

LEARNING TO PREDICT RATIONALLY WHEN BELIEFS ARE HETEROGENEOUS

By Jan Wenzelburger*

Fakultät für Wirtschaftswissenschaften
Universität Bielefeld
Postfach 100 131
D-33501 Bielefeld
Germany

Discussion Paper No. 477

Abstract

This paper develops an adaptive learning scheme for an asset market with heterogeneous traders. We introduce the concept of a perfect forecasting rule for fundamentalists which generates rational expectations in the presence of chartists whose expectations are not necessarily rational. In the case when traders have linear mean-variance preferences and chartists base their decisions on simple technical trading rules, perfect forecasting rules are approximated by successively estimating the excess supply function from historical data. Conditions are given under which trajectories generated by this adaptive learning scheme converge to trajectories with rational expectations for fundamentalists.

Keywords: Learning, bounded rationality, heterogeneity, asset markets.

JEL Classification: D83, D84, G11, O16.

First version: Dec. 1999, this version: Oct. 2001.

*ACKNOWLEDGEMENT. This research was supported by the Deutsche Forschungsgemeinschaft under contract No. He 2592/2-1. I am indebted to Volker Böhm, Jochen Jungeilges, Thorsten Pampel, and Andreas Tonn for many valuable discussions. I would like to thank Carl Chiarella for his generous hospitality and participants of the workshop on *Expectational and Learning Dynamics in Financial Markets* in Sydney, 1999 and of the 5th conference of the *Society for the Advancement of Economic Theory* in Ischia, 2001 for helpful comments.

1 Introduction

Recent search into the nature of asset prices has emphasized the importance of heterogeneity of traders in financial markets. Trade among heterogeneous agents occurs since either preferences, beliefs or endowments differ. As a consequence, their attitudes towards risk and their beliefs about the future development of asset prices constitute the key influence on the determination of actual market prices. It is a commonplace that traders, in general, will not share a common belief about future prices and it is for this reason that the evolution of actual market prices cannot be self-confirming for all traders. While this observation stresses the fact that an agent in an asset market will most likely not have rational expectations along a price path, it is clear that a descriptive and consistent theory of asset markets should pay adequate respect to the presence of heterogeneous individuals. Situations in which an agent has rational expectations should therefore be a special case of a general descriptive model.

The classical approach of the literature, however, derives asset prices from intertemporal equilibrium conditions assuming that expectations are always rational. This leaves no explanatory room for diverse and non-rational beliefs of traders (see e.g. Duffie (1996) or Pliska (1997) and references therein). Following an approach put forward in Böhm, Deutscher & Wenzelburger (2000) and Böhm & Chiarella (2000), we consider a model with a structure of myopic market participants characterized by preferences and possibly non-rational and heterogeneous beliefs. Such a structure, for instance, arises naturally in models with overlapping generations of consumers. These consumers face investment possibilities in one risk-free and K risky assets, where K is fixed but may be arbitrarily large. Given consumers' preferences and arbitrary heterogeneous beliefs, an explicit rule determining market clearing prices at each date yields an explicit map determining *endogenous random asset prices*. Combined with forecasting rules, which describe the subjective assessments of the future evolution of prices, this yields an explicit time-one map of a so-called random dynamical system (Arnold 1998) in which expectations feed back into the actual evolution of the prices.

One of the most intriguing problems concerning financial markets is to understand the structure of forecasting rules which generate rational expectations when investors are heterogeneous. Following Brock & Hommes (1998) and Chiarella & He (2000), we assume that consumers have no direct access to a forecasting technology. Instead, they are allowed to choose between financial mediators, who are knowledgeable of certain forecasting techniques, according to a discrete-choice model of the type found in Anderson, de Palma & Thisse (1992).

Extending a concept introduced in Böhm & Wenzelburger (1999, 2000a,b), a major innovation of this paper is the notion of a perfect forecasting rule for first and second moments in the presence of non-rational beliefs of other market participants. By definition, these forecasting rules generate orbits with rational expectations in the sense that they provide correct first and second moments of the price process conditioned

on the available information. From the perspective of a fundamentalist, this concept takes the investment decisions of heterogeneous non-rational agents such as chartists and noise traders into account regardless of how precise their forecasts are. Perfect forecasting rules for first and second moments are particularly important for models in which traders are characterized by mean-variance preferences.

The present paper addresses the natural question how a financial mediator, who is assumed to be boundedly rational in the usual sense, can learn a perfect forecasting rule from historical data, if initially the rule is unknown. The main informational constraint encountered by a mediator stems from the fact that neither the fraction of consumers choosing a particular mediator nor the beliefs of the mediators themselves are observable quantities. We will show that this missing information can be retrieved by estimating the excess supply function of all market participants. This will, in general, become a highly nonlinear problem. However, in the case when consumers are endowed with linear mean-variance preferences and chartists use simple linear forecasting rules, the main nonlinearity consists of the discrete-choice model which governs the switching behavior of consumers. Since rigorous techniques for estimating a discrete choice model exist (see e.g. Judge, Griffiths, Hill, Lütkepohl & Lee 1985, Chap. 18), it turns out that the problem reduces to estimating a stochastic linear functional relationship. Using a variant of the ordinary-least-square scheme which is popular in the learning literature (Evans & Honkapohja 2001), an adaptive learning scheme is developed which approximates perfect forecasting rules from historical data. Necessary and sufficient conditions for the convergence of the learning scheme are discussed.

The paper is organized as follows. After an introduction of the model in Section 2, we develop the notion of a perfect forecasting rule in Section 3. Section 4 is concerned with risk premia associated with heterogeneous beliefs and contains a generalization of the famous security market line result. The special case with linear mean-variance preferences is carried out in Sections 5 and 6, a benchmark dynamics under rational expectations is discussed in Section 7. In Section 8, we develop our adaptive learning scheme and provide a convergence theorem. Conclusions are given in Section 9, an appendix outlines the concept of a random dynamical system.

2 The model

Following Böhm, Deutscher & Wenzelburger (2000) and Böhm & Chiarella (2000), consider an overlapping generations model with a finite number of types $h = 1, \dots, H$ of young households. Each young household of type h lives for two periods, receives an initial endowment $e^h > 0$ of a non-storable commodity in the first period of life, and does not consume. In order to transfer wealth to the second period of his life, such a consumer will choose a portfolio of $K + 1 \in \mathbb{N}$ retradeable assets whose proceeds he will consume. He receives no additional endowment in the second period of his life, so that

his total consumption is equal to the return on the investment of his initial endowment when young.

The $K + 1$ assets correspond to shares of firms whose production activities induce a stochastic process of dividends which are distributed to the shareholders. There is one risk-free real asset which has a constant real rate of return $R = 1 + r > 0$ given exogenously, i.e., investing one unit of the consumption good in the risk-free asset yields R units of the consumption good in the subsequent period.

Young consumers maximize expected utility over future consumption with respect to a von-Neumann-Morgenstern utility function and subjective expectations for future asset prices. Let $0 \leq y \leq e^h$ denote the amount of the consumption good invested in the risk-free asset and $x \in \mathbb{R}^K$ be the vector of shares purchased at the *ex-dividend* price vector $p \in \mathbb{R}_+^K$ (in units of the consumption good). Since there is no consumption when young, the future wealth of consumer h when old is $w = Ry + (\hat{p} + d)^\top x$, where $d \in \mathcal{D}$ is the random dividend payment drawn from some subset $\mathcal{D} \subset \mathbb{R}_+^K$ and $\hat{p} \in \mathbb{R}_+^K$ is the future price vector of the assets. Setting $q = \hat{p} + d$ for the *cum-dividend* price vector, the budget equation $y + p^\top x = e^h$ gives

$$w = R \underbrace{(e^h - p^\top x)}_{\text{risk-free investment}} + \underbrace{q^\top x}_{\text{equity return}}.$$

Assuming that each consumer treats the price p at which he purchases as a parameter, the remaining uncertainty of the return of the portfolio rests with the cum-dividend price q of shares. Households have no direct access to a forecasting technology and select a mediator who solves their individual investment problems using subjective probability distributions.

There is a finite number $i = 0, \dots, I$ of financial mediators characterized by subjective beliefs regarding the future cum-dividend price of the assets in form of subjective probability distributions $\nu^{(i)} \in \text{Prob}(\mathbb{R}_+^K)$, $i = 0, \dots, I$ with $\text{Prob}(\mathbb{R}_+^K)$ denoting the set of all Borelian probability measures. The asset demand of a young consumer of type h based on the subjective belief $\nu^{(i)}$ provided by mediator i is defined by an optimal portfolio choice which maximizes expected utility of future wealth,

$$(1) \quad \varphi^h(\nu^{(i)}, p) = \underset{p^\top x \leq e^h}{\text{argmax}} \int_{\mathbb{R}_+^K} u^h(Re^h + (q - Rp)^\top x) \nu^{(i)}(dq).$$

For simplicity, we abstract from intermediation costs for households and suppress the constant e^h in the expression (1). To assure existence and uniqueness of the solution (1) of the consumer's decision problem, we assume first that each utility function $u^h : \mathbb{R} \rightarrow \mathbb{R}$, $h = 1, \dots, H$ is twice continuously differentiable, strictly monotonically increasing, strictly concave, and bounded (see e.g. Grandmont 1982); second we assume that each distribution $\nu^{(i)}$, $i = 0, \dots, I$ is such that no asset is redundant.

Let $\eta^{(hi)} \in [0, 1]$ with $\sum_{i=0}^I \eta^{(hi)} = 1$ denote the fraction of households of type h employing the mediator i and $\eta^{(i)} = (\eta^{(1i)}, \dots, \eta^{(Hi)}) \in \mathbb{R}_+^H$ be the distribution of households which employ i . Then $w^{(i)} = \sum_{h=1}^H \eta^{(hi)} e^h$ is the amount of capital invested by mediator i and

$$(2) \quad x^{(i)} = \Phi^{(i)}(\eta^{(i)}, \nu^{(i)}, p) := \sum_{h=1}^H \eta^{(hi)} \varphi^h(\nu^{(i)}, p)$$

denotes the aggregate demand of all households which employ i . Let $\bar{x} \in \mathbb{R}_+^K$ denote the total amount of retradable risky assets in the economy which is equal to the aggregate supply of old consumers. Then market clearing in the asset market defines an equilibrium ex-dividend price p which is a solution to

$$(3) \quad \sum_{i=0}^I \Phi^{(i)}(\eta^{(i)}, \nu^{(i)}, p) - \xi - \bar{x} = 0.$$

Here $\xi \in \mathbb{R}^K$ denotes a random quantity of assets which noise traders purchase or sell in the market.¹ If for a fixed distribution of households $\eta^{(i)}$ and fixed beliefs $\nu^{(i)}$ the aggregate excess demand for assets (3) is globally invertible with respect to p , then there exists a unique price p which clears the asset market. As pointed out in Böhm & Chiarella (2000), this implies that the market-clearing price of assets is a deterministic function of individual characteristics of young consumers and mediators.

Let $\eta_t = (\eta_t^{(i)})_{i=0}^I \in \mathbb{R}_+^{HI}$ denote the joint distributions of households, $(\nu_t^{(i)})_{i=0}^I$ describe the subjective beliefs of all these mediators at date t , and ξ_t represent the action of noise traders at that date. Then under the above invertibility assumption, the asset price vector in period t which clears the market is uniquely determined by the map $S : \mathbb{R}^K \times \mathbb{R}_+^{HI} \times \text{Prob}(\mathbb{R}_+^K)^I \rightarrow \mathbb{R}_+^K$,

$$(4) \quad p_t = S(\xi_t, \eta_t, (\nu_t^{(i)})_{i=0}^I).$$

The map (4) is an economic law with an expectations feedback in the sense of Böhm & Wenzelburger (1999, 2000a,b). Observe that S does not contain the price itself as an argument, such that the map S is of the *Cobweb type*. The dating of the expectations $\nu_t^{(i)}$ relative to the actual price p_t reveals that the economic law (4) has an expectational lead, i.e., expectations are with respect to the realization of prices one period ahead of the map S , cf. Böhm & Wenzelburger (2000a). This is illustrated in Fig. 1. The functional form of (4) shows moreover that the evolution of the asset prices is driven exclusively by the expectations formation of the financial mediators and the way in which households choose among mediators.

¹See De Long, Shleifer, Summers & Waldmann (1990, p. 709) for alternative ways of specifying noise traders.

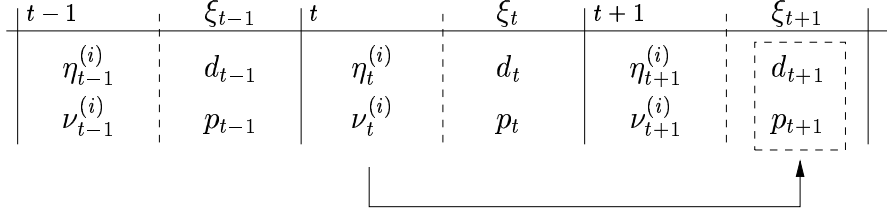


Figure 1: Timing of the price formation.

The decision of a household is based on the performance of a mediator. Let $r = R - 1$ be the interest rate paid for the risk-free asset. If $x_{t-1}^{(i)}$ denotes the portfolio held by mediator i after investing the amount $w_{t-1}^{(i)} = \sum_{h=1}^H \eta_{t-1}^{(hi)} e^h$ in period $t - 1$, then the realized return resulting from selling $x_{t-1}^{(i)}$ in period t is

$$(5) \quad R_t^{(i)} = r + [q_t - Rp_{t-1}]^\top (x_{t-1}^{(i)} / w_{t-1}^{(i)}).$$

For each i , the sample means $\hat{\mu}_t^{(i)}$ and the sample standard deviations $\hat{\sigma}_t^{(i)}$ of the time series $\{R_s^{(i)}\}_{s=0}^t$ are recursively given by

$$(6) \quad \begin{aligned} \hat{\mu}_t^{(i)} &:= \frac{1}{t+1} \sum_{s=0}^t R_s^{(i)} = \frac{1}{t+1} [R_t^{(i)} + t\hat{\mu}_{t-1}^{(i)}], \\ \hat{\sigma}_t^{(i)} &:= \left[\frac{1}{t+1} \sum_{s=0}^t (R_s^{(i)} - \hat{\mu}_s^{(i)})^2 \right]^{\frac{1}{2}} = \left[\frac{(R_t^{(i)} - \hat{\mu}_t^{(i)})^2}{t+1} + \frac{t}{t+1} \hat{\sigma}_{t-1}^{(i)2} \right]^{\frac{1}{2}}, \end{aligned}$$

where $\hat{\mu}_{-1}^{(i)} = 0$ and $\hat{\sigma}_{-1}^{(i)} = 0$. Given (6), the *Sharpe ratio* associated with the realized returns (5) of mediator i is precisely the inverse coefficient of variation of the returns (5), defined by $(\hat{\mu}_t^{(i)} - r) / \hat{\sigma}_t^{(i)}$. The fraction $\eta_t^{(hj)}$ of households of type h which employs a particular mediator j in period t is now assumed to be determined by the *discrete-choice probability*

$$(7) \quad \eta_t^{(hj)} := \frac{\exp\left(\beta^h (\hat{\mu}_{t-1}^{(j)} - r) / \hat{\sigma}_{t-1}^{(j)}\right)}{\sum_{i=0}^I \exp\left(\beta^h (\hat{\mu}_{t-1}^{(i)} - r) / \hat{\sigma}_{t-1}^{(i)}\right)}, \quad t > 0,$$

with arbitrary $\eta_0^{(hj)}$. The parameter β^h appearing in the *discrete choice model* (7) describes the *intensity of choice* of a household of type h , that is, how fast a typical consumer of type h will switch to a different mediator. The behavior of households is thus modeled using a LOGIT model², see e.g. Anderson, de Palma & Thisse (1992).

²There are numerous ways of modeling the switching behavior of boundedly rational households, see Brock & Hommes (1998, footnote 3, p. 1241). The choice probabilities (7) take into account that a boundedly rational trader with mean-variance preferences is expected to choose the portfolio corresponding to the highest estimate for the Sharpe ratio, because such a portfolio promises to be the most efficient one in the sense of the classical CAPM theory. This observation stems from a private communication with Volker Böhm.

To obtain the evolution of the asset prices, we need to specify the probabilistic assumptions on the exogenous noise and the exogenous dividend process.

Assumption 2.1 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\{\mathcal{F}_t\}_{t \in \mathbb{N}}$ an increasing family of sub- σ -algebras of \mathcal{F} . Then we assume the following.*

- (i) *The dividend payments are described by a predictable $\{\mathcal{F}_t\}_{t \in \mathbb{N}}$ -adapted stochastic process $\{d_t\}_{t \in \mathbb{N}}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in \mathbb{R}^K such that each d_t , $t \in \mathbb{N}$ is \mathcal{F}_{t-1} measurable;*
- (ii) *The behavior of noise traders is governed by a $\{\mathcal{F}_t\}_{t \in \mathbb{N}}$ -adapted stochastic process $\{\xi_t\}_{t \in \mathbb{N}}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in \mathbb{R}^K which is uncorrelated with the dividend process $\{d_t\}_{t \in \mathbb{N}}$, defined in (i).*

Given the price law (4) for ex-dividend prices and Assumption 2.1 for the noise and for the dividend process, the cum-dividend price in period t is

$$(8) \quad q_t = S(\xi_t, \eta_t, (\nu_t^{(i)})_{i=0}^I) + d_t, \quad t \in \mathbb{N}.$$

Observe that the subjective beliefs $(\nu_t^{(i)})_{i=0}^I$ and the distribution $\eta_t = (\eta_t^{(i)})_{i=0}^I$ are based on past information, that is, they are \mathcal{F}_{t-1} measurable at date t . The dividend payment d_t of the current period t is assumed to be \mathcal{F}_{t-1} measurable and thus known before trading in period t , whereas ξ_t and hence the ex-dividend price p_t is \mathcal{F}_t measurable, see Fig. 1. This implies that the uncertainty in ex-dividend prices and in the traded quantities of assets rests solely on the behavior of the noise traders. In other words, the randomness of ex-dividend prices is essentially due to the stochastic nature of noise-trader behavior.

Notice that the price law (4) as well as the resulting price process (8) would be the same for consumers with infinite lives who maximize wealth myopically as done in Brock & Hommes (1998) or Chiarella & He (2000). Thus, the underlying OLG structure is not essential for this paper. In order to apply results from the theory of random dynamical systems, we will later assume that the two processes of Assumption 2.1 are generated by a metric dynamical system in the sense of Arnold (1998) such that the resulting price process itself becomes a random dynamical system.

3 Heterogeneous beliefs

This section is concerned with the question which forecasting rules generate rational expectations along orbits of the system. The problem is particularly challenging in financial markets when investors by their very nature have heterogeneous beliefs. It is well known that the two requirements of market clearing in all periods and rational

expectations at all times can often not be fulfilled simultaneously. As shown in Böhm & Wenzelburger (2000b), the evaluation of a forecasting rule can be carried out on several levels. For the purpose of this paper, it suffices to compare moments of subjective and true distributions with a particular emphasis on mean values and covariances.

To simplify the analysis, we assume that each mediator $i = 0, \dots, I$ picks his subjective distribution $\nu^{(i)}$ from a fixed family of probability distributions parameterized in subjective means $q^{(i)}$ and subjective covariance matrices $V^{(i)}$ such that $\nu^{(i)} \equiv \nu_{q^{(i)}, V^{(i)}}$. Identify each subjective distribution $\nu_t^{(i)}$ at a particular date t with the corresponding subjective mean $q_t^{(i)}$ for the future cum-dividend price $q_{t+1} = p_{t+1} + d_{t+1}$. Let j be an arbitrary mediator and $q_t^{(-j)} = (q_t^{(0)}, \dots, q_t^{(j-1)}, q_t^{(j+1)}, \dots, q_t^{(I)}) \in \mathbb{R}^{KI}$ denote the respective subjective means of all mediators $i \neq j$. By abuse of notation, mediator j 's forecast error on the price q_t prevailing at date t is

$$(9) \quad q_t - q_{t-1}^{(j)} = S(\xi_t, \eta_t, q_t^{(j)}, q_t^{(-j)}) + d_t - q_{t-1}^{(j)},$$

where $q_{t-1}^{(j)}$ denotes the forecast for $q_t = p_t + d_t$ made at date $t - 1$. Notice that the forecast error (9) depends decisively on the current forecasts of all mediators.

Assume for a moment that the distribution of households η_t as well as the forecasts $q_t^{(-j)}$ of the other mediators are known to mediator j . Then the remaining uncertainty rests on the random transaction ξ_t of the noise-traders, because d_t was assumed to be \mathcal{F}_{t-1} measurable. Following Böhm & Wenzelburger (2000a) and Wenzelburger (2001a), all possible mean forecast errors on cum-dividend prices at a particular date t are given by the *mean error function* for mediator j , defined by

$$(10) \quad \mathcal{E}_t^{(j)} : \mathbb{R}_+^{HI} \times \mathbb{R}_+^{K(I+1)} \times \mathbb{R}_+^K \times \mathbb{R}_+^K \longrightarrow \mathbb{R}^K, \\ (\eta_t, (q_t^{(j)}, q_t^{(-j)}), d_t, q_{t-1}^{(j)}) \longmapsto \mathbb{E}_{t-1} [S(\xi_t, \eta_t, q_t^{(j)}, q_t^{(-j)})] + d_t - q_{t-1}^{(j)},$$

where the expectations operator \mathbb{E}_{t-1} is taken with respect to \mathcal{F}_{t-1} . The error function (10) describes all possible mean errors between actual cum-dividend prices and cum-dividend forecasts for mediator j conditioned on information available prior to trading in period t . An *unbiased forecasting rule* for j at date t , if it exists, is a map

$$(11) \quad \psi_t^{(j)} : \mathbb{R}_+^{HI} \times \mathbb{R}_+^{KI} \times \mathbb{R}_+^K \times \mathbb{R}_+^K \rightarrow \mathbb{R}_+^K, \quad q_t^{(j)} = \psi_t^{(j)}(\eta_t, q_t^{(-j)}, d_t, q_{t-1}^{(j)})$$

such that

$$\mathcal{E}_t^{(j)}(\eta_t, \psi_t^{(j)}(\eta_t, q_t^{(-j)}, d_t, q_{t-1}^{(j)}), q_t^{(-j)}, d_t, q_{t-1}^{(j)}) = 0$$

identically on $\mathbb{R}_+^{HI} \times \mathbb{R}_+^{K(I+1)} \times \mathbb{R}_+^K \times \mathbb{R}_+^K$. An unbiased forecasting rule (11) exists, if the error function (10) satisfies the conditions of a global implicit function theorem, which in turn stipulates the particular functional form of (11), see Böhm & Wenzelburger (2000a) and Wenzelburger (2001a). We will show below that an unbiased forecasting rule exists in the case in which all households have linear mean-variance preferences. An unbiased

forecasting rule for ex-dividend prices is obtained from (11) by setting $p_t^{(j)} = q_t^{(j)} - d_{t+1}$ as a forecast for p_{t+1} . It then follows from (10) that all forecast errors on ex-dividend prices for mediator j vanish identically, that is, $\mathbb{E}_t[p_{t+1}] - p_t^{(j)} = 0$ for all times t . In the sequel, an unbiased forecasting rule will also be referred to as a *perfect forecasting rule for first moments*.

The notion of a *perfect forecasting rule for second moments*, i.e., a forecasting rule which generates correct covariances matrices of the cum-dividend prices conditioned on the available information, could be derived using an analogous reasoning. We will postpone the discussion of such a rule to Sec. 6 when a more concrete example is available.

The natural question now is how an unbiased forecasting rule can be learned from historical data, if the rule is unknown. Let $\Phi^{(i)}(\eta^{(i)}, \nu^{(i)}, p) \equiv \Phi^{(i)}(\eta^{(i)}, q^{(i)}, p)$ for the respective aggregate demand functions (2), where we identify $\nu^{(i)}$ with $q^{(i)}$ as before. Assume, in addition, that for each $p \in \mathbb{R}_+^K$ and each $\eta^{(j)}$ the aggregate demand $\Phi^{(j)}(\eta^{(j)}, q^{(j)}, p)$ of mediator j is invertible with respect to the price forecasts $q^{(j)}$. Let $q_{t-1}^{(j)} \in \mathbb{R}_+^K$ and $\eta_t \in \mathbb{R}_+^{HI}$ be arbitrary and p_t denote the market-clearing price in period t . The market-clearing condition (3) then implies that an unbiased forecasting rule (11) for j yielding $q_t^{(j)}$ must satisfy

$$(12) \quad q_t^{(j)} = \Phi^{(j)-1}(\eta_t^{(j)}, p_t, \bar{x} + \xi_t - \sum_{i \neq j} \Phi^{(i)}(\eta_t^{(i)}, q_t^{(i)}, p_t)).$$

The functional relationship (12) shows that the existence of an unbiased forecasting rule depends exclusively on the fundamentals of the market mechanism, the subjective beliefs of all mediators who take part in the market, and on the way in which households choose their mediators. It reveals that the essential unknown quantity for a particular mediator j is the excess supply of all other market participants. Consequently, any mediator with rational expectations has to be able to replicate the demand behavior of his fellow mediators. The main informational constraint is the fact that neither the fraction of households joining a particular mediator nor the beliefs of the mediators themselves are observable quantities.

4 Risk premia

In order to evaluate the performance of a mediator and thus quality of the households' investment decisions, we analyze the risk premia obtained by mediators. Suppose that some investor, say mediator 0, has rational expectations but will not necessarily participate in the market such that the assumption of rational expectations for him is always

fulfilled.³ Let

$$(13) \quad x_t^{(0)} = \mathbb{V}_t[q_{t+1}]^{-1}[\mathbb{E}_t[q_{t+1}] - Rp_t]$$

denote the *reference portfolio* in period t . The portfolio $x_t^{(0)}$ may be held by mediator 0 after a fictitious trading in period t such that mediator 0 may be thought of being endowed with linear mean-variance preferences and a coefficient of risk aversion $\alpha^{(0)} = 1$, see Sec. 5. Investing one unit $e^{(0)} = 1$ of the consumption good into the risk-free and the risky assets, gives the portfolio $(1 - p_t^\top x_t^{(0)}, x_t^{(0)}) \in \mathbb{R} \times \mathbb{R}^K$, where $x_t^{(0)}$ is determined by (13). If $r = R - 1$ is the interest rate paid for the risk-free asset, then the return of the portfolio (13) obtained after a (fictitious) selling at prices in period $t + 1$ is

$$(14) \quad R_{t+1}^{(0)} = r + [q_{t+1} - Rp_t]^\top x_t^{(0)}.$$

The conditional variance of $R_{t+1}^{(0)}$ is $\mathbb{V}_t[R_{t+1}^{(0)}] = (x_t^{(0)})^\top \mathbb{V}_t[q_{t+1}] x_t^{(0)}$. Combined with (13), this shows that the *mean equity premium* of the reference portfolio (13) satisfies

$$(15) \quad \mathbb{E}_t[R_{t+1}^{(0)}] - r = \mathbb{V}_t[R_{t+1}^{(0)}]$$

and for this reason is always non-negative.

Let $x_t^{(i)}$ denote the portfolio held by a mediator $i = 1, \dots, I$ after investing $w_t^{(i)} = \sum_{h=1}^H \eta_t^{(hi)} e^h$ in period t . As before, the realized return resulting from selling $x_t^{(i)}$ in period $t + 1$ is

$$R_{t+1}^{(i)} = r + [q_{t+1} - Rp_t]^\top (x_t^{(i)} / w_t^{(i)}).$$

The conditional covariance $\mathbb{Cov}_t[R_{t+1}^{(i)}, R_{t+1}^{(0)}]$ between the return of mediator i and the return associated with (13) is

$$(16) \quad \mathbb{Cov}_t[R_{t+1}^{(i)}, R_{t+1}^{(0)}] = (x_t^{(i)} / w_t^{(i)})^\top \mathbb{V}_t[q_{t+1}] x_t^{(0)},$$

implying that the expected return of mediator i takes the form

$$(17) \quad \mathbb{E}_t[R_{t+1}^{(i)}] = r + \mathbb{Cov}_t[R_{t+1}^{(i)}, R_{t+1}^{(0)}].$$

Thus the mean equity premium of mediator i is equal to the covariance between i 's return and the return of a 'rational' mediator 0 with linear mean-variance preferences. Contrary to mediator 0, the equity premium of any mediator i with non-rational expectations may on average well be negative. Comparing (17) with (15) yields the following theorem.

³Mediator 0 can be treated as an outside observer with full information. As long as he does not interfere with the market by trading, the assumption of rational expectations poses no existence problem.

Theorem 4.1 *Let i be an arbitrary mediator. Assume that preferences and beliefs of all market participants are such that the asset market clears in each period t at prices p_t . Then for each $t \in \mathbb{N}$,*

$$\mathbb{E}_t[R_{t+1}^{(i)}] - r = \frac{\text{Cov}_t[R_{t+1}^{(i)}, R_{t+1}^{(0)}]}{\mathbb{V}_t[R_{t+1}^{(0)}]} [\mathbb{E}_t[R_{t+1}^{(0)}] - r].$$

Theorem 4.1 states that the mean equity premium $\mathbb{E}_t[R_{t+1}^{(i)}] - r$ of any mediator i can only be higher than the mean equity premium $\mathbb{E}_t[R_{t+1}^{(0)}] - r$ of the reference portfolio (13), if the variance of the corresponding return is greater than the variance of the reference portfolio, i.e. $\mathbb{V}_t[R_{t+1}^{(i)}] \geq \mathbb{V}_t[R_{t+1}^{(0)}]$. Observe that the result in Theorem 4.1 requires neither an assumption on the preferences of mediator i nor on the quality of the beliefs.

Using (14), the inverse coefficient of variation conditioned on information at date t associated with the return $R_{t+1}^{(i)}$ of mediator i 's portfolio $x_t^{(i)}$ is

$$(18) \quad \frac{\mathbb{E}_t[R_{t+1}^{(i)}] - r}{\sqrt{\mathbb{V}_t[R_{t+1}^{(i)}]}} = \frac{[\mathbb{E}_t[q_{t+1}] - Rp_t]^\top x_t^{(i)}}{\sqrt{(x_t^{(i)})^\top \mathbb{V}_t[q_{t+1}] x_t^{(i)}}}.$$

Observe that the coefficients (18) are invariant under rescaling of the portfolio $x_t^{(i)}$, that is, any portfolio $\tilde{x}_t^{(i)} = \lambda x_t^{(i)}$ with $\lambda \in \mathbb{R}$ will induce the same coefficient. Theorem 4.1 combined with (15) implies in particular that for each mediator $i = 0, \dots, I$, the inverse coefficients of variation satisfy

$$(19) \quad \frac{\mathbb{E}_t[R_{t+1}^{(i)}] - r}{\sqrt{\mathbb{V}_t[R_{t+1}^{(i)}]}} \leq \frac{\mathbb{E}_t[R_{t+1}^{(0)}] - r}{\sqrt{\mathbb{V}_t[R_{t+1}^{(0)}]}} = \sqrt{\mathbb{V}_t[R_{t+1}^{(0)}]} \quad \text{for all times } t.$$

Combining (13) with (15) yields

$$\mathbb{V}_t[R_{t+1}^{(0)}] = [\mathbb{E}_t[q_{t+1}] - Rp_t]^\top \mathbb{V}_t[q_{t+1}]^{-1} [\mathbb{E}_t[q_{t+1}] - Rp_t]$$

which shows that the upper bound in (19) is exclusively determined by the first two moments of the involved price process.

In our setting, Theorem 4.1 is a generalization of the famous *security market line result* (see e.g. Pliska 1997) to asset markets with heterogenous beliefs and preferences which are not necessarily of the mean-variance type. However, Theorem 4.1 shows that in a world of heterogeneous investors, perfect forecasting rules for first and second moments, if they exist, generate efficient portfolios for an investor with mean-variance preferences in the sense of classical CAPM theory. Since the Sharpe ratios given with (6) can be interpreted as estimators for the coefficients (18), households which switch according to the discrete choice model (7) may be seen as boundedly rational in the sense of CAPM theory.

5 Mean-variance preferences

Starting with the work by Markowitz (1952) and Tobin (1958), mean-variance preferences provide a popular parameterization of preferences under risk. In the linear case, it is shown in Böhm & Chiarella (2000) that the resulting aggregate demand function for shares is explicitly invertible. This allows for an explicit functional form of the market-clearing prices which is needed for a systematic analysis of the quality of an adaptive learning scheme. The relationship of mean-variance preferences to standard expected utility theory is given in the following definition (see e.g. Böhm & Chiarella 2000).

Definition 5.1 *Let $U : \mathbb{R}^2 \rightarrow \mathbb{R}$ denote a concave function which is strictly increasing in its first component and decreasing in its second. U is said to represent mean-variance preferences, if there exists a von-Neumann-Morgenstern utility function $u : \mathbb{R} \rightarrow \mathbb{R}$ and a class of probability measures $\mathcal{P} \subset \text{Prob}(\mathbb{R})$ such that*

$$\int_{\mathbb{R}} u(w) \nu(dw) = U(\mathbb{E}_{\nu}[w], \mathbb{V}_{\nu}[w]), \quad \text{for all } \nu \in \mathcal{P}.$$

Here, \mathbb{E}_{ν} and \mathbb{V}_{ν} denote the expected value and the variance with respect to ν .

As before, identify each measure $\nu_t^{(i)}$ with the respective parameterization $(q_t^{(i)}, V_t^{(i)})$ and let $(e^h - p^\top x, x) \in \mathbb{R}^{K+1}$ denote the portfolio associated with $x \in \mathbb{R}^K$. Then for each $(q_t^{(i)}, V_t^{(i)})$,

$$\mathbb{E}_t^{(i)}[w] = R e^h + (q_t^{(i)} - R p)^\top x$$

defines the subjective expected return of the portfolio conditioned on information available at time t prior to trading and $\mathbb{V}_t^{(i)}[w] = x^\top V_t^{(i)} x$ its subjective conditional variance. Let each household h be characterized by linear mean-variance preferences such that

$$(20) \quad U^h(\mathbb{E}_t^{(i)}[w], \mathbb{V}_t^{(i)}[w]) = \mathbb{E}_t^{(i)}[w] - \frac{\alpha^h}{2} \mathbb{V}_t^{(i)}[w]$$

describes the expected utility associated with the portfolio $x \in \mathbb{R}^K$ in period t , where α^h measures risk aversion.

Let $\eta_t^{(hi)}$ denote the fraction of households h which employ mediator i according to the discrete choice model (7). Then the aggregate demand (2) of all households h which employ i is given by

$$(21) \quad \Phi^{(i)}(a_t^{(i)}, q_t^{(i)}, V_t^{(i)}, p) = a_t^{(i)} V_t^{(i)-1} [q_t^{(i)} - R p],$$

where

$$(22) \quad a_t := (a_t^{(0)}, \dots, a_t^{(I)}) \quad \text{with} \quad a_t^{(i)} := \sum_{h=1}^H (\eta_t^{(hi)} / \alpha^h)$$

denotes the risk-adjusted fractions of households employing a particular mediator. Set

$$(23) \quad A_t := \left(\sum_{i=0}^I a_t^{(i)} V_t^{(i)-1} \right)^{-1}$$

and note that A_t is well defined due to the positive definiteness of all covariance matrices. From the market-clearing condition (3), one immediately obtains an explicit functional form of the ex-dividend price law (4) which for arbitrary beliefs $(q_t^{(i)}, V_t^{(i)})_{i=0}^I$ takes the form

$$(24) \quad p_t = S(\xi_t, a_t, (q_t^{(i)}, V_t^{(i)})_{i=0}^I) := \frac{1}{R} \left(\sum_{i=0}^I A_t^{(i)} q_t^{(i)} - A_t(\bar{x} + \xi_t) \right)$$

with

$$(25) \quad A_t^{(i)} := a_t^{(i)} [V_t^{(i)} A_t^{-1}]^{-1} = \begin{cases} 0 & \text{if } a_t^{(i)} = 0, \\ [Id + \sum_{j \neq i} (a_t^{(j)} / a_t^{(i)}) V_t^{(j)} V_t^{(j)-1}]^{-1} & \text{if } a_t^{(i)} > 0. \end{cases}$$

All coefficient matrices (25) and (23) in the price law (24) are determined by subjective covariances matrices $V_t^{(i)}$ and risk-adjusted distributions of households (22) and hence are \mathcal{F}_{t-1} measurable. Since (6) provides a recursive formula for the sample coefficients of returns, the distributions (22) can also be computed recursively.

The return $R_{t+1}^{(i)}$ obtained by mediator i in period $t+1$ after trading is

$$(26) \quad R_{t+1}^{(i)} = r + \frac{a_t^{(i)}}{w_t^{(i)}} [q_{t+1} - Rp_t]^\top V_t^{(i)-1} [q_t^{(i)} - Rp_t],$$

where $w_t^{(i)} = \sum_{h=1}^H \eta_t^{(hi)} e^h$ is the amount of capital provided to and invested by i . The realized returns (26) depend on the risk-adjusted fraction of households (22), the amount of provided capital, and on the subjective belief of the particular mediator. On the contrary, the inverse coefficients of variation given in (19) take the form

$$(27) \quad \frac{\mathbb{E}_t[R_{t+1}^{(i)}] - r}{\sqrt{\mathbb{V}_t[R_{t+1}^{(i)}]}} = \frac{[\mathbb{E}_t[q_{t+1}] - Rp_t]^\top V_t^{(i)-1} [q_t^{(i)} - Rp_t]}{\sqrt{[q_t^{(i)} - Rp_t]^\top \left(V_t^{(i)-1} \mathbb{V}_t[q_{t+1}] V_t^{(i)-1} \right) [q_t^{(i)} - Rp_t]}}$$

showing that these depend essentially on the beliefs of the mediator and not on the fraction of households. The behavior of households influence the corresponding coefficients of variation (27) only indirectly through the asset prices. In the special case of only one type of household, i.e. $H = 1$, one obtains a considerable simplification of (26), because each $a_t^{(i)} / w_t^{(i)} = 1 / (\alpha^1 e^1)$, $i = 0, \dots, I$ then is constant over time.

6 Perfect forecasting rules

In this section we investigate the existence of forecasting rules which generate rational expectations on cum-dividend prices for mediator 0 in the sense that the first two moments of the future asset price are correctly predicted. In what follows we assume $a_t^{(0)} > 0$ throughout this section.⁴ Recalling that d_t is assumed to be \mathcal{F}_{t-1} measurable, the expected cum-dividend price conditioned on information available at date t associated with the price law (24) is

$$(28) \quad \mathbb{E}_{t-1}[q_t] = \frac{1}{R} \left[\sum_{i=1}^I A_t^{(i)} q_t^{(i)} + A_t^{(0)} q_t^{(0)} - A_t(\bar{x} + \mathbb{E}_{t-1}[\xi_t]) \right] + d_t.$$

The condition that the forecast errors for mediator 0 vanish in the mean is

$$(29) \quad \mathbb{E}_{t-1}[q_t - q_{t-1}^{(0)}] = 0$$

for all times t . Inserting (28) into (29) and rearranging, yields an explicit expression for the new forecast $q_t^{(0)}$, given by

$$(30) \quad q_t^{(0)} = A_t^{(0)-1} \left[R(q_{t-1}^{(0)} - d_t) - \sum_{i=1}^I A_t^{(i)} q_t^{(i)} + A_t(\bar{x} + \mathbb{E}_{t-1}[\xi_t]) \right].$$

Replacing the coefficients in (30) with (25), the *unbiased forecasting rule* for mediator 0 takes the form

$$(31) \quad \begin{aligned} q_t^{(0)} &= \psi^{(0)}(\mathbb{E}_{t-1}[\xi_t], d_t, (a_t^{(i)}, q_t^{(i)}, V_t^{(i)})_{i=1}^I, q_{t-1}^{(0)}) \\ &:= R(q_{t-1}^{(0)} - d_t) \\ &\quad + a_t^{(0)-1} V_t^{(0)} \left[\bar{x} + \mathbb{E}_{t-1}[\xi_t] - \sum_{i=1}^I a_t^{(i)} V_t^{(i)-1} [q_t^{(i)} - R(q_{t-1}^{(0)} - d_t)] \right]. \end{aligned}$$

The functional form of the unbiased forecasting rule (31), also referred to as a *perfect forecasting rule for first moments*, is closely related to that of (12) in Sec. 3. The only quantity in (31) which is not observable for mediator 0 is the term in the large brackets. Since by (29) $\mathbb{E}_{t-1}[p_t] = q_{t-1}^{(0)} - d_t$, this term describes precisely the expected excess supply of all mediators $i > 0$ including the expected transactions of the noise traders $\mathbb{E}_{t-1}[\xi_t]$. Thus, the functional form (31) of the unbiased forecasting rule supports the idea that precise forecasts require good knowledge of the investment behavior of all market participants.

⁴As noted before, this imposes no restriction on the rationality of 0, because $a_t^{(0)} = 0$ simply means that no household employs mediator 0 such that his beliefs do not feed back into the price process. In this case the perfect forecasting rules coincide with the first two moments of the resulting price process.

While in the case of linear mean-variance preferences perfect forecasting rules for first moments exist generically, the requirements for forecasting rules which generate correct second moments of the price process are much stricter. Such rules will be referred to as *perfect forecasting rules for second moments*. Observe that the subjective covariance matrix $V_t^{(0)}$ of mediator 0 in the expression for (31) was as yet arbitrary. We will now try to stipulate this matrix as to provide correct second moments of the price process.

Assuming that ξ_t and d_t are uncorrelated, the conditional covariance matrix of the cum-dividend prices are obtained from the price law (24) and (28) yielding

$$(32) \quad \mathbb{V}_{t-1}[q_t] = \frac{1}{R^2} A_t \mathbb{V}_{t-1}[\xi_t] A_t.$$

Recall that $A_t = \left(\sum_{i=0}^I a_t^{(i)} V_t^{(i)-1} \right)^{-1}$ is positive definite implying that (32) is invertible, if $\mathbb{V}_{t-1}[\xi_t]$ is non-degenerate. That is, the covariance of the price process receives an expectations feedback from the subjective covariance matrices $(V_t^{(i)})_{i=0}^I$ as long as there is noise in the system, i.e. $\mathbb{V}_{t-1}[\xi_t]$ is different from the zero matrix. This implies that the volatility in the price process described by (32) is exclusively generated by the noise-trader behavior, the switching behavior of households, and the subjective covariance matrices of the mediators. The condition that mediator 0's forecast errors for second moments (32) of the price process vanish for all times t is

$$(33) \quad \mathbb{V}_{t-1}[q_t] - V_{t-1}^{(0)} = 0.$$

Inserting (32) into (33) gives the necessary condition

$$(34) \quad \left(\sum_{i=0}^I a_t^{(i)} V_t^{(i)-1} \right) V_{t-1}^{(0)} \left(\sum_{i=0}^I a_t^{(i)} V_t^{(i)-1} \right) = \frac{1}{R^2} \mathbb{V}_{t-1}[\xi_t]$$

for the subjective covariance matrix $V_t^{(0)}$. In other words, perfect forecasting rules for second moments are given by symmetric positive definite solutions $V_t^{(0)}$ to (34). Before we provide conditions under which these forecasting rules exist, we establish a technical lemma.

Lemma 6.1 *Let B, C be arbitrary symmetric positive definite matrices. Then*

$$D := \sqrt{B^{-1}} \sqrt{(\sqrt{B} C \sqrt{B})} \sqrt{B^{-1}}$$

is a symmetric positive definite matrix with $C = DBD$.

Proof. Since B is symmetric and positive definite, it can be diagonalized such that $B = O^\top \text{diag}(\lambda_1, \dots, \lambda_K) O$ for real eigenvalues $\lambda_1, \dots, \lambda_K > 0$ and some orthogonal matrix O , i.e. $O^\top O = Id$. The square root \sqrt{B} of B is a symmetric positive definite

matrix, defined by $\sqrt{B} = O^\top \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_K})O$, such that $B = \sqrt{B}\sqrt{B}$. Define $F := \sqrt{\sqrt{B^{-1}}C\sqrt{B^{-1}}}$ to see that $D := \sqrt{B^{-1}}F\sqrt{B^{-1}}$ is a symmetric positive definite solution to

$$C = DBD.$$

Q.E.D.

Proposition 6.2 *Let $\mathbb{V}_{t-1}[\xi_t]$ be non-degenerate and $a_t^{(0)} > 0$. Suppose that $V_t^{(0)}$, given by the forecasting rule $\varphi^{(0)}$ for second moments*

$$(35) \quad V_t^{(0)} = \varphi^{(0)}(\mathbb{V}_{t-1}[\xi_t], (a_t^{(i)}, V_t^{(i)})_{i=1}^I, V_{t-1}^{(0)}) := a_t^{(0)} \left(\Pi_t - \sum_{i=1}^I a_t^{(i)} V_t^{(i)-1} \right)^{-1},$$

where

$$\Pi_t = \frac{1}{R} \sqrt{V_{t-1}^{(0)-1}} \sqrt{\left(\sqrt{V_{t-1}^{(0)}} \mathbb{V}_{t-1}[\xi_t] \sqrt{V_{t-1}^{(0)}} \right)} \sqrt{V_{t-1}^{(0)-1}},$$

is well defined and positive definite. Then $\varphi^{(0)}$ provides correct second moments of the price process at date t , that is, (33) is satisfied in period t . If, in addition $\mathbb{V}_{t-1}[\xi_t] \equiv \sigma_\xi^2 Id$, then $\Pi_t = \frac{\sigma_\xi}{R} \sqrt{V_{t-1}^{(0)-1}}$.

Proof. By Lemma 6.1, Π_t as given above is a solution to the equation

$$\Pi_t V_{t-1}^{(0)} \Pi_t = \frac{1}{R^2} \mathbb{V}_{t-1}[\xi_t].$$

If $\Pi_t - \sum_{i=1}^I a_t^{(i)} V_t^{(i)-1}$ is positive definite, then (35) is well defined and the assertion follows from (34). *Q.E.D.*

If the r.h.s. of (35) is neither well defined nor positive definite, the resulting expression on the left defines no covariance matrix such that a forecasting rule (35) which is perfect for all times t may fail to exist. Hence additional requirements which guarantee that (35) is well defined in all periods are needed. Lemma 6.3 shows that the case in which all mediators agree upon subjective second moments meets such a requirement.

Lemma 6.3 *Let $\mathbb{V}_{t-1}[\xi_t]$ be non-degenerate and suppose $V_t^{(i)} \equiv V_t^{(0)}$ for all $i = 1, \dots, I$ and all times t . Then*

$$V_t^{(0)} := \left(R \sum_{h=1}^H \frac{1}{\alpha^h} \right) \sqrt{V_{t-1}^{(0)}} \sqrt{\left(\sqrt{V_{t-1}^{(0)}} \mathbb{V}_{t-1}[\xi_t] \sqrt{V_{t-1}^{(0)}} \right)^{-1}} \sqrt{V_{t-1}^{(0)}}$$

is a forecasting rule which is perfect for second moments. If $\mathbb{V}_{t-1}[\xi_t] \equiv \Xi$ is constant over time, then the constant rule $V_t^{(0)} \equiv \left(R \sum_{h=1}^H \frac{1}{\alpha^h} \right) \Xi^{-1}$ is perfect as well.

Proof. We have $A_t^{-1} = \sum_{h=1}^H \frac{1}{a^h} V_t^{(0)-1}$ and both statements follow from (34).

Q.E.D.

Another special situation in which (35) is perfect for all times t arises when all mediators $i > 0$ believe in constant covariance matrices.

Proposition 6.4 *Let $\mathbb{V}_{t-1}[\xi_t] \equiv \sigma_\xi^2 Id$ with a positive number $\sigma_\xi^2 > 0$ and $a_t^{(0)} > 0$. Assume that the following hypotheses hold true:*

- (i) *all mediators $i > 0$ use constant covariance matrices $V_t^{(i)} \equiv V^{(i)}$ such that $V^{(i)-1} = O^\top \text{diag}(\lambda^{(i1)}, \dots, \lambda^{(iK)}) O$, $i = 1, \dots, I$ for some orthogonal $K \times K$ matrix O ;*
- (ii) *there exists a constant $c < \frac{\sigma_\xi^2}{4R^2}$, such that the risk-adjusted distributions of households satisfy*

$$a_t^{(0)} \sum_{i=1}^I a_t^{(i)} \lambda^{(ik)} \leq c, \quad k = 1, \dots, K, \quad \mathbb{P}\text{-a.s. for all times } t.$$

Set $V_0^{(0)} := O^\top \text{diag}(\lambda_0^{(01)-1}, \dots, \lambda_0^{(0K)-1}) O$, where $\lambda_0^{(01)}, \dots, \lambda_0^{(0K)} > 0$ and O is given in (i). Then the forecasting rule (35) is perfect for second moments provided that the initial eigenvalues $\lambda_0^{(01)}, \dots, \lambda_0^{(0K)}$ are suitably large.

Proof. We seek a symmetric matrix $V_t^{(0)}$ which satisfies (34) and which stays positive definite for all times t . Under the given hypotheses, this is equivalent to the existence of a symmetric and positive matrix $V_t^{(0)}$ with

$$(36) \quad a_t^{(0)} V_t^{(0)-1} = \frac{\sigma_\xi}{R} \sqrt{V_{t-1}^{(0)-1}} - \sum_{i=1}^I a_t^{(i)} V_t^{(i)-1}.$$

Choose $V_0^{(0)}$ as in (iii). Then (36), the definition of the square root and induction imply that

$$V_t^{(0)-1} = O^\top \text{diag}(\lambda_t^{(01)}, \dots, \lambda_t^{(0K)}) O,$$

where

$$(37) \quad \lambda_t^{(0k)} = \frac{\sigma_\xi}{Ra_t^{(0)}} \sqrt{\lambda_{t-1}^{(0k)}} - \sum_{i=1}^I \frac{a_t^{(i)}}{a_t^{(0)}} \lambda^{(ik)}, \quad k = 1, \dots, K.$$

Equation (37) is a system of K strictly monotonically increasing and strictly concave scalar maps with random coefficients of the type $\lambda_t = b_1 \sqrt{\lambda_{t-1}} - b_2$. These determine the eigenvalues of the covariance matrix $V_t^{(0)}$. Condition (ii) now guarantees that each of the maps (37) has two positive deterministic fixed points, of which the larger one is asymptotically stable. Hence, there exists a compact set $\Lambda \subset \mathbb{R}_{++}^K$ which is forward invariant under the maps (37) such that $(\lambda_t^{(01)}, \dots, \lambda_t^{(0K)}) \in \Lambda$ for all times $t \geq 0$. *Q.E.D.*

7 Dynamics under rational expectations

In this section we discuss the dynamics of asset prices when, under the hypotheses of Sec. 6, mediator 0 has rational expectations in the sense that first moments of cum-dividend prices are correctly predicted for all times t and second moments whenever possible. Assuming that each mediator $i > 0$ uses some forecasting rules $\psi^{(i)}$ and $\varphi^{(i)}$, the prices under rational expectations for group 0 are determined by

$$(38) \quad \begin{cases} q_t &= q_{t-1}^{(0)} + \frac{1}{R} A_t (\xi_t - \mathbb{E}_{t-1}[\xi_t]) \\ q_t^{(i)} &= \psi^{(i)}(q_{t-1}, \dots, q_{t-L}) \quad \forall i = 1, \dots, I, \\ V_t^{(i)} &= \varphi^{(i)}(q_{t-1}, \dots, q_{t-L}) \quad \forall i = 1, \dots, I, \\ q_t^{(0)} &= \psi^{(0)}(\mathbb{E}_{t-1}[\xi_t], d_t, (a_t^{(i)}, q_t^{(i)}, V_t^{(i)})_{i=1}^I, q_{t-1}^{(0)}), \\ V_t^{(0)} &= \varphi^{(0)}(\mathbb{V}_{t-1}[\xi_t], (a_t^{(i)}, V_t^{(i)})_{i=1}^I, V_{t-1}^{(0)}), \end{cases}$$

where $\varphi^{(0)}$ coincides with (35) whenever (35) is well defined. The cum-dividend price q_t in (38) is essentially determined by the forecasting rule $\psi^{(0)}$ of mediator 0 which takes the form

$$(39) \quad \begin{aligned} q_t^{(0)} &= R \left[Id + \sum_{i=0}^I \frac{a_t^{(i)}}{a_t^{(0)}} V_t^{(0)} V_t^{(i)-1} \right] q_{t-1}^{(0)} - \sum_{i=1}^I \frac{a_t^{(i)}}{a_t^{(0)}} V_t^{(0)} V_t^{(i)-1} q_t^{(i)} \\ &\quad + a_t^{(0)-1} V_t^{(0)} \left[\bar{x} + \mathbb{E}_{t-1}[\xi_t] - \sum_{i=0}^I a_t^{(i)} V_t^{(i)-1} d_t \right]. \end{aligned}$$

The map (38) together with the discrete choice model (7) with (22), defines a time-one map of a random dynamical system in the sense of Arnold (1998) which describes the evolution of the asset prices. The map (38) will, in general, be nonlinear, because nonlinearities may enter the system through forecasting rules $\varphi^{(i)}$ for the covariance matrices as well as through nonlinear forecasting rules $\psi^{(i)}$, $i > 0$.

A random dynamical system is complete only, once a model for the exogenous perturbations, in our case the behavior of noise traders and the dividend process, is specified. Following Arnold (1998), we represent the exogenous perturbations by means of an ergodic metric dynamical system. This amounts to using the (equivalent) canonical realization process of an ergodic stochastic process. To be specific, let $\{\epsilon_t\}_{t \in \mathbb{Z}}$, where $\mathbb{Z} := \{0, \pm 1, \pm 2, \dots\}$ is called two-sided time, be a sequence of Σ -valued i.i.d. random variables whose probability distribution is μ supported on some subset $\Sigma \subset \mathbb{R}^K$. Then $(\Omega, \mathcal{F}, \mathbb{P}) = (\Sigma^{\mathbb{Z}}, \mathcal{B}(\Sigma)^{\mathbb{Z}}, \mu^{\mathbb{Z}})$, where Ω is the space of all sample paths of the process, $\mathcal{B}(\Sigma)^{\mathbb{Z}}$ is the Borel σ -algebra of all cylinder sets, and $\mu^{\mathbb{Z}}$ is the product measure. Each element $\omega = \{\omega(s)\}_{s \in \mathbb{Z}} \in \Omega$ is a doubly infinite series describing a sample path of the process. The map $\vartheta : \Omega \rightarrow \Omega$, $\omega \mapsto \vartheta\omega$ is defined by $(\vartheta\omega)(s) := \omega(s+1)$, $s \in \mathbb{Z}$ and is called the left shift such that $(\vartheta^t\omega)(s) = \omega(s+t)$ for all $t, s \in \mathbb{Z}$. With the help of the evaluation map $\epsilon : \Omega \rightarrow \Sigma$, $\epsilon(\omega) := \omega(0)$, we have $\epsilon_t(\omega) = \epsilon(\vartheta^t\omega)$, $t \in \mathbb{Z}$ and

thus obtain a representation of the original i.i.d. process as an *ergodic metric dynamical system* $(\Omega, \mathcal{F}, \mathbb{P}, \{\vartheta^t\}_{t \in \mathbb{Z}})$. It is well known that any stationary ergodic process $\{\epsilon_t\}_{t \in \mathbb{Z}}$ can be represented by an ergodic metric dynamical system, see Arnold (1998, Appendix A) for more details.

A brief outline of the notion of a random dynamical system is given in Appendix A. In order to obtain a linear benchmark dynamics, assume for the remainder of this section that there are only two mediators $i = 0, 1$. Let mediator 1 be a chartist who uses the simple *technical trading rule*

$$(40) \quad q_t^{(1)} = \psi^{(1)}(q_{t-1}, \dots, q_{t-L}) := \sum_{l=1}^L D^{(l)} q_{t-l}$$

as a forecasting rule, where $D^{(l)} = \text{diag}(\delta_1^{(l)}, \dots, \delta_K^{(l)})$, $l = 1, \dots, L$ are diagonal matrices whose non-zero entries denote the expected weighted trends of the asset prices, respectively.⁵ In addition, let all subjective covariance matrices be constant over time, i.e. $V_t^{(i)} \equiv V^{(i)}$, $i = 0, 1$. Suppose that the distributions of households are constant over time, that is $a_t^{(i)} \equiv a^{(i)}$, $i = 0, 1$. This case, for instance, obtains, if the intensity of choice parameters in (7) are all set to zero and no household switches between mediators. Then, under the hypotheses of Proposition 6.4, mediator 0 may have correct second moments which are constant over time, since no switching occurs. If $Q_t = (q_t^{(0)}, q_t, \dots, q_{t+1-L}) \in \mathbb{R}_+^{K+1}$ denotes the vector consisting of the last price forecast of mediator 0 and past L realized prices, then (38) defines an ARMAX process

$$(41) \quad Q_t = \mathcal{A}Q_{t-1} + B_t,$$

where the coefficient (block) matrix \mathcal{A} is given by

$$(42) \quad \mathcal{A} = \begin{pmatrix} R(Id + \frac{a^{(1)}}{a^{(0)}}V^{(0)}V^{(1)-1}) & \frac{-a^{(1)}}{a^{(0)}}V^{(0)}V^{(1)-1}D^{(1)} & \dots & \dots & \dots & \frac{-a^{(1)}}{a^{(0)}}V^{(0)}V^{(1)-1}D^{(L)} \\ Id & 0 & \dots & \dots & \dots & 0 \\ 0 & Id & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & \dots & Id & 0 \end{pmatrix}$$

and

$$(43) \quad B_t = (a^{(0)-1}V^{(0)}[\bar{x} + \mathbb{E}_{t-1}[\xi_t] - A d_t], \frac{1}{R}A(\xi_t - \mathbb{E}_{t-1}[\xi_t]), 0, \dots, 0)$$

with $A = [a^{(0)}V^{(0)-1} + a^{(1)}V^{(1)-1}]^{-1}$.

Now assume that the three processes $\{\xi_t\}_{t \in \mathbb{Z}}$, $\{\mathbb{E}_{t-1}[\xi_t]\}_{t \in \mathbb{Z}}$, and $\{d_t\}_{t \in \mathbb{Z}}$ can be represented by an ergodic metric dynamical system $(\Omega, \mathcal{F}, \mathbb{P}, \{\vartheta^t\}_{t \in \mathbb{Z}})$ as above. This implies

⁵This corresponds to trend traders in $K \geq 1$ risky assets as in Brock & Hommes (1998) and Chiarella & He (2000).

that the stochastic process (43) can be written as $B_t(\omega) = B(\vartheta^t\omega)$ with some measurable map $B(\omega)$ and that (41) is an *affine random dynamical system* in the sense of Arnold (1998, Chap. 5). A stationary solution to the stochastic difference equation (41) is given by a so-called⁶ *random fixed point* $Q^*(\omega)$ which is defined by the random variable

$$(44) \quad Q^*(\omega) = \sum_{s=0}^{\infty} \mathcal{A}^s B(\vartheta^{-(s+1)}\omega), \quad \omega \in \Omega.$$

Under certain conditions given in the next proposition, the random fixed point $Q^*(\omega)$ is said to be globally asymptotically stable, implying that \mathbb{P} -almost all solutions to the equation (41) will eventually behave like the stochastic process $\{Q^*(\vartheta^t\omega)\}_{t \in \mathbb{N}}$ induced by (44). See Appendix A for the basic definitions and for an existence theorem for the general non-linear case.

Proposition 7.1 *Assume that the following hypotheses are satisfied.*

- (i) $\{\xi_t\}_{t \in \mathbb{Z}}$, $\{\mathbb{E}_{t-1}[\xi_t]\}_{t \in \mathbb{Z}}$, and $\{d_t\}_{t \in \mathbb{Z}}$ can be represented by an ergodic metric dynamical system $(\Omega, \mathcal{F}, \mathbb{P}, \{\vartheta^t\}_{t \in \mathbb{Z}})$, where $\mathbb{E}[\xi_t] \equiv 0$ and $\mathbb{E}[d_t] \equiv \bar{d}$ with $\|\bar{d}\| < \infty$;
- (ii) $V_t^{(i)} \equiv V^{(i)}$, $i = 0, 1$ such that the coefficient matrix \mathcal{A} given in (42) is constant;
- (iii) $\det(\mathcal{A} - \lambda \mathcal{I}) \neq 0$ for all $\lambda \in \mathbb{C}$ with $|\lambda| \geq 1$, that is, all eigenvalues of \mathcal{A} lie inside the unit disk.

Then the random fixed point $Q^*(\omega)$ defined in (44) is well defined, uniquely determined, and globally asymptotically stable. The mean price level $\bar{q} \in \mathbb{R}_+^K$ is the uniquely determined solution of the linear equation

$$[[Id + (a^{(0)}/a^{(1)})V^{(1)}V^{(0)-1}]^{-1}(D - Id) - r Id]\bar{q} = A\bar{x} - \bar{d},$$

where $\mathbb{E}Q^* = (\bar{q}, \dots, \bar{q})$, $D = \sum_{l=1}^L D^{(l)}$, and $A = [a^{(0)}V^{(0)-1} + a^{(1)}V^{(1)-1}]^{-1}$.

Proof. The stability property follows from Arnold (1998, Corollary 5.6.6). The second statement is due to the fact that $\mathcal{I} - \mathcal{A}$ is invertible such that $(\mathcal{I} - \mathcal{A})\mathbb{E}Q^* = \mathbb{E}B$ has a unique solution of the form $\mathbb{E}Q^* = (\bar{q}, \dots, \bar{q})$. *Q.E.D.*

Another special case arises when the two mediators agree on the subjective covariance matrices, such that $V_t^{(0)} \equiv V_t^{(1)}$ for all times t . This implies that all $A_t^{(i)}$, $i = 0, 1$ are time-invariant diagonal matrices and the map (39) takes the form

$$(45) \quad q_t^{(0)} = R \left(1 + \frac{a^{(1)}}{a^{(0)}} \right) q_{t-1}^{(0)} - \frac{a^{(1)}}{a^{(0)}} \sum_{l=1}^L D^{(l)} q_{t-l} + a^{(0)-1} V_t^{(0)} (\bar{x} + \mathbb{E}_{t-1}[\xi_t]) - \left(1 + \frac{a^{(1)}}{a^{(0)}} \right) d_t.$$

⁶In classical terminology, a random fixed point of (41) corresponds to a ‘*steady state solution*’ of a linear stochastic system, see Hannan & Deistler (1988, Chap. 1).

Since all $D^{(l)}$, $l = 1, \dots, L$ are diagonal matrices, (45) splits into K scalar equations which are coupled only via $V_t^{(0)}$. The stability properties of (38) are now obtained from an analysis of each scalar equation for each asset $k = 1, \dots, K$.

Corollary 7.2 *Assume that $V_t^{(0)} \equiv V_t^{(1)}$ with $\mathbb{E}[\|V_t^{(0)}\|] < \infty$. Then the price process for the k -th asset is globally asymptotically stable, if all zeros of the characteristic polynomial χ^k associated with (45), given by*

$$\chi^k(\lambda) = \lambda^{L+1} - R \left(1 + \frac{a^{(1)}}{a^{(0)}}\right) \lambda^L + \frac{a^{(1)}}{a^{(0)}} \sum_{l=1}^L \delta_k^{(l)} \lambda^{L-l},$$

lie inside the unit circle. The mean price level of the k -th asset \bar{q}^k is given by

$$\bar{q}^k = \frac{\frac{y^k}{a^{(0)}+a^{(1)}} - \bar{d}^k}{\frac{a^{(1)}}{a^{(0)}+a^{(1)}} \left(\sum_{l=1}^L \delta_k^{(l)} - 1\right) - r},$$

where y^k is the k -th entry of the vector $y = \mathbb{E}[V_t^{(0)}] \bar{x} \in \mathbb{R}^K$ and \bar{d}^k is the mean of the dividend payment of the k -th asset.

Since the sum $a^{(0)} + a^{(1)} = \sum_{h=1}^H \frac{1}{\alpha^h}$ is constant, Corollary 7.2 shows in particular that an asset price decreases on average with an increasing fraction of chartists $a^{(1)}$. From Jury's test (see Elaydi 1996, p. 181) it is known that a necessary condition for the zeros of χ^k to lie inside the unit circle is

$$\chi^k(0) > 0 \quad \text{and} \quad (-1)^{L+1} \chi^k(-1) > 0.$$

If $\delta_1^k = \dots = \delta_L^k$ are equal to some δ^k , then a routine calculation shows that this is the case precisely when

$$\frac{a^{(1)}}{a^{(0)}} (\delta^k - R) > R - 1.$$

8 Adaptive learning with heterogeneous beliefs

8.1 The general idea

The discussion in Sec. 6 revealed that the primary object which mediator 0 has to know in order to obtain an unbiased forecasting rule is the aggregate supply function of all market participants. We first introduce the general idea of estimating an unbiased

forecasting rule from historical data and then turn to the particular example with the two mediators treated in Sec. 7. Using the notation of Sec. 2, let

$$(46) \quad \Phi(\xi_t, (\eta_t^{(i)}, q_t^{(i)}, V_t^{(i)})_{i=1}^I, p) := \bar{x} + \xi_t - \sum_{i=1}^I \Phi^{(i)}(\eta_t^{(i)}, q_t^{(i)}, V_t^{(i)}, p)$$

denote the aggregate excess supply of all mediators except mediator 0. Assume that each mediator $i > 0$ uses forecasting rules $\psi^{(i)}$ and $\varphi^{(i)}$ of the functional form

$$q_t^{(i)} = \psi^{(i)}(q_{t-1}, \dots, q_{t-L}) \quad \text{and} \quad V_t^{(i)} = \varphi^{(i)}(q_{t-1}, \dots, q_{t-L}).$$

Inserting these forecasting rules into (46) and suppressing the respective functional expressions for notational convenience, the excess supply (46) takes the form

$$(47) \quad \Phi(\xi_t, (\eta_t^{(i)})_{i=1}^I, p, q_{t-1}, \dots, q_{t-L}).$$

Let $x_t^{(0)}$ denote the portfolio held by mediator 0 after trading in period t . Then the market-clearing condition (3) implies that

$$(48) \quad x_t^{(0)} = \Phi(\xi_t, (\eta_t^{(i)})_{i=1}^I, p_t, q_{t-1}, \dots, q_{t-L}) \quad \text{for each } t \in \mathbb{N}.$$

Observe that the prices and the portfolios $x_t^{(0)}$ appearing in (48) are observable quantities, whereas the distribution of households $(\eta_t^{(i)})_{i=1}^I$ along with the realization ξ_t of the noise process is not observable for mediator 0. Moreover, the concrete functional form (47) including the involved forecasting rules are unknown to mediator 0.

Assume for a moment that a good approximation for the discrete choice model (7) with unknown intensity of choice parameters has been developed along with a model for the exogenous noise process $\{\xi_t\}_{t \in \mathbb{N}}$. Suppose that these models yield sufficiently good estimates $\hat{\eta}_t^{(i)}$ for $\eta_t^{(i)}$, $i = 1, \dots, I$ and $\hat{\xi}_t$ for ξ_t , respectively. Then the following adaptive estimation procedure is at hand. At an arbitrary date t , first estimate the functional relationship (48) from past observed or estimated data $\{x_s^{(0)}, \hat{\xi}_s, (\hat{\eta}_s^{(i)})_{i=1}^I, p_s, q_s\}_{s=0}^{t-1}$. This yields an approximation $\hat{\Phi}_t$ of (47) such that

$$(49) \quad \hat{\Phi}_t(\hat{\xi}_t, (\hat{\eta}_t^{(i)})_{i=1}^I, p, q_{t-1}, \dots, q_{t-L})$$

describes the approximate excess supply function for assets in period t . Second, assuming the asset demand function $\Phi^{(0)}$ of mediator 0 to be invertible with respect to expected prices, a possible candidate for an approximation of the unbiased forecasting rule (12) in period t is then given by

$$(50) \quad \begin{aligned} q_t^{(0)} &= \hat{\psi}_t^{(0)}(\hat{\xi}_t, \eta_t^{(0)}, (\hat{\eta}_t^{(i)})_{i=1}^I, q_{t-1}^{(0)}, q_{t-1}, \dots, q_{t-L}) \\ &:= \Phi^{(0)-1}(\eta_t^{(0)}, q_{t-1}^{(0)}, \hat{\Phi}_t(\hat{\xi}_t, (\eta_t^{(i)})_{i=1}^I, q_{t-1}^{(0)}, q_{t-1}, \dots, q_{t-L})). \end{aligned}$$

This forecasting scheme can be carried out at each point in time as soon as a technique to estimate (49) is available. It will be shown in the next section that this procedure may be successful in the case of linear mean-variance preferences.

8.2 The case of two mediators

Consider the case of two mediators treated in Sec. 7, where we will allow the households to switch between mediators according to (7). The proposed learning scheme now takes a particularly simple form. Assume to this end that the estimates for the fraction of households are sufficiently good such that we can abstract from prediction errors of the distribution of households. This assumption seems not to be too strong, because the sum $\eta_t^{(h0)} + \eta_t^{(h1)} = 1$ is constant over time for all $1 \leq h \leq H$ and each $\eta_t^{(h0)}$ is known to mediator 0. On the other hand, for the unknown choice intensities in (7) there are maximum-likelihood estimators available which are consistent under standard regularity assumptions. In the present case of our discrete choice model of the LOGIT type, these estimators are found in Judge, Griffiths, Hill, Lütkepohl & Lee (1985, Chap. 18, p. 765). The estimators for the choice intensities yield estimates $\hat{\eta}_t^{(h1)}$, $h = 1, \dots, H$ for the fractions of households $\eta_t^{(h1)}$, $h = 1, \dots, H$ joining mediator 1. The informational constraints faced by mediator 0 are now summarized as follows.

Assumption 8.1 *The information of mediator 0 encompasses the following:*

- (i) *the realized returns $R_t^{(1)}$ of mediator 1 as given by (26) are observable quantities; a discrete choice model for the switching behavior of households has been estimated, consistent estimates of the fraction of households are known, such that, for simplicity of exposition, prediction errors are neglected and $\hat{\eta}_t^{(h1)} = \eta_t^{(h1)}$, $h = 1, \dots, H$.*
- (ii) *the total amount of retradable assets \bar{x} is known;*
- (iii) *the correct functional form of mediator 1's forecasting rules is known; these are a simple technical trading rule of the type (40) and a constant subjective covariance matrix.*

Under Assumption 8.1, mediator 0 is left to estimate the excess supply (49) of mediator 1. In order for any learning scheme to be successful, the behavior of the noise traders may not be too arbitrary. We therefore place the following *identifiability assumption* (see, e.g. Lai & Wei 1986, p. 247) on the unobservable noise process which governs the behavior of the noise traders.

Assumption 8.2 *The noise-trader behavior is driven by the process $\{\xi_t\}_{t \in \mathbb{Z}}$, given by*

$$\xi_t = \epsilon_t + \sum_{n=1}^N C^{(n)} \epsilon_{t-n},$$

where $\{\epsilon_t\}_{t \in \mathbb{Z}}$ is an i.i.d. process on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with zero means $\mathbb{E}[\epsilon_t] = 0$ which satisfies the following:

- (1) $\sup_{t \geq 1} \mathbb{E}[\|\epsilon_t\|^\alpha] < \infty$ \mathbb{P} -a.s. for some $\alpha > 2$;
- (2) $C^{(1)}, \dots, C^{(N)}$, $N \geq 0$ are $K \times K$ matrices and the matrix polynomial $\Gamma(z) = Id + C^{(1)}z + \dots + C^{(N)}z^N$, $z \in \mathbb{C}$ is strictly positive real, that is,
- (i) $\det \Gamma(z) \neq 0$ for all $z \in \mathbb{C}$ with $|z| \leq 1$;
- (ii) For each $z \in \mathbb{C}$ with $|z| \leq 1$, the matrix $\Gamma(z) + \Gamma^\top(\bar{z})$, with $\bar{z} \in \mathbb{C}$ denoting the complex conjugate of z , is strictly positive definite, that is, $x^\top(\Gamma(z) + \Gamma^\top(\bar{z}))x > 0$ for all $0 \neq x \in \mathbb{R}^K$.

Observe that $\mathbb{E}_{t-1}[\xi_t] = \sum_{n=1}^N C^{(n)}\epsilon_{t-n}$ and $\mathbb{V}_{t-1}[\xi_t] = \mathbb{V}_{t-1}[\epsilon_t]$ and that Assumption 8.2 includes the i.i.d. case (take $N = 0$ and $\{\epsilon_t\}_{t \in \mathbb{N}}$ i.i.d.). By Assumption 8.1, mediator 0 has correct knowledge about the functional form of the forecasting rule (40) used by mediator 1 but not its concrete parameterization. Set $x_t = x_t^{(0)} - \bar{x}$ to adjust mediator 0's portfolio by the total amount of available assets. Then the market-clearing condition (48) for period t reads

$$(51) \quad x_t = - \sum_{h=1}^H \sum_{l=1}^L \frac{\eta_t^{(h1)}}{\alpha^h} V^{(1)-1} [D^{(l)} q_{t-l} - R p_t] + \epsilon_t + \sum_{n=1}^N C^{(n)} \epsilon_{t-n}.$$

Setting

$$\theta = (A^{(11)}, \dots, A^{(HL)}, B^{(1)}, \dots, B^{(H)}, C^{(1)}, \dots, C^{(N)})$$

for the unknown coefficient matrices, where

$$A^{(hl)} = \frac{-1}{\alpha^h} V^{(1)-1} D^{(l)}, \quad B^{(h)} = \frac{R}{\alpha^h} V^{(1)-1} \quad \text{for all } h = 1, \dots, H, \quad l = 1, \dots, L$$

and

$$(52) \quad y_t = ((\eta_t^{(11)} q_{t-1})^\top, \dots, (\eta_t^{(H1)} q_{t-L})^\top, (\eta_t^{(11)} p_t)^\top, \dots, (\eta_t^{(H1)} p_t)^\top, \epsilon_{t-1}^\top, \dots, \epsilon_{t-N}^\top)^\top,$$

the market-clearing condition (51) may be rewritten as

$$(53) \quad x_t = \theta y_t + \epsilon_t.$$

We use the representation (53) to estimate the unknown coefficient matrix θ . This is done by applying the so-called *approximate-maximum-likelihood (AML)* algorithm. This algorithm is an extension of the ordinary-least-square (OLS) algorithm as usually applied in the learning literature (see e.g. Evans & Honkapohja 2001) to a class of more general noise processes. The methods provides a recursive scheme which generates successive estimates

$$(54) \quad \hat{\theta}_t = (\hat{A}_t^{(11)}, \dots, \hat{A}_t^{(HL)}, \hat{B}_t^{(1)}, \dots, \hat{B}_t^{(H)}, \hat{C}_t^{(1)}, \dots, \hat{C}_t^{(N)}),$$

for θ based upon information available at date t . Since the noise process $\{\epsilon_t\}_{t \in \mathbb{N}}$ is assumed to be unobservable, for each $t \in \mathbb{N}$ set

$$(55) \quad \begin{aligned} \hat{\epsilon}_t &:= x_t - \hat{\theta}_t \hat{y}_t, \\ \hat{y}_t &:= \left((\eta_t^{(11)} q_{t-1})^\top, \dots, (\eta_t^{(H1)} q_{t-L})^\top, (\eta_t^{(11)} p_t)^\top, \dots, (\eta_t^{(H1)} p_t)^\top, \hat{\epsilon}_{t-1}^\top, \dots, \hat{\epsilon}_{t-N}^\top \right)^\top, \end{aligned}$$

where the $\hat{\epsilon}_t$ are also called *a posteriori* prediction errors. The AML algorithm⁷ is recursively defined by

$$(56) \quad \begin{cases} \hat{\theta}_{t+1} &= \hat{\theta}_t + (x_t - \hat{\theta}_t \hat{y}_t) \hat{y}_t^\top P_t, \\ P_t &= P_{t-1} - P_{t-1} (\hat{y}_t \hat{y}_t^\top) P_{t-1} (1 + \hat{y}_t^\top P_{t-1} \hat{y}_t)^{-1}, \end{cases}$$

for arbitrary initial conditions P_{-1} and $\hat{\theta}_0$. The estimated coefficients (54) obtained from applying (56) can now be used to construct an approximation of the perfect forecasting rules defined in (31) and (35) based on information available at date t . An approximation of (31) is defined by

$$(57) \quad q_t^{(0)} = \hat{\psi}_t^{(0)} \left((q_{t-1}^{(0)} - d_t), V_t^{(0)} \right) := R(q_{t-1}^{(0)} - d_t) + a_t^{(0)-1} V_t^{(0)} \hat{x}_t,$$

where

$$\hat{x}_t = \left[\sum_{h=1}^H \sum_{l=1}^L \hat{A}_t^{(hl)} (\eta_t^{(h1)} q_{t-l}) - \sum_{h=1}^H \hat{B}_t^{(h)} \eta_t^{(h1)} (q_{t-1}^{(0)} - d_t) + \bar{x} + \sum_{n=1}^N \hat{C}_t^{(n)} \hat{\epsilon}_{t-n} \right]$$

denotes the estimated excess supply of shares in period t corresponding to (49). We call (57) a least-square-based forecast for the realization q_{t+1} . Observe that $\hat{\xi}_t := \sum_{n=1}^N \hat{C}_t^{(n)} \hat{\epsilon}_{t-n}$ is an approximation for the expected behavior of noise traders $\mathbb{E}_{t-1}[\xi_t]$.

In order to obtain an approximation for (35) assume that $\mathbb{V}_{t-1}[\xi_t] = Id$. Noticing that $\frac{1}{R} \sum_{h=1}^H \hat{B}_t^{(h)}$ is an approximation for $a_t^{(1)} V_t^{(1)-1}$, set

$$\tilde{V}_t^{(0)} := \frac{1}{Ra_t^{(0)}} \left(\sqrt{V_{t-1}^{(0)-1}} - \sum_{h=1}^H \hat{B}_t^{(h)} \right).$$

An approximation of (35) is then given by

$$(58) \quad V_t^{(0)} = \hat{\varphi}^{(0)} \left(\hat{a}_t^{(1)}, V_{t-1}^{(0)} \right) := \begin{cases} \tilde{V}_t^{(0)-1} & \text{if } \tilde{V}_t^{(0)} \text{ is positive definite,} \\ V_{t-1}^{(0)} & \text{otherwise.} \end{cases}$$

⁷We adjusted the dating according to our needs and made use of the matrix inversion lemma. This algorithm is also referred to as *recursive extended-least-squares (RELS)* algorithm, see Caines (1988, p. 540 and p. 557).

The series of forecasts (57) and (58) generated by (56) defines a least-square based adaptive learning scheme for mediator 0.

The proof that the learning scheme converges to rational expectations in the sense that first and second moments are correct along each orbit of the system depends on whether or not the AML algorithm defined by (55) and (56) generates strongly consistent estimates, that is, estimates $\hat{\theta}_t$ which converge \mathbb{P} -a.s. to θ . Observe that the requirement of constant second moments of the noise process is needed only for the approximation of (35). It is important to realize that the learning scheme (56) is meaningful only if the dynamics under rational expectations as well as under the scheme itself is in some sense stable. Before we state conditions under which strongly consistent estimates obtain, recall that $\lambda_{\min}(B)$ and $\lambda_{\max}(B)$ denote the minimal and maximal eigenvalue of a symmetric matrix B , respectively.

Theorem 8.1 *Assume that the following hypotheses are satisfied:*

- (i) *mediator 1 uses a linear rule (40) and constant covariance matrices as forecasting rules;*
- (ii) *the informational constraints of mediator 0 satisfy Assumption 8.1;*
- (iii) *the noise process $\{\xi_t\}_{t \in \mathbb{N}}$ satisfies Assumption 8.2 with $\mathbb{V}_{t-1}[\xi_t] = Id$ and is uncorrelated with the dividend process $\{d_t\}_{t \in \mathbb{N}}$;*
- (iv) *the price process (24) is stable under rational expectations as given in (38) as well as under the application of the least-square-based learning scheme (56), in the sense that all sample means are bounded, i.e.*

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \|q_t\|^2 < \infty.$$

If the stochastic process $\{y_t\}_{t \in \mathbb{N}}$ defined in (52) satisfies the weak excitation condition

$$\frac{\log \left(e + \lambda_{\max} \left(\sum_{s=1}^t y_s y_s^\top \right) \right)}{\lambda_{\min} \left(\sum_{s=1}^t y_s y_s^\top \right)} \rightarrow 0 \quad \mathbb{P} - a.s. \text{ as } t \rightarrow \infty,$$

then the sequence of estimates $\{\hat{\theta}_t\}_{t \in \mathbb{N}}$ generated by the AML algorithm (56) is strongly consistent, that is

$$\hat{\theta}_t \rightarrow \theta \quad \mathbb{P} - a.s. \text{ as } t \rightarrow \infty.$$

The proof follows directly from Lai & Wei (1986, Theorem 4, p. 241), see Appendix B for more details. Observe that the essential obstacle for convergence is the weak excitation condition such that the result is independent of initial conditions as long as the system

is stable under the learning scheme. In feedback systems such as (38), it is well-known that the weak excitation condition may be violated. This caveat, however, can be cured by occasionally perturbing the inputs which in our case are the forecasts. See Lai & Wei (1986) for the concept and Wenzelburger (2001b) for an economic application.

9 Conclusions

The analysis of adaptive learning in a financial market with heterogeneous traders showed that the correct use of the structural information about the market mechanism enables a mediator to learn those forecasting rules which generate rational expectations in the presence of diverse beliefs. In a financial market good, forecasts hinge essentially on good estimates of the demand behavior of all market participants. This structural information is encoded in an error function associated with the economy which captures the true nature of the expectations feedback.

The learning scheme proposed in the present paper was based on the simple observation that the excess demand function is linear when households are endowed with linear mean-variance preferences. The remaining nonlinearity rests on the switching behavior of households for which consistent maximum likelihood estimators exist and which in the special case with only two mediators is observable. The learning scheme converges under assumptions which are standard in the literature on estimation and optimal control.

The potential of this strand of research is outlined as follows. First, the learning scheme could be extended to an algorithm which uses on-line estimations of the distributions of households instead of the true values. Second, non-linear estimation techniques as described in White, Gallant, Hornik, Stinchcombe & Woolridge (1992) could be applied to identify and estimate forecasting rules for models in which excess demand function of traders is nonlinear.

A Random Dynamical Systems

A random dynamical system in the sense of Arnold (1998) consists of two basic ingredients, a model of the exogenous noise and a model of the system which is perturbed by noise. In what follows we restrict ourselves to the case of discrete time. The noise will be modeled by a metric dynamical system in the sense of ergodic theory which essentially amounts to using the (equivalent) canonical realization process of a stochastic process. Details are found in Arnold (1998, Appendix A). To fix the notation, let $\vartheta : \Omega \rightarrow \Omega$ be a measurable invertible map on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ which is measure preserving with respect to \mathbb{P} and whose inverse ϑ^{-1} is again measurable. Assume that \mathbb{P} is ergodic with respect to ϑ and let ϑ^t denote the t -th iterate of the map ϑ . The collection $(\Omega, \mathcal{F}, \mathbb{P}, \{\vartheta^t\}_{t \in \mathbb{Z}})$ is called an *ergodic metric dynamical system*. It is well known that any

stationary ergodic process $\{\epsilon_t\}_{t \in \mathbb{Z}}$ can be represented by an ergodic dynamical system. This implies that there exists a measurable map $\epsilon : \Omega \rightarrow \Sigma$ such that for each fixed $\omega \in \Omega$, a sample path of the process $\{\epsilon_t\}_{t \in \mathbb{Z}}$ is given by $\epsilon_t(\omega) = \epsilon(\vartheta^t \omega)$, $t \in \mathbb{Z}$.

The second ingredient consists of a time-one map $f : \Omega \times X \rightarrow X$, $X \subset \mathbb{R}^m$ with

$$(59) \quad x_{t+1} = f(\vartheta^t \omega, x_t) \equiv f(\vartheta^t \omega)x_t$$

which governs the evolution of the system. If the system started at x_0 , then the iteration of the map f under the perturbation ω induces a measurable map $\phi : \mathbb{Z} \times \Omega \times X \rightarrow X$, defined by

$$(60) \quad \phi(t, \omega, x_0) = \begin{cases} (f(\vartheta^{t-1} \omega) \circ \dots \circ f(\omega))x_0 & \text{if } t > 0 \\ x_0 & \text{if } t = 0 \\ (f(\vartheta^t \omega)^{-1} \circ \dots \circ f(\omega^{-1})^{-1})x_0 & \text{if } t < 0 \end{cases}$$

such that $x_t = \phi(t, \omega, x_0)$ is the state of the system at time t . For an arbitrary initial condition $x_0 \in X$ and any perturbation $\omega \in \Omega$, the sequence of points $\gamma(x_0) := \{x_t\}_{t \in \mathbb{N}}$ with $x_t = \phi(t, \omega, x_0)$, $t \in \mathbb{N}$ is called an orbit of the random dynamical system ϕ .

The long-run behavior of a random dynamical system is described by random attractors which is the random analogue of an attractor of a deterministic system (see Arnold 1998, p. 483). Each orbit starting from the corresponding domain of attraction will then eventually end up on such an attractor. Typical candidates for these special orbits are generated by asymptotically stable *random fixed points*. The following definition of a random fixed point (Schmalfuß 1998, 1996)⁸ includes a stability notion given Definition 7.4.6 in Arnold (1998).

Definition A.1 *A random fixed point of ϕ is a random variable $x^* : \Omega \rightarrow X$ on $(\Omega, \mathcal{F}, \mathbb{P})$ such that*

$$x^*(\vartheta \omega) = \phi(1, \omega, x^*(\omega)) \quad \text{for all } \omega \in \Omega',$$

where $\Omega' \subset \Omega$ is a ϑ -invariant set $\Omega' \subset \Omega$ of full measure $\mathbb{P}(\Omega') = 1$.⁹ A random fixed point x^* is called asymptotically stable with respect to a norm $\|\cdot\|$ on X , if there exists a random neighborhood $U(\omega) \subset X$, $\omega \in \Omega$ such that \mathbb{P} -a.s.

$$\lim_{t \rightarrow \infty} \|\phi(t, \omega, x_0(\omega)) - x^*(\vartheta^t \omega)\| = 0 \quad \text{for all } x_0(\omega) \in U(\omega).$$

⁸Random fixed points can also be defined via a special class of invariant measures for the random dynamical system ϕ , also called *random Dirac measures*, see Arnold (1998, p. 25).

⁹In the context of random dynamical systems, the term almost surely (a.s.) is used in a non-standard sense: a property holds a.s. if there exists a ϑ -invariant set $\Omega' \subset \Omega$ ($\vartheta \Omega' = \Omega'$) such that the property holds for all $\omega \in \Omega'$.

The first part in Definition A.1 implies $x^*(\vartheta^{t+1}\omega) = f(\vartheta^t\omega)x^*(\vartheta^t\omega)$ for all times t . Hence, a random fixed point generates orbits $\{x^*(\vartheta^t\omega)\}_{t \in \mathbb{N}}$, $\omega \in \Omega$ which solve the random difference equation (59) induced by ϕ . If ϕ is independent of the perturbation ω , then Definition A.1 coincides with the notion of a deterministic fixed point. A random fixed point x^* is asymptotically stable, if for almost all perturbations $\omega \in \Omega$ all orbits starting in sufficiently close points $x_0(\omega) \in U(\omega)$ eventually converge to orbits of the random fixed point. By stationarity and ergodicity of ϑ , the process $\{x^* \circ \vartheta^t\}_{t \in \mathbb{N}}$ is stationary and ergodic. Let $x^*\mathbb{P}$ denote the probability distribution of x^* . If, in addition, $\mathbb{E}\|x^*\| < \infty$, then the ergodicity and the stability property of x^* imply that

$$\lim_{T \rightarrow \infty} \frac{1}{T+1} \sum_{t=0}^T 1_B(\phi(t, \omega, x_0(\omega))) = \lim_{T \rightarrow \infty} \frac{1}{T+1} \sum_{t=0}^T 1_B(x^*(\vartheta^t\omega)) = x^*\mathbb{P}(B)$$

for all $x_0(\omega) \in U(\omega)$ \mathbb{P} -a.s., with 1_B denoting the indicator function of a subset $B \subset X$. In other words, the empirical law of any ‘nearby’ orbit $\{\phi(t, \omega, x_0(\omega))\}_{t \in \mathbb{N}}$ is for \mathbb{P} -almost all $\omega \in \Omega$ well defined and converges to the true probability distribution of x^* .

We review a simplified version of an existence theorem for random fixed points due to Schmalfuß (1998, 1996). Let $G(\omega) \subset \mathbb{R}^m$, $\omega \in \Omega$ be a random set, that is, each $G(\omega)$ is closed a.s. and $\{\omega \in \Omega \mid G(\omega) \cap U \neq \emptyset\}$ is measurable for all open sets U (Arnold 1998, Prop. 1.6.2). Consider random variables $g : \Omega \rightarrow \mathbb{R}^m$ which are tempered, meaning that

$$\lim_{t \rightarrow \infty} e^{-\delta t} \|g(\vartheta^t\omega)\| = 0 \quad \text{for all } \delta > 0.$$

Let

$$\mathcal{G} := \{g : \Omega \rightarrow \mathbb{R}^m \mid g \text{ is tempered and } g(\omega) \in G(\omega), \omega \in \Omega\}$$

denote the set of all tempered random variables with values $g(\omega)$ in $G(\omega)$.

Theorem A.2 *Let $\phi : \mathbb{Z} \times \Omega \times X \rightarrow X$, $X \subset \mathbb{R}^m$ be a random dynamical system over an ergodic dynamical system $(\Omega, \mathcal{F}, \mathbb{P}, \{\vartheta^t\}_{t \in \mathbb{Z}})$ and assume, in addition, that the time-one map $x \mapsto \phi(1, \omega, x)$ is continuously differentiable a.s. Suppose there exists a random set $G(\omega) \subset X$, $\omega \in \Omega$, such that \mathcal{G} is nonempty and the following holds:*

- (i) *The map $\omega \mapsto \phi(1, \vartheta^{-1}\omega, g(\vartheta^{-1}\omega))$ is contained in \mathcal{G} for all $g \in \mathcal{G}$;*
- (ii) *$\sup_{x \in G(\omega)} \log \|D_x \phi(1, \omega, x)\| \leq c(\omega)$, $\omega \in \Omega$ for some random variable c with $\mathbb{E}c < 0$;*
- (iii) *if, for some $g \in \mathcal{G}$, $\{\phi(t, \vartheta^{-t}\omega, g(\vartheta^{-t}\omega))\}_{t \in \mathbb{N}}$ is a Cauchy sequence for each $\omega \in \Omega$, then its limit is contained in \mathcal{G} .*

Then there exists a unique random fixed point $g^ \in \mathcal{G}$ of ϕ with*

$$\lim_{t \rightarrow \infty} \|\phi(t, \omega, g(\omega)) - g^*(\vartheta^t\omega)\| = 0 \quad \text{for all } g \in \mathcal{G}, \mathbb{P} - \text{a.s.}$$

B Convergence of the AML algorithm

We briefly review a convergence theorem for the AML algorithm. A major prerequisite is that the noise process satisfies the following so-called identifiability assumption.

Assumption B.1 *The noise process $\{\xi_t\}_{t \in \mathbb{N}}$ is given by*

$$\xi_t = \epsilon_t + \sum_{n=1}^N C^{(n)} \epsilon_{t-n},$$

where $\{\epsilon_t\}_{t \in \mathbb{N}}$ is a $\{\mathcal{F}_t\}_{t \in \mathbb{N}}$ -adapted martingale difference sequence on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ which satisfies the following:

- (1) $\sup_{t \geq 1} \mathbb{E}[\|\epsilon_t\|^\alpha \mid \mathcal{F}_{t-1}] < \infty$ \mathbb{P} -a.s. for some $\alpha > 2$;
- (2) $C^{(1)}, \dots, C^{(N)}$, $N \geq 0$ are non-random $K \times K$ matrices and the matrix polynomial $\Gamma(z) = Id + C^{(1)}z + \dots + C^{(N)}z^N$, $z \in \mathbb{C}$ is strictly positive real, that is,
 - (i) $\det \Gamma(z) \neq 0$ for all $z \in \mathbb{C}$ with $|z| \leq 1$;
 - (ii) For each $z \in \mathbb{C}$ with $|z| \leq 1$, the matrix $\Gamma(z) + \Gamma^\top(\bar{z})$, with $\bar{z} \in \mathbb{C}$ denoting the complex conjugate of z , is strictly positive definite, that is, $x^\top(\Gamma(z) + \Gamma^\top(\bar{z}))x > 0$ for all $0 \neq x \in \mathbb{R}^K$.

Theorem B.1 *(Lai & Wei 1986, Thm. 4, p.241)*

Consider the stochastic regression model

$$x_t = Az_t + \epsilon_t + C^{(1)}\epsilon_{t-1} + \dots + C^{(N)}\epsilon_{t-N},$$

where A is a $M \times K$ non-random matrix and $C^{(1)}, \dots, C^{(N)}$ are non-random $M \times M$ matrices and $\{\epsilon_t\}_{t \in \mathbb{Z}}$ satisfy Assumption B.1. Let z_t be \mathcal{F}_{t-1} measurable and

$$\theta = (A, C^{(1)}, \dots, C^{(N)}) \quad \text{and} \quad y_t = (z_t^\top, \epsilon_{t-1}^\top, \dots, \epsilon_{t-N}^\top)^\top.$$

Consider the AML algorithm, given by

$$\begin{cases} \hat{\theta}_t &= \hat{\theta}_{t-1} + (x_t - \hat{\theta}_{t-1}\hat{y}_t)\hat{y}_t^\top P_t, \\ P_t^{-1} &= P_{t-1}^{-1} + (\hat{y}_t\hat{y}_t)^\top, \end{cases}$$

where

$$\begin{aligned} \hat{\epsilon}_t &:= x_t - \hat{\theta}_t\hat{y}_t, \\ \hat{y}_t &:= (z_t^\top, \hat{\epsilon}_{t-1}^\top, \dots, \hat{\epsilon}_{t-N}^\top)^\top. \end{aligned}$$

If

$$\frac{\log(e + \lambda_{\max}(\sum_{s=1}^t y_s y_s^\top))}{\lambda_{\min}(\sum_{s=1}^t y_s y_s^\top)} \rightarrow 0 \quad \mathbb{P} - a.s. \text{ as } t \rightarrow \infty,$$

then

$$\hat{\theta}_t \rightarrow \theta \quad \mathbb{P} - a.s. \text{ as } t \rightarrow \infty,$$

where $\lambda_{\min}(B)$ and $\lambda_{\max}(B)$ denote minimal and maximal eigenvalue of a symmetric matrix B , respectively.

References

- ANDERSON, S. P., A. DE PALMA & J.-F. THISSE (1992): *Discrete Choice Theory of Product Differentiation*. MIT Press, Cambridge (Mass.) a.o.
- ARNOLD, L. (1998): *Random Dynamical Systems*. Springer-Verlag, Berlin a.o.
- BÖHM, V. & C. CHIARELLA (2000): “Mean Variance Preferences, Expectations Formations, and the Dynamics of Random Asset Prices”, Discussion Paper No. 448, University of Bielefeld.
- BÖHM, V., N. DEUTSCHER & J. WENZELBURGER (2000): “Endogenous Random Asset Prices in Overlapping Generations Economies”, *Mathematical Finance*, 10(1), 23–38.
- BÖHM, V. & J. WENZELBURGER (1999): “Expectations, Forecasting and Perfect Foresight - A Dynamical Systems Approach”, *Macroeconomic Dynamics*, 3(2), 167–186.
- (2000a): “Expectational Leads in Economic Dynamical Systems”, Revised Version September 2001, Discussion Paper No. 373, University of Bielefeld, forthcoming.
- (2000b): “Perfect Predictions in Economic Dynamical Systems with Random Perturbations”, Revised Version September 2000, Discussion Paper No. 356, University of Bielefeld, forthcoming *Macroeconomic Dynamics*.
- BROCK, W. A. & C. HOMMES (1998): “Heterogeneous Beliefs and Routes to Chaos in a Simple Asset Pricing Model.”, *Journal of Economic Dynamics and Control*, 22, 1235–1274.
- CAINES, P. E. (1988): *Linear Stochastic Systems*. Wiley, New York a.o.
- CHIARELLA, C. & X. HE (2000): “Heterogenous Beliefs, Risk and Learning in a Simple Asset Pricing Model”, *Computational Economics*, forthcoming.

- DE LONG, J. B., A. SHLEIFER, L. H. SUMMERS & R. J. WALDMANN (1990): “Noise Trader risk in Financial Markets”, *Journal of Political Economy*, 98(4).
- DUFFIE, D. (1996): *Dynamic Asset Pricing Theory*. Princeton University Press, Princeton (NJ) a.o., 2nd edn.
- ELAYDI, S. N. (1996): *An Introduction to Difference Equations*. Springer-Verlag, Berlin a. o.
- EVANS, G. W. & S. HONKAPOHJA (2001): *Learning and Expectations in Macroeconomics*. Princeton University Press.
- GRANDMONT, J.-M. (1982): “Temporary General Equilibrium Theory”, in *Handbook of Mathematical Economics*, ed. by K. Arrow & M. Intrilligator. North-Holland Publishing Company, Amsterdam a.o.
- HANNAN, E. J. & M. DEISTLER (1988): *The Statistical Theory of Linear Systems*, Wiley Series in Probability and Mathematical Statistics. John Wiley, New York a.o.
- JUDGE, G. G., W. E. GRIFFITHS, R. C. HILL, H. LÜTKEPOHL & T.-C. LEE (1985): *The Theory and Practice of Econometrics*. John Wiley and Sons, 2nd edn.
- LAI, T. L. & C. Z. WEI (1986): “On the Concept of Excitation in Least Squares Identification and Adaptive Control”, *Stochastics*, 16, 227–254.
- MARKOWITZ, H. (1952): “Portfolio Selection”, *Journal of Finance*, 7, 77–91.
- PLISKA, S. (1997): *Introduction to Mathematical Finance*. Blackwell Publishers Inc, Massachusetts.
- SCHMALFUSS, B. (1996): “A Random Fixed Point Theorem Based on Lyapunov Exponents”, *Random and Computational Dynamics*, 4(4), 257–268.
- SCHMALFUSS, B. (1998): “A Random Fixed Point Theorem and the Random Graph Transformation”, *Journal of Mathematical Analysis and Application*, 225(1), 91–113.
- TOBIN, J. (1958): “Liquidity Preference as Behavior Towards Risk”, *Review of Economic Studies*, 25, 65–86.
- WENZELBURGER, J. (2001a): “Learning in Economic Systems with Expectations Feedback”, Habilitationsschrift, University of Bielefeld, in preparation.
- (2001b): “Learning in Linear Models with Expectational Leads”, Discussion Paper No. 478, University of Bielefeld.
- WHITE, H., A. GALLANT, K. HORNIK, M. STINCHCOMBE & J. WOOLRIDGE (1992): *Artificial Neural Networks: Approximation and Learning Theory*. Blackwell, Cambridge (Mass.) a.o.