On Non-Ergodic Asset Prices*

Ulrich Horst

Department of Mathematics
University of British Columbia
1984 Mathematics Road
Vancouver BC V6T 1Z2
horst@math.ubc.ca

Jan Wenzelburger

Fakultät für Wirtschaftswissenschaften
Universität Bielefeld
Postfach 100 131
D-33501 Bielefeld, Germany
jwenzelb@wiwi.uni-bielefeld.de

Abstract

We investigate the asset prices dynamics and the long-run market shares of two competing financial mediators who are selected by consumers. We demonstrate that the social interaction among consumers constitutes an endogenous path-depending source of risk in a financial market. Depending on consumers' evaluation of the mediator's investment, asset prices may behave in a non-ergodic manner: the price process converges in distribution but the limiting distribution is not necessarily uniquely determined, its multiplicity being characterized by the multiplicity of possible long-run market shares. The convergence of the process is sensitive to initial conditions and depends on the history of noise trader transactions. Long-run portfolio holdings may be in-efficient since investors holding mean-variance efficient portfolios may not be identified.

Keywords: CAPM, financial markets, social interaction, random difference equations
JEL Classification: C62, D85, G12

First version: Oct. 2005, this version: July 2006.

*We would like to thank Volker Böhm, Hans Föllmer, Roger Guesnerie, Alan Kirman, Thorsten Pampel, participants at the workshops on Interactions and Markets (Pisa), Stochastic Modelling in Mathematical Finance (Montreal), and Aggregation and Disaggregation (Banff), and seminar participants at various institutions for valuable comments and suggestions. We are indebted to Andreas Starke for computational assistance. Financial support from the Deutsche Forschungsgemeinschaft and NSERC is gratefully acknowledged. An earlier version was entitled "Non-ergodic behavior in financial markets with interacting investors".

1 Introduction

In recent years financial market models with interacting agents have increasingly attracted attention in focusing on trader heterogeneity as a main pillar of a descriptive theory of financial markets. A central issue for these markets is to model the investment strategies of agents. It is a common approach to attach some kind of technical trading strategy to an agent and to allow for switching between different strategies as time passes by. At the same time, an old conjecture which dates back to Alchian (1950) and Friedman (1953) states that agents that do not learn to make accurate predictions about the future will be driven out of the market. This conjecture is a main pillar of the rational-expectations paradigm, but has been repeatedly questioned in the recent literature.

In the context of a dynamic CAPM with interacting investors, Alchian and Friedman's conjecture suggests that only those investors who hold mean-variance efficient portfolios will survive in the long run. In a series of papers, e.g., De Long, Shleifer, Summers & Waldmann (1990, 1991) the capability of noise traders to survive in financial markets has been analyzed. However, these results only cover the static case with noisy errors. In the case of a dynamic CAPM with fully heterogeneous expectations, Wenzelburger (2004) introduces a *reference portfolio* which is (mean-variance) efficient in the classical sense of CAPM theory regardless of the diversity of beliefs. The reference portfolio generalizes the market portfolio as it coincides with the market portfolio if beliefs are homogeneous. Under Alchian and Friedman's conjecture, the *returns* realized with an efficient portfolio should empirically outperform those of any non-efficient portfolio. A simulation study in Böhm & Wenzelburger (2005) indicates that this is not always the case. Their example suggests that in the long run market shares of investors holding in-efficient portfolios may be larger than those of investors holding efficient portfolios.

This paper now provides a rigorous analysis of the long-run behavior of market shares and asset prices in a dynamic CAPM in which the demand for multiple risky assets comes from a large set of consumers. Rather than investing directly in the financial market, consumers select between two professional financial mediators who are characterized by their ability to forecast future asset prices. Following up on the seminal approach by Frankel & Froot (1986), we summarize their beliefs as 'chartist' and 'expert-trader' views. Chartists base their trading strategies upon observed historical price patterns such as trends. As in Wenzelburger (2004), we assume that an expert trader is able to correctly predict the first two moments of the price process. She has rational expectations along any price path and hence holds efficient portfolios. Consumers evaluate the mediators' performance before choosing the mediator to carry out their investment decision.

We analyze the question to what extent boundedly rational consumers are able to identify the mediator holding efficient portfolios by means of simple empirical performance measures which is either the empirical return or the Sharpe ratio associated with her trading strategy. We prove that the financial market dynamics is *ergodic* if the dependence of consumer investment decisions on the mediators' performances is sufficiently weak, so that market shares settle down to a unique equilibrium. Ergodicity breaks down if interactive complementarities become too powerful. In this case, "history matters" and the long-run market shares of competing mediators are path dependent. While market shares and asset prices still converge, their asymptotics is *random* and depends on noise trader transactions. We show that convergence of market shares implies that the price process converges in law to some random equilibrium distribution. This extends a recent result by Föllmer, Horst & Kirman (2005) beyond ergodicity. They give sufficient conditions for the convergence in distribution of asset prices in a model with interacting agents to a unique limit. In our model the long-run distribution of asset prices is path dependent if the interaction between agents is strong enough.

This result is consistent with many findings in the social interaction literature (Blume 1993, Horst & Scheinkman 2005) which establish uniqueness of equilibria if interactions are weak and which show that powerful interactive complementarities often generate non-ergodic dynamics. Our approach may also be viewed as a first step to bridge the gap between the deterministic approach initiated by Brock & Hommes (1997, 1998) with its rich dynamics but inherent inaccessibility to analytical solutions, and the probabilistic approach of Föllmer & Schweizer (1993) and Föllmer, Horst & Kirman (2005). They obtain rigorous mathematical results but their assumptions rule out many interesting phenomena including a non-ergodic dynamics. The dynamics of our model can be described by a deterministic recursion in a random environment. The environment is generated by an *exogenous* stochastic process that models the effects of noise trading and by an *endogenous* process that describes the fluctuations of mediators' market shares. The mediators' market shares depend on their relative performance and hence on asset prices. This generates a feedback from prices into the random environment. It is this feedback effect that distinguishes our model from the work of Estigneev, Hens & Schenk-Hoppé (2005) and the seminal papers of Blume & Easley (1992, 2005) and Sandroni (2000) in which investors never change their portfolio strategies. Föllmer and Schweizer (1993) and Horst (2005) allow agents to switch randomly between different portfolio rules but do not allow for a dependence of the transition dynamics on asset prices. Under the assumption that the impact of trend chaser is not too strong, we prove convergence of the model dynamics to a random limit using a stochastic approximation algorithm and a uniform moderate deviations principle for a class of linear recursions in random media. The approximation result allows for an analysis of the asymptotics of the market shares by means of the long-run behavior of a deterministic dynamical system.

The paper is organized as follows. The model is introduced in Section 2. The convergence results are stated in Section 3 and illustrated by numerical simulations in Section 4. All proofs are carried out in the appendix.

2 The Model

We investigate a dynamic financial market model in which the demand for multiple risky assets comes from many boundedly rational consumers. Instead of making direct investments in the financial markets, consumers invest through financial mediators. Mediators are characterized by mean variance preferences and heterogenous beliefs for future asset prices. Based on these beliefs they form their demand functions; the actual asset price is determined by market clearing conditions. Asset prices are influenced by an exogenous stochastic process describing noise trader activities and by an *endogenous* process that specifies the evolution of the mediators' market shares.

2.1 Interacting investors

We consider a large number of consumers who transfer wealth into the future by investing a fixed amount $e > 0$ of their exogenously given endowment into K risky assets and a bond in each period. Consumers are boundedly rational in the sense of Simon (1982). They delegate their portfolio transactions to one of two mediators $i = 1, 2$ who carry out portfolio transactions on their behalf. In terms of her respective market share $\eta_t^{(i)} \in [0, 1]$, mediator i receives $W_t^{(i)} = \eta_t^{(i)} e$ units of per capita resources from investing consumers in period t. The bond pays the risk-less rate $r > 0$. Aggregate per-capita repayment obligations from investing in a portfolio $x_{t-1}^{(i)} \in \mathbb{R}^K$ of risky assets and $y_{t-1}^{(i)} \in \mathbb{R}$ risk-less bonds to consumers in period t amount to $(1 - \delta^{(i)})\left[p_t^\top x_{t-1}^{(i)} + (1 + r)y_{t-1}^{(i)}\right]$, where $0 \leq \delta^{(i)} \leq 1$ stipulates the income share of mediator i. Here, $p_t \in \mathbb{R}_+^K$ denotes the vector of current asset prices. Thus, given a vector p of proposed asset prices, mediator i's budget constraints read

$$W_t^{(i)} = p^\top x^{(i)} + y^{(i)}.$$

Following a standard approach in the behavioral finance literature, we assume that the mediators are myopic mean-variance optimizers, e.g., see Böhm & Chiarella (2005). Their demand for the risky assets is thus solely based on their coefficient of risk aversion $\alpha^{(i)}$ along with their subjective assessments $(q_t^{(i)}, V_t^{(i)})$ of the mean value and the covariance matrix of the asset prices in the subsequent period $t + 1$. Abstracting from short-sell constraints, the per-capita aggregate demand function for risky assets of all consumers who employ i is

$$x^{(i)}(p) = \frac{\eta_t^{(i)}}{\alpha^{(i)}} V_t^{(i)-1}[q_t^{(i)} - (1 + r)p].$$

Let $\bar{x} \in \mathbb{R}_+^K$ denote the number of tradeable risky assets in per capita terms and ε_t the (per-capita) portfolio holdings of noise traders after trading in period t. Given the mediators' *beliefs* $(q_t^{(i)}, V_t^{(i)})$, the market-clearing condition in period t takes the form

$$\frac{\eta_t^{(1)}}{\alpha^{(1)}} V_t^{(1)-1}[q_t^{(1)} - (1 + r)p_t] + \frac{\eta_t^{(2)}}{\alpha^{(2)}} V_t^{(2)-1}[q_t^{(i)} - (1 + r)p_t] + \varepsilon_t \stackrel{!}{=} \bar{x}.$$

Solving for the market-clearing price p_t, we obtain a temporary equilibrium map

$$p_t := \Gamma_t^{(1)}\, q_t^{(1)} \;+\; \Gamma_t^{(2)}\, q_t^{(2)} \;-\Gamma_t(\bar{x} - \varepsilon_t), \tag{1}$$

where, for $i = 1, 2$, we put

$$\Gamma_t^{(i)} := \tfrac{\eta_t^{(i)}}{\alpha^{(i)}} \Gamma_t V_t^{(i)-1} \quad \text{with} \quad \Gamma_t := \tfrac{1}{(1+r)} \left(\tfrac{\eta_t^{(1)}}{\alpha^{(1)}} V_t^{(1)-1} + \tfrac{\eta_t^{(2)}}{\alpha^{(2)}} V_t^{(2)-1} \right)^{-1}.$$

If all covariance matrices $V_t^{(i)}$ are positive definite, then Γ_t is well defined, symmetric, and positive definite. As a result, we obtain a sequence of price equilibria driven by the evolution of the mediators' beliefs, their market shares, and noise trader transactions.

2.2 The feedback of subjective beliefs on asset prices

The mediators are assumed to be boundedly rational in sense of Sargent and use forecasting rules for first and second moments to update their subjective beliefs. Mediator 1 is assumed to be a trend chasing *chartist*. She bases her forecasts for the future asset prices on past observations and applies a simple *technical trading rule* of the form

$$q_t^{(1)} := \sum_{l=1}^{J} D^{(l)} \, p_{t-l} \tag{2}$$

where $D^{(1)}, \ldots, D^{(L)}$ denote her expected impact of the past prices $p_{t-1}, \ldots, p_{t-L}$ on p_{t+1}. For simplicity, we assume that mediator 1 never updates second moment beliefs and uses constant subjective variance-covariance matrices, denoted by $V^{(1)}$. While the beliefs of mediator 1 may well be incorrect, suppose now that mediator 2 is able to correctly predict the first two moments of the price process, conditional on all available information. As a short hand, we will use the term *rational expectations* to describe the situation in which the first two moments of mediator 2's subjective distributions of asset prices, i.e., the conditional mean values and the conditional covariance matrices coincide with the respective moments of the true distributions. Assuming that the first moment beliefs $q_t^{(2)}$ of mediator 2 are unbiased, it is shown in Wenzelburger (2004) that they are determined by an *unbiased forecasting rule* which takes the form

$$q_t^{(2)} := \Gamma_t^{(2)-1} \Big[q_{t-1}^{(2)} - \Gamma_t^{(1)} q_t^{(1)} + \Gamma_t \big(\bar{x} - \mathbb{E}_{t-1}[\varepsilon_t] \big) \Big], \tag{3}$$

where $\mathbb{E}_{t-1}$ denotes the conditional expectation with respect to all the information available in period $t-1$. The forecasting rule (3) provides unbiased forecasts of asset prices for mediator 2 in the sense that $q_{t-1}^{(2)}$ is the best least-squares prediction for p_t, given the available information. Indeed, it is straightforward to verify that almost surely $\mathbb{E}_{t-1}[p_t - q_{t-1}^{(2)}] = 0$ for all times t when the forecast $q_t^{(2)}$ is given by (3).

To focus on the effects of heterogeneity in the mediators' beliefs about expected future asset prices and on the interplay between rational expectations and trend chasing, we assume from now on that the mediators' beliefs for second moments coincide and are correct and that they share a common coefficient of risk aversion, i.e., $\alpha^{(1)} = \alpha^{(2)} = \alpha$; all results carry over to the case with differing risk aversions. If the covariance matrix $\mathbb{V}_\varepsilon$ of the noise trader transactions is assumed to be constant over time, the correct second moments beliefs take the form

$$V_t^{(i)} \equiv \left(\tfrac{1+r}{\alpha}\right)^2 \mathbb{V}_\varepsilon^{-1}. \tag{4}$$

We put $\eta_t = \eta_t^{(1)}$ and $\eta_t^{(2)} = 1 - \eta_t$ for the remainder of the paper and insert the forecasts (2)-(4) into (1). The resulting process of asset prices and forecasts is given by a list of stochastic difference equations

$$\begin{cases} p_t &= q_{t-1}^{(2)} + \frac{1+r}{\alpha}\mathbb{V}_\varepsilon^{-1}\big(\varepsilon_t - \mathbb{E}_{t-1}[\varepsilon_t]\big), \\ q_t^{(1)} &= \sum_{j=1}^{J} D^{(j)} p_{t-j}, \\ q_t^{(2)} &= \frac{1+r}{1-\eta_t}\, q_{t-1}^{(2)} - \frac{\eta_t}{1-\eta_t}\, q_t^{(1)} + \frac{(1+r)^2}{(1-\eta_t)\alpha}\, \mathbb{V}_\varepsilon^{-1}\left(\bar{x} - \mathbb{E}_{t-1}[\varepsilon_t]\big)\right., \end{cases} \tag{5}$$

where the last equation corresponds to the unbiased forecasting rule (3). Observe that the difference equations (5) are linear if the market shares η_t were constant over time. However, the consumers' decisions determine the market share of a mediator as they will evaluate a mediator's performance before making their investment decision. This generates a feedback effect from the sequence of asset prices and forecasts into the evolution of market shares. The underlying performance measures will be introduced in the following section.

Remark 2.1 *This model can easily be imbedded into an OLG framework with multi-period planning horizons without changing the key equations, cf. Hillebrand & Wenzelburger (2006)*

2.3 Performance measures and choice rules

In the spirit of De Long, Shleifer, Summers & Waldmann (1990) we assume throughout that the noise trader portfolios $\{\varepsilon_t\}_{t\in\mathbb{N}}$ are an *exogenous* i.i.d. process with mean $\bar{\varepsilon}$ and a non-degenerate variance matrix $\mathbb{V}_\varepsilon$. We can then represent the joint dynamics of asset prices and forecasts in terms of a linear difference equation in a random environment which is generated by two underlying stochastic processes $\{\varepsilon_t\}_{t\in\mathbb{N}}$ and

$\{\eta_t\}_{t\in\mathbb{N}}$, respectively. To this end, we put $X_t := (q_t^{(2)}, q_{t-1}^{(2)}, p_t, \ldots, p_{t-J-1}) \in \mathbb{R}^d$ with $d = K(J+4)$ as well as

$$\begin{aligned} a_0(\eta) &:= \tfrac{1+r}{1-\eta}, & a_j(\eta) &:= \tfrac{\eta}{1-\eta} D^{(j)},\ j = 1,\ldots,J, \\ b_0(\eta) &:= \tfrac{(1+r)^2}{(1-\eta)\alpha} \mathbb{V}_\varepsilon^{-1} \left(\bar{x} - \bar{\varepsilon}\right), & b_1(\varepsilon) &:= \tfrac{1+r}{\alpha} \mathbb{V}_\varepsilon^{-1} (\varepsilon - \bar{\varepsilon}). \end{aligned}$$

The process $X = \{X_t\}_{t\in\mathbb{N}}$ as defined by (5) then takes the linear form

$$X_t = A(\eta_t) X_{t-1} + B(\eta_t, \varepsilon_t) \qquad (t \in \mathbb{N}) \tag{6}$$

where the $d \times d$ matrix $A(\eta_t)$ and the vector $B(\eta_t, \varepsilon_t) \in \mathbb{R}^d$ are given by

$$A(\eta_t) := \begin{pmatrix} a_0(\eta_t) & 0 & a_1(\eta_t) & \cdots & a_J(\eta_t) & 0 & 0 \\ I & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ I & 0 & \ddots & & & & \vdots \\ 0 & 0 & I & \ddots & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 & I & 0 \end{pmatrix} \text{ and } B(\eta_t, \varepsilon_t) := \begin{pmatrix} b_0(\eta_t) \\ 0 \\ b_1(\varepsilon_t) \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix},$$

respectively. We will now specify a probabilistic framework for the analysis of the long-run market behavior when consumers' decisions are based on the perceived performance of mediators. To this end, let $\{\varrho_t\}_{t\in\mathbb{N}}$ be the sequence of empirical distributions associated to the process $\{X_t\}_{t\in\mathbb{N}}$, i.e.,

$$\varrho_t := \frac{1}{t} \sum_{i=0}^{t-1} \delta_{X_i} \quad \text{so that} \quad \varrho_t(f) := \int f \mathrm{d}\varrho_t = \frac{1}{t} \sum_{i=0}^{t-1} f(X_i) \tag{7}$$

for any bounded map $f : \mathbb{R}^d \to \mathbb{R}$ where δ_x denotes the Dirac measure that puts all mass on x. A consumer's propensity to invest through a specific mediator will depend on the performance associated to the mediator's investment strategy.

Definition 2.2 *Let $f^l : \mathbb{R}^d \to \mathbb{R}$ $(l = 1, 2, \ldots, L)$ be a given list of bounded measurable functions. A performance measure is a Lipschitz continuous function $\psi : \mathbb{R}^L \to \mathbb{R}^2$. The list of performances in period t is given by*

$$\pi_t = \psi\left(\varrho_t(f^1), \cdots, \varrho_t(f^L)\right).$$

Notice that the performance of a mediator at time t depends on the entire history of empirical distribution of asset prices and forecasts. This dependence generates a rich dynamics, while still being amenable to analytic solutions. Since the functions f^i are fixed, it is convenient to interpret the performance measure as a map on the set of probability distributions on $\mathbb{R}^d$ so that $\pi_t \equiv \psi(\varrho_t)$.

Example 2.3 *Suppose that the performance of mediator $i = 1, 2$ is measured by historically realized returns $\{R_s^{(i)}\}_{s=0}^t$ on investment. Having invested the amount $W_t^{(i)} = \eta_t e$, her return from selling the portfolio $x_t^{(i)} := \frac{\eta_t^{(i)}}{\alpha} V_t^{(i)-1}[q_t^{(i)} - (1+r)p_t]$ in period $t+1$ is*

$$R_{t+1}^{(i)} = r + \frac{1}{\alpha e}\,[p_{t+1} - (1+r)p_t]^\top V_t^{(i)-1}[q_t^{(i)} - (1+r)p_t]. \tag{8}$$

Since second moment beliefs are constant, realized returns at time t take the form $R_t^{(i)} = f^i(X_t)$ for suitably defined functions $f^i : \mathbb{R}^{2d} \to \mathbb{R}$, $i = 1, 2$. The specific case where ψ denotes the identity matrix yields the performance measure[1]

$$\pi_t = \left(\varrho_t(f^1), \varrho_t(f^2)\right)^\top.$$

Empirical Sharpe ratios also fit into our framework.

Example 2.4 *Based on Example 2.3, an alternative performance measure is the differences in empirical* Sharpe ratios *associated with the two times series $\{R_s^{(i)}\}_{s=0}^t$, $i = 1, 2$. Define to this end two continuous functions $f^{i+2} : \mathbb{R}^{2d} \to \mathbb{R}$, $i = 1, 2$ by*

$$\left(R_t^{(i)}\right)^2 = f^{i+2}(X_t), \quad i = 1, 2.$$

The empirical Sharpe ratios associated with the each of the mediators $i = 1, 2$ is

$$\frac{\varrho_t(f^{i+2}) - r}{\sqrt{\varrho_t(f^{i+2}) - \left[\varrho_t(f^i)\right]^2}}, \quad i = 1, 2,$$

respectively. With f^1 and f^2 as in Example 2.3 a performance measure based on mediators' empirical Sharpe ratios is given by

$$\pi_t = \left(\frac{\varrho_t(f^1) - r}{\sqrt{\varrho_t(f^3) - \left[\varrho_t(f^1)\right]^2}}, \frac{\varrho_t(f^2) - r}{\sqrt{\varrho_t(f^4) - \left[\varrho_t(f^2)\right]^2}} \right)^\top. \tag{9}$$

[1] There is no a-priori reason to assume that returns can be represented by *bounded* functions which we assume for mathematical convenience. This can be justified when consumers do not trust unusually high earnings to prevail. Similar considerations apply to Example 2.4.

Following a standard assumption in the social interaction literature, we assume that a consumer employs a mediator at random. The probabilities depend on the current performances π_t. Specifically, let consumers act conditionally independent of each other given π_t so that an individual consumer employs mediator 1 with probability

$$\Phi(\pi, \beta) := \left[\overline{\eta} - \underline{\eta}\right] \frac{\exp(\beta\pi_1)}{\exp(\beta\pi_1) + \exp(\beta\pi_2)} + \underline{\eta} = \frac{\overline{\eta} - \underline{\eta}}{\exp(\beta(\pi_2 - \pi_1)) + 1} + \underline{\eta}. \tag{10}$$

Here $\beta > 0$ specifies the dependence of the agents' choices on the mediators' performances, and $0 \le \underline{\eta} \le \overline{\eta} \le 1$ are upper and lower bounds for the realized market share η. In the limit of an infinite number of consumers the law of large numbers implies that the chartist's market share at time t is deterministic. Hence there exists a uniformly continuous "choice function" $F : \mathbb{R}^L \to \mathbb{R}$ such that

$$\eta_t = F\big(\varrho_t(f^1), \ldots, \varrho_t(f^L)\big) := \Phi\big(\psi\big(\varrho_t(f^1), \ldots, \varrho_t(f^L)\big), \beta\big). \tag{11}$$

It follows from a generalized security market line result (Wenzelburger 2004, Thm. 3.1) that the portfolios of investors with rational expectations are mean-variance efficient in the ex-ante sense of classical CAPM theory. Transposed into the present setting, this theorem implies that the conditional *Sharpe ratio* of mediator 2 will always be greater than the conditional Sharpe ratio of mediator 1. From this result one might expect that at least in the scenario of Example 2.4 the mean-variance efficient portfolios of mediator 2 will always empirically outperform the inefficient portfolios of mediator 1. A first simulation analysis in Böhm & Wenzelburger (2005), however, suggests that mediators holding efficient portfolios may not be identified.

We will provide a rigorous analysis of this phenomenon. It will turn out that the dynamics of asset prices and beliefs can be described by *path-dependent stochastic difference equation*. Its dynamics will be analyzed in the following section.

3 Convergence to equilibrium

In this section we state conditions on the behavior of consumers and mediators which guarantee that asset prices converge in distribution to some limit which is possibly random. To this end, we shall first give sufficient conditions for convergence of the market shares. Numerical simulations suggest that the distribution of long-run market shares depends on the initial condition as well as on the strength of interactions. If consumers choose the mediators more or less independently of their respective performance, the market shares settle down to a deterministic limit. If the interaction is strong enough, the limit is random. This may be viewed as an *endogenous* source of randomness originating from interaction and imitation effects. While the chartists' forecasts may be rather inaccurate, the feedback effects from consumer behavior into the dynamics of asset prices and forecasts may prevent chartists from being driven out of the market. In fact, if the interaction effects are strong, chartists and 'rational' mediators typically coexist. It is this coexistence that distinguishes our model from, for instance, Sandroni's where 'markets favor agents that make accurate predictions'.

3.1 Benchmark models driven by independent noise

Chartists may have a destabilizing affect on asset prices. Without any bound on their impact there is no reason to believe that prices and forecasts are stable in the long run. In order to guarantee long-run stability, we assume that consumers employ mediators 1, the chartist, at least with some probability $\underline{\eta}$ and at most with probability $\overline{\eta}$. This means that the choice function F in (11) is such that

$$\eta_t \in [\underline{\eta}, \overline{\eta}]. \tag{12}$$

In order specify the interval $[\underline{\eta}, \overline{\eta}]$ we introduce, for any market share η the process $X^\eta = \{X_t^\eta\}$ defined by the linear recursive relation

$$X_t^\eta = A(\eta)X_{t-1}^\eta + B(\eta, \varepsilon_t) \qquad (t \in \mathbb{N}). \tag{13}$$

The process X^η describes the evolution of asset prices and forecasts in a benchmark model with market shares "frozen" at the level η. The long-run behavior of such sequences has been extensively investigated under a contraction condition on $A(\eta)$. We need a slightly stronger condition.

Assumption 3.1 *(i) The map $\eta \mapsto A(\eta)$ is Lipschitz continuous:*

$$\|A(\eta) - A(\hat{\eta})\| \leq a|\eta - \hat{\eta}|.$$

(ii) The interval $[\underline{\eta}, \overline{\eta}]$ is chosen such that the eigenvalues of all the matrices $A(\eta)$ with $\eta \in [\underline{\eta}, \overline{\eta}]$ lie uniformly within the unit circle.

(iii) The function $B(\cdot,\cdot)$ is bounded, $|B(\eta,\varepsilon)| \le B$, and uniformly Lipschitz continuous in its first argument, i.e.,

$$\sup_{\varepsilon} |B(\eta,\varepsilon) - B(\hat{\eta},\varepsilon)| \le b|\eta - \hat{\eta}|. \tag{14}$$

When Assumption 3.1 (i) and (ii) are satisfied, the difference equation (13) has unique stationary solution, i.e., there exists a unique stationary and ergodic process $x^\eta = \{x_t^\eta\}_{t\in\mathbb{N}}$ that satisfies (13). For any starting point x, the distribution μ_t^η of X_t^η converges weakly to the distribution μ^η of x_0^η. Furthermore X^η is bounded. The following proposition, whose proof is also given in Section A, shows that Assumption 3.1 also guarantees boundedness of the sequence X. It is in this sense that (12) prevents prices from exploding by limiting the impact of chartists.

Proposition 3.2 *Under Assumption 3.1 the sequence $\{X_t\}$ is almost surely bounded. Specifically, for any initial state x there is a constant M_x such that*

$$\mathbb{P}_x\left[\sup_t |X_t| \le M_x\right] = 1. \tag{15}$$

Here $\mathbb{P}_x$ denotes the probability measure on the canonical path space induced by the process X with initial state x.

It will be convenient to write $\mu^\eta(f)$ for the integral of a bounded function f with respect to the unique stationary distribution μ^η of the process X^η defined by (13).

3.2 Characterization of equilibria

As a first step towards a general convergence result, we are now going to characterize all possible long-run distributions of the sequence $\{X_t\}_{t\in\mathbb{N}}$ by a fixed point property. The key assumption is that the empirical process $\{\eta_t\}_{t\in\mathbb{N}}$ converges almost surely as $t \to \infty$. The proof requires some preparation and will be carried out in Section A.

Theorem 3.3 *Suppose that Assumption 3.1 is satisfied and that the empirical process $\{\eta_t\}_{t\in\mathbb{N}}$ converges almost surely to some random variable η_*. Then the sequence of empirical averages $\{\varrho_t\}_{t\in\mathbb{N}}$ converges almost surely weakly to the random limiting measure μ^{η_*}. More precisely, for all bounded continuous functions f,*

$$\mathbb{P}\left[\lim_{t\to\infty} \varrho_t(f) = \mu^{\eta_*}(f)\right] = 1.$$

The previous theorem states that the distribution of X_t converges weakly to a *random* limiting measure if the sequence of market shares settles down to a random limit in the long run. The result imposes a consistency condition on limiting market shares and thus allows us to characterize the class of asymptotic market shares. The limiting market shares have to be consistent with the market shares induced by the limiting empirical distributions of X through the choice function F in (11). In order to make this more precise, we define a map $\zeta : [\underline{\eta}, \overline{\eta}] \to \mathbb{R}^L$ by

$$\zeta(\eta) := \left(\mu^\eta(f^1), \ldots, \mu^\eta(f^L)\right) \in \mathbb{R}^L. \tag{16}$$

This map assigns the *long-run empirical averages* of the Markov process X^η with fixed η to the market shares η. The question of existence and uniqueness of long-run equilibria of the process X can now be reduced to a fixed point condition.

Corollary 3.4 *Under the assumptions of Theorem 3.3, the random variable η_* takes values in the set*

$$E := \left\{\eta \in [\underline{\eta}, \overline{\eta}] : \ \eta = F \circ \zeta(\eta)\right\}, \tag{17}$$

i.e., it almost surely satisfies the fixed point condition $\eta = F \circ \zeta(\eta)$. The long-run empirical averages of the process X are given by $\zeta(\eta_*)$ and take values in the set

$$S := \left\{z \in \mathbb{R}^L : \ z = \zeta \circ F(z)\right\}. \tag{18}$$

It is well known (Brandt 1986) that the map $\zeta : [\underline{\eta}, \overline{\eta}] \to \mathbb{R}^L$ is continuous. Typically, however, no analytical expression will be available. The map requires knowledge about the structure of the stationary distributions μ^η for the Markov processes X^η for which, in general, no closed form representation will be available. However, it can easily be simulated. A purely numerical analysis of the sequence of market shares, on the other hand, is not always appropriate because the speed of convergence of the sequence $\{\eta_t\}_{t\in\mathbb{N}}$ is very slow. In fact, consider

$$\eta_t = F\left(\varrho_t(f^1), \ldots, \varrho_t(f^L)\right)$$

for bounded Lipschitz continuous functions f^l and F. For all $t, T \in \mathbb{N}$ we have that

$$|\eta_{T+t} - \eta_T| \le C \frac{t}{T+t} \tag{19}$$

for some $C < \infty$. A numerical analysis may thus become extremely time consuming and could easily be misleading. Comparing the numerical analysis in Böhm & Wenzelburger (2005) with our results, we will see below that market shares after a short period of say $t = 1000$ periods may be a poor predictor for an asymptotic asset price dynamics.

3.3 Convergence of market shares

In view of our characterization result of long-run market shares, it remains to state conditions which guarantee convergence of the sequence $\{\eta_t\}_{t\in\mathbb{N}}$. It turns out that the long-run behavior of empirical averages and hence the asymptotics of market shares can be analyzed by means of a differential equation. To this end, we recall that $z_t = \big(\varrho_t(f^1), \ldots, \varrho_t(f^L)\big)$, rewrite our stochastic difference equation (6) and (11) as

$$\begin{aligned} X_t &= A\big(F(z_t)\big)X_{t-1} + B\big(F(z_t), \varepsilon_t\big), \\ z_t &= \frac{t-1}{t} z_{t-1} + \frac{1}{t}\big(f^1(X_{t-1}), \ldots, f^L(X_{t-1})\big), \end{aligned} \tag{20}$$

and define a map $g : \mathbb{R}^L \to \mathbb{R}^L$ by setting

$$g(z) := \zeta \circ F(z) - z. \tag{21}$$

The zeros of the map g are given by the set S defined in (18). Continuity of F and ζ implies continuity of g. With this we are now in position to state our main convergence result. The proof will be given in Section B below.

Theorem 3.5 *Suppose that the map g is Lipschitz continuous such that the ODE*

$$\dot{z} = g(z) \tag{22}$$

admits a unique solution. Let

$$S^* := \{s_1, \ldots, s_N\} \subset S := \{z \in \mathbb{R}^L : \; g(z) = 0\}$$

be set of asymptotically stable steady states with corresponding basins of attraction $DA(s_i)$. If the sequence $\{z_t\}_{t\in\mathbb{N}}$ of empirical averages visits a compact subset of some $DA(s_i)$ infinitely often with probability $\mathbf{p} > 0$, then the following holds:

(i) The sequence $\{z_t\}_{t\in\mathbb{N}}$ converges to s_i with at least probability $\mathbf{p}$, i.e.,

$$\lim_{t\to\infty} |z_t - s_i| = 0 \quad \textit{with at least probability} \quad \mathbf{p}.$$

(ii) The discrete-time stochastic process $\{\eta_t\}_{t\in\mathbb{N}}$ of market shares converges with at least probability $\mathbf{p}$ *to a stationary value $F(s_i) \in E$.*

The problem of convergence of market shares and hence of asset prices can thus be reduced to establishing convergence of ordinary differential equations on the level of empirical averages. In particular, market shares converge almost surely to some constant if the ODE (21) has a unique, globally asymptotically stable steady state. This will be the case if, for instance, performances are measured by empirical returns and the dependence of consumer choices on performances is sufficiently weak. If, on the contrary, the interactive effects are too strong, ergodicity breaks down and market shares converge to a random limit as we will illustrate in the following section.

4 Convergence for returns and Sharpe ratios

We apply Theorem 3.5 to investigate whether the rational mediator who holds efficient portfolios will attain larger market shares than the chartist. Since trading of assets takes place before consumers can observe the relevant returns, the *empirical performance* of a portfolio has to rely on estimates. These estimates are reflected by the performance measure. The superiority of an efficient portfolio will only show if the estimators used for the performance measure are consistent. It is intuitively clear that for inconsistent estimators portfolios other than the efficient portfolio could appear to perform better. In fact, we find that chartists and 'rational' expert traders often coexist although we bound the probability that consumers follow the chartist.

Remark 4.1 *We assume throughout that the risk-less rate is $r = 1\%$. In this case the stability condition for our model is satisfied if the market share η is ranging in $\eta \in [\underline{\eta}, \overline{\eta}]$ with $\underline{\eta} = .06$ and $\overline{\eta} = .36$. In order to prevent prices from exploding we allow for no more than 36% chartists. All empirical densities displayed below are calculated using a sample of* 100.000 *independent repetitions.*

4.1 Empirical returns as performance measures

When the mediators' performance is measured in average returns as in Example 2.3 the difference in the performances is of the form $z_t = \varrho_t(f^1 - f^2)$ for suitable bounded continuous functions f^i on $\mathbb{R}$.

4.1.1 Analytical results

A numerical approximation of the function g defining the ODE (22) is depicted in Figure 1(a). It indicates that g has at least three steady states, two asymptotically stable ones and one unstable in the middle. To see if additional steady states exist, notice that in the present case $\eta_t = F(z_t)$ is a diffeomorphism, i.e., an invertible map with differentiable inverse F^{-1}. In this case the ODE (22) is topological conjugate (Arrowsmith & Place 1994) to the ODE for market shares

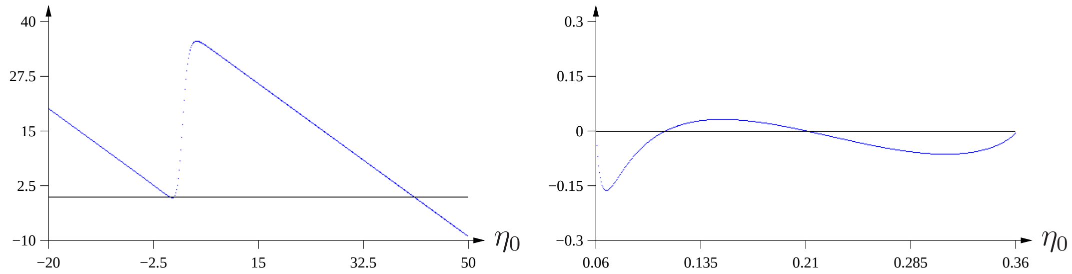
$$\dot{\eta} = h(\eta) \tag{23}$$

where $h := (F' \circ F^{-1})(g \circ F^{-1}) = (F' \circ F^{-1})\big(\zeta - F^{-1}\big)$. The conjugacy implies that taking empirical averages as performance measures the behavior of (22) and (23) is qualitatively the same. In particular, the long-run behavior of empirical averages is precisely described by the long-run behavior of market shares. We infer from

$$h(\underline{\eta}) = h(\overline{\eta}) = 0 \quad \text{because} \quad (F' \circ F^{-1})(\eta) = -\beta(\overline{\eta} - \eta)(\eta - \underline{\eta}), \tag{24}$$

the conjugacy of the two ODEs and Figure 1 that h has *five* steady states when $\beta = 2$; the three steady states corresponding to g along with $\underline{\eta}$ and $\overline{\eta}$. Four equilibria are clearly visible; the fifth lies close to $\underline{\eta}$. The solution to (22) exists for all times, and standard monotonicity arguments show that all solutions converge to one of the two asymptotically stable fixed points. The respective basins of attraction are simply separated by the unstable fixed point. Now Theorems 3.3 and 3.5 guarantee convergence of both market shares and asset prices. The respective limits, however, are not necessarily unique. They may depend on the random noise trader transactions. Furthermore, the lowest and highest possible market share of chartists are always unstable under the dynamics of the ODE. This implies that the long-run market share of the chartists will always be strictly above $\underline{\eta}$ and strictly smaller than $\overline{\eta}$.

Remark 4.2 *The ODE (23) has a unique globally asymptotically stable steady if the map $\eta \mapsto \zeta(\eta) - F^{-1}(\eta)$ has a unique zero. Thus, our financial market dynamics are ergodic if the dependence of consumers' choices on performances is sufficiently weak; for small β the inverse choice function F^{-1} is essentially a vertical line. Ergodicity* *breaks down if the dependence of consumers' choices on the mediators' performance is too strong. In this case the limiting behavior of market shares and price distributions is random.*



(a) ODE $\dot{z} = g(z)$ (b) ODE $\dot{\eta} = h(\eta)$

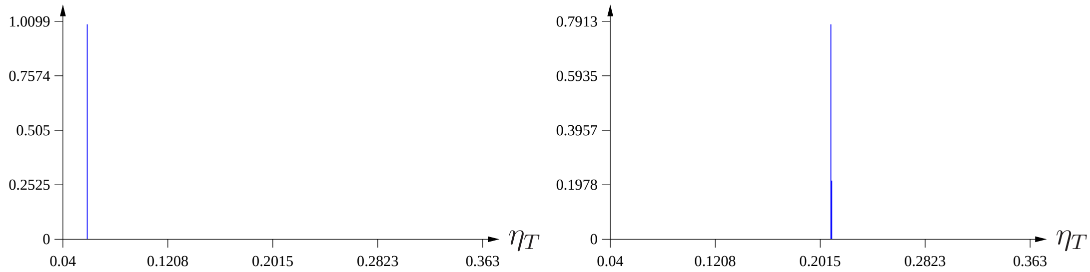
Figure 1: Approximated ODEs for $T = 10000$, $\beta = 2$

4.1.2 Numerical results

To further illustrate the result of Theorem 3.5 we simulate the non-linear model (20) with empirical averages as the performance measure using the program package MACRODYN . We find that the long-run market shares are in fact random. The probability with which they converge to the possible steady states depends both on the initial condition η_0 and the intensity of choice. Figure 2 shows the empirical density of market shares after $T = 10.000$ periods for $N = 100.000$ independent samples of η_T when chartists initially have a market share of 6.5% and 15%, respectively. The densities are roughly concentrated at the left- and rightmost steady state of the ODE in Figure 1(b), respectively.

4.2 Empirical Sharpe ratios as performance measures

The analysis of asymptotic market shares becomes more involved if the mediators' performance is measured by historical Sharpe ratios rather than average returns: in this case the consumers' choice function as given by (11) is no longer invertible.



(a) Initial market share $\eta_0 = 0.065$

(b) Initial market share $\eta_0 = 0.15$

Figure 2: Empirical distributions of market shares for $\beta = 2$.

4.2.1 Analytical results

If the mediators' performance is measured in historical Sharpe ratios as in Example 2.4, then $z_t = (\varrho_t(f^1), \ldots, \varrho_t(f^4))$ is a 4-dimensional vector. Contrary to the previous Example 4.1, the dynamics of the corresponding ODE cannot be described in terms of market shares alone because the choice function F is no longer invertible. The set of asymptotic market shares as given in Corollary 3.4 allows the representation

$$E = \left\{\eta \in [\underline{\eta}, \overline{\eta}] : \Phi^{-1}(\eta, \beta) = \Psi \circ \zeta(\eta)\right\}, \tag{25}$$

where $\Phi^{-1}(\cdot, \beta)$ is the inverse of the logit function (10) and $\Psi \circ \zeta$ describes the stationary difference in Sharpe ratios of the two mediators. While $\Phi^{-1}(\cdot, \beta)$ is analytically available, the map $\Psi \circ \zeta$ can only be obtained by simulating the benchmark models (13). Figure 3 indicates that the two functions in (25) have three intersection points, provided that the intensity of choice β is sufficiently large, whereas the leftmost intersection point ($\beta = 2$) is hardly visible but exists; $\Phi^{-1}(\cdot, \beta)$ has vertical asymptotes at $\underline{\eta}$ and $\overline{\eta}$. These intersection points characterize the possible long-run market shares of the chartist. In particular, market shares and hence asset prices converge to a unique limit if the dependence of consumer decisions on the mediators' performance is weak.

4.2.2 Numerical results

When simulating the non-linear model (20) we find again that the asymptotics of market shares are random and depend on the initial condition and the intensity of choice. Figure 4 shows the empirical density for N=100.000 independent samples of

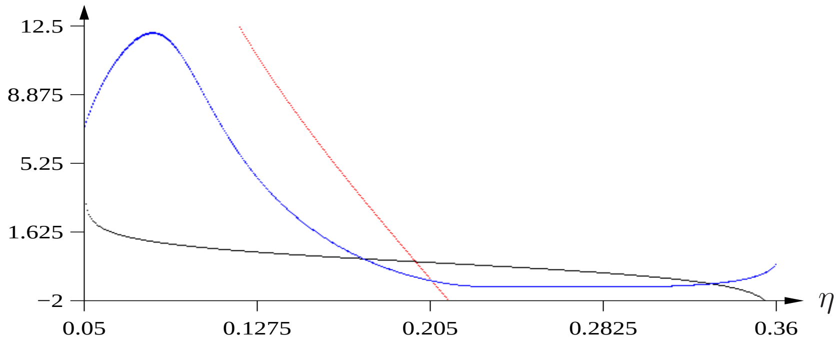
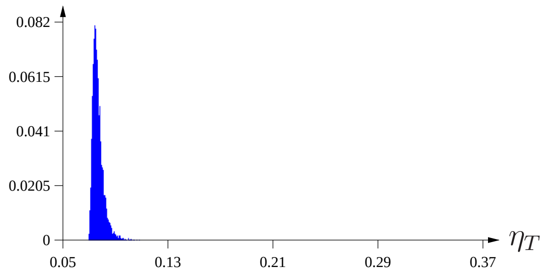
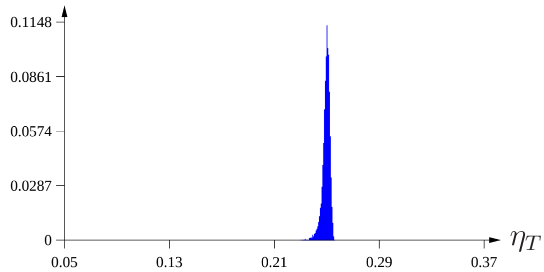


Figure 3: Long-run Sharpe ratios versus logit function; $\beta = 0.5$ (red) and $\beta = 2$ (blue).

market shares after $T = 10.000$ periods for $\beta = 1$ and $\eta_0 = 6.5\%$ and $\eta_0 = 35.5\%$, respectively. Its support roughly coincide with the right- and leftmost intersection point of the two functions depicted in Figure 3, respectively. We infer that the long-run market shares are sensitive to initial conditions. Similar observations are made



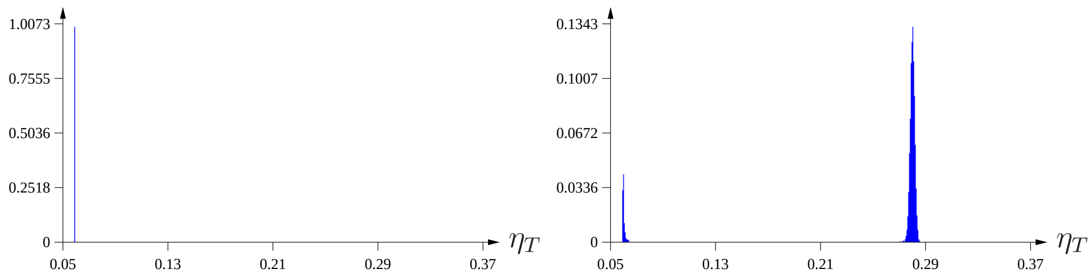
(a) Initial market share $\eta_0 = 0.065$



(b) Initial market share $\eta_0 = 0.355$

Figure 4: Empirical distributions of market shares for $\beta = 1$.

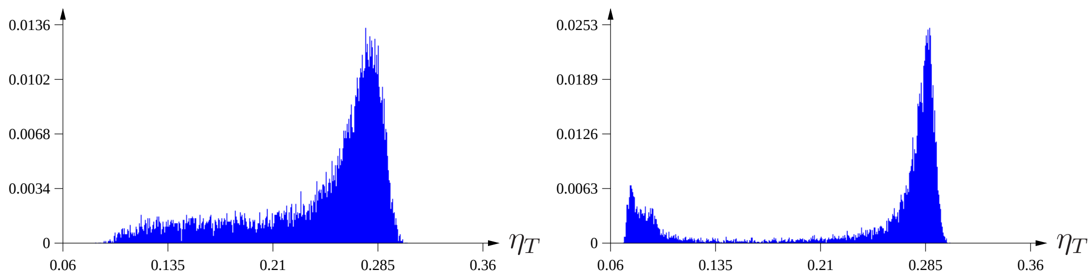
for $\beta = 2$. In this case, however, the distribution of market shares is either unimodal or bimodal depending on the initial market shares. Figure 5(a) shows the empirical distribution for $\eta_0 = 6.5\%$. In this case the chartists almost die out while for $\eta_0 = 35.5\%$ the empirical distribution of asymptotic market shares is bimodal with two peaks which are approximately located at the outer intersection points of the two functions depicted in Figure 3. If the chartists initially have a sufficiently high market share, they 'survive' with high probability. If the interaction between consumer is strong, the long-run market shares not only depend on the initial market share, but also on the specific history of noise trader transactions which is determined by the exogenous noise. It



(a) Initial market share $\eta_0 = 0.065$ (b) Initial market share $\eta_0 = 0.355$

Figure 5: Empirical distributions of market shares for $\beta = 2$.

should also be pointed out that the empirical distribution often converges only very slowly. This is witnessed in Figure 6 which displays the empirical densities of market shares after $T = 500$ and $T = 1000$ periods, respectively.



(a) Distribution after 500 periods. (b) Distribution after 1000 periods.

Figure 6: Empirical distribution of market shares for $\beta = 2$ and $\eta_0 = 0.355$.

5 Conclusions

We investigated the long-run behavior of asset prices in a financial market with interacting agents demonstrating that the price process looses the property of ergodicity if interactive complementarities of agents become too strong. Asset prices may behave in a non-ergodic manner as the price processes converge in distribution, but the limiting distribution is path dependent. Possible long-run market shares along with their limiting distributions were characterized by steady states of a limiting ordinary differential equation. It was shown that the long-run market shares of two competing financial mediators depend strongly on the random environment of the market which is created by the historic behavior of consumers. Taking either empirical average returns or Sharpe ratios as a performance measure, it was shown that efficient portfolios may fail to empirically outperform inefficient portfolios. Chartists may attain higher market shares than mediators holding efficient portfolios, so that rational mediators may not be identified or even be driven out of the market. Economically, this result implies that social interaction constitutes an endogenously generated source of risk which may lead to inefficient long-run portfolio holdings as empirical performance measures may be highly misleading. The result also implies that instead of looking for efficient portfolios, it may be more profitable for professional financial mediators to look for larger market shares.

A Proof of Theorem 3.3

We denote by $\mathbb{P}^\eta_\mu$ and $\mathbb{P}_\mu$ the law of the Markov processes X^η and X with initial distribution μ, respectively, and put $\mathbb{P}^\eta_x := \mathbb{P}^\eta_{\delta_x}$ and $\mathbb{P}_x := \mathbb{P}_{\delta_x}$. The respective expectations are denoted $\mathbb{E}^\eta_\mu$ and $\mathbb{E}^\eta_x$. The empirical distributions associated with X^η and X are denoted $\{\varrho^\eta_t\}_{t\in\mathbb{N}}$ and $\{\varrho_t\}_{t\in\mathbb{N}}$.

Extending $\{\varepsilon_t\}_{t\in\mathbb{N}}$, to a sequence of i.i.d. random variables on $\mathbb{Z}$, the stationary solution $\{x^\eta_t\}_{t\in\mathbb{N}}$ of (13) may be viewed as the Markov chain x^η with initial value

$$x^\eta_0 = \sum_{j=1}^{\infty} A^j(\eta)B(\eta, \varepsilon_{-j}). \tag{26}$$

Under Assumption 3.1 any solution of (13) converges to almost surely to the stationary solution in the sense that $\lim_{t\to\infty} |X^\eta_t - x^\eta_t| = 0$ almost surely. In particular X^η has a unique stationary distribution μ^η. The distribution depends continuously on η and under the assumptions of our main theorem

$$\lim_{t\to\infty} \int f d\varrho^\eta_t = \int f d\mu^\eta \qquad \mathbb{P}^\eta_\mu\text{-a.s.}$$

for any bounded continuous function f, each initial distribution μ, and each $\eta \in [\underline{\eta}, \overline{\eta}]$. We refer the reader to Brandt (1986) for a more detailed discussion of linear stochastic recursive equations.

A.1 Proof of Proposition 3.2

The process X admits the explicit representation

$$X_{t+1} = \sum_{j=0}^{t} \left(\prod_{i=t-j+1}^{t} A(\eta_i) \right) B(\eta_{t-j}, \varepsilon_{t-j}) + \prod_{i=0}^{t} A(\eta_i) X_0. \tag{27}$$

To prove that X is bounded we need to show that the involved matrix products converge fast enough to zero.

Proof of Proposition 3.2: To establish boundedness of (27), we first provide an estimate for the random products of the matrices $A(\eta)$. Since all the eigenvalues of the matrices $A(\eta)$ lie uniformly within the unit circle there exists an $\alpha < 1$ and a constant $N(\eta) \in \mathbb{N}$ which satisfies

$$\|A^{N(\eta)}(\eta)\| < \alpha \quad \text{and so} \quad \|A^{nN(\eta)}(\eta)\| < \alpha^n \quad \text{for all} \quad n \in \mathbb{N};$$

see Varga (1962), Theorem 3.2 for details. Thus, $X \mapsto A(\eta)X$ is a contraction of order $N(\eta)$. Since the entries of the matrices $A(\eta)$ are uniformly bounded, we obtain constants $C(\eta)$ such that

$$\|A^n(\eta)\| < C(\eta)\alpha^{\left[\frac{n}{N(\eta)}\right]} \quad \text{for all} \quad n \in \mathbb{N} \tag{28}$$

where $[x]$ denotes the largest integer less than or equal to $x \in \mathbb{R}^+$. Uniform continuity of the map $\eta \mapsto \|A(\eta)^n\|$ yields

$$N := \sup_{\eta \in [\underline{\eta}, \overline{\eta}]} N(\eta) < \infty. \tag{29}$$

From this and (19) we see that there exists a constant $\tilde{C} < \infty$ such that

$$\left\| \prod_{i=0}^{t} A(\eta_i) \| \right\| \le \tilde{C}\alpha^{\left[\frac{t}{N}\right]}. \tag{30}$$

Since the random variables $B(\eta, \varepsilon)$ are uniformly bounded this shows that

$$\sup_t \sum_{j=0}^{t} \left\| \prod_{i=t-j+1}^{t} A(\eta_i) \right\| \le C$$

for some $C < \infty$. Hence the assertion follows from the explicit representation (27). □

The same arguments as in the proof of Proposition 3.2 can also be applied to prove that family of Markov chains $\{X^\eta\}_{\eta \in [\underline{\eta}, \overline{\eta}]}$ is uniformly bounded.

Corollary A.1 *(i) For any compact set D of initial values there exists a constant M_D such that*

$$\mathbb{P}\left[\sup_{t,\eta} |X_t^\eta| \leq M_D \mid X_0^\eta \in D\right] = 1.$$

(ii) For any two processes X^η and Y^η with initial values x and y, respectively,

$$\sup_{\eta,t} |X_t^\eta - Y_t^\eta| \leq C\alpha^{\left[\frac{t}{N}\right]}|x-y|.$$

The proof of Proposition 3.2 shows that all the stationary distributions μ^η are concentrated on a common compact set K. Thus, for any compact set of initial values D, the Markov chains may be viewed as Markov chains on a compact state space K_D. That is, we may assume that the transition kernels Π_η of X^η satisfy

$$\Pi_\eta(x; K_D) = 1 \quad \text{for all} \quad x \in K_D.$$

A.2 Proof of Theorem 3.3

In this section we show that almost sure convergence of market shares implies convergence in distribution of asset prices and forecasting rules. From the proof of Proposition 3.2 we know that there exists a constant $N \in \mathbb{N}$ such that $\sup_\eta \|A^N(\eta)\| < 1$. To ease some of the notational complexity, we prove Theorem 3.3 under the simplifying assumption that $N = 1$, i.e., that

$$\alpha := \sup\{\|A(\eta)\| : \underline{\eta} \leq \eta \leq \overline{\eta}\} < 1. \tag{31}$$

The general case follows from straightforward modification of the arguments given below. At this point it is also convenient to recall that the *Vasserstein metric*

$$d(\mu,\nu) := \sup\left\{|f(\mu) - f(\nu)| \ : \ \|f\|_\infty \leq 1, \ f \text{ Lipschitz with constant } 1\right\} \tag{32}$$

induces the weak topology on the class of all probability measures on $\mathbb{R}^d$. In terms of this metric, almost sure convergence of empirical distributions to their expected value under the unique stationary measure translates into

$$\lim_{t\to\infty} d\left(\varrho_t^\eta, \mu^\eta\right) = 0 \qquad \mathbb{P}_x^\eta\text{-a.s.}$$

Proof of Theorem 3.3: Let us introduce, for any $T \in \mathbb{N}$ a benchmark processes $\overline{X}^T$ by

$$\overline{X}_t^T = X_t \quad \text{for} \quad t \le T \quad \text{and} \quad \overline{X}_{t+1}^T = A(\eta_T)\overline{X}_t^T + B(\eta_T, \varepsilon_t) \quad \text{for} \quad t > T.$$

In view of our simplifying condition (31) and Assumption 3.1 we obtain

$$\begin{aligned} |X_{T+t} - \overline{X}_{T+t}^T| &\le \alpha |X_{T+t-1} - \overline{X}_{T+t-1}^T| + C \sup_{t \ge T} |\eta_t - \eta_T| \\ &\le \frac{C}{1-\alpha} \sup_{t \ge T} |\eta_t - \eta_T|. \end{aligned}$$

Notice now that for any bounded Lipschitz continuous function g with constant 1,

$$\sup_{t \ge T} |\varrho_t(g) - \overline{\varrho}_t^T(g)| \le \sup_{t \ge T} \frac{1}{t-T} \sum_{i=T+1}^{t} |X_i - \overline{X}_i^T|.$$

This implies

$$\lim_{T \to \infty} \mathbb{P}_x \left[\sup_{t \ge T} d(\varrho_t, \overline{\varrho}_t^T) \ge \varepsilon \right] = 0. \tag{33}$$

Moreover, almost sure convergence of the sequence of market shares $\{\eta_t\}_{t \in \mathbb{N}}$ to η_* yields

$$\mathbb{P}_x \left[\lim_{T \to \infty} d(\mu^{\eta_T}, \mu^{\eta_*}) = 0 \right] = 1. \tag{34}$$

Since the random variables η_T and $\varepsilon_{T+1}, \varepsilon_{T+2}, \ldots$ are independent and all the eigenvalues of $A(\eta_T)$ lie inside the unit circle, $\overline{X}^T$ is an ergodic Markov chain with invariant distribution μ^{η_T}. The associated sequence of empirical distributions $\{\overline{\varrho}_t^T\}_{t \ge T}$ converges almost surely weakly to μ^{η_T}:

$$\mathbb{P}_x \left[\lim_{t \to \infty} d(\overline{\varrho}_t^T, \mu^{\eta_T}) = 0 \right] = 1. \tag{35}$$

An application of the triangle inequality shows that the assertion follows from (33)-(35).

□

B Proof of the convergence result

This section proves our convergence result for market shares stated in Theorem 3.5. The arguments will be based on a stochastic approximation result for stochastic difference equations and a uniform large deviations principle for stable autoregressive processes.

B.1 Stochastic approximation

Our goal is to apply a stochastic approximation result of Kushner & Yin (2003). To this end we first rewrite the second equation in (20). For each $T \in \mathbb{N}$, set

$$z_T^t := \frac{1}{t} \sum_{s=T+1}^{T+t} \big(f^1(X_{s-1}), \ldots, f^L(X_{s-1})\big),$$

such that empirical averages up to time $T+t$ take the form

$$z_{T+t} = \frac{T}{T+t} z_T + \frac{t}{T+t} z_T^t. \tag{36}$$

Using (21), we obtain the following representation of z_{T+t}:

$$z_{T+t} = z_T + \frac{t}{T+t}\big[g(z_T) + \beta_T^t\big] \quad \text{with} \quad \beta_T^t = z_T^t - \zeta\big(F(z_T)\big). \tag{37}$$

Let us also introduce a different time scale by defining $\{t_n\}_{n\in\mathbb{N}}$ and $\{T_n\}_{n\in\mathbb{N}}$ by

$$t_n := n^4, \quad T_n := \sum_{i=1}^{n} t_{i-1}.$$

Setting $\hat{\theta}_n := z_{T_n}$ and $\beta_n := \beta_{T_n}^{t_n}$ the random sequence $\{\hat{\theta}_n\}_{n\in\mathbb{N}}$ satisfies

$$\hat{\theta}_{n+1} = \hat{\theta}_n + \varepsilon_n\big[g(\hat{\theta}_n) + \beta_n\big] \quad \text{where} \quad \varepsilon_n := \frac{t_n}{T_n + t_n}. \tag{38}$$

Since T_n is of the order n^5, the quantity ε_n is of the order n^{-1}. This allows us to apply a stochastic approximation algorithm to the sequence $\{\hat{\theta}_n\}_{n\in\mathbb{N}}$ defined in (38), if the "error terms" β_n converge to zero sufficiently fast. The ODE method for approximating the dynamics of the discrete time process (38) uses a continuous time interpolation of the sequence $\{\hat{\theta}_n\}_{n\in\mathbb{N}}$. A natural time scale for the interpolation is defined in terms of the step-size sequence $\{\varepsilon_n\}_{n\in\mathbb{N}}$. Specifically, let us define

$$\tau_0 = 0 \quad \text{and} \quad \tau_n := \sum_{i=0}^{n-1} \varepsilon_i$$

and the continuous time interpolation $\theta^0 = (\theta^0(t))_{t\geq 0}$ of the discrete time process $\{\hat{\theta}_n\}_{n\in\mathbb{N}}$ by $\theta^0(0) = \hat{\theta}_0$ and

$$\theta^0(t, \omega) = \hat{\theta}_n(\omega) \quad \text{for} \quad \tau_n \leq t < \tau_{n+1}.$$

Furthermore, we introduce the left-shifts θ^n of θ^0 by $\theta^n(t,\omega) = \theta^0(\tau_n + t,\omega)$. If the sequence $\{\varepsilon_n\}_{n\in\mathbb{N}}$ satisfies

$$\sum_{n\geq 0} \varepsilon_n = \infty \quad \text{and} \quad \sum_{n\geq 0} \varepsilon_n^2 < \infty \tag{39}$$

and if the "error terms" β_n are asymptotically negligible in the sense that

$$\sum_{n\geq 0} \varepsilon_n |\beta_n| < \infty \quad \text{almost surely} \tag{40}$$

then the functions $\theta^n(\cdot,\omega)$ are equicontinuous for almost every ω. If, in addition the discrete time sequence $\{\hat{\theta}_n\}_{n\in\mathbb{N}}$ is bounded with probability one, then any limit $\theta(\cdot,\omega)$ of some convergent subsequence $\{\theta^{n_k}(\cdot,\omega)\}_{k\in\mathbb{N}}$ of $\{\theta^n(\cdot,\omega)\}_{n\in\mathbb{N}}$ is a trajectory of the ordinary differential equation

$$\dot{\theta} = g(\theta). \tag{41}$$

More precisely, the functions $\{\theta^{n_k}(\cdot,\omega)\}_{k\in\mathbb{N}}$ converge to the unique solution of the ODE (41) with initial condition $\theta(0,\omega)$, given by

$$\theta(t,\omega) = \theta(0,\omega) + \int_0^t g(\theta(s,\omega))\, ds,$$

uniformly on compact time intervals. For all $T < \infty$ we have

$$\lim_{k\to\infty} \sup_{0\leq t\leq T} |\theta^{n_k}(t,\omega) - \theta(t,\omega)| = 0. \tag{42}$$

This approximation result allows us to analyze the asymptotics of the sequence $\{\hat{\theta}_n\}_{n\in\mathbb{N}}$ by means of the long-run behavior of the ODE (41). Theorem 2.1 in Kushner & Yin (2003, Chapter 5) states that if $\{\hat{\theta}_n(\omega)\}_{n\in\mathbb{N}}$ visits some compact subset C_i of a basin of attraction $DA(s_i)$ of some asymptotically stable steady state s_i of the system $\dot{\theta} = g(\theta)$ infinitely often with probability $\mathbf{p} > 0$, then

$$\lim_{n\to\infty} |\hat{\theta}_n - s_i| = 0 \quad \text{with at least probability} \quad \mathbf{p}.$$

Proof of Theorem 3.5: In order to apply the convergence result of Kushner & Yin (2003) we only need to verify (40). This will be done in Lemma B.5 below. If (40) holds, the discrete-time process $\{\hat{\theta}_n\}_{n\in\mathbb{N}}$ defined in (38) converges to some asymptotically stable fixed point $s_i \in S^*$ of the associated ODE with at least probability $\mathbf{p}$. This proves the assertion because (19) and uniform continuity of the choice functions yields

$$\max\left\{|z_{T_n} - z_t| \,:\, t = T_n + 1, T_n + 2, \ldots, T_{n+1}\right\} = O(n^{-1}).$$

□

B.2 A uniform large deviations principle

For a given market share η the Markov chain X^η is ergodic. The large deviation principle provides a measure for the speed of convergence of time averages to their expected values under the stationary measure.

Definition B.1 *A sequence $\{M_t\}_{t\in\mathbb{N}}$ of random variables defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ satisfies a large deviations principle with speed $a_t \uparrow \infty$ $(t \to \infty)$ and rate function I if*

$$\limsup_{t\to\infty} \frac{1}{a_t} \log \mathbb{P}\left[|M_t| \in F\right] \le -\inf\{I(u) : u \in F\}$$

for any closed set F and

$$\liminf_{t\to\infty} \frac{1}{a_t} \log \mathbb{P}\left[|M_t| \in U\right] \ge -\inf\{I(u) : u \in U\}$$

for any open set U. The sequence $\{M_t\}_{t\in\mathbb{N}}$ satisfies a moderate deviations principle with speed $a_t = o(t)$ and rate function $\hat{I}$ if the sequence $\left\{\sqrt{\frac{t}{a_t}} M_t\right\}_{t\in\mathbb{N}}$ satisfies a large deviations principle with speed a_t and rate function $\hat{I}$.

If the sequence of random variables $\{M_t\}_{t\in\mathbb{N}}$ satisfy a moderate deviations principle with speed a_t and rate function $\hat{I}$, then for all sufficiently large $t \in \mathbb{N}$:

$$\mathbb{P}\left[|M_t| \ge \sqrt{\frac{a_t}{t}}\right] = \mathbb{P}\left[\sqrt{\frac{t}{a_t}}|M_t| \ge 1\right] \le e^{-a_t \inf\{\hat{I}(u):|u|\ge 1\}} \tag{43}$$

B.2.1 A uniform moderate deviations principle for empirical averages

In the sequel we derive a *uniform* moderate deviations principle for the random variables

$$M_t^\eta := \varrho_t^\eta(f) - \mu^\eta(f) \quad \text{for a bounded Lipschitz function } f. \tag{44}$$

Proposition B.2 *There exists $c > 0$ such that for any compact set D of initial values*

$$\limsup_{t\to\infty} \frac{1}{\sqrt{t}} \log \sup_{\eta, x\in D} \mathbb{P}_x^\eta \left[|M_t^\eta| \geq t^{-1/4} \right] \leq -\frac{1}{2c}. \tag{45}$$

As an immediate corollary we then obtain that uniformly in all the possible market shares the random variable $M_{t^4}^\eta$ exceeds the values t^{-1} only finitely often.

Corollary B.3 *For any compact set of initial values D and all $\eta \in [\underline{\eta}, \overline{\eta}]$, we have that*

$$\mathbb{P}_x^\eta \left[|M_{t^4}^\eta| \geq \frac{1}{t} \text{ infinitely often} \right] = 0.$$

Proof: The assertion follows from the Borel-Cantelli Lemma because

$$\sup_{\eta, x\in D} \mathbb{P}_x^\eta \left[|M_{t^4}^\eta| \geq \frac{1}{t} \right] \leq e^{-\frac{t^2}{2c}} \quad \text{for all sufficiently large } t.$$

□

The proof of Proposition B.2 requires some preparation. From

$$M_t^\eta \leq \sqrt{N} \sum_{j=0}^{N-1} |M_t^{\eta,j}| \quad \text{where} \quad M_t^{\eta,j} := \frac{1}{\sqrt{[t/N]}} \sum_{i=0}^{[t/N]} \left(f(X_{Ni+j}^\eta) - \mu^\eta(f) \right)$$

we obtain

$$\left\{ |M_t^\eta| \leq t^{-1/4} \right\} \supset \left\{ |M_t^{\eta,j}| \leq \frac{1}{N^{3/2}} t^{-1/4} \text{ for } j = 1, 2, \ldots, N-1 \right\}.$$

As a result, it suffices to prove Proposition B.2 for the Markov chain $\{X_{Nt}^\eta\}_{t\in\mathbb{N}}$. In view of the discussion at the end of Section A.1 we may as well assume that the Markov chains X^η are contractions uniformly in $\eta \in [\underline{\eta}, \overline{\eta}]$, i.e., that $\sup_\eta \|A(\eta)\| < 1$. It then follows from Worms (1999, Thm. 2), that the sequence

$$\left\{ \sqrt{\frac{t}{a_t}} M_t^\eta \right\}_{t\in\mathbb{N}}$$

satisfies a large deviation principle with speed $\{a_t\}_{t\in\mathbb{N}}$ if $a_t = o(t)$. Furthermore, the rate function $\hat{I}^\eta$ can be given in closed form:

$$\hat{I}^\eta(u) = \sup_\theta \left\{ u\theta - \frac{1}{2}\theta^2 c_\eta \right\} = \frac{1}{2} \frac{u^2}{c_\eta}. \tag{46}$$

The constant c_η is given in terms of the solution G_η to the *Poisson equation* associated to f and the transition operator Π_η of the Markov chain X^η. More specifically, there exist functions G_η which solve

$$f - \mu^\eta(f) = G_\eta - \Pi_\eta G_\eta, \tag{47}$$

and Lipschitz continuity of f yields Lipschitz continuity of G_η and $G_\eta - \Pi_\eta G_\eta$ with the *same* constant, cf. Duflo (1997, Chap. 6). The normalized functions $G_\eta(x) - G_\eta(0)$ also satisfy the Poisson equation (47) and share the same Lipschitz constant. Thus, we may with no loss of generality assume that $G_\eta(0) = 0$. In this case, the functions G_η are uniformly bounded and equicontinuous because, for any compact set D of initial values, the processes X^η may be viewed as a Markov processes on a common compact state space. In terms of G_η the constant c_η and the random variable M_t^η are given by, respectively,

$$c_\eta = \int \left[G_\eta^2 - (\Pi_\eta G_\eta)^2\right] d\mu^\eta = \mathbb{E}_{\mu^\eta}\left[G_\eta^2 - (\Pi_\eta G_\eta)^2\right] \tag{48}$$

and

$$M_t^\eta = \frac{1}{t}\sum_{s=1}^{t}\left[G_\eta(X_s^\eta) - \Pi_\eta G_\eta(X_{s-1}^\eta)\right] + \frac{1}{t}G_\eta(X_0) - \frac{1}{t}G_\eta(X_t^\eta). \tag{49}$$

Since all the functions G_η are uniformly bounded, the constants c_η are uniformly bounded by some constant $c < \infty$. Moreover, if $\{\mathcal{F}_t^\eta\}_{t\in\mathbb{N}}$ denotes the filtration generated by the Markov chain X^η, then

$$\mathbb{E}_x^\eta\left[G_\eta(X_t^\eta) - \Pi_\eta G_\eta(X_{t-1}^\eta)|\mathcal{F}_{t-1}^\eta\right] = 0.$$

Thus, the deviation of empirical averages from their expected values under the stationary measure can be described in terms of a *martingale difference* sequence; see Duflo (1997, Theorem 6.3.20) for details.

Remark B.4 *The Markov property of X^η implies that*

$$\begin{aligned}
\mathbb{E}_x^\eta\left[G_\eta(X_{t+j}^\eta)\Pi_\eta G_\eta(X_{t+j-1}^\eta)|\mathcal{F}_j^\eta\right] &= \mathbb{E}_{X_j^\eta}^\eta\left[G_\eta(X_t^\eta)\Pi_\eta G_\eta(X_{t-1}^\eta)\right] \\
&= \mathbb{E}_{X_j^\eta}^\eta\left[\mathbb{E}_{X_j^\eta}^\eta\left[G(X_t^\eta)\Pi_\eta G_\eta(X_{t-1}^\eta)|X_{t-1}^\eta\right]\right] \\
&= \mathbb{E}_{X_j^\eta}^\eta\left[(\Pi_\eta G_\eta)^2(X_{t-1}^\eta)\right] \\
&= \mathbb{E}_x^\eta\left[(\Pi_\eta G_\eta)^2(X_{t+j-1}^\eta)|\mathcal{F}_j^\eta\right].
\end{aligned}$$

From this we obtain

$$\mathbb{E}_x^\eta\left[G_\eta^2(X_{t+j}^\eta) - (\Pi_\eta G_\eta)^2(X_{t+j-1}^\eta)|\mathcal{F}_t^\eta\right] = \mathbb{E}_x^\eta\left[\left(G_\eta(X_{t+j}^\eta) - \Pi_\eta G_\eta(X_{t+j-1}^\eta)\right)^2|\mathcal{F}_t^\eta\right].$$

Since μ^η is the stationary distribution for X^η, a similar argument shows that

$$c_\eta = \mathbb{E}_{\mu^\eta}\left[G_\eta^2(x_j^\eta) - (\Pi_\eta G_\eta)^2(x_{j-1}^\eta)\right] = \mathbb{E}_{\mu^\eta}\left[G_\eta(x_j^\eta) - (\Pi_\eta G_\eta)(x_{j-1}^\eta)\right]^2.$$

This allows us to estimate the speed of convergence by means of a uniform moderate deviation principle for martingale difference sequences established in Gao (1996).

With the specific choice $a_t = \sqrt{t} = o(t)$, it follows from the results in Worms (1999) along with (43) and the fact that the constants c^η are bounded by some c that

$$\sup_x \mathbb{P}_x^\eta\left[|M_t^\eta| \geq \frac{1}{t^{1/4}}\right] \leq e^{-\frac{1}{2c}\sqrt{t}} \quad \text{for all} \quad t \geq T_\eta. \tag{50}$$

For our purposes we need the latter estimate to be uniform in all the possible market shares. To this end, recall first that the stationary solution x^η may be viewed as the Markov chain X^η with initial value x_0^η. The arguments given in the proof of Proposition 3.2 show that x_0^η is bounded uniformly in $\eta \in [\underline{\eta}, \overline{\eta}]$. In view of Corollary A.1 (ii) this yields a constant C depending on D, but not on η such that

$$\sup_{x\in D}|X_t^\eta - x_t^\eta| \leq C\alpha^t$$

almost surely. Since the functions $G_\eta - \Pi_\eta G_\eta$ are uniformly bounded and uniformly Lipschitz continuous, so are G_η^2 and $(\Pi_\eta G_\eta)^2$. Thus, in view of Remark B.4 there exists a constant $L < \infty$ such that

$$\begin{aligned} & \left|\mathbb{E}_x^\eta\left[\left(G_\eta(X_{t+j}^\eta) - (\Pi_\eta G_\eta)(X_{t+j-1}^\eta)\right)^2|\mathcal{F}_t^\eta\right] - c_\eta\right| \\ = & \left|\mathbb{E}_x^\eta\left[G_\eta^2(X_{t+j}^\eta) - (\Pi_\eta G_\eta)^2(X_{t+j-1}^\eta)|\mathcal{F}_t^\eta\right] - \mathbb{E}_{\mu^\eta}\left[G_\eta^2(x_j^\eta) - (\Pi_\eta G_\eta)^2(x_{j-1}^\eta)\right]\right| \\ \leq & L\alpha^{-(j-1)} \end{aligned}$$

uniformly in all the possible market shares and in $x \in D$. This yields

$$\lim_{j\to\infty, \frac{j}{T}\to 0} \sup_{\eta, x\in D}\left\|\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}_x^\eta\left[\left(G_\eta(X_{t+j}^\eta) - \Pi_\eta G_\eta(X_{t+j-1}^\eta)\right)^2|\mathcal{F}_t^\eta\right] - c_\eta\right\|_{L^\infty(\mathbb{P}_x^\eta)} = 0.$$

Hence it follows from Theorem 1.1 in Gao (1996) that

$$\limsup_{T\to\infty} \frac{1}{\sqrt{T}} \log \sup_{\eta, x\in D} \mathbb{P}_x^\eta \left[\frac{1}{T^{3/4}} \left| \sum_{t=1}^T G_\eta(X_t^\eta) - \Pi_\eta G_\eta(X_{t-1}^\eta) \right| \geq 1 \right] \leq -\frac{1}{2c}.$$

In view of (49) and because the functions G_η are uniformly bounded we obtain

$$\limsup_{T\to\infty} \frac{1}{\sqrt{T}} \log \sup_{\eta, x\in D} \mathbb{P}_x^\eta \left[|M_T| \geq T^{-1/4} \right] \leq -\frac{1}{2c}.$$

This finishes the proof of Proposition B.2.

B.2.2 Bounding error terms by a uniform moderate deviations principle

Our uniform large deviations principle allows us to prove that the error terms β_n in the representation (38) of the sequence $\{\hat{\theta}_n\}_{n\in\mathbb{N}}$ vanish sufficiently fast.

Lemma B.5 *We have that*

$$\mathbb{P}\left[\sum_{n\geq 0} \frac{t_n}{T_n + t_n} |\beta_n| < \infty \right] = 1. \tag{51}$$

PROOF: Recall that $\{\overline{\varrho}_t^T\}_{t\geq T}$ denotes the sequence of empirical distributions of a process $\overline{X}^T$ starting at time T in X_T which evolves according to the linear recursive relation (13) with fixed market shares η_T. Write

$$\overline{z}_t^T := \left(\overline{\varrho}_t^T(f^1), \ldots, \overline{\varrho}_t^T(f^L) \right)$$

and recall (37) to get

$$\beta_n = \left(z_{T_n}^{t_n} - \overline{z}_{t_n}^{T_n} \right) + \left(\overline{z}_{t_n}^{T_n} - \zeta(\eta_{T_n}) \right).$$

With regards to the second term in the decomposition of β_n note first that independence of the random variables η_{T_n} and $\varepsilon_{T_n+1}, \varepsilon_{T_n+2}, \ldots$ yields

$$\mathbb{P}_x \left[\left| \overline{z}_{t_n}^{T_n} - \zeta(\eta_{T_n}) \right| \geq \frac{1}{n} \right] \quad = \quad \mathbb{P}_{X_{T_n}}^{\eta_{T_n}} \left[\left| z_{t_n}^{\eta_{T_n}} - \zeta(\eta_{T_n}) \right| \geq \frac{1}{n} \right].$$

Given an initial value x, the process $\{X_t\}_{t\in\mathbb{N}}$ is almost surely bounded, due to Proposition 3.2. In particular, there exists a compact set D such that

$$\mathbb{P}_x \left[X_t \in D \text{ for all } t \in \mathbb{N} \right] = 1.$$

Thus, Proposition B.2 applied to the Lipschitz continuous functions f^i $(i = 1, 2, \ldots, L)$ yields a constant $N \in \mathbb{N}$ such that

$$\begin{aligned}\mathbb{P}_x\left[\left|\overline{z}_{t_n}^{T_n} - \zeta(\eta_{T_n})\right| \geq \frac{1}{n}\right] &\leq \sup_{x\in D}\sup_{\eta\in[\underline{\eta},\overline{\eta}]} \mathbb{P}_x^{\eta}\left[\max_{i=1,2,\ldots,L}\left|\varrho_{t_n}^{\eta}(f^i) - \mu^{\eta}(f^i)\right| \geq \frac{1}{n}\right] \\ &\leq Le^{-\frac{1}{2c}n}\end{aligned}$$

for all $n \geq N$. Hence the lemma by Borel and Cantelli yields

$$\sum_{n\geq 1}\frac{t_n}{T_n + t_n}\left|\overline{z}_{t_n}^{T_n} - \zeta(\eta_{T_n})\right| < \infty \qquad \mathbb{P}_x\text{-a.s.}$$

To study the first term of the decomposition we apply similar arguments as in the second and third part of the proof of Theorem 1.4 to deduce that up to multiplicative constants

$$\begin{aligned}\left|X_{T_n+t_n} - \overline{X}_{T_n+t_n}^{T_n}\right| &\leq \alpha_0^{\left[\frac{t_n}{N}\right]}\sup_{0\leq k\leq N}\left|X_{T_n+k} - \overline{X}_{T_n+k}^{T_n}\right| + \sup_{T_n\leq t\leq T_n+t_n}\left|\eta_t - \eta_{T_n}\right| \\ &\leq \alpha_0^{\left[\frac{t_n}{N}\right]}\sup_{0\leq k\leq N}\left|X_{T_n+k} - \overline{X}_{T_n+k}^{T_n}\right| + \frac{t_n}{T_n + t_n}.\end{aligned}$$

The second inequality follows from (19). Since $\sup_{0\leq k\leq N}\left|X_{T_n+k} - \overline{X}_{T_n+k}^{T^n}\right|$ is almost surely bounded by some constant that depends only on the starting point of X, we see that

$$\left|X_{T_n+t_n} - \overline{X}_{T_n+t_n}^{T_n}\right| = O\left(n^{-1}\right) \qquad \mathbb{P}_x\text{-a.s.}$$

Hence uniform continuity of the maps $f^1, \ldots, f^l$ and the choice function F yields

$$\left|\overline{z}_{t_n}^{T_n} - z_{t_n}^{\eta_{T_n}}\right| = O\left(n^{-1}\right) \qquad \mathbb{P}_x\text{-a.s.}$$

and so

$$\sum_{n\geq 1}\frac{t_n}{T_n + t_n}\left|\overline{z}_{t_n}^{T_n} - z_{t_n}^{\eta_{T_n}}\right| < \infty \qquad \mathbb{P}_x\text{-a.s.}$$

□

References

ALCHIAN, A. (1950): "Uncertainty, Evolution and Economic Theory", *Journal of Political Economy*, 58, 211–221.

Arrowsmith, D. K. & C. M. Place (1994): *Introduction to Dynamical Systems*. Cambridge University Press, Cambridge.

Blume, L. (1993): "The statistical mechanics of strategic interaction", *Games and Economic Behavior*, 5, 387–424.

Blume, L. & D. Easley (1992): "Evolution and Market Behavior", *Journal of Economic Theory*, 58, 9–40.

——— (2005): "Optimality and natural selection in markets", *Journal of Economic Theory*, in press.

Böhm, V. (2003): "MACRODYN – The Handbook", Discussion Paper 498, Department of Economics, University of Bielefeld.

Böhm, V. & C. Chiarella (2005): "Mean Variance Preferences, Expectations Formations, and the Dynamics of Random Asset Prices", *Mathematical Finance*, 15(1), 61–97.

Böhm, V. & J. Wenzelburger (2005): "On the Performance of Efficient Portfolios", *Journal of Economic Dynamics and Control*, 29(4), 721–740.

Brandt, A. (1986): "The Stochastic Equation $Y_{t+1} = A_t Y_t + B_t$ with Stationary Coefficients", *Journal of Applied Probability*, 18, 211–220.

Brock, W. & C. Hommes (1997): "A Rational Route to Randomness," *Econometrica*, 65 (5), 1059–1096.

——— (1998): "Heterogeneous Beliefs and Routes to Chaos in a Simple Asset Pricing Model.", *Journal of Economic Dynamics and Control*, 22, 1235–1274.

De Long, J.B., A. Shleifer, L.H. Summers & R. J. Waldmann (1990): "Noise Traders Risk in Financial Markets", *Journal of Political Economy*, 98, 703–738.

——— (1991): "The Survival of Noise Traders", *Journal of Business*, 64, 1–19.

Duflo, M. (1997): *Random Iterative Models*. Springer, Berlin.

Estigneev, I., Hens, T. & K.R. Schenk-Hoppé (2005): "Evolutionary Stabile Stock Markets," *Economic Theory*, in press.

Föllmer, H., U. Horst & A.P. Kirman (2005): "Equilibria in financial markets with heterogeneous agents: A probabilistic perspective", *Journal of Mathematical Economics*, 41 (1-2), 123-155.

FÖLLMER, H. & M. SCHWEIZER (1993): "A Microeconomic Approach to Diffusion Models for Stock Prices," *Mathematical Finance*, 3, 1–23.

FRANKEL, J.A. & K. FROOT (1986): "The Dollar as an Irrational Speculative Bubble: A Tale of Fundamentalists and Chartists," *The Marcus Wallenberg Papers on International Finance*, 1, 27–55.

FRIEDMAN, M. (1953): *Essays in Positive Economics*. University of Chicago Press, Chicago.

GAO, F. (1996): "Moderate Deviations for Martingales and Mixing Sequences", *Stochastic Processes and Their Applications*, 61, 263–277.

HILLEBRAND, M. & J. WENZELBURGER (2006): "The Impact of Multiperiod Planning Horizons on Portfolios and Asset Prices in a Dynamic CAPM", forthcoming *Journal of Mathematical Economics*.

HORST, U. (2005): "Financial price fluctuations in a stock market model with many interacting agents", *Economic Theory*, 25(4), 917–932.

HORST, U. & J. SCHEINKMAN (2005): "Equilibria in systems of social interaction", *Journal of Economic Theory*, in press.

KUSHNER, H.J. & G.G. YIN (2003): *Stochastic Approximation and Recursive Algorithms and Applications*, Stochastic Modelling and Applied Probability. Springer-Verlag, 2nd. edn.

SANDRONI, A. (2000): "Do Markets Favor Agents Able to Make Accurate Predictions?", *Econometrica*, 68, 1303–1341.

SIMON, H.A. (1982): *Models of Bounded Rationality*. MIT Press, Cambridge Mass.

VARGA, R.S. (1962): *Matrix Iterative Analysis*. Prentice-Hall, Englewood Cliffs.

WENZELBURGER, J. (2004): "Learning to Predict Rationally when Beliefs are Heterogeneous", *Journal of Economic Dynamics and Control*, 28(10), 2075–2104.

WORMS, J. (1999): "Moderate Deviations for Stable Markov Chains and Regression Models", *Electronic Journal of Probability*, 4, 1–28.