x_{k-1})}{f(x_{k-1}) + f(y_{k-1})} \times ... \times \frac{f(x_2)}{f(x_2) + f(y_2)} \times$$
$$\times \left[\frac{-f'(E - x_T - ... - x_2) f(E - y_T - ... - y_2)}{[f(E - x_T - ... - x_2) + f(E - y_T - ... - y_2)]^2} (Z_0 - Z_1) \right] = 0.$$

In symmetric equilibrium, $x_T = y_T, ..., x_2 = y_2$ and we have

$$\frac{f'(x_k)}{4f(x_k)} \left[\frac{V_{k-2} + Z_{k-1}}{2} - Z_k \right] = \left(\frac{1}{2}\right)^{k-1} \left[\frac{f'(E - x_T - ... - x_2)}{4f(E - x_T - ... - x_2)} (Z_0 - Z_1) \right].$$

Finally, we get

$$\frac{f'(x_k)}{f(x_k)} \left(2^{k-2} Z_{k-1} + 2^{k-3} Z_{k-2} + ... + (Z_1 + Z_0) - 2^{k-1} Z_k \right) =$$
$$= \frac{f'(E - x_T - ... - x_2)}{f(E - x_T - ... - x_2)} (Z_0 - Z_1). \tag{7}$$

Assumption (4) guarantees that the left-hand side (LHS) in equation (7) is a strictly decreasing function of x_k on the interval $[0, E]$, and the right-hand side (RHS) in the same equation is a strictly increasing function of x_k on the interval $[0, E]$. There are two cases. Either equation (7) does not have a solution or it has a unique solution, since it defines the intersection of a decreasing and an increasing functions. Denote

$x = x_T + ... + x_{k+1} + x_{k-1} + ... + x_2$. Player i must distribute endowment part $(E-x)$ between period k and the last period. Note that $x_k \geq x_1$ if and only if $LHS\left(\frac{E-x}{2}\right) \geq RHS\left(\frac{E-x}{2}\right)$. If x_k is equal to $\frac{E-x}{2}$, or endowment parts in period k and the last period are equal, then

$$LHS\left(\frac{E-x}{2}\right) = \frac{f'\left(\frac{E-x}{2}\right)}{f\left(\frac{E-x}{2}\right)}\left(2^{k-2}Z_{k-1} + 2^{k-3}Z_{k-2} + ... + (Z_1 + Z_0) - 2^{k-1}Z_k\right)$$

and

$$RHS\left(\frac{E-x}{2}\right) = \frac{f'\left(\frac{E-x}{2}\right)}{f\left(\frac{E-x}{2}\right)}(Z_0 - Z_1).$$

Note that from assumption (1)

$$\left(2^{k-2}Z_{k-1} + 2^{k-3}Z_{k-2} + ... + (Z_1 + Z_0) - 2^{k-1}Z_k\right) =$$

$$(Z_0 - Z_1) + \left(2^{k-2}Z_{k-1} + 2^{k-3}Z_{k-2} + ... + 2Z_1 - 2^{k-1}Z_k\right) \geq$$

$$\geq (Z_0 - Z_1),$$

with equality if and only if $Z_T = ... = Z_1$. Hence, in the symmetric equilibrium $x_k \geq x_1$ for *any prize scheme* $(Z_0, Z_1, ...Z_T)$, with equality if and only if $Z_T = ... = Z_1$.

Using the same logic, for any $k \geq 3$, it can be shown that $x_k \geq x_2$ for *any prize scheme* $(Z_0, Z_1, ...Z_T)$, with equality if and only if $Z_T = ... = Z_1$, and so on. Finally, we have that the optimal endowment distribution in the symmetric equilibrium must be $x_1 \leq x_2 \leq ... \leq x_T$, for any prize structure $(Z_0, Z_1, ..., Z_T)$. Equal distribution among all rounds $x_1 = x_2 = ... = x_T$ takes place only in the winner takes it all case. **End of proof.**

Corollary 1. *When the differences in prizes $Z_2 > Z_3 > ... > Z_T$ is positive, the players' endowment distribution decreases across rounds.*

Assumptions (1) and (3) are common and I will illustrate the role of assumption (4) by the following example, where function $f(x)$ is "very convex".

Example. Suppose that there are two rounds, $T = 2$; the endowment is equal to one, $E = 1$; $f(x) = e^{(x+1)^2}$; and the following prize structure, $Z_2 = 0$, $Z_1 = 1$ and $Z_0 = 3$. From Proposition 1, there exists a symmetric equilibrium (x_1, x_2). Note that $f'(x) = 2(x+1)e^{(x+1)^2}$ and find the equilibrium (x_1, x_2). Condition (7) will be

$$2(x_2 + 1) = (2 - x_2)$$

or

$$x_2 = 0.$$

Hence, in the symmetric equilibrium $(x_1, x_2) = (1, 0)$, every player spends *all her* endowment in the final round.

3. Discussion

I consider T-round elimination cup tournaments where players have an endowment. The symmetric equilibrium is shown to exist and be unique. Moreover, in this equilibrium, all players spend most of their endowments in round 1 and least in the last round, T. The intuition is straightforward; the expected payoffs are much higher in round 1 than in all other rounds. This simple model helps understand why tennis players or football teams are not at their best in the finals: they spend too much effort in the previous rounds.

All players are usually not equal however[1]. Some of them know that they have a very small, if any chance, of winning the tournament. These players will spend most of their endowments in the first round(s). Thereby, they have a real chance of passing one or two rounds and be eliminated at a later stage. This situation is interesting to analyze and is observed in many tournaments, when an unknown player beats a famous player in the first or second round and is easily eliminated by somebody else in the next round.

How can the situation be improved? The main result of the paper is that equal distribution of the endowment over all, T, rounds implies the winner takes it all prize scheme. This is contrary to the result in Rosen (1986), who considers independent individual's effort choice in every round of a tournament. He finds that prizes must increase over rounds. In Friebel and Matros (2001), we made the next step and investigate the impact of the assumption of a fixed initial effort endowment in elimination tournaments on the allocation of effort and optimal prize scheme in Rosen's (1986) framework.

References

[1] Baik, Kyung H. and Shogren, Jason F., "Strategic Behavior in Contests: Comment." *American Economic Review*, March 1992, 82(1), pp. 359-62.

[2] Baye, Michael R. and Shin, Onsong, "Strategic Behavior in Contests: Comment." *American Economic Review*, June 1999, 89(3), pp. 691-93.

[1]People often have different endowments in different tournaments. I consider a symmetric case, since participants in the tournament are usually treated as symmetric players.

[3] Dixit, Avinash, “Strategic Behavior in Contests.” *American Economic Review*, December 1987, 77, pp. 891-98.

[4] Dixit, Avinash, “Strategic Behavior in Contests: Reply.” *American Economic Review*, June 1999, 89(3), pp. 694.

[5] Friebel, Guido and Matros, Alexandr, "A Theory of Exhaustion in Tournaments", *Mimeo*, Stockholm School of Economics, April 2001.

[6] Lazear, Edward P. and Rosen, Sherwin, “Rank-Order Tournaments as Optimum Labor Contracts.” *Journal of Political Economy*, October 1981, 89, pp. 841-64.

[7] Rosen, Sherwin, “The Economics of Superstars.” *American Economic Review*, December 1981, 71, pp. 845-58.

[8] Rosen, Sherwin, “Prizes and Incentives in Elimination Tournaments.” *American Economic Review*, September 1986, 76, pp. 701-15.

[9] Weibull, Jörgen W., "*Evolutionary Game Theory*", Massachusetts, MIT Press, 1995.

EFI
The Economic Research Institute

Reports since 1995
A complete publication list can be found at www.hhs.se/efi

Published in the language indicated by the title

2001

Björklund, C., Work Motivation – Studies of its Determinants and Outcomes.
Eriksson, R., Price Responses to Changes in Costs and Demand.
Hill, M., Essays on Environmental Policy Analysis: Computable General Equilibrium Approaches Applied to Sweden.
Lange, F. och Wahlund, R., Category Management – När konsumenten är manager.
Liljenberg, A., Customer-geared competition – A socio-Austrian explanation of Tertius Gaudens -
Löf, M., On Seasonality and Cointegration.
Skoglund, J., Essays on Random Effects model and GARCH.

2000

Berg-Suurwee, U., Styrning före och efter stadsdelsnämndsreform inom kultur och fritid – Resultat från intervjuer och enkät.
Bergkvist, L., Advertising Effectiveness Measurement: Intermediate Constructs and Measures.
Brodin, B., Lundkvist, L., Sjöstrand, S-E., Östman, L., Koncernchefen och ägarna.
Bornefalk, A., Essays on Social Conflict and Reform.
Charpentier, C., Samuelson, L.A., Effekter av en sjukvårdsreform.
Edman, J., Information Use and Decision Making in Groups.
Emling, E., Svenskt familjeföretagande.
Ericson, M., Strategi, kalkyl, känsla.
Gunnarsson, J., Wahlund, R., Flink, H., Finansiella strategier i förändring: segment och beteenden bland svenska hushåll.
Hellman, N., Investor Behaviour – An Empirical Study of How Large Swedish Institutional Investors Make Equity Investment Decisions.
Hyll, M., Essays on the Term Structure of Interest Rates.
Håkansson, P., Beyond Private Label – The Strategic View on Distributor Own Brands.
I huvudet på kunden. Söderlund, M., (red).
Karlsson Stider, A., Familjen och firman.
Ljunggren, U., Styrning av grundskolan i Stockholms stad före och efter stadsdelsnämndsreformen – Resultat från intervjuer och enkät.
Ludvigsen, J., The International Networking between European Logistical Operators.
Nittmar, H., Produktutveckling i samarbete – Strukturförändring vid införande av nya Informationssystem.
Robertsson, G., International Portfolio Choice and Trading Behavior.
Schwarz, B., Weinberg, S., Serviceproduktion och kostnader – att söka orsaker till kommunala skillnader.
Stenström, E., Konstiga företag.

Styrning av team och processer – Teoretiska perspektiv och fallstudier. Bengtsson, L., Lind, J., Samuelson, L.A., (red).
Sweet, S., Industrial Change Towards Environmental Sustainability - The Case of Replacing Chloroflouorocarbons.
Tamm Hallström, K., Kampen för auktoritet – Standardiseringsorganisationer i arbete.

1999

Adler, N., Managing Complex Product Development.
Allgulin, M., Supervision and Monetary Incentives.
Andersson, P., Experto Credite: Three Papers on Experienced Decision Makers.
Ekman, G., Från text till batong – Om poliser, busar och svennar.
Eliasson, A-C., Smooth Transitions in Macroeconomic Relationships.
Flink, H., Gunnarsson, J., Wahlund, R., Svenska hushållens sparande och skuldsättning– ett konsumentbeteende-perspektiv.
Gunnarsson, J., Portfolio-Based Segmentation and Consumer Behavior: Empirical Evidence and Methodological Issues.
Hamrefors, S., Spontaneous Environmental Scanning.
Helgesson, C-F., Making a Natural Monopoly: The Configuration of a Techno-Economic Order in Swedish Telecommunications.
Japanese Production Management in Sunrise or Sunset. Karlsson, C., (red).
Jönsson, B., Jönsson, L., Kobelt, G., Modelling Disease Progression And the Effect of Treatment in Secondary Progressive MS. Research Report.
Lindé, J., Essays on the Effects of Fiscal and Monetary Policy.
Ljunggren, U., Indikatorer i grundskolan i Stockholms stad före stadsdels-nämndsreformen – en kartläggning.
Ljunggren, U., En utvärdering av metoder för att mäta produktivitet och effektivitet i skolan – Med tillämpning i Stockholms stads grundskolor.
Lundbergh, S., Modelling Economic High-Frequency Time Series.
Mägi, A., Store Loyalty? An Empirical Study of Grocery Shopping.
Mölleryd, B.G., Entrepreneurship in Technological Systems – the Development of Mobile Telephony in Sweden.
Nilsson, K., Ledtider för ledningsinformation.
Osynlig Företagsledning. Sjöstrand, S-E., Sandberg, J., Tyrstrup, M., (red).
Rognes, J., Telecommuting – Organisational Impact of Home Based – Telecommuting.
Sandström, M., Evaluating the Benefits and Effectiveness of Public Policy.
Skalin, J., Modelling Macroeconomic Time Series with Smooth Transition Autoregressions.
Spagnolo, G., Essays on Managerial Incentives and Product-Market Competition.
Strauss, T., Governance and Structural Adjustment Programs: Effects on Investment, Growth and Income Distribution.
Svedberg Nilsson, K., Effektiva företag? En studie av hur privatiserade organisationer konstrueras.
Söderström, U., Monetary Policy under Uncertainty.
Werr, A., The Language of Change The Roles of Methods in the Work of Management Consultants.
Wijkström, F., Svenskt organisationsliv – Framväxten av en ideell sektor.

1998

Andersson, M., On Testing and Forecasting in Fractionally Integrated Time Series Models.

Berg-Suurwee, U., Styrning av kultur- och fritidsförvaltning innan stadsdelsnämnds-reformen.
Berg-Suurwee, U., Nyckeltal avseende kultur- och fritidsförvaltning innan stadsdels-nämndsreformen.
Bergström, F., Essays on the Political Economy of Industrial Policy.
Bild, M., Valuation of Takeovers.
Charpentier, C., Samuelson, L.A., Effekter av en sjukvårdsreform – en analys av Stockholmsmodellen.
Eriksson-Skoog, G., The Soft Budget Constraint: The Emergence, Persistence and Logic of an Institution. The Case of Tanzania 1967-1992.
Gredenhoff, M., Bootstrap Inference in Time Series Econometrics.
Ioannidis, D., I nationens tjänst? Strategisk handling i politisk miljö – en nationell teleoperatörs interorganisatoriska, strategiska utveckling.
Johansson, S., Savings Investment, and Economic Reforms in Developing Countries.
Levin, J., Essays in Company Valuation.
Ljunggren, U., Styrning av grundskolan i Stockholms stad innan stadsdelsnämnds-reformen.
Mattsson, S., Från stat till marknad – effekter på nätverksrelationer vid en bolagise-ringsreform.
Nyberg, A., Innovation in Distribution Channels – An Evolutionary Approach.
Olsson, P., Studies in Company Valuation.
Reneby, J., Pricing Corporate Securities.
Roszbach, K., Essays on Banking Credit and Interest Rates.
Runsten, M., The Association Between Accounting Information and Stock Prices. Model development and empirical tests based on Swedish Data.
Segendorff, B., Essays on Bargaining and Delegation.
Sjöberg, L., Bagozzi, R., Ingvar, D.H., Will and Economic Behavior.
Sjögren, A., Perspectives on Human Capital: Economic Growth, Occupational Choice and Intergenerational Mobility.
Studier i kostnadsintäktsanalys. Jennergren, P., (red)
Söderholm, J., Målstyrning av decentraliserade organisationer. Styrning mot finansiella och icke-finansiella mål.
Thorburn, K., Cash Auction Bankruptcy and Corporate Restructuring
Wijkström, F., Different Faces of Civil Society.
Zethraeus, N., Essays on Economic Evaluation in Health Care. Evaluation of Hormone Replacement Therapy and Uncertainty in Economic Evaluations.

1997

Alexius, A., Essays on Exchange Rates, Prices and Interest Rates.
Andersson, B., Essays on the Swedish Electricity Market.
Berggren, N., Essays in Constitutional Economics.
Ericsson, J., Credit Risk in Corporate Securities and Derivatives. Valuation and Optimal Capital Structure Choice.
Charpentier, C., Budgeteringens roller, aktörer och effekter. En studie av budget-processerna i en offentlig organisation.
De Geer, H., Silfverberg, G., Citizens'Trust and Authorities' Choices, A Report from The Fourth International Conference on Ethics in the Public Service, Stockholm June 15-18, 1994.
Friberg, R., Prices, Profits and Exchange Rates.
Från optionsprissättning till konkurslagstiftning. Bergström, C., Björk, T., (red)
Hagerud, G.E., A New Non-Linear GARCH Model.

Haksar, A., Environmental Effects of Economywide Policies: Case Studies of Costa Rica and Sri Lanka.
He, C., Statistical Properties of Garch Processes.
Holmgren, M., Datorbaserat kontrollrum inom processindustrin; erfarenheter i ett tidsperspektiv.
Jennergren, P., Tutorial on the McKinsey Model for Valuation of Companies.
Lagerlöf, J., Essays on Political Economy, Information, and Welfare.
Lange, F., Wahlund, R., Planerade och oplanerade köp - Konsumenternas planering och köp av dagligvaror.
Löthgren, M., Essays on Efficiency and Productivity; Contributions on Bootstrap, DEA and Stochastic Frontier Models.
Nilsson, H.E., Using Conceptual Data Modelling in Management Accounting: A Case Study.
Sjöberg, L., Ramsberg, J., En analys av en samhällsekonomisk bedömning av ändrade säkerhetsföreskrifter rörande heta arbeten.
Säfvenblad, P., Price Formation in Multi-Asset Securities Markets.
Sällström, S., On the Dynamics of Price Quality.
Södergren, B., På väg mot en horisontell organisation? Erfarenheter från näringslivet av decentralisering och därefter.
Tambour, M., Essays on Performance Measurement in Health Care.
Thorén, B., Berg-Surwee, U., Områdesarbete i Östra Hökarängen - ett försök att studera effekter av decentralisering.
Zhang Gang., Chinese Rural Enterprises Between Plan and Market.
Åhlström, P., Sequences in the Profess of Adopting Lean Production.
Åkesson, G., Företagsledning i strategiskt vakuum. Om aktörer och förändrings-processer.
Åsbrink, S., Nonlinearities and Regime Shifts in Financial Time Series.

1996

Advancing your Business. People and Information Systems in Concert.
Lundeberg, M., Sundgren, B (red).
Att föra verksamheten framåt. Människor och informationssystem i samverkan.
red. Lundeberg, M., Sundgren, B.
Andersson, P., Concurrence, Transition and Evolution - Perspectives of Industrial Marketing Change Processes.
Andersson, P., The Emergence and Change of Pharmacia Biotech 1959-1995. The Power of the Slow Flow and the Drama of Great Events.
Asplund, M., Essays in Industrial Economics.
Delmar, F., Entrepreneurial Behavior & Business Performance.
Edlund, L., The Marriage Market: How Do You Compare?
Gunnarsson, J., Three Studies of Financial Behavior. Research Report.
Hedborg, A., Studies of Framing, Judgment and Choice.
Holgersson, C., Höök, P., Ledarutveckling för kvinnor - Uppföljning av en satsning på Volvo - Research Report.
Holmberg, C., Stores and Consumers - Two Perspectives on Food Purchasing.
Håkansson, P., Wahlund, R., Varumärken. Från teori till praktik.
Karlsson, A., The Family Business as an Heirloom. Research Report.
Linghag, S., Man är handelsstudent. Research Report.
Molin, J., Essays on Corporate Finance and Governance.
Mägi, A., The French Food Retailing Industry - A Descriptive Study.
Mölleryd, B., Så byggdes en världsindustri - Entreprenörskapets betydelse för svensk mobiltelefoni. Research Report.

Nielsen, S., Omkostningskalkulation for avancerede produktions-omgivelser - en sammenligning af stokastiske og deterministiske omkost-ningskalkulationsmodeller.
Normark, P., Danielsson, L., Larsson, A., Lundblad, P., Kooperativa nyckeltal. Research Report.
Sandin, R., Heterogeneity in Oligopoly: Theories and Tests.
Sandén-Håkansson, U., Från kampanjmål till mediemix - en studie av samarbete mellan annonsörer, reklambyråer och mediebyråer. Research Report.
Stein, J., Söderlund, M., Framgång i arbetet, strategier för att utföra arbetsuppgifterna, arbetsuppgifternas karaktär och utbildningskvalitet. En empirisk studie av civil-ekonomer. Research Report.
Strömberg, P., Thorburn, K., An Empirical Investigation of Swedish Corporations in Liquidation Bankruptcy. Research Report.
Söderlund, M., Och ge oss den nöjda kunden. En studie av kundtillfredsställelse och dess orsaker och effekter. Research Report.
Thodenius, B., Användningen av ledningsinformationssystem i Sverige: Lägesbild 1995. Research Report.
Ulfsdotter, U., Internationalisering för expansion eller hemmamarknadsförsvar? De nordiska marknaderna för fruktyoghurt 1982-1994.
Westelius, A., A Study of Patterns of Communication in Management Accounting and Control Projects.
Wijkström, F., Den svenska ideella sektorn och pengarna. Research Report.
Örtendahl, M., Health and Time - A Problem of Discounting.

1995

Becker, T., Essays on Stochastic Fiscal Policy, Public Debt and Private Consumption.
Blomberg, J., Ordning och kaos i projektsamarbete - en socialfenomenologisk upplösning av en organisationsteoretisk paradox.
Brodin, B., Lundkvist, L., Sjöstrand, S-E., Östman, L., Styrelsearbete i koncerner
Brännström, T., Bias Approximation and Reduction in Vector Autoregressive Models. Research Report.
Ekonomisk politik i omvandling. Jonung, L (red).
Gunnarsson, J., Wahlund, R., Hushållens finansiella strategier. En explorativ studie.
Höök, P., Chefsutveckling ur könsperspektiv - Mentorskap och nätverk på Vattenfall - Research Report.
Levin, J., Olsson, P., Looking Beyond the Horizon and Other Issues in Company Valuation. Research Report.
Mägi, A., Customer Statisfaction in a Store Performance Framework. Research Report.
Nittmar, H., Produktutveckling i samarbete.
Persson, P-G., Modeling the Impact of Sales Promotion on Store Profits.
Roman, L., Institutions in Transition. A Study of Vietnamese Banking.
Sandberg, J., Statistisk metod - dess vetenskapliga hemvist, grundläggande principer och möjligheter inom samhällsvetenskapen. Research Report.
Sandberg, J., How Do We Justify Knowledge Produced by Interpretative Approaches? Research Report.
Schuster, W., Redovisning av konvertibla skuldebrev och konvertibla vinstandelsbevis - klassificering och värdering.
Söderberg, K., Farmartjänsten - Ny kooperation inom lantbruket. Research Report.
Söderqvist, T., Benefit Estimation in the Case of Nonmarket Goods. Four Essays on Reductions of Health Risks Due to Residential Radon Radiation.
Tamm Hallström, K., Kampen för auktoritet – standardiseringsorganisationer i arbete.
Thorén, B., Användning av information vid ekonomisk styrning - månadsrapporter och andra informationskällor.