UvA-DARE (Digital Academic Repository)

The time-variation of volatility and the evolution of expectations

van der Weide, R.

Publication date
2012

[Link to publication](#)

Citation for published version (APA):
van der Weide, R. (2012). *The time-variation of volatility and the evolution of expectations*. [Thesis, fully internal, Universiteit van Amsterdam]. University of Amsterdam.

Chapter 6

Herding, A-synchronous updating and heterogeneity in memory

6.1 Introduction

The current literature on agent based economic models can be roughly divided into two types: (i) computational finance models and (ii) economic dynamic models. Both approaches aim to discover what is behind the stylized facts commonly exhibited by empirical financial time series.

In computational finance, artificial stock markets are used as a tool for simulating time series of aggregate variables, such as prices, volumes, etcetera (for an overview, see LeBaron, 2000). To improve the understanding of markets with a large number of interacting heterogeneous agents, LeBaron *et al.* (1999) for example developed the Santa Fe artificial stock market. In the Santa Fe market, the actions of agents, which are based upon their expectations, are explicitly modelled and traced for each individual agent. With computers becoming cheaply available and faster, these artificial markets allow for more and more detailed modelling. Although detailed models may possibly generate more realistic price dynamics, their increasing complexity also makes simulation less attractive as a tool for the analytic study of the mechanisms linking the heterogeneity of beliefs with the stylized facts of prices.

The economic dynamics literature follows a somewhat different route. There agent based modelling explicitly aims at an analysis of the joint dynamics of beliefs and prices

(see e.g. Brock and Hommes, 1997, 1998). To achieve the necessary analytical tractability, several simplifying assumptions are usually made, such as a continuum of agents who at each time can choose from a finite number of belief types. Typically a low dimensional deterministic dynamical system is derived, allowing the modeller to perform a detailed bifurcation analysis, providing insights in the long-run behaviour of the dynamics for different economic and behavioral parameters. The natural trade-off of this approach is that it at best provides an approximation to the *deterministic* aspects of the price process of interest. Indeed, the 'holy grail' in economic dynamics is beyond any doubt *the* heterogeneous agent model that is just as realistic in terms of its price dynamics as it is tractable analytically.

In previous work (Diks and van der Weide, 2003) we introduced the concept of a *continuous beliefs system* (CBS). This framework is built around a continuous *beliefs space* representing the possible point predictors agents can choose from. On this space a time dependent *beliefs distribution* is defined, which is updated according to a continuous choice model. As new market prices become available, the beliefs distribution is updated depending on past performances of strategies. Coevolution of the beliefs distribution with price dynamics thus arises as newly realized prices feed the ongoing evaluation of the strategies by the agents. As is natural in a continuous choice setting, individual choices are considered to be random variables, distributed according to the beliefs distribution. The preferences of agents provide an endogenous noise source which affects price dynamics. Because the randomness in prices can be associated naturally with the diversity of beliefs, the latter has been labelled a 'natural source of randomness'.

Diks and van der Weide (2003) discuss several scenarios under which the natural randomness in prices does not average out in the limit as the number of agents tends to infinity. These include: (i) a disproportionate distribution of market impacts; (ii) particular combinations of utility functions and predictors, and (iii) dependence among agents, which may arise for example when agents partly coordinate on a random variable, not necessarily related to economic fundamentals ('sunspots'). All these examples are based on a violation of at least one of the conditions of the law of large numbers. In the last case (dependence) it is often impossible to distinghuish between an infinite number of dependent agents, or a finite number of independent agents. For convenience we therefore introduced the notion of an effective number of agents, n_{eff}, defined as the number of *independent* agents required to obtain stochastic price dynamics with similar stochastic dynamics. For a finite effective number of agents the dynamics is random by construction. In general, a CBS

gives rise to a random dynamical system, allowing for the stochastic price dynamics to be described explicitly. It therefore offers an analytic alternative to the simulation based studies commonly performed in computational finance.

In sum, a CBS shares several advantages of both computational finance models and economic dynamic models. Rather than obtaining the stochastic macroscopic dynamics through laborious simulation, the macroscopic behaviour of a CBS with agents who can choose among a continuum of alternative strategies, can often be derived analytically. The aggregate behaviour can typically be captured in a few simple dynamic equations, for an arbitrary number of agents. The analytic tractability can be used to study features such as the conditional mean and variance of prices analytically, and relate these to behavioural characteristics.

The objective of this chapter is to examine the effects on price dynamics of a number of behavioral assumptions. We introduce a simple benchmark CBS and expand it in the following directions: (i) herd behavior; (ii) a-synchronous updating of beliefs; and (iii) heterogeneity in the memory of agents. The behavioral models obtained in this way give rise to aggregate dynamics which are represented by well-known classes of econometric models. At the center is the random walk obtained for our benchmark CBS. We will show that that: (i) herd behaviour introduces moving average (MA) structure yielding an ARIMA model; (ii) an ARIMA model with GARCH disturbances is obtained if beliefs are not updated synchronously; and (iii) a fractionally integrated (ARFIMA) process with long-range dependence is obtained if the memory of agents exhibits heterogeneity. As an immediate result the parameters of these standard econometric models can be related directly to the parameters of the behavioral model.

Near the end of the chapter we will only carefully scratch the surface of the possible empirical applications. This is done by estimating some of the stylized models using daily exchange rate data. We wish to emphasize honestly at the outset, however, that while interpreting the empirical results in terms of behavioral parameters, one should keep in mind that the models are stylized and, for clarity of exposition, deliberately stripped from a range of features that may become important for a realistic summary of the data. For example, we are aware of the fact that our natural source of randomness is not the only relevant source of randomness, and that more factors than those included in our behavioral model may be responsible for MA and/or GARCH structure in the empirical data. Although omitting additional relevant model aspects will surely bias our parameter estimates, we consider comparing the signs of the estimated parameters with those expected

from our theory a sensible 'reality check'. If successful, such an empirical validation may hopefully open doors to more detailed empirical assessments in the future.

Finally note that inherent in our approach is the assumption that the agents believe in structure in prices that can be profitably exploited by 'technical trading rules' summarized in terms of point predictors, which we will refer to interchangeably as strategies or beliefs. Empirical evidence supporting this belief in predictability in the foreign exchange markets is provided by e.g. Sweeney (1986), Taylor and Allen (1992), Levich and Thomas (1993) and LeBaron (1999). Furthermore, our approach is entirely built around the heterogeneity of beliefs. Some arguments for expectations being heterogeneous have been provided by, among others, Frankel and Froot (1990). In a theoretical exchange rate model, Frankel and Froot (1988) particularly employ the heterogeneity of expectations to explain some of their empirical findings.

In the following section we briefly review the basic concepts of a CBS. The introduction of the benchmark CBS and its extensions are described in section 6.3. In section 6.4 the models will be subjected to a basic empirical validation. Section 6.5 concludes.

6.2   Continuous beliefs systems

In this section we briefly review the concept of a continuous beliefs system. The main ingredients for describing the co-evolution of prices and the beliefs distribution in a CBS are the following. As is common in economic dynamics models, agents are assumed to predict future prices by using a point predictor which is a function of the information available to them (we typically have in mind past realized prices). The point predictors agents can choose from are represented in a beliefs space Ω, parameterized by a continuous parameter θ, called the *belief parameter*. Each value of $\theta \in \Omega$ represents a predictor $x^e_{t+1}(\theta)$ of the next price in terms of the information $\mathcal{F}_t$ available to the agents. For example if agents believe in a linear world and expect tomorrow's price x_{t+1} to be a linear function $x^e_{t+1}(\theta) = \theta_0 + \theta_1 x_t + \ldots + \theta_d x_{t-d+1}$ of the last d prices,then the beliefs space consists of the $d + 1$-dimensional Euclidean space $\mathbf{R}^{d+1}$.

Given the individual beliefs $\{\theta_{i,t-1}\}$, a realization of the price x_t is obtained via some market clearing mechanism. Agents' utility functions and predictors imply individual demand functions, and the market price is defined to be the unique price x_t for which the market clears.[1] Of course the appropriate market clearing price equation depends on the

---

[1] Uniqueness of market clearing prices is not self-evident. However, if demands are strictly decreasing

context, and we will have to provide some additional arguments to justify our particular choice made for the benchmark model later.

As new prices become available, agents update their evaluations of parameter values by confronting them with the newly arrived information. As agents have differences in tastes and interpret identical information differently they will typically choose different predictors. This is captured by a continuous choice model, which defines a probability distribution on the beliefs space Ω. The probability density function of this time dependent *beliefs distribution* is denoted by $\phi_t(\theta)$. The actual belief parameters $\theta_{i,t}$ individual agents use to base their next prediction on are modelled as random variables distributed according to $\phi_t(\theta)$. In this way endogenous randomness enters the price dynamics in a natural way. In the simplest case the $\theta_{i,t}$ are assumed to be independent conditional on information available up to and including time t.

To see how the beliefs distribution evolves, let us consider the time when x_t has just been quoted. After x_t becomes public, agents re-evaluate the strategies using a *performance measure*, $\pi_t(\theta)$. The performance depends on past realized prices and might for example be based on *ex post* prediction errors, or the *ex post* profits realized by a strategy. A typical choice is to take $\pi_t(\theta)$ to consist of minus the *ex post* squared prediction error:

$$\pi_t(\theta) = -(x_t^e(\theta) - x_t)^2. \tag{6.1}$$

This choice is less restrictive as it may seem, since for $\phi_t(\theta)$ only *differences* in performance play a role. Using this property it can be shown that minus squared prediction errors are equivalent to risk adjusted profits when agents hold identical beliefs about conditional variances (see e.g. Hommes, 2001). In practice, decisions to change strategies are not based on yesterday's performances only, but rather on the success over a larger history. Indeed, financial analysts typically test their candidates on a large sample of past price history, also known as back-testing. Strategies that have performed best over some history of the sample are more likely to be selected for future trading. The simplest way to implement, or mimic, this back-testing is to consider a geometrically down-weighted average of past performances rather than only the last performance, so that the weighted average utility becomes:

$$U_t(\theta) = (1 - \alpha) \sum_{i=0}^{\infty} \alpha^i \pi_{t-i}(\theta),$$

in x_t, as will be the case in our benchmark CBS, uniqueness of clearing prices is guaranteed.

which can be recast into the more practical form

$$U_t(\theta) = \alpha U_{t-1}(\theta) + (1 - \alpha)\pi_t(\theta). \tag{6.2}$$

Given the average past utility the continuous logit model provides the following pdf for the beliefs distribution just after time t:[2]

$$\phi_t(\theta) = \frac{e^{\beta U_t(\theta)}}{Z_t}, \tag{6.3}$$

where Z_t is a normalization constant independent of θ. The parameter β is referred to as the *intensity of choice*. The inverse $1/\beta$ is also known as the *propensity to err*. If agents are sensitive to small differences in utility, this will be reflected by a large value of β. Notice that Eq. (6.3) implies that only differences in performance play a role. Modifications of $\pi_t(\theta)$ that are independent of θ will be absorbed by the normalization factor Z_t, leaving $\phi_t(\theta)$ unchanged. With this in mind it is not surprising that risk adjusted profits and minus squared prediction errors turn out to be equivalent under certain conditions.

Substitution of Eq. (6.2) into Eq. (6.3) gives

$$\phi_t(\theta) = \frac{e^{\alpha\beta U_{t-1}(\theta) + (1-\alpha)\beta\pi_t(\theta)}}{Z_t} = \frac{[\phi_{t-1}(\theta)]^\alpha e^{(1-\alpha)\beta\pi_t(\theta)}}{Z'_t}, \tag{6.4}$$

which expresses how a new beliefs distribution can be obtained from the previous beliefs distribution and the last performance measure. The variable Z'_t is again a normalization constant. Although $\phi_t(\theta)$ is an infinite dimensional state variable, the time evolution of $\phi_t(\theta)$ can often be fully specified by a finite number of variables such as its first k moments, in which case the dynamics becomes finite dimensional. In summary, the expectations feedback can be represented schematically as

$$\ldots x_{t-1} \to \phi_{t-1}(\theta) \rightsquigarrow x_t \to \phi_t(\theta) \rightsquigarrow x_{t+1} \to \ldots,$$

where "$\to$" indicates a deterministic step and "$\rightsquigarrow$" a step involving endogenous noise.

It is important to note that the beliefs distribution $\phi_t(\theta)$ is determined only by the performance measure and past realised prices, and otherwise evolves independently of the individual choices of agents. The beliefs distribution is a convenient concept introduced

---

[2]In Diks and van der Weide (2003) the right-hand-side was multiplied by a so-called opportunity function $\varphi(\theta)$ which we take equal to unity here.

in the choice literature to model the fact that agents have idiosyncrasies in tastes, and that their choices are not identical given identical information. The fact that identical information is interpreted differently by agents is confirmed empirically by Kandel and Pearson (1995). The strategies $\theta_{i,t}$ used by the agents, from the point of view of the econometrician, are random variables, representing the individual choices of agents, so that x_{t+1} is a stochastic variable, even in the absence of exogenous shocks.

6.3 A simple benchmark CBS and its extensions

In this section we introduce a benchmark CBS and examine the effects on the dynamics of extending the model in the following behavioral directions: (i) herding, (ii) asynchronous updating, and (iii) heterogeneity in memory. The CBS formulation enables one to obtain insights into how these behavioral assumptions affect the interaction between the stylized facts and the unobserved beliefs distribution. The benchmark model is an extremely simple CBS deliberately stripped from any features that are not of direct interest. Although taking such a highly stylized model as a reference point may seem overly restricted, it clarifies the exposition and helps to isolate the effects of each of the extensions.

We wish our benchmark CBS to be as simple as possible while preserving economic relevance and bringing out the essentials of the CBS concept. Typically a CBS will be used for modelling the joint dynamics of beliefs and prices of well-specified economic units, such as shares, which are traded sequentially in a large market. Prices in such markets are determined by the expectations held by agents regarding future prices and cash flows, as well as their preferences concerning returns and risk. Agents interact by submitting demand functions depending on expected utility, while their realized utility is determined by realized market clearing prices. Regardless of preferences, it is clear that better abilities to predict future prices given the information available will help agents in realizing a higher utility. In this sense the setting resembles an ever ongoing version of Keynes' beauty contest.

Motivated by these arguments we consider a very stylized market mechanism which assumes prices to equal the average prediction over all agents in the market, i.e. the *market expectation*:

$$x_t = \frac{1}{n} \sum_{i=1}^n x_{t+1}^e(\theta_{i,t-1}), \tag{6.5}$$

where n denotes the number of agents. Price equations used in economic dynamic models

often involve similar market expectations. Depending on the type of market, also the risk free interest rate and expectations concerning dividend payments may enter the price equation. We have in mind agents who are myopic in the sense that they are concerned only with maximizing their one-step-ahead utility. At each time then, agents act as if they are entering the market today while knowing that they will leave the market tomorrow. As a result, only their expectations concerning the next price and additional cash flows (such as dividends) play a role. Typically, dividend payments (if any) are rare and announced well in advance, which means that expectations regarding short run profits are driven by expectations concerning future prices, rather than dividends. If, moreover, time intervals are sufficiently short (e.g. one to five days, see Lehmann, 1991) the interest rates over a single interval are sufficiently small for agents to consider their predicted future price to be a 'fair price' today as well. Under those conditions, if we assume agents to be mean-variance optimizers the market clearing price is given by the right hand side of Eq. (6.5).[3]

The simplest continuous class of point predictors consists of a continuous range of different constants. We incorporate this by saying that agents with belief θ predict tomorrow's price to be equal to θ:

$$x_{t+1}^e(\theta) = \theta.$$

Heterogeneity will arise if agents employ different beliefs in that they have in mind different parameter values $\theta_{i,t}$. The conditional variance $\text{Var}_t[\theta_{i,t}]$ can then be considered to be a measure for the degree of heterogeneity. Forecast surveys of foreign exchange market participants, for example, indicate that such a dispersion of beliefs indeed exists, as found by Frankel and Froot (1990). An argument they provide in favour of heterogeneity is the high trading volume observed in the foreign exchange market, since it takes differences of opinion to trade.

For the performance measure we take minus squared prediction errors (Eq. 6.1), and memory is included in the utility function through Eq. (6.2). Having observed and compared overall performance, all agents subsequently adapt their beliefs. The new distribution of beliefs is determined by means of the continuous choice model, and is given by Eq. (6.3). Because the utility function is a quadratic function in the belief parameter θ, it follows

---

[3]A mean-variance optimizing agent with belief θ expects a risk-adjusted short-term profit of $d_t[x_{t+1}^e(\theta) - x_t] - \frac{1}{2}cd_t^2$, where d_t is his demand and c is a risk aversion parameter times the perceived conditional variance of x_{t+1}. Maximizing this with respect to demand d_t, gives the demand function $d_t = (x_{t+1}^e(\theta) - x_t)/c$. If c is homogeneous among agents, the unique equilibrium price is given by Eq. (6.5). Heterogeneity in risk aversion and or perceived risk would lead to a weighted average $x_t = \sum w_i x_{t+1}^e(\theta_{i,t+1})/\sum w_i$ with weights $w_i = c_i^{-1}$, which in a large market is equivalent to Eq. (6.5).

that, for all t, the distribution of beliefs is described by a normal distribution:

$$\phi_t(\theta) = \frac{1}{\sqrt{2\pi}\sigma_t} \exp\left[-\frac{1}{2}\left(\frac{\theta - \mu_t}{\sigma_t}\right)^2\right]. \tag{6.6}$$

Normality implies that the beliefs distribution is fully specified by its mean μ_t and variance σ_t^2. Using Eq. (6.4) it can be seen that the dynamics of μ_t and σ_t is governed by:

$$\begin{aligned}
\mu_t &= \alpha\left(\frac{\sigma_t^2}{\sigma_{t-1}^2}\right)\mu_{t-1} + 2(1-\alpha)\beta\sigma_t^2 x_t \\
\frac{1}{\sigma_t^2} &= \frac{\alpha}{\sigma_{t-1}^2} + 2(1-\alpha)\beta.
\end{aligned} \tag{6.7}$$

In the long run σ_t^2 will converge to $\sigma^2 = \frac{1}{2\beta}$, and the dispersion of beliefs thus tends to a constant proportional to the propensity to err. If the propensity to err is high, agents are less sensitive to differences in performances of strategies. As a consequence, a large propensity to err implies a larger dispersion of beliefs. If we substitute $\sigma^2 = \frac{1}{2\beta}$ into the dynamics for the average belief μ_t, we obtain:

$$\mu_t = \alpha\mu_{t-1} + (1-\alpha)x_t. \tag{6.8}$$

Having derived the dynamics for beliefs, we now shift focus to its interaction with the price equation Eq. (6.5) which in the constant predictor case becomes:

$$x_t = \frac{1}{n}\sum_{i=1}^n x_t^e(\theta_{i,t-1}) = \frac{1}{n}\sum_{i=1}^n \theta_{i,t-1},$$

For simplicity, we consider the system in the limit as the number of agents tends to infinity. In the absence of dependence among agents, the law of large numbers applies, and it follows that the observed price quoted as the average prediction will converge to μ_{t-1}:

$$x_t = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^n \theta_{i,t-1} = E[\theta_{i,t-1}] = \mu_{t-1}. \tag{6.9}$$

Together with the 'dynamics' of the mean μ_t described by Eq. (6.8) this gives:

$$x_t = \alpha x_{t-1} + (1-\alpha)x_{t-1} = x_{t-1}.$$

Thus in the absence of uncertainty, the price will be constant at all times.

Exogenous shocks, such as news, are assumed to affect the price equation only through the expectations of agents. This assumption appears natural, since the only way news can possibly affect prices is indeed through expectations of individuals. The dynamics become

$$
\begin{aligned}
x_t &= \mu_{t-1}, \\
\mu_t &= \alpha\mu_{t-1} + (1-\alpha)x_t + \xi_{t+1},
\end{aligned}
\tag{6.10}
$$

where ξ_{t+1} is a noise term representing news shocks arriving between time t and $t+1$. Upon substituting the first of those equations in the second, it can be seen that μ_t follows a random walk driven by ξ_t. The first equation states that x_t simply equals μ_{t-1}, the average belief under agents given all observables up to and including time $t-1$, and the news that has become available before time t.[4] In terms of increments we thus have, in the presence of exogenous shocks $\{\xi_t\}$:

$$
\Delta x_t \equiv x_t - x_{t-1} = \xi_t.
$$

The random walk model is consistent with the efficient market hypothesis (EMH) and has been traditionally viewed as *the* model for financial variables, such as asset prices and exchange rates. Indeed, the random walk has long been considered unbeatable, a view epitomized by Malkiel (1990). Only recently it seems that it has to give in to a growing list of statistical evidence against the EMH, some of which is collected in Lo and MacKinlay (1997). As a consequence we see that more sophisticated models have begun to make their appearances. Kilian and Taylor (2001), for example, introduced an exchange rate model which is close to a unit root process near the fundamental, but in which the exchange rate is driven back to the fundamental when their difference becomes too large. They report empirical evidence supporting predictability of exchange rates at long horizons, but at short horizons exchange rates are hard to distinguish from a random walk.

6.3.1   Herd behaviour

If the conditions for the law of large numbers (LLN) are not satisfied, the price dynamics may be stochastic in a large (many trader) market even in the absence of exogenous

---

[4]The random walk price dynamics can alternatively be derived with agents forming expectations of the type $x_t^e(\theta) = R_t + \theta$, where R_t is a publicly known estimate of the fundamental exchange rate with increments $R_t - R_{t-1} = \xi_t$. In that case μ_t is stationary, while the exchange rate follows a random walk, $x_t = R_t + \text{cnst}$.

uncertainty. In the introduction we briefly mentioned a number of such 'sources of randomness' that might cause the LLN to fail. These are addressed in more detail in Diks and van der Weide (2003). Here we will focus on the dependence among agents arising from herd behaviour, as a source of 'natural randomness'. It turns out that sufficient correlation between the agents will indeed violate the conditions for the LLN to apply.

In the literature many different definitions of herding have been described. The type of herding we wish to address is that where conditional on public information, the strategies used by the agents exhibit dependence. In particular we have in mind situations where agents herd to publically announced opinions of others. During each period some 'early' agents reveal their expectations, allowing other agents to adapt their expectations towards those of the 'leaders' before submitting their market orders. This type of herding is 'concurrent' in that it involves agents coordinating simultaneously on common signals. Concurrent herding is different from studies that associate herding with for example informational cascades or learning (see e.g. Bikhchandani *et al.*, 1992; Avery and Zemsky, 1998 and Nelson, 2002). In informational cascades uninformed agents try to mimic, or learn from, the more informed by extracting information from realized prices. This type of herding therefore has a 'sequential' nature. As one would expect, having in mind 'concurrent' or 'sequential' herding makes a difference. For example, Nelson (2002) finds that 'high correlation among agents' actions does not necessarily imply herding'. As in our model dependence among actions must originate from dependence in strategies, we would refer to such correlation among actions as herding by definition.

Possible motivations for (concurrent) herding are numerous. First note that according to Kandel and Pearson (1995), even though each individual typically has his own interpretation of public signals only few behave completely independent. A motivation is that being too different from the rest can be risky and might jeopardize career perspectives or reputation (see e.g. Trueman, 1994 and Scharfstein and Stein, 1990). Hong *et al.* (2000) have recently strengthened these arguments by providing some empirical evidence in favour of them. When delving into experience, they also find that younger analysts forecast closer to the average forecast, and that they are more likely to be terminated when they deviate from the consensus.

Even when accurate (short term) forecasting is the only objective agents have, herding might still be rational. Scharfstein and Stein (1990), as an oblique remark not that: "...investing is also driven by group psychology, which weakens the link between information and market outcomes." Indeed, note that in our benchmark CBS prices are entirely driven

by speculation (group psychology).[5] Froot *et al.* (1992) have given the thought more attention, and concluded the following. In the short term, speculative forces may be stronger than those of the 'underlying economic system'. As a consequence, price series generally show fluctuations around the fundamental, where one can always find short periods in time where prices exhibit divergence. In such periods it is thus rational to *coordinate* on speculative behaviour (herding) rather than on fundamentals.

Kirman (1993) describes herding as arising from a process, which involves agents that make choices first, 'recruit' other agents to make the same decision. There are several possible forms in which this recruiting takes place, which we will not delve into here. What is important for our model is that there may be mechanisms that introduce dependence among the strategies selected by the agents. The 'recruiting' scenario is an example mechanism which gives just this type of dependence in our model.

Herding towards a leading agent can be implemented easily in a CBS by introducing interaction terms into the utility function.[6] For simplicity we assume that there is only a single leading agent. Let $\theta_{0,t}$ denote the parameter value announced publically by the leading agent, which we assume to be distributed according to the original beliefs distribution given by Eq. (6.6). Upon taking a new utility function $\tilde{U}_t(\theta) = U_t(\theta) - J(\theta - \theta_{0,t})^2$, the beliefs distribution of the other agents becomes

$$\phi_t(\theta) = Z_t^{-1} \exp\left[\beta U_t(\theta) - \beta J(\theta - \theta_{0,t})^2\right]. \tag{6.11}$$

The interaction strength $J > 0$ can be considered as a measure for the desire of agents to select a parameter value close to that of the leader. Straightforward comparison of coefficients of powers of θ in the exponent with those in Eq. (6.6) gives the following values for the mean $\tilde{\mu}_t$ and variance $\tilde{\sigma}_t^2$ of the beliefs distribution under herding:

$$\tilde{\mu}_t = \mu_t + \frac{J(\mu_t - \theta_{0,t})}{J + 1} \equiv \mu_t,$$

and

$$\tilde{\sigma}^2 = \frac{\sigma^2}{1 + \sigma^2 J}.$$

---

[5]Lux (1995) is another paper that considers the possibility where agents pay more attention to each other than to the fundamentals. He assumes a market in which initially only optimists and pessimists can be distinguished. The evolution of the number of optimists is completely determined by the numbers of optimists and pessimists that surround individuals. At a later stage he assumes the presence of a fundamentalist.

[6]See Diks and van der Weide (2003), where social interaction is incorporated in a CBS, in analogy with Brock and Durlauf (2001) who consider social interaction in a dynamic discrete choice setting.

In the latter relation the time indices have been dropped as σ_t^2 was already known to be constant over time (see Eq. (6.7) and the discussion thereafter).

Under the assumption that market weights of individual agents (including the leading agent's) are negligible, realized prices are given by the mean of the pdf in Eq. (6.11):

$$x_t = \tilde{\mu}_{t-1} = \mu_{t-1} + \sigma\lambda u_t, \tag{6.12}$$

where $\{u_t\} \equiv \left\{(\mu_{t-1} - \theta_{0,t-1})/\sigma\right\}$ is a sequence of independent standard normal random variables.

If we incorporate the dynamics for the average belief μ_t, which is given in Eq. (6.8), it follows that the market's price dynamics is described by the following ARIMA model:

$$x_t = x_{t-1} + \frac{J}{J+1}\sigma\left[u_t - \alpha u_{t-1}\right].$$

The amount of endogenous uncertainty entering the price dynamics is proportional to $J/(J+1)$. Since this function is increasing in J and takes values between 0 and 1, it is convenient to define a measure of herding λ, via

$$\lambda \equiv \frac{J}{J+1},$$

so that the price dynamics in the absence of news can be concisely stated as

$$\Delta x_t = \sigma\lambda\left(u_t - \alpha u_{t-1}\right), \tag{6.13}$$

where $\Delta x_t \equiv x_t - x_{t-1}$. As herd behaviour acts to affect the *effective* number of agents, it should come as no surprise that a change in the extent of herding λ can not be distinguished from a reduction of the number of agents in the CBS. Only their joint effect can be identified.[7]

At this stage all stochasticity in the model can be associated with endogenous uncertainty. That is, the randomness observed by the econometrician reflects the random nature with which individual agents select beliefs from the distribution of beliefs. If in addition

---

[7]Diks and van der Weide (2003) related the amount of endogenous noise to both the market impact of agents, and the correlation between agent's decisions. The effective number of agents (which is inversely proportional to the variance of endogenous noise) is given by $n_{\text{eff}} = \left(\sum_{i=1}^n w_i^2 + \rho\sum_i\sum_{j\neq i} w_i w_j\right)$ where n is the total number of agents, $\{w_i\}$ represent their market impacts, and where ρ is the correlation between agent's beliefs. Taking equal weights, $w_i = \frac{1}{n}$, and letting n tend to infinity as we do here, only leaves $n_{\text{eff}} = \rho^{-1}$.

we include independent, identically distributed (IID) exogenous shocks ξ_t to the system, we find:

$$\begin{aligned} \Delta x_t &= \sigma\lambda\left(u_t - \alpha u_{t-1}\right) + \rho\xi_t \\ &\sim \eta\left(u'_t - \psi u'_{t-1}\right), \end{aligned} \tag{6.14}$$

where $\{u'_t\}$ is a sequence of IID standard normal random variables, η a scaling parameter, and $-\psi$ a moving average coefficient. The relation between our original parameters $(\sigma, \lambda, \alpha, \rho)$ and the new parameters (η, ψ) is the following:

$$\begin{aligned} \eta &= \left(\sigma^2\lambda^2 + \rho^2\right)^{\frac{1}{2}} \\ \psi &= \frac{\alpha\sigma\lambda}{\left(\sigma^2\lambda^2 + \rho^2\right)^{\frac{1}{2}}}. \end{aligned} \tag{6.15}$$

The equivalence of the two representations in Eq. (6.14) follows directly from comparing the auto-covariance functions of both representations. While both representations describe the same ARIMA(0,1,1) model, the reduction from four to two parameters shows that the first representation was over-parameterized. The original (behavioral) model parameters can therefore not be identified all independently. For an introduction to ARIMA and related models we refer to Mills (1999).

It is worth noting that we have established a link between behavioral finance and econometric modelling. As a direct result, one can associate the parameters from the econometric model with economically relevant and/or behavioral quantities. In this particular example, the scale of the endogenous noise appears in the econometric model as the product $\sigma\lambda$ of the dispersion σ of the beliefs distribution and the herding parameter λ, while ρ measures the size of the exogenous noise.

6.3.2   A-synchronous updating

So far, as is also common in economic dynamic models based on discrete choice, we have assumed implicitly that the time scale on which agents re-evaluate strategies is short with respect to the trading time. It seems reasonable to assume that in the absence of costs for evaluating and updating beliefs, every agent will indeed decide to update every period. However, if updating is either time consuming or costly, agents are likely to update less frequently. Hong *et al.* (2000) found empirical evidence supporting this view. In their empirical study they concluded that inexperienced analysts revise their forecasts more frequently than experienced analysts.

In general, if the time scales on which agents review strategies is longer in practice than the clock rate at which prices/exchange rates are quoted, the dynamics of both prices and the distribution of beliefs will be affected. Indeed, Brock and LeBaron (1996) find that time series with more realistic properties, such as strong persistence of volatility, are obtained in the case of less frequent updating. The effects of the time scale for evaluations can be incorporated in the model in several ways. One might take the approach of Brock and LeBaron (1996) where all agents update their beliefs all at once with a certain fixed probability per time unit.

An alternative approach, which we will follow here, is to assign a probability δ that individual agents do not update their beliefs, and a probability $1 - \delta$ that individual agents do update their belief, at each time step. The latter is generally more difficult to implement, since it implies that the beliefs distribution becomes a mixture with weight δ on the old utility function, and $1 - \delta$ on the new, that is,

$$
\begin{aligned}
\phi_t(\theta) &= \delta\phi_{t-1}(\theta) + (1 - \delta)\tilde{\phi}_t(\theta) \\[2mm]
\tilde{\phi}_t(\theta) &\sim e^{\beta U_t(\theta)} \sim e^{\beta(1-\alpha)\pi_t(\theta)}[\tilde{\phi}_t(\theta)]^{1-\alpha}.
\end{aligned}
\tag{6.16}
$$

Here $\tilde{\phi}_t(\theta)$ represents the up-to-date distribution of choices that would be made based on information up to the recent past, while the actual beliefs distribution $\phi_t(\theta)$ is still partly based on old information. In simple cases, however, where the mean and variance of the mixture depend on the past utility only through the previous mean and variance, a closed form updating rule for the new mean and variance can still be given. This is demonstrated in the following example.

The average belief and the dispersion of beliefs among the agents that have decided to update are denoted by $\{\mu_t, \sigma_t^2\}$, where $\sigma_t^2 = \sigma^2$ is still constant over time. The average and the dispersion of beliefs for the entire population is now denoted by $\mu_{\delta,t}$ and $\sigma_{\delta,t}^2$. Using Eq.(6.16) it can be verified that the dynamics of the first two moments of the beliefs distribution is now governed by:

$$
\begin{aligned}
\mu_{\delta,t} &= \delta\mu_{\delta,t-1} + (1 - \delta)\,\mu_t \\[2mm]
\left(\mu_{\delta,t}^2 + \sigma_{\delta,t}^2\right) &= \delta\left(\mu_{\delta,t-1}^2 + \sigma_{\delta,t-1}^2\right) + (1 - \delta)\left(\mu_t^2 + \sigma^2\right),
\end{aligned}
\tag{6.17}
$$

where as before (Eq. 6.8), we have:

$$\mu_t = \alpha \mu_{t-1} + (1 - \alpha) \, x_t. \tag{6.18}$$

If we take squares on both sides of the first relation in Eq. (6.17) and subtract both relations to eliminate $\mu_{\delta,t}^2$, we find the following dynamics for the dispersion of beliefs $\sigma_{\delta,t}^2$:

$$\sigma_{\delta,t}^2 = \sigma^2 + \delta \left(\sigma_{\delta,t-1}^2 - \sigma^2 \right) + \delta \left(1 - \delta\right) \left(\mu_t - \mu_{\delta,t-1} \right)^2. \tag{6.19}$$

Thus the dispersion of beliefs exhibits fluctuations when agents do not move together, even when the dispersion of beliefs among those who adapt is constant in each period. The movements are driven by the differences in average beliefs between those who decide to adapt and those who wait.

The time evolution for the price is now described by the following analogue of Eq. (6.12):

$$x_t = \mu_{\delta,t-1} + \lambda \sigma_{\delta,t-1} u_t, \tag{6.20}$$

where λ is our measure for herding and $\sigma_{\delta,t-1}$ is the time varying standard deviation of the beliefs distribution. Let us first focus on the average beliefs $\{\mu_t, \mu_{\delta,t}\}$. Note that Eqs (6.17) and (6.18) give, respectively,

$$(1 - \delta) \, \mu_t = \mu_{\delta,t} - \delta \mu_{\delta,t-1}$$

and

$$(1 - \delta) \, \mu_t = \alpha \left(1 - \delta\right) \mu_{t-1} + (1 - \alpha) \left(1 - \delta\right) x_t,$$

which together yield:

$$\mu_{\delta,t} - \delta \mu_{\delta,t-1} = \alpha \left(\mu_{\delta,t-1} - \delta \mu_{\delta,t-2} \right) + (1 - \alpha) \left(1 - \delta\right) x_t.$$

If we substitute $\mu_{\delta,t} = x_{t+1} - (\lambda \sigma_{\delta,t} u_{t+1})$ (Eq. 6.20), and rearrange terms, we obtain:

$$\Delta x_t = \alpha \delta \Delta x_{t-1} + \lambda \left(\sigma_{\delta,t-1} u_t - (\alpha + \delta) \sigma_{\delta,t-2} u_{t-1} + \alpha \delta \sigma_{\delta,t-3} u_{t-2} \right),$$

where, as before, $\{u_t\}$, is again a sequence of IID standard normal random variables.

For positive memory ($\alpha > 0$) and a-synchronous updating ($\delta > 0$), the dynamics of the differences $\{\Delta x_t\}$ is described by an ARMA(1,2) process. Interestingly, and in

contrast with our previous example, the (conditional) variance of the disturbances is no longer constant. The conditional variance is proportional to the dispersion of beliefs, which now exhibits endogenous fluctuations. Intuitively, these fluctuations in the degree of heterogeneity can be ascribed to the differences in opinion between those who decide to update their beliefs and those who believe that updating can wait. According to Eq. (6.19), the dynamics of heterogeneity is described by an autoregressive process of order one, from which one may correctly expect a GARCH(1,1) model to emerge for the conditional variances. For clarity of exposition we set memory to zero i.e. $\alpha = 0$, so that the initial ARMA(1,2) process for $\{\Delta x_t\}$ is stripped down to a MA(1) process:

$$
\begin{aligned}
\Delta x_t &= \lambda \sigma_{\delta,t-1} u_t - \lambda \delta \sigma_{\delta,t-2} u_{t-1} \\
&\equiv \epsilon_t - \delta \epsilon_{t-1}
\end{aligned}
\tag{6.21}
$$

where the conditional variance $h_t = \text{Var}_t(\epsilon_t) = \lambda^2 \sigma_{\delta,t-1}^2$ is updated according to:

$$
h_t = (1-\delta)\lambda^2\sigma^2 + \delta(1-\delta)\lambda^2\epsilon_{t-1}^2 + \delta h_{t-1},
\tag{6.22}
$$

where we have used Eq. (6.19) together with the fact that for $\alpha = 0$, Eq. (6.18) implies $\mu_t = x_t$, so that Eq. (6.20) gives $\mu_t - \mu_{\delta,t-1} = x_t - \mu_{\delta,t-1} = \epsilon_{t-1}^2$. Because $0 < \lambda^2 < 1$ a given GARCH parameter δ implies an upper limit of $\delta(1-\delta)$ for the ARCH parameter. Note that the sum $s = \delta + \lambda^2\delta(1-\delta)$ of the ARCH and GARCH parameters satisfies $0 < s < 1$ for $0 < \delta < 1$, so that the resulting GARCH process is stable.

6.3.3 Heterogeneity in memory

In this section we consider the third and final behavioral extension of the benchmark model: heterogeneity in memory. As the diversity in future time horizons seems relevant, also the diversity in past time horizons may play an important role in the co-evolution of prices and beliefs. With past time horizons we refer to the history of price observations used to test candidates for future trading. Agents who are more concerned with price behaviour on the short-term are more likely to compare performances over a smaller sample of recent observations than agents who have a clear interest in the long-term dynamics. Another reason for agents to use small memory parameters might be that they believe that the world they live in is inherently non-stationary, and only recent performances can be considered relevant.

Recently, a number of studies addressed the issue of different (future) time horizons

among agents (see e.g. Muller *et al.*, 1997; Dacorogna, 2002). This generalization is supported by the observation that agents trade at different frequencies. That is, the population of traders often consists of both long-term traders and short-term traders. Heterogeneity in agents' time scale are believed to be responsible for a number of stylized facts. Long term traders naturally focus on long-term behavior of prices thereby neglecting fluctuations at the smallest time scale, whereas short-term traders are not concerned with price movements on the long-run but rather aim to exploit short term predictability. The effects of the diversity in time horizons on price dynamics have been studied by LeBaron (2001) in an artificial stock market model. He concluded that the presence of heterogeneity in horizons may lead to an increase in return variability, as well as volatility-volume relationships similar to those of actual markets. Dacorogna (2002) found evidence supporting trading horizon heterogeneity to be responsible for the slow decay of correlations found empirically.

In the case where the memory parameter α is distributed among agents according to the pdf $\nu(\alpha)$ with $0 < \alpha < 1$, the most recent beliefs distribution for each memory type can be denoted by $\tilde{\phi}_{\alpha,t}(\theta)$. The overall beliefs distribution is then given by

$$\phi_t(\theta) = \int_0^1 \tilde{\phi}_{\alpha,t}(\theta)\nu(\alpha)\,\mathrm{d}\alpha.$$

Again, in general it can be very difficult to derive the updating rules for $\phi_t(\theta)$ analytically. However, if closed form expressions for the moments $E[\alpha^k]$ with respect to $\nu(\alpha)$ are known, for example, it is possible to derive analytic expressions for some of the moments, such as the mean μ_t and variance σ_t^2, of the beliefs distribution. This is demonstrated in some examples below.

Let α_i denote the memory of agent i, and $\mu_{i,t}$ and $\sigma_{i,t}^2$ represent the first two moments of the beliefs distribution corresponding to agents with memory parameter α_i. In the case of synchronous updating, it can be verified that $\sigma_{i,t}^2 = \sigma^2$ for all i, and that the dynamics of the average $\{\mu_t\}$ over the entire population is described by:

$$\mu_t = \int_0^1 \mu_t(\alpha)\nu(\alpha)\mathrm{d}\alpha. \tag{6.23}$$

For the mean over agents with memory parameter α_i we find

$$
\begin{aligned}
\mu_{i,t} &= \alpha_i \mu_{i,t-1} + (1 - \alpha_i) x_t \\
&= (1 - \alpha_i) \sum_{k=0}^{\infty} \alpha_i^k x_{t-k} \\
&= \sum_{k=0}^{\infty} \alpha_i^k x_{t-k} - \sum_{k=0}^{\infty} \alpha_i^{k+1} x_{t-k}.
\end{aligned}
$$

Taking expectations on both sides yields:

$$
\mu_t = \sum_{k=0}^{\infty} E_\nu \left[\alpha^k \right] x_{t-k} - \sum_{k=0}^{\infty} E_\nu \left[\alpha^{k+1} \right] x_{t-k}, \tag{6.24}
$$

where $E_\nu[\cdot]$ stands for the expectation with respect to the distribution of the memory parameter among agents, which has pdf $\nu(\alpha)$. For a large but finite number of agents, endogenous noise enters the price equation via

$$
x_t = \mu_{t-1} + \sigma_t \lambda u_t \tag{6.25}
$$

where λ is a measure for the scale of endogenous noise which can either result from the presence of herding or a finite number of agents, and $\{u_t\}$ a sequence of independent random variables with mean zero and unit variance (cf. Eqs (6.12) and (6.20)). From Eqs (6.24) and (6.25) we obtain:

$$
\begin{aligned}
x_t &= \sum_{k=0}^{\infty} E_\nu \left[\alpha^k \right] x_{t-k-1} - \sum_{k=0}^{\infty} E_\nu \left[\alpha^{k+1} \right] x_{t-k-1} + \sigma_t \lambda u_t \\
&= x_{t-1} - \sum_{k=1}^{\infty} E_\nu \left[\alpha^k \right] (x_{t-k} - x_{t-k-1}) + \sigma_t \lambda u_t.
\end{aligned}
$$

It follows that the increments $\Delta x_t = x_t - x_{t-1}$ are described by:

$$
\Delta x_t = - \sum_{k=1}^{\infty} E_\nu [\alpha^k] \Delta x_{t-k} + \sigma_t \lambda u_t.
$$

Upon defining L as the lag operator i.e. $Lx_t = x_{t-1}$, we can reformulate to obtain:

$$\left(\sum_{k=0}^{\infty} E_{\nu} \left[\alpha^k \right] L^k \right) \Delta x_t = \sigma_t \lambda u_t. \tag{6.26}$$

If α is assumed to be uniformly distributed on $[0, 1]$, then we have $E_{\nu} \left[\alpha^k \right] = \frac{1}{1+k}$ (see e.g. Linden, 1999, who studies the aggregated AR(1) process with uniformly distributed coefficients). As the latter does not decay exponentially when k tends to infinity, we observe long memory behavior for $\{\Delta x_t\}$. Thus for $\alpha \sim \text{UNIF}(0, 1)$, we have:

$$\left(\sum_{k=0}^{\infty} \frac{L^{k+1}}{k + 1} \right) \Delta x_t = L \sigma_t \lambda u_t,$$

or equivalently:

$$- \ln \left(1 - L \right) \Delta x_t = \sigma_{t-1} \lambda u_{t-1},$$

where the operator $\ln \left(1 - L \right)$ on the left hand side should be interpreted in terms of its power series in L.

More generally, we may consider a memory parameter distributed according to a BETA(a, b) distribution on $(0, 1)$, with pdf

$$\nu(\alpha) = \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} \alpha^{a-1} (1 - \alpha)^{b-1}, \qquad 0 < \alpha < 1.$$

The expectation $E_{\nu}[\alpha^k]$ can then be expressed as

$$E_{\nu}[\alpha^k] = \frac{\Gamma(a + b)}{\Gamma(a)} \frac{\Gamma(a + k)}{\Gamma(a + b + k)}.$$

Although we will leave a detailed analysis of the dynamics for some of these more general cases for future research, we will make some remarks here.

For $a = d$, $b = 1 - d$, with $d \in (0, 1)$, one finds

$$E_{\nu}[\alpha^k] L^k = \frac{\Gamma(k + d)}{\Gamma(d)\Gamma(k + 1)} L^k.$$

Since these correspond to the terms of a Taylor expansion of $(1 - L)^{-d}$ around $L = 0$, the

sum in Eq. (6.26) can be written in terms of the following fractionally integrated process:

$$(1 - L)^{-d} \Delta x_t = \sigma_t \lambda u_t.$$

Similar processes, when driven by IID normal innovations, are known as fractionally integrated brownian motion. For $d \in (0, \frac{1}{2})$ the latter is invertible and stationary, and the autocorrelation coefficients are negative (see Hosking, 1981), a property which is commonly observed in exchange rate returns, for example. Although Hosking refers to these processes as being short-memory, both the autocorrelation function and partial autocorrelation function decay hyperbolically to zero. Therefore fractionally integrated brownian motion is often referred to as being long-range dependent (see e.g. Campbell *et al.*, 1997). The fact that we find a related process is in line with the results of Granger (1980), who showed that aggregation of simple processes with different characteristic time scales can easily lead to long memory.

However, considering σ_t to be time independent would be an oversimplification. Although the variance σ_t^2 of beliefs among agents using the same memory parameter α converges to a constant σ^2, the belief parameters θ_t for agents with different memory parameters α are centered around different values of μ_t, which depend on α. The conditional variance of the θ_t for the population as a whole is therefore given by

$$\sigma_t^2 = \sigma^2 + \operatorname{Var}_\nu(\mu_t(\alpha))$$

where the variance of $\mu_t(\alpha)$ is taken with respect to the distribution $\nu(\alpha)$ of the memory parameter α among the population. For this variance one finds

$$\begin{aligned}
\operatorname{Var}_\nu(\mu_t(\alpha)) &= \operatorname{Var}_\nu\left((1 - \alpha) \sum_{k=0}^{\infty} \alpha^k x_{t-k}\right) \\
&= E_\nu\left[\left((1 - \alpha) \sum_{k=0}^{\infty} \alpha^k x_{t-k}\right)^2\right] - \mu_t^2,
\end{aligned}$$

where, as before, μ_t is given by Eq. (6.23). The first term on the right hand side can alternatively be expressed as

$$E_\nu\left[\left((1 - \alpha) \sum_{k=0}^{\infty} \alpha^k x_{t-k}\right)^2\right] = E_\nu\left[(1 - \alpha)^2 \sum_{m=0}^{\infty} \alpha^m \sum_{n=0}^{m} x_{t-n} x_{t-m+n}\right].$$

Although this is a closed form expression of the second moment of $\mu_t(\alpha_i)$ with respect to $\nu(\alpha)$, it seems that this term cannot be expressed in a form which permits simple recursive

updating as new prices x_t become available. It does however show that the dispersion of beliefs among agents obeys a non-trivial time dependence in the presence of heterogeneity in memory, and that the second moment of the innovations governing the price dynamics will exhibit a likewise complexity.

6.4   Empirical validation

As a framework for modelling the stochastic dynamics of prices and beliefs explicitly, the CBS aims to make insightful how the stylized facts of observed prices may be traced back to the unobservable diversity of beliefs, and vice versa. In the previous section we examined the specific role of three different behavioural aspects. By keeping the basic CBS as simple as possible, the three extended behavioral models were shown to relate directly to familiar econometric time-series models. The basis of the models was found to be the ARIMA model. Interestingly, a-synchronous updating of beliefs was found to translate into GARCH-type structure, and long range dependence (or long memory) emerged for agents employing different time horizons in evaluating past performance of beliefs (i.e. heterogeneity in memory).

As an immediate result, we find that the parameters of these well-known models from econometrics suddenly have an economic and behavioral interpretations attached. Although tempting, interpreting these hidden behavioral parameters directly from the data seems optimistic in this stage. We are aware of the fact that several relevant behavioral model aspects and economic variables have been ignored. The addition of relevant other features to the model is likely to affect the GARCH structure, or to alter the long range dependence. Our models do not capture, for example, longer range effects as described by Taylor and Allen (1992). Note, however, that the idea that the degree of fractional integration is a proxy for the degree of heterogeneity in memory finds support in Dacorogna (2002).

In this light there is little interest in focusing on the *exact* values of the parameter estimates. However, it will be interesting to see whether the signs of the MA coefficients and roughly the size of the GARCH parameters are both consistent with what is predicted by the theory i.e. the stylized CBSes. Finding negative values for the memory parameter, for example, is likely not to win support for our CBS. The predicted effects of the behavioral assumptions are put to the test in this empirical validation. Focus is in particular on the model obtained in the presence of a-synchronous updating (see Eqs. 6.21 and 6.22). The

Country	MA(1)	ARCH	GARCH	ARCH$_{\max}^{Implied}$
Belgium	−0.05	0.17	0.24	0.18
Germany	0.00	0.15	0.60	0.24
Finland	−0.05	0.05	0.95	0.05
France	−0.03	0.05	0.91	0.08
Austria	−0.03	0.06	0.90	0.09
Singapore	−0.08	0.08	0.91	0.08

Table 6.1: *Estimates of an MA(1) model with GARCH(1,1) disturbances for six exchanges rate returns, and implied ARCH upper limit (last column). The predicted MA coefficient was $-\psi$ with ψ non-negative (Eq. 6.15), implying non-positive MA coefficients, as found in all cases.*

ARCH and GARCH coefficients that emerge in that example are given by $\delta(1-\delta)$ and δ, respectively, where $(1-\delta)$ denotes the probability for each agent to update his belief at any given period in time. This behavioural aspect of the model thus predicts a relation between the ARCH and the GARCH coefficients, which can be checked empirically. Moreover, the model predicts a negative MA(1) coefficient equal to $-\delta$. Hereby we have set memory to zero, leaving us with a model that is more tractable, but also with one degree of freedom less. Keep in mind, however, that in the case of herding we also predicted a negative MA(1) coefficient, equal to $-\alpha$ (see Eq. 6.13).

The data used to illustrate the model are quotes from the foreign exchange market. We consider the daily exchange rate of the US dollar against the local currencies of six different countries. Five of these countries are European: Belgium, Germany, Finland, France, and Austria, while the last is from South East Asia, namely Singapore. Let p_t denote the amount of US dollars one can buy for one unit of the local currency at time t. Instead of absolute prices, however, we consider the daily logarithms i.e. $x_t = \log p_t$, so that Δx_t denotes the daily log return on the exchange rate. The sample period, from 5-11-1987 until 5-11-2002, includes a total of 3914 observations.

Table 6.1 reports the estimates of an MA(1) model with GARCH(1,1) disturbances for all six of the exchanges rate returns. The first three columns list the actual estimates, whereas the fourth column reports the *implied* upper limit for the ARCH coefficient i.e. the upper limit implied by our behavioral model which predicts a relation between the ARCH and GARCH coefficients (see Eq. 6.22, and the discussion following it). According to the latter equation, a GARCH coefficient of δ implies an ARCH coefficient no larger than $\delta(1-\delta)$. A first look at these results shows that the size of the estimated ARCH effect indeed satisfies this criterion obtained in the case of a-synchronous updating of beliefs.

With only one degree of freedom, namely the extent to which beliefs are updated in an a-synchronous fashion, it should not be surprising that the exact size of the estimated MA coefficient is not consistent with the (G)ARCH parameters. Apart from having set α to zero, the setup is also stylized in other respects. With that in mind, it is encouraging to see that the signs of the estimated MA coefficient are also consistent with what we would expect from the theory. In sum, in light of all limitations, we consider the empirical validation successful and hope that it may open doors to more informative empirical assessments by means of more sophisticated CBSes.

6.5   Concluding remarks

The objective of this chapter was to make insightful how certain behavioral features may render some of the well-established stylized facts of observed prices. In order of treatment, we considered: (i) herd behaviour; (ii) a-synchronous updating of beliefs; and (iii) heterogeneity in the time horizons (memory) of agents. The adopted framework is that of the Continuous Beliefs System (CBS) (see Diks and van der Weide, 2003). We used two qualities of a CBS that are particularly useful. First, it enables modelling the evolution of the distribution of beliefs in an explicit fashion. Second, it does so by relating the unpredictable nature of prices with that of the unpredictable nature of individual preferences. In sum, we used the fact that a CBS allows for describing the stochastic dynamics of both prices and beliefs explicitly by means of a Random Dynamical System.

By keeping our benchmark CBS illustrative but deliberately simple, the stochastic dynamics of each of the three behavioral models was shown to correspond directly to a model well-known from time-series econometrics. We came across ARIMA, ARFIMA, and GARCH. This marks our first result, as the stylized facts for which these models are celebrated may now be traced back to the hidden features of beliefs which we hold responsible. Here, the ultimate simplicity of the benchmark CBS helped us to isolate the role of each of the three behavioural features considered.

As we summarize the results for the three models seperately, keep in mind that our point of departure was a model in which the degree of heterogeneity is constant, and where price dynamics is described by a random walk. For each behavioral extension we investigated how it modifies both the dynamics of prices and that of heterogeneity.

The effect of herd behaviour is roughly that of reducing the number of agents in the market. As such, it increases the amount of endogenous uncertainty, and thereby the

level of volatility. Moreover, it adds moving average structure, leading to an ARIMA specification in levels or MA in differences. Notably, it does not alter the dynamics of heterogeneity. The latter remains constant.

Fluctuations do enter the dynamics of heterogeneity when we allow agents to update their beliefs at different frequencies. These fluctuations are found to be driven by the differences in average beliefs between those who decide to evaluate and update, and those who believe that updating can wait. If endogenous uncertainty persists in the market, for example as a result of herd behaviour, a time-varying degree of heterogeneity results into a time-varying level of volatility. The extent of persistence in volatility is governed by the likelihood that individual agents delay the updating of their beliefs. This result finds support in an earlier study by Brock and LeBaron (1996), who saw the persistence in volatility of simulated price returns rise to a more realistic level when allowing for less frequent updating. Interestingly, we obtained an explicit description of how the updating of beliefs interacts with the volatility process. For our stylized but yet intuitive example we found the dynamics that to be exactly that of a GARCH(1,1) model.

Heterogeneity in the time horizons (memory) of agents was found to enrich the dynamics of both the first and second moment. Assuming a distribution in memory rather than identical memory for each agent, we encountered one more stylized fact, namely that of long-range dependence. Leaving aside the volatility process for a moment, taking a $BETA(d, 1-d)$ distribution for memory, we obtained the well-known fractionally integrated process documented by e.g. Hosking (1981) for the first moment. When delving into greater detail, we quickly stumbled upon greater complexity. In particular, the degree of heterogeneity was found to move in mysterious ways. Although the heterogeneity among those who employ identical memory is constant over time, additional dispersion of opinions is present which can be attributed exclusively to the fact that not all agents operate at the same time horizon. The latter does not only add to the overall degree of heterogeneity, it was also found to be time-varying.

To understand the complexity, note that the supply of fluctuations is provided by the dispersion of average beliefs across groups of agents that are operating at different time horizons. As memory plays an important role in the updating of beliefs, we find that the evolution of average beliefs for each of those groups of agents interacts differently with the price dynamics, given that their memories are different. Heterogeneity being driven by the diversity of these evolutions, it requires aggregation of these different dynamics to obtain the dynamics of heterogeneity. The latter, although interesting, easily becomes convoluted.

At the end of the chapter we included a modest empirical validation to address the intimate relation between our behavioural models and the more familiar models from econometrics. As a corollary, the parameters of these time-series models now have an economic and/or behavioural interpretation attached. For example, our stylized model predicts a relation between the ARCH and GARCH parameters. Merely as a 'reality check', we verified whether both that relation, and the signs of the other coefficients, finds support in the empirical data. In light of all simplifications, we consider the empirical validation successful, and hope that more informative empirical assessments by means of more sophisticated CBSes can be made in the future.

As a final remark we would like to comment on the difference with the recently proposed LTL approach of Brock *et al.* (2005). Although a CBS is more explicit in terms of the beliefs distribution, the price dynamics of a CBS and an LTL are closely related. (For a more elaborate comparison, see Diks and van der Weide (2003).) In case the opportunity function $\varphi(\theta)$ (see footnote 2) is normalizable, the LTL coincides with the deterministic part of the corresponding CBS. However, there are two key differences. For one, it is not clear how an LTL can be set up with a non-normalizable opportunity function (which for an LTL corresponds to the pdf according to which strategies are drawn at random). More importantly, an LTL is a deterministic dynamical system by construction, and up until now not tailored to dealing with endogenous randomness. It is exactly the associated stochastic dynamics that plays a leading role in this chapter.

Samenvatting (Summary in Dutch)

Het proefschrift bestaat, naast het inleidende hoodstuk, uit vijf hoofdstukken. Het eerste gedeelte introduceert een model met als doel om empirische data van financiële markten beter te beschrijven, het tweede gedeelte introduceert een model met als doel om dezelfde data beter te begrijpen. Hoofdstuk twee en drie concentreren zich metname op het econometrisch modelleren van (multivariate) financiële volatiliteit. Hoofdstuk vier fungeert als een brug tussen de twee verschillende delen van mijn proefschrift. Dit hoofdstuk doet een poging om marktverwachtingen over toekomstige financiële volatiliteit te schatten uit empirische optieprijs data. In de laatste twee hoofdstukken ontwikkelen we een structureel economisch model waarin de manier waarop economische agenten verwachtingen vormen over toekomstge prijzen centraal staat. Hoofdstuk vijf introduceert een analytisch raamwerk dat is ontworpen om de co-evolutie van marktverwachtingen en prijzen te modelleren. Hoofdstuk zes behandelt een aantal speciale gevallen van dit raamwerk door de nadruk te leggen op een selectie van verschillende gedragskenmerken van de economische agenten (in het bijzonder de manier waarop ze verwachtingen vormen).

In hoofdstuk twee wordt een nieuw multivariaat volatiliteitmodel geintroduceert, naar welke gerefereerd zal worden als Gegeneraliseerd Orthogonaal GARCH (GO-GARCH). Het is een lid van de ARCH-familie, waar ARCH voor Auto-Regressieve Conditionele Heteroskedasticiteit staat. Het ARCH-model is tot op heden waarschijnlijk het meest populaire model dat wordt gebruikt voor het beschrijven van de bewegelijkheid van financiële volatiliteit over de tijd. Het model werd geïntroduceerd door Engle in 1982 en is kort daarna gegeneraliseerd tot GARCH door Bollerslev (in 1986). De multivariate uitbreiding, multivariate GARCH, voorziet in een modelspecificatie voor de tijd-variatie van de covariantie matrix (met tijd-variërende varianties op- en covarianties van de diagonaal). Schattingen van multivariate GARCH modellen kunnen echter problematisch zijn omdat het aantal onbekende model parameters die moeten worden geschat snel toenemen met de dimensie.

De eerste algemene multivariate GARCH modellen, welke naar voren zijn gebracht in de literatuur, gebruiken een relatief groot aantal parameters welke tot convergentie moeilijkheden leiden bij het schatten (zie bv. Bauwens e.a., 2006). Nieuwe specificaties worden vaak bepaald door middel van praktische overwegingen. De uitdaging is om een parameterisatie van de covariantie matrix te vinden die uitvoerbaar is in schattingstermen tegen een minimaal verlies van algemeenheid.

Het GO-GARCH model verwijst naar een natuurlijke generalisatie van het O-GARCH model, en nestelt zich als een speciaal geval in het meer algemene BEKK model. Het geeft ruimte aan de belangrijkste gestileerde feiten van multivariate volatiliteit terwijl de uitvoerbaarheid behouden blijft. Zowel kunstmatige als empirische voorbeelden zijn opgenomen om het model te illustreren. Voor de gepubliceerde versie van het hoofdstuk zie van der Weide (2002).

In hoofdstuk drie introduceren we een nieuwe schattingsmethode voor voor het GO-GARCH model. De methode maakt gebruik van de eigenvectoren van een geschikte selectie van steekproef autocorrelatie matrices van gekwadrateerde en kruisvermenigvuldigingen van de rendementen data. De methode is numeriek aantrekkelijker dan de originele schattings methode. In het bijzonder vereist de nieuwe methode geen strikte veronderstellingen omtrent de volatiliteitmodellen van de latente factoren, en is daarom minder gevoelig voor model misspecificatie. Het hoofdstuk voorziet zowel in Monte-Carlo simulaties, en in een studie van empirische data van Europese sector rendementen. Voor de gepubliceerde versie zie Boswijk en van der Weide (2011).

Ingebed in optieprijzen zitten marktverwachtingen over toekomstige volatiliteit. Hoewel de veronderstelling van rationele verwachtingen een populair paradigma is, is het moeilijk om de subjectieve aard van verwachtingen te negeren. Het doel van hoofdstuk vier is om marktverwachtingen zichtbaar te maken terwijl zij zich ontwikkelen over de tijd, en om opties te prijzen consistent met de heersende verwachtingen, of deze nu rationeel of niet-rationeel zijn. We werken met een analytisch passend raamwerk voor het prijzen van opties dat zowel voorziet in stochastische volatiliteit als a-symmetrische volatiliteit. Dagelijkse schattingen van de geïmpliceerde verdeling van volatiliteit zijn verkregen door, dag voor dag, het model op optie prijzen te schatten. We leggen geen structuur op hoe verwachtingen worden bijgesteld door de tijd heen, maar laten toe dat marktverwachtingen vrij kunnen bewegen. Zie Peters en van der Weide (2011) voor de working paper versie.

Met hoofdstuk vijf concentreren we ons op het economisch modelleren van het proces waarmee agenten hun verwachtingen vormen. We introduceren een nieuw analytisch

raamwerk voor het bestuderen van de co-evolutie van verwachtingen en prijzen. Agenten baseren hun keuzes op resultaten uit het verleden en her-evalueren strategieën zodra nieuwe informatie beschikbaar komt. Door individuele keuzes te beschouwen als stochastische variabelen, wat natuurlijk is in een stochastische preferentie raamwerk, kan heterogeniteit gezien worden als een natuurlijke bron van stochasticiteit in de dynamica van prijzen. We beschouwen enkele voorbeelden waarvoor deze stochasticiteit niet verdwijnt zelfs als het aantal agenten naar oneindigheid gaat. Het hoofdstuk is gepubliceerd als een working paper, zie Diks en van der Weide (2003).

In hoofdstuk zes beschouwen we een specifiek eenvoudig voorbeeld van het raamwerk dat geïntroduceerd is in hoofdstuk vijf met het doel om de effecten te onderzoeken op prijsdynamieken van enkele gedrags veronderstellingen: (i) kuddegedrag; (ii) a-synchroon herzien van verwachtingen; en (iii) heterogeniteit in tijds-horizonnen (geheugen) onder agenten. Het benchmarkmodel met veel handelaren geeft een "random walk" gedreven door nieuws. Door kuddegedrag te introduceren wordt dit proces aangepast tot een ARIMA (0,1,1) proces. In termen van rendementen voorspelt het model MA(1) structuur met een negatieve coëfficiënt. A-synchroon herzien van verwachtingen leidt tot een MA(1) model voor rendementen met GARCH(1,1) innovaties, en voorspelt een relatie tussen de ARCH en GARCH coëfficiënten. Heterogeniteit in geheugen leidt tot lange termijn afhankelijkheid in rendementen. In het empirische deel van het hoofdstuk voeren we een bescheiden 'reality-check' uit met betrekking tot het verwachte teken van de MA coëfficiënt en de relatie tussen de ARCH en GARCH coëfficiënten voor wisselkoersdata. Voor de gepubliceerde versie van het hoofdstuk, zie Diks en van der Weide (2005).