Equilibria in Financial Markets with Heterogeneous Agents: A Probabilistic Perspective*

Hans Föllmer†, Ulrich Horst‡, and Alan Kirman§

13th October 2004

Abstract

We analyse financial market models in which agents form their demand for an asset on the basis of their forecasts of future prices and where their forecasting rules may change over time, as a result of the influence of other traders. Agents will switch from one rule to another stochastically, and the price and profits process will reflect these switches. Among the possible rules are "chartist" or extrapolatory rules. Prices can exhibit transient behaviour when chartists predominate. However, if the probability that an agent will switch to being a "chartist" is not too high then the process does not explode. There are occasional bubbles but they inevitably burst. In fact, we prove that the limit distribution of the price process exists and is unique. This limit distribution may be thought of as the appropriate equilibrium notion for such markets. A number of characteristics of financial time series can be captured by this sort of model. In particular, the presence of chartists fattens the tails of the stationary distribution.

JEL subject classification: C62,D84,G12

Key words: financial markets, stochastic price processes, limit distributions, forecasting rules

---

*We would like to thank Buz Brook, Thorsten Hens, Cars Hommes, Thomas Lux, Romain Ricciotti, Gilles Teyssiere, Richard Topol, and seminar participants at various institutions for valuable comments, and Christian Rothe for assistance with the numerical simulations. Part of this research was carried out while Hans Föllmer and Ulrich Horst were visiting the Department of Operations Research and Financial Engineering and the Bendheim Center for Finance at Princeton University, respectively. Grateful acknowledgement is made for hospitality. Financial support through the DFG Research Center "Mathematics for Key Technologies" (FZT 86), the German Academic Exchange Service and the NSF grant No. 0001647 to Princeton University is gratefully acknowledged.

†Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany

‡Institut für Mathematik, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany, E-mail:horst@mathematik.hu-berlin.de

§Institute Universitaire de France, EHSS, Université d'Aix-Marseille III, 2 Rue de la Charité, 13002 Marseille, France.

1 Introduction

One of the major difficulties with standard models of asset markets and particularly of foreign exchange markets is that it is difficult to explain the appearance of "bubbles" which grow and then burst. There is a substantial literature on "rational bubbles"; see, e.g., Evans (1991). Such models suffer, in general, from the lack of an underlying market framework and also from the fact that such bubbles will be permanently explosive. Other papers have shown how prices can be correlated with apparently irrelevant exogenous variables, "sunspots" for example (Cass and Shell (1983)), and Woodford (1990) has shown that individuals can rationally learn to believe that sunspots play the role of fundamentals and, as a consequence these beliefs will be fulfilled. Typically, the analysis in such examples wishes to show that prices will settle to some equilibrium, albeit a "wrong" one.

Our aim here is to take a rather different approach and to show how agents may change their expectations as they learn and, given the self reinforcing aspect of their expectations, prices will move with the expectations. It may also be the case that agents will cease for some periods to believe in any fundamentals and that they will become "chartists" or trend chasers. In this case markets detach themselves from any fundamentals for a period but then return; for a good overview of theoretical aspects and empirical evidence for this see Hunter, Kaufman, and Pomerleano (2003). Such a vision fits well with the burgeoning econometrics literature on "switching regimes" (Hamilton (1989)) and provides an economic framework for this phenomenon.

As agents switch from one belief to another there will be periods where there is considerable heterogeneity of expectations, consistent with empirical observations; see Chionis and MacDonald (2002). Yet, over time, agents cannot be said to be behaving irrationally. Heterogeneous expectations are also necessary to explain other phenomena. Without them, one could not explain as Bacchetta and van Wincoop (2003) point out, the enormous volume of trade on financial markets. Spot trading on foreign exchange markets in 2001 was approximately \$1.2 trillion per day, for example.

What we suggest then, is that changes in the distribution of expectations play a key role in explaining the evolution of prices. Yet, in the standard "representative agent" model there is no place for any heterogeneity of expectations and many authors have suggested that this is the reason for the poor predictive power of such models. This argument has been made in particular for the foreign exchange market and evidence for this is given by Meese and Rogoff (1983), Frankel and Rose (1995) and Cheung, Chinn, and Pascual (2002).

Once we accept the idea that expectations can be heterogeneous, it seems that we have to move away from the standard "rational expectations" framework. One idea, pioneered by Long et al. (1989, 1990), is simply to introduce agents who systematically have "wrong" expectations but who may survive nevertheless. Kurz (1994, 1997) introduced the notion of *rational belief equilibria* where agents' beliefs about the probabilistic structure of some stock prices may be wrong as long as they are compatible with the stationary distribution suggested by the available data. In his model agents can "learn" this distribution in the sense that empirical averages converge to their expected values under the stationary measure.

Another alternative is to introduce dispersed information into the model. One approach suggested by Townsend (1983) is to have symmetrically dispersed information and to analyse the consequences of "higher order expectations," i.e., expectations about others' expectations. The idea here is that a small amount of non-fundamental trade may generate considerable volatility since traders perceive movements in asset prices as conveying information about future values of fundamentals; see Allen, Morris, and Shin (2003). Again, despite the more sophisticated reasoning attributed to agents, a certain degree of irrational behaviour is needed to generate the results. An alternative explanation of the dynamics of asset market prices is that based on the idea of a dependence on market organization. Bottazi et al. (2004) analyse the consequences of different types of organization, for example Walrasian or "order book," for the resulting price dynamics. We will remain within the Walrasian temporary equilibrium framework although it would certainly be interesting to anlayse other market forms.

The aim of this paper is to analyse a model in which

- expectations can be heterogeneous, consistent with empirical observations,
- where agents learn from their experience of using different rules, how to form their expectations,
- where agents are not systematically wrong,
- and where significant departures from fundamentals can occur but where there is always a return to fundamentals.

However, this return should not be, and is not, in our model, permanent. If agents learn about the true fundamentals and the prices then adhere to those fundamentals as in Bacchetta and van Wincoop (2003), bubble-like departures would cease to occur. In our model such bubbles will continue to develop and to burst. This gives a more realistic meaning to the widely held view, that "in the long run fundamentals matter;" see Mark (1995). Prices will inevitably return to those implied by the fundamentals but will just as inevitably leave them again. This is, of course consistent with the fact, that on average, in the short run fundamentals are poor predictors of prices.

But these reflections leave us with a puzzle. Is there any appropriate notion of equilibrium in such markets? If there is, it cannot be thought of as a deterministic state around which there are fluctuations. Nor can it be thought of as a single price path towards which the market price will converge after some time. Instead, we use a probabilistic notion of equilibrium where the distribution of state variables does not change through time. If the equilibrium is unique, then the dynamics is even ergodic. This means that time averages converge to their expected values under the unique stationary distribution.

A number of models have been constructed in which agents change their views as time passes and in some of these the changes can be self reinforcing. A particularly relevant contribution is that of Brock and Hommes (1997) in which agents can use one of two predictors. One of them is costly but when all agents use it the process is stable. The other is cheaper but when used by many individuals induces unstable behaviour of the price process. Thus their model has periods of stability interspersed with bubble like behaviour. They present analytical results to characterise this possibly chaotic behaviour of their system. Brock et al. (2002) study the evolution of a model with many types of agents. These authors, like us, focus on the limit distribution of an evolutionary process. In their case the limit distribution is over types of agents whereas our limit distribution is over prices and the performance of predictors. They have as their fundamental benchmark the rational expectations equilibrium whereas we have in mind the idea that there is an equilibrium value of some exogenous economic variable that plays this role.

The papers by Alós-Ferrer and Ania (2004), Hens and Schenk-Hoppé (2003), and Lux and Schornstein (2004) are also related to our approach. Alós-Ferrer and Ania (2004) analyse a game in which players have to decide how to allocate their portfolios among assets with a given matrix of returns. There is a unique Nash equilibrium in pure strategies in which agents allocate their wealth across assets proportional to their expected returns. It is shown that there can be prolonged departures from this equilibrium when a strategy which involves a wrong view of the world has a "lucky run." Nevertheless the system always comes back after sufficient time to the fundamentals. However, their system does not exhibit the self reinforcing properties of herding that characterise our model. Hens and Schenk-Hoppé (2003) analyse portfolio selection rules in incomplete financial markets where the dynamics of investors' wealth shares can be described by a random dynamical system, i.e., by a difference equation in a random environment which is given exogenously and thus does not share the self reinforcing contagion properties of our model. Lux and Schornstein (2004) present a model of a two country foreign exchange market which shares some of the features of our framework. These authors also use a learning rule to study the evolution of different types of behaviour. In their model agents learn by using genetic algorithms. We justify the changes in the proportions by varying the probability of agents choosing a rule as a function of the efficiency of the rule in the past.

There are many other models in which agents have different forecasting rules, and these are often summarized as "chartist" and "fundamentalist" views, following on from the seminal article of Frankel and Froot (1986). For instance, Böhm and Wenzelburger (2002) consider the performance of efficient portfolios in a setting where current prices depend on the heterogenous expectations about future prices. Many such models, when simulated, generate time paths of prices which switch from one expectations regime to another and which generate "realistic" time series, but no analytical characterization of their properties is available.

What we will show is that, despite the fact that the balance between the numbers of agents holding different expectations is constantly changing we can nevertheless characterise the structure of this process in the long term. Thus we do not have to rely on simulations to understand the nature of the process and we do not simply have to argue as some have done that, in such situations, "anything can happen".

Our model involves agents who may use one of a number of predictors which they might obtain from financial "gurus", for example, to forecast future prices of an asset. They learn in the following sense. They evaluate the performance of the gurus over time and their probability of choosing a given guru will be related to the latter's success. There is a substantial literature on such adaptive learning rules which are associated with "discrete choice" models; see, e.g., Anderson et al. (1992). Different forms of these rules have been used and accounts are given in Weisbuch, Chenevez, Kirman, and Nadal (1998) and Brock and Durlauf (2001a, 2001b). In addition, there is a self reinforcing aspect of the model which leads to agents concentrating on one of the gurus for a period. Nevertheless, after some time the guru in question may fall out of fashion and another will replace him. It is clear that the prices will be modified as the distribution of expectations changes. It is also clear that if each guru is characterised by a forecast, then the price will tend to return to a particular fundamental, that the prices will shift from being near one of these fundamentals to another. In such a case it is reasonable to expect the price process to be ergodic, and in fact this will follows as a special case from our main result. An appropriate equilibrium notion would be the limit distribution of this process, and one could think of the average time spent in each subset of the price space as characterizing the structure of the process.

However, once we introduce chartists who extrapolate from previous prices there is no a priori reason that the process should not explode. Our main result shows that, if we limit the probability that individuals will follow chartist rules appropriately, we can still show that our price process is ergodic. Thus, the process has a unique stationary distribution, and time averages converge to their expected value under the stationary measure. In our model prices will not stay in some compact set. They will blow up every once in a while but will always return to the underlying fundamentals. Thus, chartists will be temporarily destabilizing but their influence will not take over the market for ever. The presence of the chartists is revealed by the nature of the limit distribution and this is illustrated in figures 1 and 2. In the first we show the approximately normal distribution which results when there are no chartists and, in the second, we see evidence of fatter tails.

The structure of the paper is as follows. In Section 2 we introduce our discrete-time model for the microstructure of financial price fluctuations. The logarithmic stock price in period t will be determined by the market clearing condition of zero total excess demand. Individual demand involves liquidity demand and an assessment of what an adequate price in the following period should be. The individual assessments are based on the recommendations of the gurus. Assuming a log-linear structure for the demand, we obtain a class of log-linear stock price processes in a random environment. The random environment is generated by an underlying stochastic process describing the agents' liquidity demand and the agents' and the gurus' profiles, and by a stochastic process which describes the agents' individual choices of an expert. The latter process is assumed to be *endogenous*, in contrast to Föllmer and Schweizer (1993) and Horst (2004) where the agents' choices are exogenous, without any feedback from past prices. In our setting agents evaluate the gurus' performance before making their investment decision. An agent's propensity to follow a specific guru depends on the guru's current performance level. Performances are measured by a weighted sum of past profits which would have been generated by the gurus' recommendations. This dependence of the agents' choices on the gurus' performances introduces a feedback from past asset prices into the random environment.

Mathematically, the logarithmic price process can be viewed as one component of a Markov chain $\{\xi_t\}$ associated to an iterated function system. Iterated function systems have been extensively analyzed in the mathematical literature (Barnsley and Elton (1988), Steinsaltz (1999)), and in connection with learning models (Norman (1972), Karlin (1953)). Under a mean contraction condition Barnsley, Demko, Elton, and Geronimo (1988) and Norman (1972) give sufficient conditions for existence and uniqueness of stationary distributions for the Markov chain $\{\xi_t\}$. This mean contraction property, however, is not satisfied in our case because a guru's profit depends in a highly non-linear manner on stock prices. The key to our existence result is a partial average contraction property which only holds for the price component, but not for the entire Markov chain. We show that the price-performance process $\{\xi_t\}$ has a stationary distribution if we limit the probability that the agents adopt a trend chasing strategy. The proof of ergodicity turns out to be more involved. The key observation is that we can measure the experts' performance on a different scale without altering the dynamics of stock prices. This allows us to apply a comparison argument: we compare our price-performance process with some reference process whose convergence can be established using standard methods from the theory of iterated function systems. The results are stated in Section 3 and the proofs are given in Section 4.

2 A temporary equilibrium model for asset prices

We consider a financial market model with a finite set $\mathbb{A}$ of economic agents trading a single risky asset. The price evolution of the asset will be described by a sequence of temporary price equilibria. In reaction to a proposed price p in period t the agent $a \in \mathbb{A}$ forms a random excess demand $e_t^a(p,\omega)$. The actual asset price $P_t(\omega)$ will be determined by the market clearing condition of zero total excess demand. Thus, we obtain a random sequence $\{P_t(\omega)\}_{t\in\mathbb{N}}$ of temporary price equilibria on an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ defined by the implicit equation

$$\sum_{a\in\mathbb{A}} e_t^a(P_t(\omega),\omega) = 0. \tag{1}$$

Individual excess demand will involve some exogenous liquidity demand and an endogenous amount obtained by comparing, on a logarithmic scale, the proposed price with some reference level. As explained below, an agent will use a reference level recommended by some expert, and the choice of an expert will depend on the expert's performance.

2.1 Reference levels

We take a pragmatic approach to specify the demand. As in, e.g., Garman (1976) or Lux (1995), we do not formulate an individual optimization problem under budget constraints for the agents. Instead, we start right away with individual demand functions. We assume that the excess demand of the agent $a \in \mathbb{A}$ in period t is the sum of two components. The first component, denoted by $\eta_t^a(\omega)$, is an exogeneous random liquidity demand. The second component depends on the deviation of the proposed price p from some benchmark, or reference level, which the agent adopts for period t. In this paper, we consider the simplest case where the dependence is linear on a logarithmic scale. Thus, we assume that the individual excess demand of the agent $a \in \mathbb{A}$ takes the log-linear form

$$e_t^a(p,\omega) := c_t^a(\omega)\left(\hat{S}_t^a(\omega) - \log p\right) + \eta_t^a(\omega) \tag{2}$$

with nonnegative coefficients $c_t^a(\omega)$, where $\hat{S}_t^a$ denotes the logarithmic *reference level* chosen for period t. We assume that

$$c_t(\omega) := \sum_{a \in \mathbb{A}} c_t^a(\omega) > 0,$$

and define

$$\eta_t(\omega) := \frac{1}{c_t(\omega)} \sum_{a \in \mathbb{A}} \eta_t^a(\omega).$$

The logarithmic equilibrium price

$$S_t(\omega) := \log P_t(\omega)$$

is then determined via (1) and (2) as an aggregate,

$$S_t(\omega) = \frac{1}{c_t(\omega)} \sum_{a \in \mathbb{A}} c_t^a(\omega) \hat{S}_t^a(\omega) + \eta_t(\omega), \tag{3}$$

of individual price assessments and liquidity demands.

As in Kirman (1998) the choice of the reference level will be based on the recommendations R_t^i of some financial experts or *gurus* $i \in I := \{1, 2, \ldots, m\}$. More precisely, we assume that the choice of agent $a \in \mathbb{A}$ in a given period t consists in following one of the gurus, i.e.,

$$\hat{S}_t^a(\omega) \in \{R_t^1(\omega), \ldots, R_t^m(\omega)\}.$$

Later on, we will specify how the successive choices are generated by a probabilistic learning scheme based on the gurus' performance.

We denote by $x_t^a \in I$ the *individual choice* of the agent $a \in \mathbb{A}$ at date $t \in \mathbb{N}$. Thus, $x_t^a = i$ means that $\hat{S}_t^a = R_t^i$. Given the current configuration of individual states $x_t := (x_t^a)_{a \in \mathbb{A}}$, the associated weighted *empirical distribution* $\pi_t := (\pi_t^i)_{i \in I}$ is defined by

$$\pi_t^i := \frac{1}{c_t(\omega)} \sum_{a \in \mathbb{A}} c_t^a(\omega) \mathbf{1}_{\{i\}}(x_t^a) \qquad (i \in I). \tag{4}$$

In terms of these empirical distributions, the logarithmic equilibrium price in period $t+1$ defined via (3) takes the form

$$S_t(\omega) = \sum_{i \in I} \pi_t^i R_t^i(\omega) + \eta_t(\omega). \tag{5}$$

Let us now be more specific about the experts' recommendations R_t^i. We associate to each guru $i \in I$ his subjective perception $F_t^i(\omega)$ of the underlying fundamental value. The expert's recommendation is based on his expectation that, on a logarithmic scale, the asset price will move towards a weighted average of his individual assessment of the fundamental value and an observed past price trend. We focus on the simplest case where the past price trend is represented by the logarithmic price increment. Thus, the reference level suggested by expert i for the following period t is given by

$$R_t^i(\omega) := S_{t-1}(\omega) + \alpha_t^i(\omega)[F_t^i(\omega) - S_{t-1}(\omega)] + \beta_t^i(\omega)[S_{t-1}(\omega) - S_{t-2}(\omega)]. \tag{6}$$

The random coefficients α_t^i and β_t^i may be viewed as the gurus' current estimates of the speed of price adjustments. These quantities are not necessarily non-negative. If, for example, $\beta_t^i < 0$, then guru i's recommendation incorporates an "anti-cyclic" component as in, e.g., Brock and Hommes (1997).

Remark 2.1 *An example that fits into our framework is that developed in Kirman, Ricciotti, and Topol (2003). In this case there are two countries and the individuals must decide as to how to split their portfolio between safe domestic assets and foreign assets which carry the exchange rate risk. There is a fundamentalist guru in each country each of whom forecasts the exchange rate with some fundamental but each has noise in its forecast. In addition there is a chartist guru who extrapolates. This generates an exchange rate series with periods of transient behavior as observed in real exchange data; see, e.g., Bartolini and Giorgianni (1999).*

The following example shows how the specific model in Kirman (1998) fits into our framework.

Example 2.2 *Consider a situation where three experts are active on the financial market. The gurus $i = 1, 2$ adopt a purely fundamentalist point of view with constant level F^i. This means that their recommendations are solely based on the expectation that the logarithmic asset price will move closer to their assessment of the stocks' fundamental value. Thus, $\beta_t^i = 0$ $(i = 1, 2)$, and so*

$$R_t^i := S_{t-1} + \alpha_t^i[F^i - S_{t-1}] \qquad (i = 0, 1). \tag{7}$$

The third guru is described as a trend chaser *or* chartist. *Thus $\alpha_t^3 = 0$, and*

$$R_t^3 := S_{t-1} + \beta_t^3[S_{t-1} - S_{t-2}]. \tag{8}$$

In our general framework, we introduce the aggregate quantities

$$\overline{\alpha}_t := \sum_{i \in I} \alpha_t^i \pi_t^i, \quad \overline{\beta}_t := \sum_{i \in I} \beta_t^i \pi_t^i \quad \text{and} \quad \overline{F}_t := \sum_{i \in I} \alpha_t^i \pi_t^i F^i. \tag{9}$$

In view of (5) and (6) the dynamics of the sequence of temporary logarithmic price equilibria is governed by a linear recursive relation of the form

$$S_t = F(S_{t-1}, S_{t-2}, \tau_{t+1}) := \left[1 - \overline{\alpha}_t + \overline{\beta}_t\right] S_{t-1} - \overline{\beta}_t S_{t-2} + \overline{F}_t + \eta_t, \tag{10}$$

in a random environment described by the sequence

$$\tau_t := (x_t, \gamma_t, \eta_t) \tag{11}$$

where $\gamma_t := ((c_t^a)_{a\in\mathbb{A}}, (\alpha_t^i, \beta_t^i)_{i\in I}, (F_t^i)_{i\in I})$. The random environment summarizes the stochastic evolution $\{x_t\}_{t\in\mathbb{N}}$ of the agents' choices, the experts' and the agents' profiles as described by the process $\{\gamma_t\}_{t\in\mathbb{N}}$, and the noise trading $\{\eta_t\}_{t\in\mathbb{N}}$.

Remark 2.3 *If all the adjustment factors α_t^i and β_t^i vanish, then we have $\hat{S}_t^a = S_{t-1}$ for all agents $a \in \mathbb{A}$, and the price dynamics on a logarithmic scale reduces to the equation*

$$S_t = S_{t-1} + \eta_t.$$

In particular, the dynamics of logarithmic prices reduces to a simple random walk model if the aggregate liquidity demand is modelled as a sequence of independent and identically distributed random variables. This is nothing else than the discretized version of the Black-Scholes model.

Example 2.4 *Let us return to the situation described in Example 2.2 assuming for simplicity that $F^1 = 0$ and $F^2 = 1$. If only the two fundamentalist gurus are active on the market, then the resulting fluctuation of logarithmic asset prices is described by the linear equation*

$$S_t = [1 - \overline{\alpha}_t] S_{t-1} + \overline{F}_t + \eta_t \quad \textit{with} \quad \overline{F}_t := \pi_t^2 \alpha_t^2. \tag{12}$$

In this case the sequence of temporary price equilibria may be viewed as an Ornstein-Uhlenbeck process in a random environment. If $\alpha_t^i \in (0,1)$, then $0 < 1 - \overline{\alpha}_t < 1$, and stock prices behave in a mean-reverting manner. In the presence of chartists then the sequence $\{S_t\}_{t\in\mathbb{N}}$ follows the recursive dynamics

$$S_t = \left[1 - \overline{\alpha}_t + \overline{\beta}_t\right] S_{t-1} - \overline{\beta}_t S_{t-2} + \overline{F}_t + \eta_t. \tag{13}$$

Thus, the asset price process behaves in a mean-reverting manner in periods when the majority of the agents adopts a fundamentalist benchmark, i.e., whenever $\overline{\beta}_t < \overline{\alpha}_t$. In periods where the impact of trend chasers becomes so strong that the multiplier of S_t is larger than 1, i.e., if $\overline{\beta}_t > \overline{\alpha}_t$, then stock prices may behave in a transient manner.

As illustrated in the preceding example, the price process defined through (10) behaves in a mean-reverting manner in periods where the impact of technical trading on the formation of price equilibria is weak enough. In periods with a distinct trend chasing effect, however, stock prices may become transient. Economically this feature can be viewed as the temporary occurrence of bubbles or crashes generated by trend chasing. Nevertheless we shall see that the overall behavior of the price process will be ergodic if the destabilizing effects resulting from the presence of chartists are on average not too strong.

Remark 2.5 *The specific assumptions of a log-linear structure of the demand functions in (2) and the gurus' recommendations in (6) were made primarily for analytical convenience. Our convergence result for the resulting log-linear price dynamics will be based on a general mean contraction argument, and this argument carries over to more general asset price processes induced by a more general class of demand functions and reference levels.*

2.2 Performance measures

In our financial market model, asset prices are induced by an underlying stochastic process $\{\eta_t\}_{t\in\mathbb{N}}$ describing the agents' liquidity demand, by the process $\{x_t\}_{t\in\mathbb{N}}$ describing the agents' individual choices of an expert, and by the experts' and the agents' profiles as described by the process $\{\gamma_t\}_{t\in\mathbb{N}}$. The process $\{\gamma_t\}_{t\in\mathbb{N}}$ is assumed to be *exogenous*. If, in addition, the process $\{x_t\}_{t\in\mathbb{N}}$ is modelled as an *exogenous* stochastic processes, then we are in the situation of Föllmer and Schweizer (1993); see also Föllmer (1994) or Horst (2004). In this case the dynamics of the logarithmic stock price process (10) can be described by an affine equation in an *exogenous* random environment. Such equations have been extensively investigated in the mathematical literature; see, e.g., Brandt (1986), Kesten (1973), Vervaat (1979) or Horst (2001). An analysis of the diffusion limit can be found in Föllmer and Schweizer (1993) and Horst (2004). From an economic point of view, however, it is desirable to study situations where the process of individual choices is generated *endogenously*. In this section we specify a probabilistic learning mechanism where the agents' decisions are based on the perceived performance of the financial experts.

A guru's performance index is defined as a weighted average of the profits his recommendations would have generated in the past. To make this more precise, we introduce the one-period profits obtained by multiplying the price increment with that part of the excess demand which is based on the guru's recommendation. Specifically, the profit of guru i's trading strategy generated in period t is given by

$$\begin{aligned} P^i_t &:= (R^i_t - S_{t-1})(e^{S_t} - e^{S_{t-1}}) \\ &= \left[\alpha^i_t(F^i_t - S_{t-1}) + \beta^i_t(S_{t-1} - S_{t-2})\right]\left(e^{S_t} - e^{S_{t-1}}\right) \\ &=: P^i(S_{t-1}, S_{t-2}, \tau_t); \end{aligned} \tag{14}$$

note that the notation makes sense in view of (10). The sequence $\{P_t\}_{t\in\mathbb{N}}$ will be called the *profit process* where $P_t := (P^i_t)_{i\in I}$. An agent's propensity to choose the benchmark R^i_t will depend on the discounted sums of past profits associated with the experts' recommendations.

Definition 2.6 *Let $0 < \alpha < 1$ be a discount factor. The process $\{U^i_t\}_{t\in\mathbb{N}}$ defined by $U^i_0 = 0$ and*

$$U^i_t = \alpha U^i_{t-1} + (1-\alpha)P^i_t = (1-\alpha)\sum_{j=0}^{t} \alpha^{t-j} P^i_j \tag{15}$$

is called the performance measure *associated with the expert $i \in I$. The stochastic process $\{U_t\}_{t\in\mathbb{N}} = \{(U^i_t)_{i\in I}\}_{t\in\mathbb{N}}$ is called the* performance process*.*

Remark 2.7 *Alternatively, the performance an agent $a \in \mathbb{A}$ associates with a guru $i \in I$ could be based on his individual experience with the expert i. Such a situation is captured by a performance index of the form*

$$U^{a,i}_t = \alpha U^{a,i}_{t-1} + (1-\alpha)P^i_t \boldsymbol{1}_{\{i\}}(x^a_t). \tag{16}$$

In the sequel we focus on performance measures of the form (15). However, all our results carry over to performance measures of the type (16).

We assume that an agents' propensity to follow the recommendation of the guru $i \in I$ depends on the current performance vector $U_t = (U^i_t)_{i\in I}$. The conditional probability

$$Q(U_t; x_{t+1}) \tag{17}$$

of the new configuration of individual states $(x^a_{t+1})_{a\in\mathbb{A}}$ given the actual performance U_t is described by a stochastic kernel Q from $\mathbb{R}^m$ to $I^{\mathbb{A}}$. We denote the conditional probability that the agent $a \in \mathbb{A}$ adopts the advice of the guru $i \in I$ in the following period $t+1$ by

$$q^{a,i}(U_t) \in (0,1) \quad \text{where} \quad \sum_{i\in I} q^{a,i} \equiv 1.$$

In a situation where the agents adopt the advice of the guru $i \in I$ independently of the choices of the other agents, the kernel Q is determined by the agents' choice functions as the following example illustrates. In this case the agents' choice function have an interactive structure, but the transition to a new configuration itself is made independently by different agents. Our framework is flexible enough to allow for an additional correlation in the transition probabilities.

Example 2.8 *Suppose that the agent $a \in \mathbb{A}$ takes his decision independently of the choices of the other agents.In this case the distribution $Q(U_t; \cdot)$ takes the product from*

$$Q(u; \cdot) = \prod_{a \in \mathbb{A}} Q^a(u; \cdot), \tag{18}$$

where $Q^a(u; \cdot)$ denotes the probability measure $(q^{a,i}(u))_{i \in I}$ on the set $I = \{1, \dots, m\}$. The agents' choice functions $q^{a,i}$ could, for instance, take the logit form

$$q^{a,i}(u) = \frac{e^{u^i}}{e^{u^1} + \dots + e^{u^m}}; \tag{19}$$

see Weisbuch et al. (1998) for a systematic discussion of such choice functions.

2.3 The associated Markov chain

We view the logarithmic price process $\{S_t\}_{t \in \mathbb{N}}$ as the first component of the Markov chain

$$\xi_t := (S_t, S_{t-1}, U_t)$$

with state space $E := \mathbb{R}^{m+2}$. The dynamics of the process $\{\xi_t\}_{t \in \mathbb{N}}$ can be described by a recursive relation of the form

$$\xi_{t+1} = V(\xi_t, \tau_{t+1}) \tag{20}$$

in the random environment $\{\tau_t\}_{t \in \mathbb{N}}$ defined by (10) where the map V is given by

$$V(\xi_t, \tau_{t+1}) := \begin{bmatrix} F(S_t, S_{t-1}, \tau_{t+1}) \\ S_t \\ \alpha U_t + P(S_t, S_{t-1}, \tau_{t+1}). \end{bmatrix}$$

In order to obtain the Markov property of the process $\{\xi_t\}_{t \in \mathbb{N}}$, we introduce the following assumption on the exogeneous part of the random environment.

Assumption 2.9 *The random variables (γ_t, η_t) $(t \in \mathbb{N})$ are integrable and identically distributed with law ν. Moreover, $(\gamma_{t+1}, \eta_{t+1})$ is independent of $\xi_0, \dots, \xi_t$ and x_{t+1}.*

This assumption implies that, given the history up to time t and in particular the present state $\xi_t = (S_t, S_{t-1}, U_t)$, the random variable $\tau_{t+1} = (x_{t+1}, \gamma_{t+1}, \eta_{t+1})$ has the distribution

$$Z(\xi_t; \cdot) := Q(U_t; \cdot) \otimes \nu. \tag{21}$$

Thus, $\{\xi_t\}_{t\in\mathbb{N}}$ can be viewed as a Markov chain whose transition operator Π acts on bounded measurable functions f according to

$$\Pi f(\xi) = \int f \circ V(\xi, \tau) Z(\xi, d\tau). \tag{22}$$

In other words, the transition probability $\Pi(\xi; \cdot)$ is defined as the distribution of $V(\xi, \tau)$ where τ is a random vector with distribution $Z(\xi; \cdot)$.

Assumption 2.10 *The stochastic kernel Q is Lipschitz continuous in the sense that*

$$\sup_x |Q(u_1; x) - Q(u_2; x)| \leq \widehat{L} \|u_1 - u_2\| \tag{23}$$

for some $\widehat{L} < \infty$.

It is easy to check that the previous condition holds if, for instance, the agents adopt the advices of the experts independently of each other and if their choice functions $q^{a,i}$ are Lipschitz continuous. As will be shown in Lemma 4.2, this Lipschitz condition for Q guarantees that the Markov chain $\{\xi_t\}_{t\in\mathbb{N}}$ has the Feller property. This means that the function Πf is continuous for any bounded continuous f. In other words, the measure $\Pi(\xi; \cdot)$ is continuous in ξ with respect to the weak topology for probability measures on E.

For an initial state $\xi_0 = \xi$, we denote by $\mathbb{P}_\xi$ the corresponding probability measure on the canonical path space for the Markov chain $\{\xi_t\}_{t\in\mathbb{N}}$.

3 Ergodicity of the Price-Performance Process

Our aim is to establish existence and uniqueness of a stationary distribution of the Markov chain $\{\xi_t\}_{t\in\mathbb{N}}$. A probability measure μ is a *stationary distribution* for the price-performance process if $\mu\Pi = \mu$. In this case the sequence $\{\xi_t\}_{t\in\mathbb{N}}$ is stationary under the law

$$\mathbb{P}_\mu(\cdot) := \int_E \mathbb{P}_\xi(\cdot)\mu(d\xi).$$

In Section 3.1 we show the existence of a stationary distribution for the process $\{\xi_t\}_{t\in\mathbb{N}}$ in situations where the impact of trend chasing on the formation of equilibrium prices is not too strong. Under additional conditions on the experts' profiles we prove in Section

3.2 that this measure is unique, and this implies that the Markov chain $\{\xi_t\}_{t\in\mathbb{N}}$ is *ergodic*; see, e.g., Breiman (1968), Chapter 7.2. Ergodicity means that time averages converge to expectations under the unique stationary distribution μ, i.e.,

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1} f(\xi_t) = \int f(\xi)\mu(d\xi) \qquad \mathbb{P}_\mu\text{-a.s.} \tag{24}$$

and hence $\mathbb{P}_\xi$-a.s. for μ-a.e. initial state $\xi \in E$. In fact, we will obtain convergence for *every* initial state ξ and every bounded continuous function $f : E \to \mathbb{R}$.

3.1 Existence of Stationary Distributions

In order to prove the existence of an invariant distribution for the Markov chain $\{\xi_t\}_{t\in\mathbb{N}}$, we first introduce a condition which limits the impact of trend chasing. For a given performance level u, the expected impact of fundamentalist and trend chasing behavior, respectively, is measured by the quantities

$$\alpha(u) := \mathbb{E}[\overline{\alpha}_{t+1}|U_t = u] = \sum_{i\in I}\sum_{a\in\mathbb{A}} \alpha^{a,i} q^{a,i}(u)$$

and

$$\beta(u) := \mathbb{E}[\overline{\beta}_{t+1}|U_t = u] = \sum_{i\in I}\sum_{a\in\mathbb{A}} \beta^{a,i} q^{a,i}(u)$$

where

$$\alpha^{a,i} := \mathbb{E}[\alpha_t^i c_t^{-1} c_t^a] \quad \text{and} \quad \beta^{a,i} := \mathbb{E}[\beta_t^i c_t^{-1} c_t^a]$$

Note that the coefficients $\alpha^{a,i}$ and $\beta^{a,i}$ only depend on the distribution of γ and not on the time t, and that, due to Assumption 2.9

$$\mathbb{E}[\overline{\alpha}_{t+1}|\xi_0,\ldots,\xi_t] = \alpha(U_t) \quad \text{and} \quad \mathbb{E}[\overline{\beta}_{t+1}|\xi_0,\ldots,\xi_t] = \beta(U_t).$$

In order to guarantee the existence of a stationary distribution for the price-performance process, we need to put a quantitative bound on the impact of trend chasers on the formation of equilibrium prices.

Assumption 3.1

$$\sup_u |1-\alpha(u)+\beta(u)| + \sup_u |\beta(u)| < 1 \tag{25}$$

Remark 3.2 *Suppose that the experts' profiles $(\alpha_t^i, \beta_t^i)_{i\in I}$ are independent of the agents' profiles $(c_t^a)_{a\in\mathbb{A}}$. Then*

$$\alpha(u) = \sum_{i\in I}\sum_{a\in\mathbb{A}} \mathbb{E}[\alpha_t^i]\mathbb{E}[c_t^{-1}c_t^a] q^{a,i}(u) \geq \min_i \mathbb{E}[\alpha_t^i]$$

since $\sum_{i \in I} q^{a,i}(u) = 1$ *and* $\sum_{a \in \mathbb{A}} \mathbb{E}[c_t^{-1} c_t^a] = 1$. *This shows that condition (25) is always satisfied if the agents only adopt fundamentalist beliefs, i.e., if*

$$\beta_t^i = 0 \quad \textit{and} \quad \mathbb{E}[\alpha_t^i] > 0 \quad \textit{for all } i \in I.$$

The following example illustrates what we mean by saying that Assumption (25) puts a quantitative bound on the impact of chartists.

Example 3.3 *Let us return to the situation described in Example 2.2. Suppose that* $c_1^a \equiv 1$, *that* $\mathbb{E}\alpha_t^i = \alpha$ *for* $i = 1, 2$ *and* $\mathbb{E}\beta_t^3 = \beta$ *and define* $q := \sup_{a,u} q^{a,3}(u)$. *Then (25) reduces to*

$$2\beta q < \alpha(1 - q).$$

If, for example, $\alpha = \frac{3}{4}$ *and* $\beta = \frac{1}{2}$, *this condition is satisfied for*

$$q < \frac{3}{7}$$

i.e., if an individual agent adopts with sufficiently high probability a fundamentalist benchmark, independently of the performance of the experts.

We are now ready to state our existence result. The proof of the following theorem will be given in Section 4 below.

Theorem 3.4 *Under Assumptions 2.9, 2.10, and 3.1, the Markov chain* $\{\xi_t\}_{t \in \mathbb{N}}$ *has an invariant distribution.*

3.2 Uniqueness of Stationary Distributions

So far, we have established the existence of stationary distributions for the price utility process $\{\xi_t\}_{t \in \mathbb{N}}$ under the assumption that the impact of trend chasing on the dynamics of asset prices is not too strong. In this section we state additional conditions which guarantee uniqueness of the stationary distribution.

Assumption 3.5 *(i) The probability measures* $Q(u; \cdot)$ *are uniformly positive in the sense that* $Q(u; x) \geq \delta$ *for some* $\delta > 0$ *and all* $u \in \mathbb{R}^m$ *and every configuration of individual states* x.

(ii) The law of the random variables η_t *is defined through a strictly positive density on* $\mathbb{R}$ *and the random variables* γ_t *are uniformly bounded.*

(iii) The measure $\Pi^2(\xi; \cdot)$ *is absolutely continuous with respect to Lebegues measure.*

(iv) For any two sequences $\{\xi_n\}_{n\in\mathbb{N}}$ and $\{\hat{\xi}_n\}_{n\in\mathbb{N}}$ converging to the same limit, the measures $\Pi^2(\xi_n;\cdot)$ and $\Pi^2(\hat{\xi}_n;\cdot)$ are not singular to each other for all sufficiently large n.

We are now going to state our main result. Its proof requires some preparation and will be carried out in Section 4.2 below.

Theorem 3.6 *Under Assumptions 2.9, 3.1 and 3.5 the Markov chain $\{\xi_t\}_{t\in\mathbb{N}}$ has a unique stationary distribution μ, and*

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1} f(\xi_t) = \int f(\xi)\mu(d\xi) \qquad \mathbb{P}_\xi\text{-a.s.}$$

for all bounded continuous functions f and each initial state $\xi \in E$.

Let us comment on our Assumption 3.5. Part (i) is satisfied if, for instance, the agents take their decisions independently of each other given the actual performance vector u and if the choice functions take the modified logit form

$$q^{a,i}(u) = (1-\hat{\delta})\frac{e^{u^i}}{e^{u^1}+\cdots+e^{u^m}} + \frac{\hat{\delta}}{m} \quad \text{for some} \quad \hat{\delta} > 0. \tag{26}$$

Part (iii) implies that the joint distribution of (S_t, S_{t-1}, U_t) given the past is absolutely continuous to Lebesgue measure. This would be easy if we would add independent and identically distributed "measurement errors" ϵ_t^i with absolutely continuous distributions in the definition of the performance measures, i.e, if we would consider performance processes of the form

$$U_t^i = \alpha U_{t-1}^i + (1-\alpha)P_t^i + \epsilon_t^i.$$

If, in addition, we would require that the density of ϵ_t^i is strictly positive, then $\Pi^2(\xi;\cdot)$ would even be equivalent to Lebesgue measure on E, and the proof of uniqueness would be straightforward. We prefer, however, to retain the precise definition (15) of the performance process. In such a situation, Assumption 3.5 (iii) can be justified in terms of conditions on the experts' profiles. The conditional distribution of U_t given $S_t \neq S_{t-1}$ and the past is absolutely continuous if the conditional distribution of the random vector $(R_t^i)_{i\in I}$ is absolutely continuous, given the past up to time $t-1$ if $S_{t-1} \neq S_{t-2}$. This is clearly the case under mild conditions on the distribution of $(\alpha_t^i, \beta_t^i, F_t^i)_{i\in I}$. By Assumption 3.5 (i), each empirical distribution of individual states has strictly positive probability. Hence it follows from (5) and Assumption 3.5 (ii) that the conditional distribution of $(R_t^i)_{i\in I}$, given the past up to time $t-1$ and also the logarithmic stock price S_t, is still absolutely continuous if $S_t \neq S_{t-1} \neq S_{t-2}$. Thus, (14) shows that the random vector $(U_t^i)_{t\in\mathbb{N}}$ has an absolute continuous distribution, given S_t, S_{t-1} and the past up to time $t-1$ if $S_t \neq S_{t-1} \neq S_{t-2}$. The latter condition holds almost surely, and so $\Pi^2(\xi; \cdot)$ has a density. Assumption 3.5 is satisfied if, for instance, at least some components of the experts' profiles have continuous densities.

Example 3.7 *Let us return to the situation described in Example 2.2 where the agents can choose between two fundamentalist and a chartist trading strategy, and assume that* $c_t^a \equiv 1$ *and that the random variable* $\gamma_0 = (\alpha_0^1, \alpha_0^2, \beta_0^3)$ *has a continuous density* g. *It follows from the transformation formula for Lebesgue integrals that the conditional distribution of* ξ_{t+2}, *given* ξ_t *has a density* f_{ξ_t} *and that the set* $\{f_{\xi_t} > 0\}$ *is open. By direct calculation using Assumption 3.5 (ii) and the fact that* γ_0 *has a continuous density we see that there exists* $\zeta \in \{f_{\xi_t} > 0\}$ *and* $\epsilon > 0$ *such that*

$$\zeta \in \{f_{\hat{\xi}} > 0\} \quad \text{for all} \quad \|\xi_t - \hat{\xi}\| < \epsilon.$$

Thus, for any two sequences $\{\xi_n\}_{n\in\mathbb{N}}$ *and* $\{\hat{\xi}_n\}_{n\in\mathbb{N}}$ *converging to the same limit* ξ_∞ *there exists* $\zeta \in \{f_{\xi_\infty} > 0\}$ *and* $N \in \mathbb{N}$ *such that*

$$\zeta \in \{f_{\xi_n} > 0\} \cap \{f_{\hat{\xi}_n} > 0\} \qquad \text{for all} \quad n \geq N.$$

This shows that Assumption 3.5 (iv) is satisfied because the sets $\{f_{\xi_n} > 0\}$ *and* $\{f_{\hat{\xi}_n} > 0\}$ *are open.*

Assumption 3.5 does not yet guarantee that the price-performance measure admits a unique stationary distribution. Hence additional arguments will be needed in order to establish the uniqueness result.

Remark 3.8 *Some of the assumptions in Theorems 3.4 and 3.6 were made primarily for analytical convenience and to reduce notational complexity. They can be relaxed without altering the model's qualitative features. For instance, we can allow more general patterns of trend chasing: with a suitable modification of our stability condition (25) we may as well assume that the experts' benchmarks are of the form*

$$\hat{R}_{t+1}^i := S_t + \alpha_t^i[F^i - S_t] + \sum_{j=0}^{l} \beta_t^{i,j}[S_{t-j} - S_{t-j-1}]$$

for some $l \in \mathbb{N}$. *Our results also carry over to models in which the dynamics of the experts' assessments of fundamental values are described by mean reverting random walks fluctuating around long run fundamentals* $\overline{F}^i$.

4 Proof of Existence and Uniqueness

The key to our results on existence and uniqueness of stationary distributions will be a contraction property in two steps for the two-dimensional logarithmic price process $(S_t, S_{t-1})_{t\in\mathbb{N}}$; one step is not enough, due to the deterministic dependence in the second component. Given a state $\xi_t = (S_t, S_{t-1}, U_t)$ and two successive values (τ_{t+1}, τ_{t+2}) of the environment, the resulting prices (S_{t+2}, S_{t+1}) after two iterations are specified by the map

$$W(\xi_t, \tau_{t+1}, \tau_{t+2}) := (F(\cdot, S_t, \tau_{t+2}) \circ F(S_t, S_{t-1}, \tau_{t+1}), F(S_t, S_{t-1}, \tau_{t+1}))\,.$$

Lemma 4.1 *Under Assumption 3.1 the map W is a mean contraction in the sense that*

$$\int\int \|W(\xi, \tau_1, \tau_0) - W(\hat{\xi}, \tau_1, \tau_0)\| Z(V(\xi, \tau_0); d\tau_1) Z(\xi; d\tau_0) \le r\|\xi - \hat{\xi}\| \tag{27}$$

for some $r < 1$ and all $\xi, \hat{\xi} \in E$.

Proof. Let us introduce a 2×2 matrix A_t and a 2-dimensional vector B_t by

$$A_t := \begin{pmatrix} 1 - \overline{\alpha}_t + \overline{\beta}_t & -\overline{\beta}_t \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad B_t := \begin{pmatrix} \overline{F}_t + \eta_t \\ 0 \end{pmatrix},$$

respectively, where the quantities $\overline{\alpha}_t$, $\overline{\beta}_t$, and $\overline{F}_t$ are defined in (9). With this notation,

$$W(\xi_t, \tau_t, \tau_{t-1}) = A_{t+1}A_t \begin{pmatrix} S_t \\ S_{t-1} \end{pmatrix} + A_{t+1}B_t + B_t.$$

Let us now put

$$\varrho_0 := \sup_u |1 - \alpha(u) + \beta(u)| \quad \text{and} \quad \varrho_1 := \sup_u |\beta(u)| \tag{28}$$

and

$$A := \begin{pmatrix} \varrho_0 & \varrho_1 \\ 1 & 0 \end{pmatrix}$$

Then the absolute value of each entry in the matrix $\int A_t Z(\xi_t; d\tau)$ is bounded from above by the corresponding entry of the matrix A because

$$\left|\mathbb{E}_{\xi_t}[1 - \overline{\alpha}_{t+1} + \overline{\beta}_{t+1}]\right| = |1 - \alpha(U_t) + \beta(U_t)| \le \varrho_0 \quad \text{and} \quad \left|\mathbb{E}_{\xi_t}[\overline{\beta}_t]\right| = |\beta(U_t)| \le \varrho_1.$$

The absolute value of each entry of the matrix $\int A_{t+1}A_t Z(V(\xi_t, \tau_t); d\tau_{t+1}) Z(\xi_t; d\tau_t)$ is then bounded by the respective entry of the matrix

$$A^2 = \begin{pmatrix} \varrho_0^2 + \varrho_1 & \varrho_0\varrho_1 \\ \varrho_0 & \varrho_1 \end{pmatrix}.$$

This proves our assertion because, due to Assumption 3.1 the row-sum norm of the latter matrix is strictly less than 1. $\square$

For the existence result we need the Feller property of the stochastic kernel Π and here we use our Assumption 2.10 on the stochastic kernel Q.

Lemma 4.2 *Under the assumptions of Theorem 3.4 the Markov process* $\{\xi_t\}_{t\in\mathbb{N}}$ *has the Feller property.*

PROOF. Let $\{\xi_n\}_{n\in\mathbb{N}}$ be a sequence in E such that $\lim_{n\to\infty} = \xi$. It is enough to show that

$$\lim_{n\to\infty} \Pi f(\xi_n) = \Pi f(\xi)$$

for every bounded Lipschitz continuous function $f : E \to \mathbb{R}$. To this end, we first apply the Lipschitz property (23) of the product kernel Q in order to obtain

$$\begin{aligned} |\Pi f(\xi_n) - \Pi f(\xi)| &= \left| \int f \circ V(\xi_n, \tau) Z(\xi_n, d\tau) - \int f \circ V(\xi, \tau) Z(\xi, d\tau) \right| \\ &\leq \int |f \circ V(\xi_n, \tau) - f \circ V(\xi, \tau)| \, Z(\xi, d\tau) + L\|f\|_\infty \|\xi_n - \xi\| \end{aligned}$$

where $\|\cdot\|_\infty$ denotes the usual sup-norm. Hence it suffices to show that

$$\lim_{n\to\infty} \int |f \circ V(\xi_n, \tau) - f \circ V(\xi, \tau)| \, Z(\xi, d\tau) = 0.$$

To this end, we fix $\epsilon > 0$ and a compact set K such that $Z(\xi; K^c) < \epsilon$. Denoting by $L(f)$ the Lipschitz constant of the map f, we obtain

$$\begin{aligned} \int |f \circ V(\xi_n, \tau) - f \circ V(\xi, \tau)| \, Z(\xi, d\tau) &\leq L(f) \int_K |V(\xi_n, \tau) - V(\xi, \tau)| \, Z(\xi, d\tau) + 2\epsilon \|f\|_\infty \\ &\leq L(f) \sup_{n, \tau \in K} \|V(\xi_n, \tau) - V(\xi, \tau)\| + 2\epsilon \|f\|_\infty. \end{aligned}$$

Hence the assertion follows from uniform continuity of V on compact sets. $\square$

4.1 Proof of Existence of Stationary Distributions

Let $\mathbb{P}_\xi$ denote the distribution of the Markov chain $\{\xi_t\}_{t\in\mathbb{N}}$ with initial state ξ. We are now going to show that the sequence $\{\xi_t\}_{t\in\mathbb{N}}$ is *tight* under $\mathbb{P}_\xi$, i.e., that

$$\lim_{c\to\infty} \sup_t \mathbb{P}_\xi[\|\xi_t\| \geq c] = 0.$$

This tightness property together with the Feller property of the stochastic kernel Π will imply that the process $\{\xi_t\}_{t\in\mathbb{N}}$ admits an invariant measure.

Proposition 4.3 *If the assumptions of Theorem 3.4 (i) are satisfied, then the sequence* $\{\xi_t\}_{t\in\mathbb{N}}$ *is tight under* $\mathbb{P}_\xi$.

PROOF. Let us first show that the sequence $\{S_t\}_{t\in\mathbb{N}}$ is tight, i.e., that

$$\lim_{c\to\infty}\sup_t \mathbb{P}_\xi[|S_t|\geq c]=0. \tag{29}$$

To this end, we put $\varrho := \sup_u \mathbb{E}[|\overline{F}_t||U_t=u]$. Using the notation introduced in (28) we define an almost surely non-negative stochastic process $\{\overline{S}_t\}_{t\in\mathbb{N}}$ by $\overline{S}_i=|S_i|$ $(i=0,1)$ and

$$\overline{S}_{t+1}=\sum_{i=0}^{1}\varrho_i\overline{S}_{t-i}+\varrho+|\eta_t| \qquad (t=2,3,\ldots).$$

Since $\mathbb{E}[|1-\alpha(u)+\beta(u)||S_t]\leq\varrho_0$, $\mathbb{E}[\beta(u)|S_{t-1}]\leq\varrho_1$, and $\mathbb{E}|S_i|=\mathbb{E}\overline{S}_i$ for $i=0,1$ it follows by induction that

$$\begin{aligned}
\mathbb{E}|S_{t+1}| &\leq \mathbb{E}\left[|1-\alpha(u)+\beta(u)||S_t|\right]+\mathbb{E}\left[\beta(u)|S_{t-1}|\right]+\varrho+\mathbb{E}|\eta_t| \\
&\leq \mathbb{E}\left[|S_t|\cdot\mathbb{E}[|1-\alpha(u)+\beta(u)||S_t]\right]+\mathbb{E}\left[|S_{t-1}|\cdot\mathbb{E}[|\beta(u)||S_{t-1}]\right]+\varrho+\mathbb{E}|\eta_t| \\
&\leq \varrho_0\mathbb{E}|S_t|+\varrho_1\mathbb{E}|S_{t-1}|+\varrho+\mathbb{E}|\eta_t| \\
&\leq \mathbb{E}\overline{S}_{t+1}
\end{aligned}$$

for all $t\in\mathbb{N}$. Thus, (29) follows from Tchebychev's inequality if we can show that

$$\sup_t \mathbb{E}\overline{S}_t<\infty. \tag{30}$$

In order to prove (30), we put $\overline{\zeta}_t=(\overline{S}_t,\overline{S}_{t-1})$, and introduce a sequence $\{B_t\}_{t\in\mathbb{N}}$ of independent and identically distributed 2-dimensional random vectors and a 2×2 matrix A by

$$B_t:=\begin{pmatrix}\varrho+|\eta_t|\\0\end{pmatrix}\quad\text{and}\quad A:=\begin{pmatrix}\varrho_0&\varrho_1\\1&0\end{pmatrix},$$

respectively. For any initial price vectors $x=(S_1,S_0)$ and $y=(\hat{S}_1,\hat{S}_0)$, the $\mathbb{R}^2$-valued stochastic sequences $\{\overline{\zeta}^x_t\}_{t\in\mathbb{N}}$ and $\{\overline{\zeta}^y_{2t}\}_{t\in\mathbb{N}}$ starting in x and y, respectively, satisfy almost surely the recursive relations

$$\overline{\zeta}^x_{t+1}=A\overline{\zeta}_t+B_t\quad\text{and}\quad\overline{\zeta}^y_{2t}=A^2\overline{\zeta}_{2t-2}+AB_{2t-2}+B_{2t-1}\qquad(t\in\mathbb{N}).$$

In particular, we have the explicit representation

$$\overline{\zeta}^x_{t+1}=\sum_{j=0}^{t-1}\left(\prod_{i=t-j}^{t-1}A\right)B_{t-j-1}+A^{t-1}x. \tag{31}$$

By Assumption 3.1, the row-sum norm of the matrix A^2 are strictly smaller than one. Moreover, due to Assumption 2.9 the random variables $AB_{2t-2} + B_{2t-1}$ $(t \in \mathbb{N})$ are independent, identically distributed and integrable. Thus, independently of the initial value $\overline{\zeta}_0 = y \in \mathbb{R}^2$, the sequence $\{\overline{\zeta}^y_{2t}\}_{t\in\mathbb{N}}$ converges in distribution and in L^1 to

$$\overline{\zeta} := \sum_{t=1}^{\infty} (AB_t + B_{t+1}) \prod_{j=1}^{t-1} A^{2j}$$

due to a fundamental convergence theorem by Kesten (1973), see also Embrechts, Klüppelberg and Mikosch (1997), Proposition 8.4.3. Now, we wish to show that the entire sequence $\{\mathbb{E}|\overline{\zeta}^x_t|\}_{t\in\mathbb{N}}$ converges to $\mathbb{E}|\overline{\zeta}|$ independently of the starting point. To this end, it suffices to show that

$$\lim_{t\to\infty} \mathbb{E}|\overline{\zeta}^y_{2t+1}| = \mathbb{E}|\overline{\zeta}|$$

This, however, follows from the convergence of the processes $\{\overline{\zeta}^x_{2t}\}_{t\in\mathbb{N}}$ because

$$\overline{\zeta}^y_{2t+1} = \overline{\zeta}^{Ay+B_1}_{2t} \qquad \mathbb{P}\text{-a.s.}$$

Thus,

$$\lim_{t\to\infty} \mathbb{E}|\overline{\zeta}^x_t| = \mathbb{E}|\overline{\zeta}|,$$

and this shows (29). We are now ready to prove tightness of the sequence $\{\xi_t\}_{t\in\mathbb{N}}$. Due to (29) and because of Theorem 3.7.2 in Ethier and Kurtz (1986), there exists, for any $\epsilon > 0$, a compact set $K \subset \mathbb{R}^\infty$ such that

$$\mathbb{P}_\xi[\{S_t\}_{t\in\mathbb{N}} \in K] \geq 1 - \epsilon. \tag{32}$$

Since the random variables $(\alpha^i_t, \beta^i_t, F^i_t)$ $(t \in \mathbb{N})$ is independent of the price process and because P is continuous, it follows from (32) and from the compact containment condition (Ethier and Kurtz (1986), Remark 3.7.3) that there exists a compact set $\widetilde{K}$ such that

$$\mathbb{P}[\{(S_t, S_{t-1}, U_t)\}_{t\in\mathbb{N}} \in \widetilde{K}] \geq 1 - \epsilon.$$

This proves our assertion. □

We are now ready to prove existence of a stationary distribution for the Markov chain $\{\xi_t\}_{t\in\mathbb{N}}$.

Proof of Theorem 3.4: Since the sequence of probability measures $\{\Pi^t(\xi; \cdot)\}_{t\in\mathbb{N}}$ is tight for any starting point $\xi \in E$, the sequence $\{\nu^\xi_t\}_{t\in\mathbb{N}}$ defined by

$$\nu^\xi_t := \frac{1}{t} \sum_{j=1}^{t} \Pi^j(\xi; \cdot) \tag{33}$$

is tight, too. Hence, it follows from Prohorov's theorem (Shiryaev (1989), Theorem 3.2.1) that the sequence $\{\nu_t^\xi\}_{t\in\mathbb{N}}$ has an accumulation point ν with respect to the topology of weak convergence for probability measures on E. Since the stochastic kernel Π has the Feller property, it follows from standard arguments that ν is a stationary distribution for our Markov chain; see, e.g. Meyn and Tweedie (1993), Chapter 12 for details. □

4.2 Proof of Uniqueness of Stationary Distributions

In this section we establish the ergodicity result stated in Theorem 3.6. Our proof is based on a comparison argument. In a first step we use results from the theory of iterated function systems in order to show that a suitably transformed Markov chain $\{\hat{\xi}_t\}_{t\in\mathbb{N}}$ with state space

$$\hat{E} := [-1,1]^2 \times \mathbb{R}^m$$

converges in distribution to a unique stationary measure. In a second step we show that convergence of the sequence $\{\hat{\xi}_t\}_{t\in\mathbb{N}}$ implies that the price-performance process $\{\xi_t\}_{t\in\mathbb{N}}$ has a *unique* stationary distribution.

In order to describe the dynamics of the sequence $\{\hat{\xi}_t\}_{t\in\mathbb{N}}$ we fix a strictly increasing continuously differentiable *transformation function* $G : \mathbb{R} \to [-1,1]$ with continuously differentiable inverse whose derivative G' satisfies $|G'| \le 1$, and introduce a vector of modified profit functions $\hat{P}(\cdot,\tau) : \hat{E} \to \mathbb{R}^m$ by

$$\widehat{P}(S_1, S_0, \tau) := P(\cdot, S_1, \tau) \circ G \circ F(S_0, S_1, \tau)$$

where $\tau = (x, \gamma, \eta)$ and F is defined in (10). The map $\widehat{P}$ may be viewed as the experts' profit function if logarithmic stock prices are transformed by G. Since the random variables γ_t are uniformly bounded, the modified profit functions $\widehat{P}(\cdot,\tau) : [-1,1]^2 \to \mathbb{R}^{m+1}$ are *globally* Lipschitz continuous:

$$\left|\widehat{P}(S_1, S_0, \tau) - \widehat{P}(\tilde{S}_1, \tilde{S}_0, \tau)\right| \le R\|(S_1, S_0) - (\tilde{S}_1, \tilde{S}_0)\|. \tag{34}$$

The evolution of the sequence

$$\hat{\xi}_t = (\hat{S}_t, \hat{S}_{t-1}, \hat{U}_t)$$

is now described by the stochastic difference equation

$$\hat{\xi}_{t+1} = \hat{V}(\hat{\xi}_t, \tau_{t+1}) \quad \text{where} \quad \tau_{t+1} \sim Z(\hat{\xi}_t; \cdot),$$

and the map $\hat{V}$ is given by

$$\hat{V}(\hat{\xi}_t, \tau_{t+1}) := \begin{bmatrix} G \circ F(\hat{S}_t, \hat{S}_{t-1}, \tau_{t+1}) \\ \hat{S}_t \\ \alpha\hat{U}_t + (1-\alpha)\widehat{P}(\hat{S}_t, \hat{S}_{t-1}, \tau_{t+1}) \end{bmatrix}.$$

Let us now show that the transformed Markov chain $\{\hat{\xi}_t\}_{t\in\mathbb{N}}$ is ergodic.

Proposition 4.4 *The Markov process $\{\hat{\xi}_t\}_{t\in\mathbb{N}}$ converges in law to a unique stationary measure.*

As a first step towards the proof we can apply similar arguments as in the proof of Proposition 4.3 to show that the Markov chain $\{\hat{\xi}_t\}_{t\in\mathbb{N}}$ is tight, and that it has a stationary distribution $\hat{\mu}$. In order to prove uniqueness we measure an expert's performance on a different scale. More specifically, we consider a suitably "rescaled" Markov chain $\{\hat{\xi}_t^L\}_{t\in\mathbb{N}}$ which coincides in law with the Feller process $\{\hat{\xi}_t\}_{t\in\mathbb{N}}$, and whose convergence in distribution can be shown by standard arguments from the theory of iterated function systems. In order to define the $\{\hat{\xi}_t^L\}_{t\in\mathbb{N}}$ for a fixed $L > R$, we introduce the maps

$$\widehat{P}^L(\hat{S}_1, \hat{S}_0, \tau) := \frac{1}{L}\widehat{P}(\hat{S}_1, \tilde{S}_0, \tau).$$

In view of (34) these maps are *contractions* uniformly in τ. For $\theta := \frac{R}{L} < 1$ we have

$$\left|\widehat{P}^L(\hat{S}_1, \hat{S}_0, \tau) - \widehat{P}^L(\tilde{S}_1, \tilde{S}_0, \tau)\right| \leq \theta \|(\hat{S}_1, \hat{S}_0) - (\tilde{S}_1, \tilde{S}_0)\|. \tag{35}$$

Let us now put

$$\hat{\xi}_t^L = (\hat{S}_t^L, \hat{S}_{t-1}^L, \hat{U}_t^L),$$

and consider the Markov chain $\{\hat{\xi}_t^L\}_{t\in\mathbb{N}}$ on $\hat{E}$ that is defined through the non-linear recursive relation

$$\hat{\xi}_{t+1}^L = \hat{V}^L(\hat{\xi}_t^L, \tau_{t+1}) \quad \text{where} \quad \tau_{t+1} \sim \hat{Z}^L(\hat{\xi}_t^L; \cdot) := Q(L\hat{U}_t^L; \cdot) \otimes \nu(\cdot), \tag{36}$$

and the map $\hat{V}^L$ is defined by

$$\hat{V}^L(\hat{\xi}_t^L, \tau_{t+1}) := \begin{bmatrix} G \circ F(\hat{S}_t^L, \hat{S}_{t-1}^L, \tau_{t+1}) \\ \hat{S}_t \\ \alpha \hat{U}_t^L + (1-\alpha)\widehat{P}^L(\hat{S}_t^L, \hat{S}_{t-1}^L, \tau_{t+1}) \end{bmatrix}.$$

Given the initial performance levels $\frac{1}{L}u$ and u for the processes $\{\hat{U}_t^L\}_{t\in\mathbb{N}}$ and $\{\hat{U}_t\}_{t\in\mathbb{N}}$, respectively, and given the evolution of both the modified logarithmic price process $\{(\hat{S}_t^L, \hat{S}_{t-1}^L)\}_{t\in\mathbb{N}}$ and the environment $\{\tau_t\}_{t\in\mathbb{N}}$, an induction argument shows that

$$\hat{U}_{t+1}^L = \alpha \hat{U}_t^L + \frac{1-\alpha}{L}\widehat{P}(\hat{S}_t, \hat{S}_{t-1}, \tau_t) = \alpha \frac{1}{L}\hat{U}_t + \frac{1-\alpha}{L}\widehat{P}(\hat{S}_t, \hat{S}_{t-1}, \tau_t) = \frac{1}{L}\hat{U}_{t+1}.$$

As shown by the following lemma, our rescaling does not alter the distribution of the price process, if we modify the agents choice probabilities according to (36).

Lemma 4.5 *Given the initial vectors (S_1, S_0, Lu) and (S_1, S_0, u) for the processes $\{\hat{\xi}_t\}_{t\in\mathbb{N}}$ and $\{\hat{\xi}_t^L\}_{t\in\mathbb{N}}$, respectively, we have*

$$\left(\hat{S}_t, \hat{S}_{t-1}, \frac{1}{L}\hat{U}_t\right) \stackrel{\text{Law}}{=} \left(\hat{S}_t^L, \hat{S}_{t-1}^L, \hat{U}_t^L\right) \qquad (t \in \mathbb{N}). \tag{37}$$

PROOF. The assertion, is obvious for $t = 1$. Let us therefore assume that (37) holds for all $k \leq t$ and denote by $\hat{\mu}_t$ and $\hat{\mu}_t^L$ the distribution of the random vectors $\left(\hat{S}_t, \hat{S}_{t-1}, \frac{1}{L}\hat{U}_t\right)$ and $\left(\hat{S}_t^L, \hat{S}_{t-1}^L, \hat{U}_t^L\right)$, respectively. For any bounded continuous function f we obtain

$$\begin{aligned}
\int f d\hat{\mu}_{t+1}^L &= \int\int f \circ \hat{V}^L(\hat{S}_t^L, \hat{S}_{t-1}^L, \hat{U}_t^L, \tau_{t+1}) Z^L(\hat{S}_t^L, \hat{S}_{t-1}^L, \hat{U}_t^L; d\tau_{t+1}) d\hat{\mu}_t^L \\
&= \int\int f \circ \hat{V}^L(\hat{S}_t, \hat{S}_{t-1}, \frac{1}{L}\hat{U}_t, \tau_{t+1}) Z(\hat{S}_t, \hat{S}_{t-1}, \hat{U}_t; d\tau_{t+1}) d\hat{\mu}_t \\
&= \int f d\hat{\mu}_{t+1}
\end{aligned}$$

where the second equality follows from the induction hypothesis along with (36). This shows

$$(\hat{S}_{t+1}^L, \hat{S}_t^L, \hat{U}_{t+1}^L) \stackrel{\text{Law}}{=} (\hat{S}_{t+1}, \hat{S}_t, \hat{U}_{t+1}).$$

□

In order to prove convergence of the transformed Markov chain $\{\hat{\xi}_t\}_{t\in\mathbb{N}}$ it is now enough to show that the rescaled process $\{\hat{\xi}_t^L\}_{t\in\mathbb{N}}$ converges in distribution as $t \to \infty$. Thus, Proposition 4.4 is a direct consequence of the following lemma.

Lemma 4.6 *The Markov chain $\{\hat{\xi}_t^L\}_{t\in\mathbb{N}}$ converges in law to a unique limiting distribution.*

The proof uses results from the theory of iterated function systems. An *iterated function system* is of the form

$$\Sigma = (M_1, M_2, V, Z)$$

where M_1 and M_2 are measurable spaces, where $V : M_1 \times M_2 \to M_1$ is a measurable map and where Z is a stochastic kernel from M_1 to M_2. Such a system Σ induces an *associated Markov chain* $\{\xi_t\}_{t\in\mathbb{N}}$ on the state space M_1 by the recursive relation

$$\xi_{t+1} = V(\xi_t, \varrho_{t+1}) \quad \text{where} \quad \varrho_{t+1} \sim Z(\xi_t; \cdot).$$

Note that the conditional distributions of the driving stochastic process $\{\varrho_t\}_{t\in\mathbb{N}}$ are allowed to depend on the current states of the induced process $\{\xi_t\}_{t\in\mathbb{N}}$. This feedback effect is crucial for our purposes. Without it, our notion of an iterated function system would reduce to a *random dynamical system* in the sense of Arnold (1998) where "noise is something given to us and not at our disposal".

An iterated function system is called *mean contractive* if M_1 is a metric state space with metric d and if there exists $\theta < 1$ such that

$$\int d(V(\xi,\varrho), V(\hat{\xi},\varrho))Z(\xi; d\varrho) \leq \theta d(\xi,\hat{\xi}) \tag{38}$$

for all $\xi, \hat{\xi} \in M_1$. Convergence results for Markov chains associated to *mean-contractive* iterated function systems have been established under a Lipschitz condition on the transformation V and under the Lipschitz condition

$$\|Z(\xi;\cdot) - Z(\hat{\xi};\cdot)\|_{TV} \leq Rd(\xi,\hat{\xi})$$

on the stochastic kernel Z with respect to the total variation norm $\|\cdot\|_{TV}$ by Barnsley and Elton (1988), Barnsley et al. (1988) and Norman (1972) among others. The results by Barnsley et al. (1988) and Norman (1972) easily extend to systems which are mean-contractive in two steps as in (27). However, they do not apply to the iterated function system defined through (20) and (21). The first reason is that the profit functions in (14) and thus the map V in (20) are not *global* Lipschitz continuous. The profit functions are only *locally* Lipschitz continuous, i.e., Lipschitz continuous on compact sets. Secondly, Barnsley et al. (1988) study iterated function systems where M_2 is a finite set, while Norman (1972) allows for an uncountable set M_2 but assumes that the state space M_1 of the associated Markov chain is bounded. Neither condition is satisfied in our case. This is why we first analyze the modified chain $\{\hat{\xi}_t\}_{t\in\mathbb{N}}$, and then compare $\{\xi_t\}_{t\in\mathbb{N}}$ to $\{\hat{\xi}_t\}_{t\in\mathbb{N}}$.

Proof of Lemma 4.6: By (35) and (36), the process $\{\hat{\xi}_t^L\}_{t\in\mathbb{N}}$ can be viewed as the Markov chain associated to the *mean-contractive* iterated function system

$$\hat{\Sigma}^L := (\hat{E}, \mathbb{R}\times\Gamma\times\mathbb{A}^m, \hat{V}^L, \hat{Z}^L).$$

In order to show that the Markov chain $\{\hat{\xi}_t^L\}_{t\in\mathbb{N}}$ has a unique invariant measure, we denote by $\hat{\mu}_t^{L,\xi}$ the conditional distribution of the random variable $\hat{\xi}_t^L$, given $\hat{\xi}_0^L = \xi$, and $f : \hat{E} \to \mathbb{R}$ is a bounded continuous function.

In view of (37) the Feller process $\{\hat{\xi}_t^L\}_{t\in\mathbb{N}}$ on the state space $\hat{E}$ has an stationary distribution $\hat{\mu}^L$. For each $\epsilon > 0$ there exists a compact set $K \subset \hat{E}$ such that $\hat{\mu}^L(K^c) < \epsilon$. Since

$$\left|\int f d\hat{\mu}_t^{L,\xi} - \int f d\hat{\mu}^L\right| \leq \sup_{\xi_1,\xi_2\in K}\left|\int f d\hat{\mu}_t^{L,\xi_1} - \int f d\hat{\mu}_t^{L,\xi_2}\right| + 2\|f\|_\infty \hat{\mu}^L(K^c)$$

it suffices to prove existence of a constant $T \in \mathbb{N}$ that satisfies

$$\sup_{t \geq T} \sup_{\xi_1, \xi_2 \in K} \left| \int f d\hat{\mu}_t^{L,\xi_1} - \int f d\hat{\mu}_t^{L,\xi_2} \right| < \epsilon. \tag{39}$$

In fact, it is enough to prove vague convergence of the sequence $\{\hat{\mu}_t^{L,\xi_i}\}_{t \in \mathbb{N}}$ to $\hat{\mu}^L$, i.e., to prove (39) for all Lipschitz continuous functions with compact support.

Since the modified profit functions are uniformly bounded and because $\alpha < 1$ there exists a constant $\hat{K} < \infty$ such that

$$\sup_{\xi \in K} \mathbb{P}_\xi \{ |\hat{\xi}_t^L| \leq \hat{K} \text{ for all } t \in \mathbb{N} \} = 1.$$

For all starting points $\xi \in K$, the Markov chain $\{\hat{\xi}_t\}_{t \in \mathbb{N}}$ may thus be viewed as a Feller process associated to a mean-contracting iterated function system on the *bounded* state space $\widetilde{E}$. Hence (39) follows from Theorem 4.1.1 in Norman (1972).

The precise argument goes as follows: Let $\{\tilde{\xi}_t^L\}_{t \in \mathbb{N}}$ the Markov chain on the compact state space $\widetilde{E}$ which is defined through the recursive relation

$$\tilde{\xi}_{t+1}^L = \widetilde{V}(\tilde{\xi}_t^L, \tau_{t+1}) \quad \text{where} \quad \tau_{t+1} \sim \hat{Z}^L(\tilde{\xi}_t^L; \cdot)$$

and the map $\widetilde{V}$ is given by

$$\widetilde{V}(\tilde{\xi}_t^L, \tau_{t+1}) := \begin{bmatrix} G \circ F(\tilde{S}_t^L, \tilde{S}_{t-1}^L, \tau_{t+1}) \\ \tilde{S}_t \\ \max\left\{ \min\left\{ \left(\alpha \tilde{U}_t^L + (1-\alpha) \widehat{P}^L(\tilde{S}_t^L, \tilde{S}_{t-1}^L, \tau_{t+1}) \right), \hat{K} \right\}, -\hat{K} \right\} \end{bmatrix}.$$

Given a common initial value $\xi \in K$ for the processes $\{\hat{\xi}_t^L\}_{t \in \mathbb{N}}$ and $\{\tilde{\xi}_t^L\}_{t \in \mathbb{N}}$ we have

$$\hat{\xi}_t^L \stackrel{\text{Law}}{=} \tilde{\xi}_t^L \quad \text{for all } t \in \mathbb{N}.$$

The Feller process $\{\tilde{\xi}_t^L\}_{t \in \mathbb{N}}$ on the bounded state space $\widetilde{E}$ may be viewed as the Markov chain associated to the iterated function system

$$\widetilde{\Gamma} = \left(\widetilde{E}, \mathbb{R} \times \Gamma \times \mathbb{A}^m, \widetilde{V}, \hat{Z}^L \right).$$

This system is mean contractive of order two in the sense of (27). Moreover (23) yields Lipschitz continuity of the stochastic kernel $\hat{Z}^L$:

$$\|\hat{Z}^L(\xi_1; \cdot) - \hat{Z}^L(\xi_2; \cdot)\|_V \leq L\widehat{L} \|\xi_1 - \xi_2\|.$$

Since $Q(u; \cdot) \geq \delta$, the iterated function system $\widetilde{\Gamma}$ satisfies the assumptions of Theorem 4.1.1 in Norman (1972). Hence

$$\lim_{t \to \infty} \sup_{\xi_1, \xi_2 \in \tilde{E}} \left| \int f \tilde{\mu}_t^{\xi_1} - \int f \tilde{\mu}_t^{\xi_2} \right| = 0$$

for every Lipschitz continuous function $f : \tilde{E} \to \mathbb{R}$. Thus, (39) is satisfied. This shows that the sequence $\{\hat{\xi}^L\}_{t\in\mathbb{N}}$ and hence the modified Markov chain $\{\xi_t^L\}_{t\in\mathbb{N}}$ converges in law to a unique limiting distribution.

We are now ready to prove our uniqueness result for the price-performance process. The proof will be based on a general comparison argument for Feller kernels which is given in the appendix.

Proof of Theorem 3.6: In order to show uniqueness of the stationary distribution we have to prove that the transition kernel $\hat{\Pi}$ of the Markov chain $\{\hat{\xi}_t\}_{t\in\mathbb{N}}$ with closed state space $\hat{E} = [-1,1]^2 \times \mathbb{R}^{m+1}$ satisfies the assumptions of Theorem A.1. In view of Proposition 4.3 and Proposition 4.4 we need to show that

$$\Pi^2(\xi;\hat{E}) > 0 \quad \text{for all } \xi \in E \text{ and that} \quad \hat{\Pi}^2(\xi;\cdot) \ll \Pi^2(\xi;\cdot) \quad \text{for all } \xi \in E.$$

The first condition holds because the law of the random variable η_t has a strictly positive density on $\mathbb{R}$. The second condition follows from Assumption 3.5 (iii) along with strict monotonicity of the transformation function G. Indeed, let us fix a state $\xi_t \in \hat{E}$. By Assumption 3.5 (ii) and because G is strictly increasing, the transformation formula for Lebesgue integrals shows that the conditional distributions of the vectors $\zeta_t := (S_{t+1}, S_{t+2})$ and $\hat{\zeta}_t := (\hat{S}_{t+1}, \hat{S}_{t+2})$ are defined through strictly positive densities f_{ξ_t} and $\hat{f}_{\xi_t}$ on $\mathbb{R}^2$ and $[-1,1]^2$, respectively. By Assumption 3.5 (iii),

$$\Pi^2(\xi_t;A) = \int_A g_{\xi_t,\zeta}(u) f_{\xi_t}(\zeta) d\zeta \quad \text{and} \quad \hat{\Pi}^2(\xi_t;A) = \int_A g_{\xi_t,\zeta}(u) \hat{f}_{\xi_t}(\zeta) d\zeta$$

where g_{ξ_t,ζ_t} denotes the density of the conditional distribution of the performance level U_{t+2}, given ξ_t and ζ_t. For all $\hat{\xi} \in \hat{E}$, this yields

$$\hat{\Pi}^2(\xi;\cdot) \approx \Pi^2(\xi;\cdot) \quad \text{on} \quad \hat{E}.$$

This proves uniqueness of the stationary distribution. In view of the tightness result in Proposition 4.3, uniqueness implies convergence in (24) for any initial state ξ and every bounded continuous function f because any weak limit of the sequences $\{\nu_t^\xi\}_{t\in\mathbb{N}}$ in (33) is a stationary distribution for the transition kernel Π. □

5 Numerical Simulations

What we have shown up to this point is that with both fundamentalists and chartists present we can guarantee the existence of a unique stationary distribution if the probability that agents become chartists is not too high. However, in addition to knowing that their presence does not jeopardize the existence of a limit distribution, it is also interesting to see how the characteristics of that distribution are affected by the presence of chartists. We will see, using simulations of a simple example, that chartists add to the noise in the system and that bubbles and crashes appear. The translation of this in terms of the stationary distribution is increased kurtosis and variance. Thus, loosely speaking chartists increase the volatility of the price process. The switching of agents' expectations between fundamentalist and chartist provides an answer to the "excess volatility" puzzle mentioned in the introduction. The simulations enable us to see this clearly.

We consider a model with two fundamentalists and one trend chaser. Specifically, we put

$$R_t^1 := S_{t-1}+\frac{3}{4}(F^1-S_{t-1}), \quad R_t^2 := S_{t-1}+\frac{3}{4}(F^2-S_{t-1}) \quad \text{and} \quad R_t^3 := S_{t-1}(\omega)+\frac{1}{2}(S_t-S_{t-1})$$

where $F^1 = -F^2 = \frac{1}{2}$. The discount factor is $\alpha = 0.9$, and the random variables η_t follow a normal distribution with mean 0 and variance 0.5:

$$\eta_t \sim N\left(0, \frac{1}{2}\right).$$

There are 10 market participants adopting their investment strategies independently of

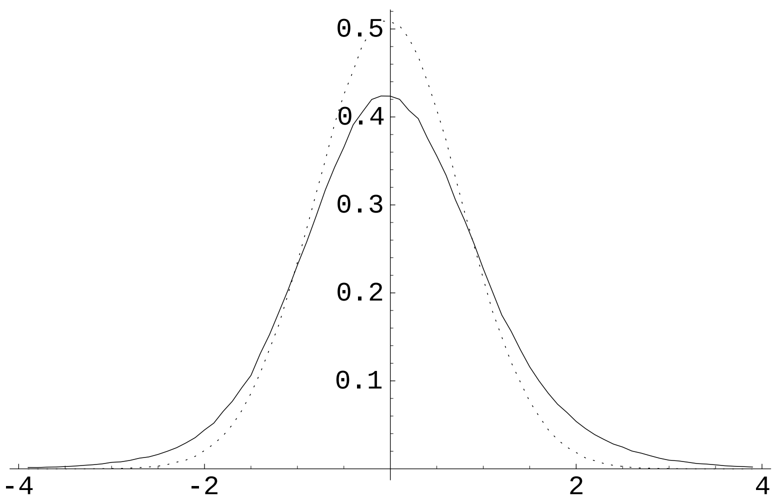

Figure 1: *Empirical stationary distribution of logarithmic asset prices: benchmark model without a chartist (dotted line); model with chartist (solid line).*

each other, given the gurus' current performances. The choice functions take the logit form

$$q^{a,i}(u) = \frac{e^{\gamma^i u^i}}{e^{\gamma^1 u^1} + e^{\gamma^2 u^2} + e^{\gamma^3 u^3}}$$

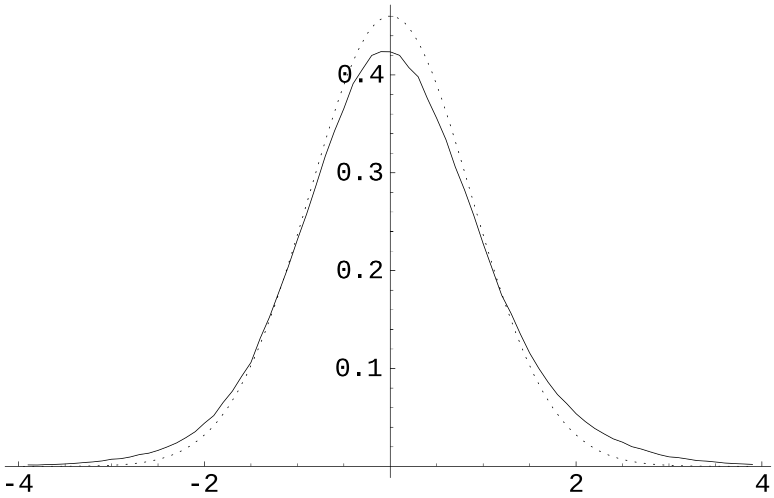

Figure 2: *Empirical stationary distribution of logarithmic asset prices (solid line) and a normal distribution with the same mean and variance (dotted line).*

where $(\gamma^1, \gamma^2, \gamma^3) = (1, 1, 5)$ denotes the relative weights of the experts' performance. In order to suitably limit the impact of trend chasing, we bound the probability with which the agents follow a trend chaser by $\frac{3}{7}$. In view of Example 3.3 this guarantees that the price performance process has a stationary distribution. Our simulations suggest that the stationary distribution is also unique though the experts' profiles are not random as in Example 3.7.

The presence of chartists is revealed in the structure of the limiting empirical distribution μ^* of logarithmic asset prices. It turns out that the trend chasers increase both the variance and the kurtosis of μ^*. Loosely speaking, the chartist amplifies the noise generated by random liquidity trading. Specifically, in a benchmark model without a chartist we found an empirical mean and variance of 0 and 0.66, respectively, and a kurtosis of approximately 3. This suggest that in the absence of trend chasing, empirical averages of logarithmic stock prices are asymptotically $N(0, 0.61)$ distributed. The presence of trend chasers increases both the variance ($\sigma^2 \approx 0.89$) and the kurtosis ($\kappa \approx 3.97$) and thickens the tails of the distribution. In view of the considerable excess kurtosis we do not expect μ^* to be normal. This is confirmed by Figure 5 which shows the right tails of the stationary distributions in a reference model without a chartist (dashed line), in a model with a trend chaser (solid line) and the tails of an $N(0, 0.89)$-distribution (dotted line).

The presence of chartists also leads to the temporary occurrence of bubbles and crashes. A typical bubble is displayed in Figure 3. Bubbles do not evolve gradually, but occur rather spontaneously. This suggest that they are triggered by an unusually high liquidity

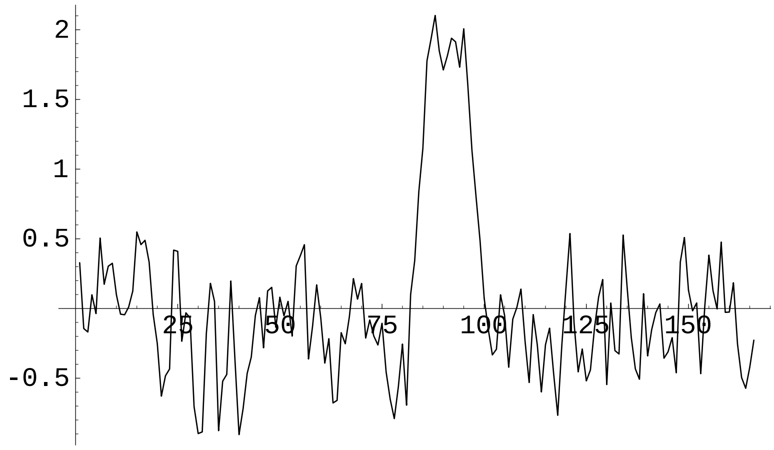

Figure 3: *A* typical bubble path.

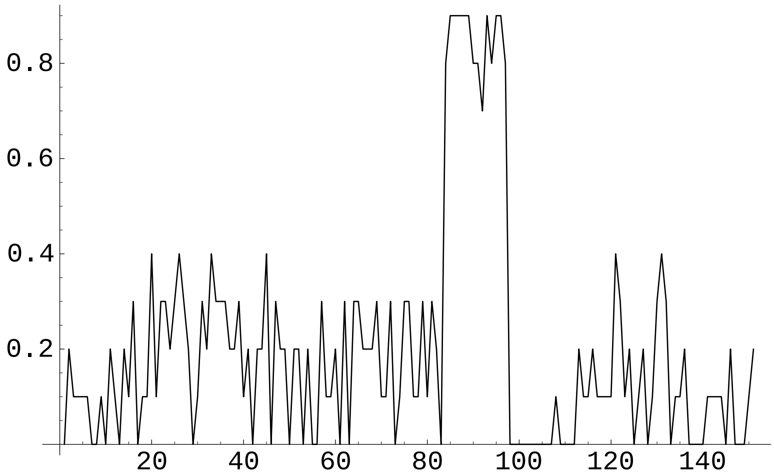

Figure 4: *F*raction of chartists associated to the bubble in Figure 3.

demand leading to an unusually high increase in prices. An increase in the price above the "optimistic fundamental value" decreases the performance of the fundamentalists, but may increase a chartist's performance. This can trigger spontaneous herd behavior. During a bubble the profit generated by the fundamentalists is negative, and so the agents have an increasing probability of following the chartist, and typically they do so. Thus, the fraction of trend chasers is high during a bubble, while it is normally quite low; see Figure 4. Since we limit the probability with which the agents adopt a trend chasing strategy, persistent bubbles lasting for many periods are unlikely, and this is what we observe in our simulation. On average, a bubble lasts for 15-20 periods.

The presence of bubbles which burst fairly rapidly is due to the restriction that we

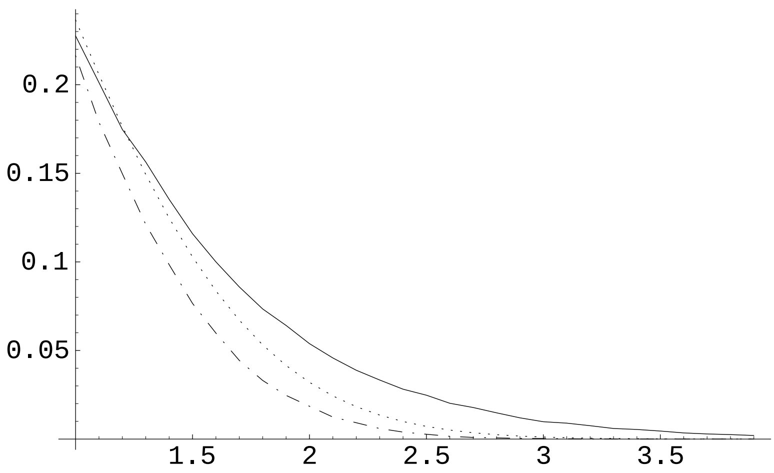

Figure 5: *Right tails of the empirical stationary distributions in a model with (solid line) without (dotted line) a chartist and the tails of a centered normal distribution (dashed line) with variance* 0.89 *(dotted line).*

put on the probability of following the chartist rule. There is, as we have observed, a self reinforcing trend towards chartism as soon as the price moves out of the limits defined by the fundamental values. However there is always a sufficiently high probability that agents will again adopt fundamentalist rules, and this prevents the process from exploding without limit.

6 Conclusion

Many explanations have been offered for the existence of bubbles, crashes, and excess volatility of the prices of financial assets. These vary from endowing individuals with substantial memories and complicated forecasting rules with sophisticated lag structures, to the presence of a complicated process generating external shocks to the system. Both such approaches have been used in representative agent models with rational expectations. The failure of these models to capture some of the salient features of financial time series is well documented. An alternative approach has been to allow for heterogeneous agents. A classic example of this is given by the contributions of De Long et al. (1989). The objection to this sort of analysis is that it involves assuming that some agents are systematically wrong. A further avenue is that opened by Brock and Hommes (1997), who show that the process generating prices can become unstable if agents switch to simple forecasting rules once prices are close to fundamental values. The reason for their switch to the simple rule is that it is less costly than the more complicated rule that led the process to converge in the first place.

Our approach is rather different. The driving feature of the price process is the switching of agents from one forecasting rule to the other. This switching can be attributed to the relative success of the rules. The switching process has the characteristic that agents can, at any point in time, herd on one rule. When this happens agents forecasts are self reinforcing. Thus, agents are not systematically wrong, and their behaviour cannot be described as irrational. Furthermore, there is freedom in specifying the function which maps past profits from the different forecasting rules into the probability of choosing those rules. This makes our specification rather general.

Using this simple but general framework, we are able to capture certain stylized facts of empirical financial price time series and to show the existence of an equilibrium which is different from those commonly used in the literature.

An important feature of the model is the inclusion of chartists or trend chasers who may induce temporary bubbles and crashes. Whilst this introduces such characteristics as "excess volatility" and "long memory" in the price process, see Kirman and Teyssiere (2002), it presents a major difficulty from an analytical point of view. When chartists predominate they may generate bubbles but there is no a priori bound on the levels that prices may attain during a bubble. Thus it is not obvious that the price process will not explode. Nevertheless we show that the stochastic process that prices follow is ergodic and has a unique invariant distribution Thus, although the process does not converge pathwise to any steady state there is a certain overall stability in its behavior. The invariant distribution reflects the presence of the chartists since it has, for example, fatter tails than the distribution with only fundamentalists present. Viewing the problem in this way shows how one can reconcile such commonly reiterated and empirically confirmed statements as "in the short run fundamentals have no predictive power," whereas "in the long run fundamentals play an important role."

A An ergodicity criterion for Markov chains

Our proof of Theorem A.1 is based on a comparison of the Markov chain $\{\xi_t\}_{t\in\mathbb{N}}$ with a modified chain $\{\hat{\xi}_t\}_{t\in\mathbb{N}}$ whose convergence is established by standard methods. In this section we formulate this comparison argument in a general framework.

Theorem A.1 *Let Π be a Feller kernel on a complete separable metric state space (E, d) such that $d\Pi^m(x;\cdot) = f_x d\lambda$ for some $m \in \mathbb{N}$ and some reference measure λ. Let $\hat{\Pi}$ be a Feller kernel on a closed subspace $\hat{E} \subset E$, and assume that the following conditions hold:*

(i) The kernel $\hat{\Pi}$ has a unique invariant distribution $\hat{\mu}$.

(ii) The law $\hat{\Pi}^m(x;\cdot)$ is absolutely continuous with respect to $\Pi^m(x;\cdot)$.

(iii) The Markov chain Π satisfies $\Pi^m(x;\hat{E}) > 0$ for all $x \in E$.

(iv) For any two sequences $\{x_n\}_{n\in\mathbb{N}}$ and $\{y_n\}_{n\in\mathbb{N}}$ in $\hat{E}$ converging to the same limit, the measures $\Pi^m(x_n;\cdot)$ and $\Pi^m(y_n;\cdot)$ are not singular to each other for all sufficiently large n.

Then Π has at most one stationary distribution μ.

Proof. Let $\mathcal{I} := \{\mu : \mu\Pi = \mu\}$ be the convex set of all invariant probability measures for Π. Each $\mu \in \mathcal{I}$ has a unique representation as a mixture over the set $\mathcal{I}_e$ of all ergodic invariant measures. That is, each μ allows a unique representation of the form

$$\mu(\cdot) = \int_{\mathcal{I}_e} \nu(\cdot)\pi(\mu; d\nu),$$

and distinct probability measures in $\mathcal{I}_e$ are mutually singular; see, e.g., Theorem A.1.1 in Kifer (1986). Now suppose that Π has more than one stationary distribution. Then Π has at least two ergodic stationary measures μ_1 and μ_2. In this case Π is decomposable. That is, there exist measurable sets C_i $(i = 1, 2)$ such that

$$\Pi(\xi; C_i) = \delta_{i,j} \quad \text{for all } \xi \in C_j \text{ and such that} \quad C_1 \cap C_2 = \emptyset; \tag{40}$$

see, e.g., Breiman (1968), Theorem 7.6. Let $\overline{C}_i$ be the topological closure of the set C_i. Since Π has the Feller property, the Portmonteau theorem yields

$$\Pi(x^i; \overline{C}_i) \geq \limsup_{n\to\infty} \Pi(x_n^i; C_i) = 1$$

for each sequence $\{x_n^i\}_{n\in\mathbb{N}}$ in C_i that converges to $x^i \in \overline{C}_i$. Thus, the sets $\overline{C}_i$ are stochastically closed with respect to Π.

We put $D_i := \overline{C}_i \cap \hat{E}$. By (iii) $\mu_i(\hat{E}) > 0$, and so $\mu_i(D_i) > 0$. In particular, $D_i \neq \emptyset$ $(i = 1, 2)$. For $x \in D_i \subset \overline{C}_i$ we have $\Pi(x; \overline{C}_i^c) = 0$. Hence it follows from (ii) that

$$\hat{\Pi}(x; D_i) = 1,$$

and so D_1 and D_2 are stochastically closed with respect to $\hat{\Pi}$. For any choice of $x^i \in D_i$ (i=1,2), the weak limit ν_i of the sequence $\{\hat{\Pi}^n(x^i;\cdot)\}_{n\in\mathbb{N}}$ is a measure on the closed set D_i. But $\nu_1 = \nu_2 = \hat{\mu}$, due to (i). Thus, $D_1 \cap D_2 \neq \emptyset$. In particular, $\overline{C}_1 \cap \overline{C}_2 \neq \emptyset$. Hence, there exist sequences $\{x_n^i\}_{n\in\mathbb{N}}$ in C^i for $i = 1, 2$ which converge to the same limit $x \in \overline{C}_1 \cap \overline{C}_2$. For any $n \in \mathbb{N}$, the measures $\mu_n^1 := \Pi^m(x_n^1;\cdot)$ and $\mu_n^2 := \Pi^m(x_n^2;\cdot)$ are singular because $x_n^i \in C^i$. This, however, contradicts assumption (iv), and this shows that there is at most one stationary distribution for Π. $\square$

References

ALLEN, F., S. MORRIS, AND H-S. SHIN (2003): "Beauty Contests, Bubbles and Iterated Expectations in Asset Markets," *Cowles Foundation Discussion Paper No 1406*, Yale University.

ALÓS-FERRER, C. AND A.B. ANIA (2004): "The Stock Market Game," *Preprint.*

ANDERSON, M.F., A. DE PALMAAND J.-F. THISSE (1992): *Discrete Choice Theory of Product Differentiation,* MIT Pres, Cambridge.

ARNOLD, L. (1998): *Random dynamical Systems*, Springer, Berlin.

BACCHETTA, P. AND E. VAN WINCOOP (2003): "Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle?," *Preprint.*

BARNSLEY, M.F., S.G. DEMKO, J.H. ELTON, AND J. S. GERONIMO (1988): "Invariant Measures for Markov Processes Arising from Iterated Function Systems with Place-Dependent Probabilities," *Ann. Inst. Henri Poincare*, 24, 367–394.

BARNSLEY, M. F., AND J. H. ELTON (1988): "A new Class of Markov Processes for Image Encoding," *Adv. Appl. Prob.*, 20, 14–32.

BARTOLINI, L., AND L. GIORGIANNI (1999): "Excess Volatility and the Asset-Pricing Exchange Rate Model with Unobserved Fundamentals," IMF Discussion Paper, No. 99/71, International Monetary Fund, Washington D.C.

BÖHM, V., AND J. WENZELBURGER (2002): "On the Performance of Efficient Portfolios," *J. Economic Dynamics and Contral*, to appear.

BOTTAZZI, G., G. DOSI AND I. REBESCO (2003): "Institutional Architectures and Behavioral Ecologies in the Dynamics of Financial Markets," *Preprint.*

BRANDT, A. (1986): "The Stochastic Equation $Y_{n+1} = A_n Y_n + B_n$ with Stationary Coefficients," *Adv. Appl. Probab.*, 18, 211–220.

BREIMAN, L. (1968): *Probability.* Addison-Wesley, Reading.

BROCK, W., AND S. DURLAUF (2001a): "Discrete Soice with Social Interactions," *The Review of Economic Studies*, 68, 235–260.

——— (2001b): "Interaction-Based Models," in J. Heckman and E. Leamer (eds), *Handbook of Econometrics*, Vol V, North-Holland, Amsterdam.

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

Brock, W., C. Hommes, and F. Wagener (2002): "Evolutionary Dynamics in Markets with Many Trader Types," *Preprint*.

Cass, D. and K. Shell (1983): "Do Sunspots Matter?," *Journal of Political Economy*, 91 (2), 295–329.

Cheung, Y-W., M.D. Chinn and A.G. Pascual (2002): "Empirical Exchange Rate Models in the Nineties: Are they fit to Survive?," *NBER Working Paper*, 9393.

Chionis, D. and R. MacDonald (2002): "Aggregate and Disaggregate Meausres of the Foreign Exchange Risk Premium," *International Review of Economics and Finance*, 11, 57–84.

De Long, J.B., A. Shleifer, L.H. Summers, and R.J. Waldmann (1989): "The Size and Incidence from Noise Trading," *Journal of Finance*, 44 (3), 681–699.

——— (1990): "Noise Trader Risk in Financial Markets," *Journal of Political Economy*, 98 (4), 703-738.

Embrechts, P., C. Klüppelberg, and T. Mikosch (1997): *Modelling Exremal Events*. Springer-Verlag, Berlin.

Ethier, S.N. and T.G. Kurtz (1997): *Markov Processes: Characterisation and Convergence*. John Wiley & Sons, New York.

Evans, G.W. (1991): "Pitfalls in Testing for Explosive Bubbles in Asset Prices," *American Economic Review*, 81, 922–930.

Föllmer, H. (1994): "Stock Price Fluctuation as a Diffusion in a Random Environment," *Philos. Trans. R. Soc. Lond., Ser. A*, 1684, 471–483.

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

Frankel, J.A., and 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.

Frankel, J.A., and A.K. Rose (1995): "Emirical Research on Nominal Exchange Rates," in G.M. Grossman and K. Rogoff (eds.) *Handbook of International Economics*, Vol III, North-Holland, Amsterdam.

GARMAN, M.(1976): "Market Microstructure," *Journal of Financial Economics*, 3, 257–275.

GOIURINCHAS, P-O. AND A. TORNELL (2002): "Exchange Rate Dynamic. Learning and Misperception," *NBER Working Paper* 9391.

HAMILTON, J.D. (1989): "A new Approach to Economic Analysis of Nonstationary Time Series and the Buisiness Cycle," *Econometrica*, 16, 357–384.

HENS, T. AND K.R.. SCHENK-HOPPÉ (2003): "Evolutionary Stability of Portfolio Rules," *Preprint*.

HORST, U. (2001): "The Stochastic Equation $Y_{t+1} = A_t Y_t + B_t$ with Non-Stationary Coefficients," *J. Appl. Prob.*, 38, 80–95.

——— (2004): "Financial Price Fluctuation in a Stock Market Model with many Interacting Agents," *Economic Theory*; to appear.

HUNTER, W.C., G.G. KAUFMAN, AND M. POMERLEANO (2003): *Asset Prices Bubbles.* MIT Cambridge, Massachussets.

KARLIN, S. (1953): "Some Random Walks Arising in Learning Models," *Pac. J. Math.*, 3, 725–756.

KESTEN, H. (1973): "Random Difference Equations and Renewal Theory for Products of Random Matrices," *Acta Math.*, 131, 207–248.

KIFER, Y. (1986): *Ergodic Theory of Random Transformations.* Birkhäuser, Boston.

KIRMAN, A. (1998): "On the Transitory Nature of Gurus," Working paper, EHESS and Université de Marseille III.

KIRMAN, A., R. RICCIOTTI, AND R. TOPOL (2003): "Bubbles and Foreign Exchange Markets; it takes two to tango," Working paper, GREQAM.

KIRMAN, A. AND G. TEYSSIERE (2002): "Microeconomic Models for Long-Memory in the Volatility of Financial Time Series," *Studies in Nonlinear Dynamics and Econometrics*, 5, 281–302.

KURZ, M. (1994): "On Rational Belief Equilibria," *Economic Theory*, 4, 859–876.

——— (1997): "Asset Prices with Rational Beliefs," in M. Kurz (ed.) *Endogenous Economic Fluctuations: Studies in the Theory of Rational Belief*, Springer Series in Economic Theory (6), Springer, Berlin.

Lux, T. (1995): "Herd behavior, bubbles and crashes," *The Economic Jornal*, 105, 881–896.

Lux, T. and S. Schornstein (2004): "Genetic Learning as an Explanation of Stylized Facts of Foreign Exchange Markets," *Preprint*.

Lyons, R.K. (2001): *The Microstructure Approach to Exchange Rates*. MIT Press, Cambridge, Massachusetts.

Mark, N.C. (1995): "Exchange Rate and Fundamentals: Evidence on Long-horizon Predictability," *American Economic Review*, 85, 210–218.

Meese, R.A. and K. Rogoff (1983): "Empirical Exchange Rate Models of the Seventies: do they fit out-of-sample," *Journal of International Economics*, 14, 3–24.

Meyn, S.P. and R.L. Tweedie (1993): *Markov Chains and Stochastic Stability*. Springer-Verlag, Berlin.

Norman, F.M. (1972): *Markov Processes and Learning Models*. Academic Press, New York.

Shiryaev, A.N. (1989): *Probability*. Sringer-Verlag, Berlin.

Steinsaltz, D. (1999): "Locally Contractive Iterated Function Systems," *Annals of Probability*, 4, 546–588.

Townsend, R.M. (1983): "Forecasting the Forecast of Others," *Journal of Political Economy*, 91, 546–588.

Vervaat, W. (1979): "On a Stochastic Difference Equation and a Representation of non-negative Infinitely Divisible Random Variables," *Adv. Appl. Probab.*, 11, 750–783.

Weisbuch, G., O. Chenevez, A. Kirman, and J-P. Nadal (1998): "A Formal Approach to Market Organization: Choice Functions, Mean Field Approximations and Maximum Entropy Principle," in *Advances in Self-Organization and Evolutionary Economics*, J. Lesourne and A. Orlean (eds.), Economica, 149–159.

Woodford, M. (1990): "Learning to Believe in Sunspots," *Econometrica*, 58 (2), 277–307.