Continuous Beliefs Dynamics*

Cees Diks[†] & Roy van der Weide[‡]

CeNDEF, Faculty of Economics and Econometrics, University of Amsterdam, and Tinbergen Institute, Roetersstraat 11, 1018 WB, Amsterdam, The Netherlands

October 28, 2002

Abstract

We propose a general framework for studying the evolution of heterogeneous beliefs in a dynamic feedback setting. Beliefs distributions are defined on a continuous space representing the possible strategies agents can choose from. Agents base their choices on past performances. As new information becomes available strategies are re-evaluated and the beliefs distribution is updated using a continuous choice model. This approach gives rise to price dynamics in which the beliefs distribution evolves together with realized prices. The statistical properties of the endogenous random price fluctuations are fully determined by the model. The structure of the macroscopic model depends on the class of predictors and on the performance measure used by the agents. Whenever a well-known econometric model is obtained, an economic interpretation of the model parameters can be given, as is shown here for an ARCH model. The approach is illustrated with several examples and empirical applications.

Keywords: Expectation formation; Heterogeneity; Continuous beliefs; Expectations feedback; Endogenous noise;

JEL classification: C00, D84, G12

*The authors wish to thank the participants of the CeNDEF workshop, June 2002, in Leiden for stimulating discussions. Research supported by the Netherlands Organization for Scientific Research (NWO) under a MaGW-Pionier and a MaGW-Vernieuwingsimpuls grant.

[†] *E-mail address*: diks@fee.uva.nl
[‡] *E-mail address*: rvdweide@fee.uva.nl

1 Introduction

As little as it takes one to realize that expectations about future prices are essential for investment decisions made by market participants, as difficult it is to study the interaction between prices and expectations empirically. Expectations, let alone the beliefs on which they are based, are hardly ever directly observable in practice. For stock markets, for example, one typically observes realized prices, reflecting market aggregated expectations rather than individual expectations.

Although realized economic observables such as prices can be viewed as some (discounted) market aggregate of expectations among market parties, this does not imply that the time series properties of the aggregates can be modeled easily in terms of market aggregates only. On the contrary, recent results indicate that, to a large extent, the dispersion of beliefs among market participants bears important consequences to the behavior of the aggregates over time. In the recent literature on heterogeneity several theoretical models have been proposed and a number of relations between heterogeneity and stylized facts were derived and validated empirically. Shalen (1993) and Michaely and Vila (1996), for example, marked that dispersion of beliefs enhances both trading volume and volatility. This supports the empirical evidence reported by Gallant *et al.* (1992) of a positive correlation between the two. Recent work also indicates that heterogeneity of expectations might be responsible for several other empirical observations. For example, Diether *et al.* (2002) presented evidence that the dispersion of opinions on future returns affects future earnings negatively, while Ziegler (2002) found that heterogeneity of beliefs can lead to the well-known "smile effect" in implied option volatility.

Considering the fact that the effects of heterogeneity have been studied so extensively, surprisingly little is known about the dynamical feedback between market aggregates and heterogeneity among agents. Although the effects of heterogeneity on observables are fairly well understood, current theory does not seem to include a detailed description of the response of heterogeneity to incoming observations, and the implications of endogenously evolving heterogeneity on the time series properties of observables. The aim of this paper is to develop a general framework for studying these issues. This is done by constructing an expectations based economic dynamical model which includes a full description of the nature of the dynamical feedback between realized prices and the distribution of beliefs. Insights into the nature of this feedback can help explain several observations made in the literature on heterogeneity in expectations. Moreover, at a more fundamental level, the long-run objective motivating the approach developed here is to enable one to study the time evolution of heterogeneity among agents via the interaction between publically available information and the distribution of beliefs of market participants. Future empirical applications of the concept might eventually allow one to visualize the evolution of beliefs and heterogeneity from observed time series data. The current paper should be considered a first step in this direction, focusing on the development of the general methodology while illustrating implications and potential applications with examples.

The context we have in mind is one in which the beliefs of individual agents are updated when new public information (e.g. prices) becomes available. The investment decisions made by the agents, based on their expectations, in turn determine the new price. To this end we introduce the concept of a continuous beliefs system (CBS), which, in short, contains the following basic ingredients.

A central role is played by the *beliefs space*, Ω, in which a class of possible predictors, given a public *information set*, is represented by a continuous parameter θ. At any time t agents form beliefs regarding future economic variables by deciding on a belief parameter θ in the beliefs space. They do so by evaluating the possible strategies, given the information available prior to time t, using a *performance measure*. This performance measure might be based, for example, on the history of past prediction errors, or on profits a strategy would have realized in the past. Individual agents choose the strategy which optimizes their expected utility, but since individual agents differ in tastes, they will typically differ in the strategy they will choose for the next period. The distribution of beliefs over the beliefs space is described by a probability density function (pdf) $\phi_t(\theta)$, called the *beliefs distribution*. The beliefs distribution can be thought of as representing the empirical fact that even when agents have identical information, they still have differences of opinion regarding future revenues (see e.g. Frankel and Froot, 1990; Kandel and Pearson, 1995). The specific form of the beliefs distribution given the history of realized past observables, can be expressed in terms of past performances using a continuous choice model. The belief, (or predictor, or strategy) $\theta_{i,t}$ actually employed by agent i at time t, from the modeler's point of view, is considered to be a random variable distributed according to the beliefs distribution. Given the individual beliefs of agents prior to time t, their net demand functions determine the next observable (price) to be quoted publically. This involves some *aggregation mechanism*, such as a market maker setting the price so as to obtain zero aggregated net demand. After this observable becomes public, the performances of strategies can be re-evaluated, a new beliefs distribution arises, etc. From a dynamical point of view, the ongoing evolution of strategies by agents as new information becomes available results in an evolution of the distribution of beliefs over time, while the beliefs distribution in turn determines the properties of the aggregate observables.

The CBS approach has several theoretical implications. A dynamical system is obtained in which the beliefs distribution co-evolves together with aggregate observables such as prices. An explicit feedback relation between aggregate economic observables and the beliefs distribution is obtained, and often analytically tractable equations are found for the aggregate observables and the first few moments of the beliefs distribution. In this respect the approach is more convenient for studying the evolution of beliefs than the approach of Brock *et al.* (2001), although we should say that their motivation was also completely different from ours. Their 'large type limit' (LTL) approach was developed to investigate whether the aggregate behaves more realistically in a market where agents can choose between a large number of alternative strategies. While Brock *et al.* (2001) consider the dynamical aspects of a continuum of agents who are allowed to choose from an increasing number of strategies, we instead start with a continuum of strategies and examine the dynamics for a finite number of agents. By construction this approach provides more explicit insights into the nature of the feedback between the aggregate observables and the beliefs distribution, for any number of agents.

Another essential difference between the LTL approach and ours can be found in the stochastic aspects of the models obtained. In a CBS the predictors used by the agents are modeled by the continuous choice model. Given the model parameters, the beliefs distribution can be determined explicitly by the econometrician in terms of past realized observables. That is, the beliefs distribution is a deterministic function of past observables. However, the fact that the strategies used by the agents are modeled as random variables, distributed according to the beliefs distribution, provides a natural mechanism for endogenous randomness. By construction, the dynamics for a finite number of agents is stochastic. Even if the number of agents tends to infinity, there may be conditions under which the individual choices of agents do not average out to give a deterministic law for the aggregate observables. If the dependence among the agent's choices is sufficient, aggregates may remain random for an infinity of agents. Also, for some combinations of the class of predictors and the performance measure, the conditions for the law of large numbers may not be satisfied, providing another source of randomness which does not vanish if the number of agents increases.

The natural sources of randomness of a CBS provide theoretical mechanisms which might help explain part of the market fluctuations we observe in daily life. For example, the randomness might provide an extra source of excess volatility in markets as described by Shiller (1981), while the usual agent based models, such as that proposed by Brock and Hommes (1997), only attempt to uncover the deterministic part of excess volatility. The stochastic properties of the CBS approach might also prove useful for examining the statistical aspects of agent models analytically, rather than through the laborious simulations usually performed in computational agent models (for an overview, see LeBaron, 2000).

Apart from theoretical consequences, the natural randomness in the model also provides some extra grip on agents based empirical research. The stochastic aspects of the feedback dynamics opens the door to possible empirical applications, since the magnitude and distribution of the noise terms in the dynamic feedback equations can be derived analytically. The latter is of direct importance to model estimation, since it implies that this can be done by standard maximum likelihood methods. To the best of our knowledge, in the literature, the effects of endogenous noise have only been studied with random shocks with ad-hoc distributional properties, modeled as additive noise on the price equation without further theoretical justification. In fact, in some of the applications presented later it will become clear that most of the nontrivial structure arising from heterogeneous beliefs feedback loops resides in the noise rather than in the deterministic skeleton. In a stylized example of a CBS in which agents have AR(1) beliefs regarding returns, we show that the returns dynamics is equivalent to a linear random coefficient model, in which the mean and variance of the beliefs distribution appear as time-varying coefficients. Since the class of random coefficient models obtained contains the ARCH model as a specific case, a straightforward interpretation of ARCH parameters can be given in terms of economic quantities. To illustrate a possible application we estimate the (effective) number of active market parties across markets and over time, and briefly touch upon the issue of visualizing the evolution of beliefs and heterogeneity from observed time series data.

This paper is organized as follows. In section 2 the concept of a beliefs distribution in a continuous beliefs space is introduced, as well as its form given by the continuous choice model. Section 3 introduces the concept of a continuous beliefs system in which the co-evolution of beliefs and public information such as market prices is taken into account. As an illustration the implied dynamics in a standard asset pricing context is examined. In section 4 the mechanism by which endogenous noise arises from the dynamic continuous choice dynamics is described, and the role of the number of agents and their dependence considered. In section 5 potential empirical implications are discussed in the context of an extremely simple stylized model. Section 6 summarizes and discusses the results.

2 Continuous beliefs distributions

This section describes how a continuous beliefs distribution is obtained in a continuous choice setting in which agents are faced with a single choice. The dynamic aspects, which involve expectations feedbacks, are postponed to the next section. The continuous choice model has some analogy with the discrete choice model. With the proper additional measure theoretic structure the discrete choice model, and mixed discrete/continuous choice models, can in fact be considered as special cases of the continuous choice model. For clarity of exposition we will not delve deeply into measure theoretic arguments, but rather present both models starting with the more familiar discrete choice model, referring the interested reader to the literature cited for more detailed analogies.

Agent based models represent market participants as (a typically large number of) agents, who can select among a number of alternative strategies. If the strategies among which the agents can choose consists of a finite set of strategies, $s_1, \ldots, s_m$ say, then agents employing strategy s_i, $i = 1, \ldots, m$ are said to be of type i. McFadden (1973) derived an expression for the probability p_i that an individual will select strategy i, starting from the concept of random utility functions. It is assumed that the utility function of agent j can be written as

$$V_j(s) = U(s) + \epsilon_j(s)$$

where $U(s)$ is a non-stochastic "common" utility function representing the tastes of the population, and $\epsilon_j(s)$ is stochastic and reflects the idiosyncrasies of individuals in tastes. The individuals choose the alternative which optimizes their utility. Under the assumption that the disturbances of the utility function follow an extreme value distribution, it can be shown that this leads to the *multinomial logit model*:

$$p_i = \frac{\mathrm{e}^{\beta U(s_i)}}{\sum_{l=1}^{m} \mathrm{e}^{\beta U(s_l)}},$$

where $U(s_l)$ is the utility associated with alternative l. The parameter β is referred to as the *intensity of choice*, and is related to the scale of the noise term $\epsilon_j(s)$. The larger the value of β, the smaller the noise, and the larger the probability that an agent chooses the option which actually optimizes $U(s)$. This is why $1/\beta$ is sometimes interpreted as the propensity of agents to err, presuming they actually all wish to optimize $U(s)$.

In the presence of a continuum of belief types it is convenient to introduce a finite dimensional parameter space Ω, containing all possible alternatives that can be selected by the agents. In our context, each possible choice, that is, each element θ in Ω uniquely represents a possible strategy, and Ω will be referred to as the *beliefs space*. Note that the choice of the beliefs space is not unique, since any one-to-one transformation of the beliefs space Ω into another space, Ω', say, will again yield a suitable beliefs representation. In analogy with the discrete choice model, we wish to represent the diversity of belief types by a probability distribution function over the beliefs space. The distribution of strategies can be obtained from the generalization of the discrete choice model referred to as mixed discrete/continuous choice models. Let us denote the pdf associated with the beliefs distribution by $\phi(\theta)$. As in the discrete choice setting, it is convenient to adopt a random utility approach (Hanemann, 1984; Dagsvik, 1994; Resnick and Roy, 1994). The random part of the utility function of an agent will eventually determine which strategy that particular agent considers optimal. Therefore, the strategies employed by individual agents in a random utility framework are random variables (note that this does not imply that individual agents perceive their own utility functions to be random, only that they are random to the econometrician). If we let θ_i denote the strategy adopted by agent i, then the pdf $\phi(\theta)$ is the probability density function associated with the random variables θ_i, or the relative likelihood that agent i selects a strategy in an infinitesimal neighborhood of θ. The continuous choice analogue of the multinomial logit model is the continuous logit model (see e.g. Ben-Akiva and Watanatada 1981; Dagsvik, 1994):

$$\phi(\theta) = \frac{\mathrm{e}^{\beta U(\theta)}\varphi(\theta)}{Z}, \tag{1}$$

where Z is a normalization constant, given by

$$Z = \int_\Omega \mathrm{e}^{\beta U(\vartheta)}\varphi(\vartheta)\mathrm{d}\vartheta.$$

As in the discrete choice setting, β represents the intensity of choice. The function $\varphi(\theta)$ is nonnegative, and can be used to put different weights on different parts of the beliefs space. We refer to $\varphi(\theta)$ as the *opportunity function*. The introduction of $\varphi(\theta)$ is not an additional expansion of the model, but necessary for obtaining a representation that is consistent in that all averages determined from the beliefs distribution should remain invariant under a re-parameterization of the beliefs space Ω. Following the mathematical convention, the beliefs distributions obtained with choice models that are equivalent in this sense will be called *equivariant*. From a modeling point of view, a choice model is fixed as soon as it is specified up to a certain equivariance class. Formally, $\varphi(\theta)$ represents the integration measure in the beliefs space Ω. In economic applications the opportunity function can be thought of as reflecting the *a priori* faith of individuals in parameters within certain regions of the parameter space. If agents have a tendency to avoid strategies in certain parts of the parameter space, for example because extreme parameter values are implausible to agents, this will be reflected by small values of $\varphi(\theta)$ in those regions of the parameter space. In that case, such “extreme” strategies need to outperform more common strategies to a large extent in order to convince the agents of their use.

The beliefs distribution is equivariant under one-to-one transformations of Ω, provided that the opportunity function is transformed properly when moving from one parameterization to another, as stated in the following proposition.

Proposition 1 *The beliefs representation is independent of the parameterization of Ω, if the opportunity function $\varphi(\theta)$ transforms as a density function under one-to-one measurable coordinate transformations $g : \Omega \mapsto \Omega'$, that is, $\varphi'(\theta') =\mid J_{g^{-1}}(\theta') \mid \varphi(g^{-1}(\theta'))$, where $J_{g^{-1}}(\theta')$ is the Jacobian of $g^{-1}(\cdot)$ evaluated at θ'.*

Proof: Requiring that the pdf $\phi(\theta)$ transforms as a density leads to

$$\phi'(\theta') =\mid J_{g^{-1}}(\theta') \mid \phi(g^{-1}(\theta')).$$

From the definition of $\phi(\theta)$, after substituting $U(g^{-1}(\theta')) = U'(\theta')$, one obtains

$$\mathrm{e}^{\beta U(g^{-1}(\theta'))}\varphi'(\theta') =\mid J_{g^{-1}}(\theta') \mid \mathrm{e}^{\beta U(g^{-1}(\theta'))}\varphi(g^{-1}(\theta')).$$

which implies $\varphi'(\theta') =\mid J_{g^{-1}}(\theta') \mid \varphi(g^{-1}(\theta'))$. □

Generally speaking, we neither need to require that $\varphi(\theta)$ be a pdf, nor that it is integrable in order for the distribution of beliefs $\phi(\theta)$ to be well-defined. For example, if $\Omega = \mathbf{R}^m$, $\phi(\theta) = 1$, and $U(\theta)$ is a quadratic form in θ with a single maximum, then $\phi(\theta)$ is a multivariate normal probability density function.

As stated above the distribution of beliefs does not always exist. The following proposition provides a necessary and sufficient condition for the existence of the probability density of beliefs over Ω.

Proposition 2 *The beliefs distribution $\phi(\theta)$ given in Eq. (1) is well-defined as a pdf if and only if $Z = \int_\Omega \mathrm{e}^{\beta U(\vartheta)}\varphi(\vartheta)\,\mathrm{d}\vartheta$ is positive and finite.*

Proof: Since $\mathrm{e}^{\beta U(\theta)}$ and $\varphi(\theta)$ are non-negative, the function $\phi(\theta)$ is nonnegative if and only if $\int_\Omega \mathrm{e}^{\beta U(\vartheta)}\varphi(\vartheta)\,\mathrm{d}\vartheta$ is positive and finite, in which case $\int_\Omega \phi(\vartheta)\,\mathrm{d}\vartheta$ equals one. □

An equivalent way of stating this proposition is that $\phi(\theta)$ is well-defined if and only if $\mathrm{e}^{\beta U(\theta)}$ is φ-integrable, and the integral over Ω is nonzero. It is possible to give some sufficient conditions for $\mathrm{e}^{\beta U(\theta)}\varphi(\theta)$ to be integrable. For example, if either $\varphi(\theta)$ or $\mathrm{e}^{\beta U(\theta)}$ is bounded from above and away from zero, and the other has a positive and finite Lebesgue integral, $\mathrm{e}^{\beta U(\theta)}\varphi(\theta)$ is integrable and the integral nonzero. In that case Z is finite and $\phi(\theta)$ is well-defined.

3 Continuous beliefs systems

In this section we discuss the co-evolution of economic observables and the beliefs distribution in a CBS. Next to the beliefs space and the beliefs distribution, two additional ingredients are required. First, we assume that agents evaluate strategies according to some *performance measure*, which might for example be based on past prediction errors, or past profits a strategy would have realized in case it would have been used in the past. Secondly, a market mechanism, or more generally, an *aggregation mechanism* is required which translates the individual beliefs of agents into publically available information.

For clarity of exposition the presentation in this section is biased toward an asset pricing framework, the modifications required for applications in other dynamic settings in which beliefs distributions are dynamically updated being straightforward. As is also common in discrete choice agent models, agents are assumed to be myopic mean variance optimizers. More generally, one could consider agents who make choices depending on the entire future, but for the purposes of illustrating the methodology the myopic setting suffices. Throughout, time will be considered to be discrete in the models. For a recently proposed generalization of discrete choice models to continuous time, we refer the interested reader to Dagsvik (2002).

3.1 The evolution of the beliefs distribution

We consider the evolution of a time dependent economic variable, p_t, say. The discussion here focuses around asset prices here, but in general p_t might represent any other economic quantity, such as an exchange rate, bond price, etc., which depends on aggregate expectations. For simplicity we discuss the case where the agents are myopic in that, at time t, they only form expectations, about the price of the asset at time $t+1$ (the next time they can sell the asset) including possible dividends payed in this period. The information available to agents just before time t is denoted by $\mathcal{F}_t$. In simple cases the information set could consists of a historic record of prices p_s up to and including p_{t-1}, but in general it might also include exogenous variables, such as the interest rate. The possible strategies from which the agent can choose to predict future prices are represented by a function of the observables in the information set, parameterized by θ, $f_\theta(p_{t-1}, p_{t-2}, \ldots)$ say. For a fixed price history we can consider this predictor as a function of θ. The predictors for a given information set are denoted by

$$p^e_{t+1}(\theta) = f_\theta(p_{t-1}, p_{t-2}, \ldots).$$

Thus, the prediction of p_{t+1} made by agent i, based on price information up to and including p_{t-1}, using strategy $\theta_{i,t-1}$, is denoted by $p^e_{t+1}(\theta_{i,t-1})$. To give an example of a beliefs space, the class of d-th order linear predictors consists of all predictors of the form

$$p^e_{t+1}(\theta) = \theta_0 + \theta_1 p_{t-1} + \ldots + \theta_d p_{t-d}.$$

In this case the beliefs are represented in $\mathbf{R}^{d+1}$ and the beliefs distribution is a probability distribution on this space.

The beliefs distribution based on $\mathcal{F}_t$ is denoted by $\phi_{t-1}(\theta)$. Schematically, the expectations feedback can then be represented as

$$\ldots \to p_{t-1} \to \phi_{t-1}(\theta) \to p_t \to \phi_t(\theta) \to \ldots.$$

First we consider the formation of the observables p_t, for a finite number, n, of agents, each of which is assigned a strategy according to the previous beliefs distribution $\phi_{t-1}(\theta)$, possibly allowing for dependence between agents. We assume that each predictor has a unique associated optimal trading strategy. Also, we assume that the price p_t will be determined by some *aggregation mechanism* (one might think e.g. of market clearing), the outcome of which depends on the agents' expectations, at time t, then, given the agents' beliefs, or predictors $\theta_{i,t-1}$ the state variable p_t is completely determined by the market mechanism. In dynamic asset pricing models for example, todays price can be determined by the present value of aggregate beliefs concerning future prices and dividends:

$$(1+r)p_t = \frac{1}{n}\sum_{i=1}^{n} p^e_{t+1}(\theta_{i,t-1}) + y^e_{t+1}$$

where n is the number of agents and $p^e(\theta_{i,t-1})$ denotes the expected price of individual i based on past observables available up to and including t, and y^e_{t+1} expected future dividends. More generally, one might introduce market weights $w_{i,t}$, denoting the market weight of the i'th agent at time t, such that $0 \le w_{i,t} \le 1$ and $\sum_i w_{i,t} = 1$, which gives: $(1+r)p_t = \sum_{i=1}^{n} w_{i,t} p^e_{t+1}(\theta_{i,t-1}) + y^e_{t+1}$. This will be discussed in more detail in section 4.2.

Alternatively, in a cobweb framework the price equation typically has the form

$$p_t = D^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} S(p^e_t(\theta_{i,t-1}))\right),$$

where $S(p^e_t(\theta_{i,t-1}))$ is firm i's supply and $\frac{1}{n}\sum_{i=1}^{n} S(p^e_t(\theta_{i,t-1}))$ is the aggregate supply, and $D(p)$ aggregate demand at a given price p. Note that we have assumed homogeneity in supply curves, but more generally one might consider cases where firms have different supply curves $S_i(p)$.

Notice that we have only considered point predictors here. Some price formation mechanisms can not be formulated in terms of point predictors only, however. For example, Guesnerie (2002) considers agents who have subset predictors (e.g. an interval) in mind rather than point predictors. Under certain conditions, the agents can, by a rationality assumption and a common knowledge argument, agree on a unique trading price. Such a mechanism could be incorporated in a CBS framework, provided that the more general predictors can be represented by a finite number of parameters, and the price agents eventually agree to trade on can be expressed explicitly in terms of the agents' individual beliefs parameters.

Next we consider how the new price affects the beliefs distribution. After p_t becomes part of the information set, agents can re-evaluate the possible strategies. At this point we wish to invoke the continuous choice model to obtain an expression for the new beliefs distribution. Although we discussed the continuous choice model in general in the previous section, a particular CBS requires a specification of the the utility function. The common (nonrandom) part of the utility function, conditioned upon the information $\mathcal{F}_{t+1}$ available just after time t, is denoted by $U_t(\theta) = U(\theta; \mathcal{F}_{t+1})$. Typically, we consider cases where $U_t(\theta)$ is based on a fitness measure of strategies θ, such as the last net ex post profits or squared prediction errors. More generally, one can introduce dependence on further past performances by introducing memory in the model. The evolution of the fitness measure for strategy θ can for example be modeled as:

$$U_t(\theta) = \alpha U_{t-1}(\theta) + (1-\alpha)\pi_t(\theta), \tag{2}$$

where $\pi_t(\theta)$ is the performance in period t, and $\alpha \in [0, 1)$ is a memory parameter. The utility function then becomes a geometrically weighted sum of ex post performances of strategy θ. An interpretation of the memory parameter α is the following; when analysts or traders consider new strategies it is common to perform back-testing, that is, to test the candidate strategies using historic data. The better performing strategies are then more likely to be selected for actual trading.

Given the observed price p_t, $U_t(\theta)$ can be updated, and the pdf of the beliefs distribution is determined by the continuous choice model:

$$\phi_t(\theta) = \frac{\varphi(\theta)\mathrm{e}^{\beta U_t(\theta)}}{Z_t}$$

the time dependent analogue of Eq. (1). By substituting Eq. (2) the effect of memory can be written as

$$\phi_t(\theta) \propto \varphi(\theta)\mathrm{e}^{\alpha\beta U_{t-1}(\theta)+(1-\alpha)\beta\pi_t(\theta)} \propto [\varphi(\theta)]^{1-\alpha}[\phi_{t-1}(\theta)]^{\alpha}\mathrm{e}^{(1-\alpha)\beta\pi_t(\theta)}, \tag{3}$$

which gives an update of the beliefs distribution in terms of the previous beliefs distribution and the last performance measure. Note that in general, since $\phi_t(\theta)$ is an infinite dimensional state variable, this may lead to an infinite dimensional dynamical system. However, in many cases the evolution of $\phi_t(\theta)$ can be completely described by a finite number of variables such as its first k moments, in which case it becomes finite dimensional.

Notice that the beliefs distribution $\phi_t(\theta)$ represents the population distribution, or marginal distribution of beliefs. One may think of this distribution as the sample distribution that would be obtained in the limit where the number of agents tends to infinity. The beliefs distribution itself is determined only by the performance measure and past observations, and otherwise evolves independently of the individual choices of agents. The statistical properties, however, of the feedback dynamics, do depend on the details of how, given the marginal beliefs distribution, the strategies actually used by the agents are drawn from this distribution. There are many beliefs selection mechanisms compatible with a given beliefs distribution, differing in dependence among strategies used among agents and over time. Dependence among traders could arise for example by the exchange of private information or coordination on publically announced predictions. For the moment we have in mind a situation where, given $\mathcal{F}_t$, the strategies $\theta_{i,t-1}$ are independent and distributed according to $\phi_{t-1}(\theta)$. Later we will relax this assumption.

In interpreting the price equations it is essential to realize that the $\theta_{i,t}$ are random variables, representing the individual strategy choices of agents. The dynamical system, at least for a finite number of agents is stochastic. In some of the later sections we will examine the dynamics obtained in limits where the number of traders tends to infinity. In doing so, one can appeal to the central limit theorem and the law of large numbers to obtain some asymptotic properties of the dynamics. As stated above, we only focus on examples in asset pricing contexts here, hoping to consider other contexts in future research.

3.2 Example: asset pricing with continuous AR(1) predictors

We consider two examples CBSes with AR(1) predictors. The role of the performance measure is investigated by focusing on two cases: squared prediction errors and squared logarithmic prediction errors, leading to normal and log-normal beliefs distributions respectively.

The equilibrium price in a dynamic present value asset pricing model is given by

$$(1+r)\,p_t = \frac{1}{n}\sum_{i=1}^{n} p^e_{t+1}(\theta_{i,t-1}) + \bar{y}. \tag{4}$$

where $\bar{y}$ denotes the (public) expected dividend payment y_{t+1} during the next period. The market clearing equation thus states that today's price is the sum of of the average expected future price and expected dividend, discounted by the risk free interest rate. Note that again each predictor has equal weight. For the effect of non-uniform weights we refer to section 4.2.

For illustrational purposes we will focus on a simple class of linear predictors, in which agents choose a strategy θ that represents their perceived future growth rate. The expected price at time $t+1$ associated with predictor θ is then

$$p^e_{t+1}(\theta) = \theta p_{t-1}.$$

The utility function associated with belief type θ is given by Eq. (2). The only ingredient missing from the CBS so far is the performance measure used to evaluate strategies. In the following subsections we examine two common performance measures, based on past squared prediction errors and squared logarithmic prediction errors. The price equation in both cases reads:

$$(1+r)p_t = \frac{1}{n}\sum_{i=1}^{n} \theta_{i,t-1} p_{t-1} + \bar{y} \tag{5}$$

where $\theta_{i,t-1}$ is the strategy chosen by agent i just after p_{t-1} became public. We postpone a discussion of the statistical properties of the equilibrium price to the next section, and first consider the updating of the beliefs distribution given realized prices, for two different performance measures.

3.2.1 Updating with squared prediction errors

First we consider the case where the performance measure is minus the squared prediction error:

$$\pi_t(\theta) = -\left(p_t^e(\theta) - p_t\right)^2 = -\left(\theta p_{t-2} - p_t\right)^2. \tag{6}$$

After the price p_t has been established, the distribution of beliefs is updated according to the continuous choice model. The new distribution characterizing the dispersion of belief types is then given by Eq. (3). Substituting the expression for $\pi_t(\theta)$ in Eqn. (6) into this equation, with a constant opportunity function, $\varphi(\theta) = 1$ (representing the case with agents without aversion against using extreme parameter values), gives:

$$\phi_t(\theta) \propto [\phi_{t-1}(\theta)]^\alpha \exp\left[-\beta(1-\alpha)\left(\theta p_{t-2} - p_t\right)^2\right]. \tag{7}$$

Since the exponent contains only up to second order forms in θ with a negative coefficient for the quadratic term in θ, the distribution of beliefs in each period can be described by a normal distribution. Here we implicitly have assumed that $\phi_{t-1}(\theta)$ is also normal, which can be justified by assuming that the dynamics has been running since the infinite past.

If we denote the mean and variance of $\phi_t(\theta)$ by μ_t and σ_t^2 respectively, we have

$$\phi_t(\theta) = \frac{1}{\sqrt{2\pi}\sigma_t} \exp\left[-\frac{1}{2\sigma_t^2}\left(\theta - \mu_t\right)^2\right]. \tag{8}$$

Eq. (7) gives

$$\phi_t(\theta) \propto \exp\left[-\frac{\alpha}{2\sigma_{t-1}^2}(\theta - \mu_{t-1})^2 - \beta(1-\alpha)(\theta p_{t-2} - p_t)^2\right]. \tag{9}$$

By comparing the coefficients of θ^2 and θ in the exponents in Eqs. (8) and (9), the mean μ_t and variance σ_t^2 can be seen to evolve according to

$$\begin{aligned}
\frac{\mu_t}{\sigma_t^2} &= \alpha\frac{\mu_{t-1}}{\sigma_{t-1}^2} + 2\beta(1-\alpha)p_t p_{t-2} \\
\frac{1}{\sigma_t^2} &= \frac{\alpha}{\sigma_{t-1}^2} + 2\beta(1-\alpha)p_{t-2}^2.
\end{aligned}$$

The remaining terms independent of θ are accounted for by the normalization factor Z_t. The mean μ_t and variance σ_t^2 are fully determined by past prices, and completely describe the evolution of beliefs. The mean μ_t can be interpreted as the average belief, while the variance σ_t^2 is a measure of heterogeneity. The full CBS also involves the price equation, Eq. (5), which is independent of the performance measure used by the agents.

3.2.2 Updating with squared logarithmic prediction errors

Next we examine the case where the performance measure is taken to be minus the squared logarithmic prediction error:

$$\pi_t(\theta) = -\left(\ln p_t^e(\theta) - \ln p_t\right)^2 = -\left(\ln\theta - \ln p_t + \ln p_{t-2}\right)^2. \tag{10}$$

An argument for using logarithmic prediction errors rather than just mean squared prediction errors is that this error measure is independent of the price level. This leads to dynamics which scales with the price level and is equivalent before and after possible stock splits.

Using a uniform opportunity function again, i.e. $\varphi(\theta) = 1$, substitution of the expression for $\pi_t(\theta)$ in Eqn. (10) into Eq. (3) gives:

$$\phi_t(\theta) = \frac{[\phi_{t-1}(\theta)]^\alpha \exp\left[-\beta(1-\alpha)\left(\ln\theta - \ln p_t + \ln p_{t-2}\right)^2\right]}{Z_t'} \tag{11}$$

In this case the exponent contains up to second order forms in $\ln\theta$ with a negative coefficient for the quadratic term in $\ln\theta$, which indicates that the distribution of beliefs in each period can be described by a log-normal distribution of the form

$$\phi_t(\theta) = \frac{1}{\sqrt{2\pi}\sigma_t\theta} \mathrm{e}^{-\frac{(\ln\theta - \mu_t)^2}{2\sigma_t^2}}. \tag{12}$$

The evolution of the parameters of the beliefs distribution follow again from Eq. (11), which results in

$$\phi_t(\theta) \propto \exp\left[-\alpha\ln\theta - \frac{\alpha}{2\sigma_{t-1}^2}(\ln\theta - \mu_{t-1})^2 - \beta(1-\alpha)(\ln\theta - \ln p_t + \ln p_{t-2})^2\right]. \tag{13}$$

A comparison of the coefficients of $\ln\theta$ and $(\ln\theta)^2$ in the exponents in Eqs. (12) and (13), gives the evolution rules for μ_t and σ_t^2:

$$\begin{aligned}
\frac{m_t}{\sigma_t^2} &= \alpha\frac{m_{t-1}}{\sigma_{t-1}^2} + 2\beta(1-\alpha)(\ln p_t - \ln p_{t-2}) \\
\frac{1}{\sigma_t^2} &= \frac{\alpha}{\sigma_{t-1}^2} + 2\beta(1-\alpha),
\end{aligned}$$

where $m_t = \mu_t - \sigma_t^2$. Note that σ_t^2 does not interact with the other variables, and simply tends to its steady state value $1/(2\beta)$ at an exponential rate.

The two examples discussed here are analytically tractable due to the quadratic terms in the exponent. A closed form analytic derivation of the dynamics might become cumbersome or even impossible if: (i) the beliefs distributions can not be represented within a finite parameter class, closed under updating according to Eqn 3 (it is e.g. if $\pi_t(\theta)$ is a finite order polynomial in θ), or (ii) agent's decisions, rather than by myopic mean variance optimization, are based on more complicated optimization procedures such as those typical in a dynamic programming contexts. In those cases simulation and estimation of the models might still be possible, with the appropriate generalization for the methods known in discrete choice simulation and estimation (see e.g. Keane and Wolpin, 1994). These generalizations are beyond the scope of this paper and left for future research.

4 Natural sources of randomness

Traditionally, randomness in asset pricing has always been associated with exogenous shocks, for example due to news that affects a company's future earnings, or fluctuations of the interest rate. The CBS framework provides several possible natural sources of randomness, through possible randomness of the market aggregate. As stated above, in a simple asset pricing model, the market price is simply determined by the aggregate expectation $\frac{1}{n}\sum_{i=1}^{n} p_{t+1}^{e}(\theta_{i,t-1})$, through the relation

$$(1+r)p_t = \frac{1}{n}\sum_{i=1}^{n} p_{t+1}^{e}(\theta_{i,t-1}) + \bar{y},$$

while in a cobweb context the price is the inverse demand function of the aggregate of supply functions:

$$p_t = D^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} S\left(p_t^e(\theta_{i,t-1})\right)\right).$$

In this section we consider the more general case where the observables are determined by a function, $H(\cdot)$, of an aggregate of a function $G(\cdot)$ of expectations, that is:

$$p_t = H\left(\frac{1}{n}\sum_{i=1}^{n} G\left(p_{t+k}^{e}(\theta_{i,t-1})\right)\right), \tag{14}$$

where k is the number of steps the agents predict ahead of t based on information up to and including time $t-1$. Thus, $k=1$ in a myopic asset pricing model, while $k=0$ in a cobweb context. For simplicity, $H(\cdot)$ will be assumed to be one-to-one, so that p_t given $\mathcal{F}_t$ is deterministic if and only if $\frac{1}{n}\sum_{i=1}^{n} G(p_{t+k}^{e}(\theta_{i,t-1}))$ is. For convenience, we introduce the following short-hand notation: $G_{i,t} = G(p_{t+k}^{e}(\theta_{i,t-1}))$ and $\bar{G}_t = \frac{1}{n}\sum_{i=1}^{n} G_{i,t}$. The conditional mathematical expectation and variance, given the information set $\mathcal{F}_t$, are denoted by $E_t[\cdot]$ and $\mathrm{Var}_t[\cdot]$ respectively.

Next we will elaborate on four natural sources of randomness. Firstly, a finite number of traders gives rise to stochasticity, because traders are assigned a belief $\theta_{i,t-1}$ at random from the beliefs distribution $\phi_{t-1}(\theta)$. The aggregate $\bar{G}_t$ for finite n then is stochastic unless all predictions are identical with probability one. Secondly, so far, we have implicitly assumed that no agents have a dominant market impact. This might not be the case, for example, if the wealth distribution among agents is fat-tailed. In that case the market impact of the wealthiest agents might not become negligible when the number of traders tends to infinity. The conditions for the law of large numbers are then not satisfied, providing yet another natural source of uncertainty. Thirdly, dependence among agents might prevent the dynamics from becoming deterministic when the number of agents tends to infinity. If, for example, agents coordinate (at least partly) on a random variable, such as an exogenous variable, or on predictors announced by some 'leading' agents who quote their predictors publically, then the stochastic properties of that variable show up in the price dynamics. In the extreme case where all agents have identical idiosyncrasies, they would all use the same (random) strategy. Finally, for certain combinations of the utility function and the predictor function $p^e_{t+k}(\theta_{i,t-1})$, the law of large numbers may not apply because the mean of the individual terms $G_{i,t}$ need not exist. In those situations, the limiting price dynamics might still be defined as a stochastic dynamical system, in the limit where the number of agents tends to infinity.

Before looking at the various sources of randomness, we start by examining some conditions under which the dynamics become deterministic. Consider the following assumption:

Assumption 1 *(Cross-sectional Independence) The strategies $\theta_{i,t}$ employed by agent i at time t, for each fixed time t are independent random variables, distributed according to the population distribution of beliefs at time t, $\phi_t(\theta)$.*

Assuming independence over agents seems reasonable, since it is always possible to consider expectations of groups of correlated agents as expectations of a single agent representative of this group. The effect of dependence then is merely a reduction in the effective number of agents. A stronger assumption is made if one additionally assumes temporal independence of the idiosyncratic noise terms of each agent over time. This stronger assumption is reasonable if the time interval corresponding to one time step in the model is large compared to the time scale on which idiosyncratic preferences of single agents change over time.

For the following theorem, which is concerned with the almost sure behavior of the model in the limit where the number of agents tends to infinity, the weaker assumption suffices. Again we use the notation $p^e_{t+k}(\theta) = f_\theta(p_{t-1}, p_{t-2}, \ldots)$ for the expected future price p_{t+k} given the information $\mathcal{F}_t$.

Theorem 1 *(Law of Large Numbers) Given $\mathcal{F}_t$, under Assumption 1, if n tends to infinity, the aggregate*

$$\bar{G}_t = \frac{1}{n}\sum_{i=1}^{n} G_{i,t} = \frac{1}{n}\sum_{i=1}^{n} G\left(p^e_{t+k}(\theta_{i,t-1})\right)$$

converges a.s. to

$$E_t[G_{i,t}] = E_t\left[G\left(p^e_{t+k}(\theta_{i,t-1})\right)\right],$$

if and only if $E_t\left[\left|G\left(p^e_{t+k}(\theta_{i,t-1})\right)\right|\right] < \infty$.

Proof: By Assumption 1, the strategies $\theta_{i,t-1}$ conditionally on $\mathcal{F}_t$ are IID random variables in Ω, which implies that the corresponding predictions $p^e_{t+1}(\theta_{i,t-1})$ of agents are also IID. The result is immediate from Kolmogorov's strong law of large numbers for the IID random variables $G\left(p^e_{t+k}(\theta_{i,t-1})\right)$, given $\mathcal{F}_t$ (see e.g. Resnick, 1998, p. 220). □

Note that $E_t[G_{i,t}] = E_t\left[G\left(p^e_{t+k}(\theta_{i,t-1})\right)\right]$ can be expressed as $\int_\Omega G\left(p^e_{t+k}(\vartheta)\right)\phi_{t-1}(\vartheta)\mathrm{d}\vartheta$. Theorem 1 states that, given the information public at time t, a necessary and sufficient condition for the aggregate expectation about p_{t+k} to converge a.s. to the mean expectation over the beliefs distribution, is the existence of this mean. This result has the following corollary concerning the dynamics.

Corollary 1 *(Functionally Deterministic Dynamics) Under Assumption 1 the observable* p_t *in the limit* $n \to \infty$ *almost surely tends to a deterministic functional of the information* $\mathcal{F}_t$ *available at time* t *if and only if* $E_t[G(p^e_{t+1}(\theta_{i,t-1}))] < \infty$.

We coin the term functional determinism here to make a distinction between the concept of determinism we describe here and the usual definition of determinism for time series processes. The term determinism is usually reserved for deterministic processes of finite order, the reason being that without this restriction any observed time series would allow a deterministic explanation. That is, functionally deterministic processes can still be non-deterministic from an empirical point of view, if there is dependence on the infinite past. For the moment, we note only that functional determinism is a necessary, but generally insufficient condition for the dynamics to be finite order deterministic.

Of course the dynamics in the limit where the number of agents tends to infinity need not be realistic. Particularly when the statistical properties of the dynamics are of interest it would be unnecessarily restrictive to confine attention only to the limiting dynamics. However, the limiting dynamics can be interpreted as a first order approximation to the price dynamics for a system with a large but finite number of agents. Later it will become clear that the random deviations around the deterministic limit in general can not be represented by a random variable with time-independent distributional properties.

Notice that discrete choice models can be considered as special cases of a CBS (when allowing for discrete opportunity distributions) in which agents can only choose among a finite number of alternative strategies θ_l, $l = 1, \ldots, m$. Provided that the expected future prices of these strategies are all finite, the average expected price is well-defined, so that the strong law of large numbers applies, and the dynamics converges to a deterministic dynamical system with probability one as the number of traders tends to infinity. A discrete choice setting with a continuum of agents, as considered by Brock and Hommes (1997), should thus always lead to functional determinism in the absence of exogenous noise provided that $G(p^e_{t+k}(\theta_l))$ is finite for each $l = 1, \ldots, m$.

4.1 Finite number of agents

In general, for a finite number of agents, the CBS leads to a stochastic dynamical system, the only exceptions being cases where all $G_{i,t} = G(p^e_{t+k}(\theta_{t,i-1}))$ are identical with probability one. In general it might not be straightforward to derive the distributional properties of the aggregate, unless the $G_{i,t}$, for each fixed t, are jointly normally distributed random variables. However, in case the mean and variance of the aggregate are finite, one can appeal to the central limit theorem which states that the distribution of the aggregate is asymptotically normally distributed. The following theorem and corollary are direct applications of the central limit theorem.

Theorem 2 *(Central Limit Theorem for the Aggregate) Given* $\mathcal{F}_t$*, under Assumption 1, if* n *tends to infinity, the random variable*

$$\sqrt{n}\left(\bar{G}_t - E_t[G_{i,t}]\right) = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} G(p^e_{t+k}(\theta_{i,t-1})) - E_t[G(p^e_{t+1}(\theta_{i,t-1}))]\right)$$

converges in distribution to

$$W \sim N(0, \sigma_t^2)$$

with

$$\sigma_t^2 = \mathrm{Var}_t[G(p^e_{t+k}(\theta_{i,t-1}))]$$

if and only if $E_t[|G(p^e_{t+1}(\theta_{i,t-1}))|] < \infty$ *and* $\mathrm{Var}_t[G(p^e_{t+1}(\theta_{i,t-1}))] < \infty$.

Corollary 2 *(Asymptotic Normality) If $H(s)$ is differentiable and has a derivative $H'(s)$ which is non-vanishing at $s = E_t[G_{i,t}]$, then $\sqrt{n}\left(H(\bar{G}_t) - H(E_t[G_{i,t}])\right)$ converges in distribution to*

$$W' \sim N(0, H'(E_t[G_{i,t}])\sigma_t^2).$$

Next we will illustrate these results for one of the asset pricing models discussed above.

Example: Logarithmic prediction errors

In this example we consider the asset pricing example discussed in section 3.2.2, with a performance measure based on logarithmic squared prediction errors. We assume that agents make decisions regarding their strategy independently of the other agents. In the asset pricing framework, the aggregation mechanism is given by Eq. (5). The predictors $p^e_{t+1}(\theta)$ are linear in the parameter θ, and the price equation is a special case of Eq. (14).

Infinite number of agents: deterministic dynamics Since the conditions of Theorem 1 are satisfied for the distribution of beliefs, if the number of agents tends to infinity, the stochastic random variable $\bar{G}_t = \bar{\theta}_{t-1}p_{t-1}$ tends to a deterministic limit almost surely. The average $\bar{\theta}_{t-1}$ of the $\theta_{i,t-1}$ converges in probability to

$$M_{t-1} = E_t[\theta_{i,t-1}] = \int \phi_{t-1}(\vartheta)\,\mathrm{d}\vartheta = \exp\left[\mu_{t-1} + \frac{1}{2}\sigma^2_{t-1}\right],$$

so that the dynamical system in the limit of an infinite number of agents becomes

$$\begin{aligned}
(1+r)p_t &= \exp\left[\mu_{t-1} + \frac{1}{2}\sigma^2_{t-1}\right]p_{t-1} + \bar{y} \\
\frac{m_t}{\sigma_t^2} &= \alpha\frac{m_{t-1}}{\sigma^2_{t-1}} + 2\beta(1-\alpha)(\ln p_t - \ln p_{t-2}) \\
\frac{1}{\sigma_t^2} &= \frac{\alpha}{\sigma^2_{t-1}} + 2\beta(1-\alpha),
\end{aligned}$$

again with $m_t = \mu_t - \sigma_t^2$.

This deterministic dynamical system, or *deterministic skeleton*, has a fixed point solution $p_t = p^*$, $\mu_t = \mu^*$ and $\sigma_t^2 = \sigma^{2*}$ with $p^* = \bar{y}/(1+r-\exp[1/(4\beta)])$ and $\mu^* = \sigma^{2*} = 1/(2\beta)$. Note that for all finite positive values of β the fixed point price p^* is larger than the fundamental value $\bar{y}/r$. Only when β tends to infinity (i.e. the propensity to err tends to zero) the fixed point price tends to the fundamental price.

Finite number of agents: stochastic dynamics Similarly, the central limit theorem provides a normal approximation for a finite but large numbers of agents. The conditional variance of $\theta_{i,t-1}$, given $\mathcal{F}_t$ is

$$S_t^2 = \mathrm{Var}_t[\theta_{i,t-1}] = \mathrm{e}^{2\mu_{t-1}+\sigma_{t-1}^2}(\mathrm{e}^{\sigma_{t-1}^2} - 1),$$

so that the price equation for n agents becomes

$$(1+r)p_t \simeq \exp\left[\mu_{t-1} + \frac{1}{2}\sigma_{t-1}^2\right] p_{t-1} + \bar{y} + \frac{S_t}{\sqrt{n}}\epsilon_t,$$

with $\epsilon_t \sim N(0,1)$. Notice that one might consider the "skeleton" value of the price, i.e. the discounted value of the first two terms on the right hand side, as a first order approximation to the price obtained for a large but finite number of agents. Notice that an analogous asymptotic normal approximation will be exact for the squared prediction error case, even for a finite number of agents n, since there the $\theta_{i,t}$ are normally distributed (see section 3.2.1).

Volume-volatility relationship A relation between price changes and trading volume can be derived straightforwardly in the case of normally distributed predictors. The agent's demand functions underlying the linear price equation are linear in deviations from expected revenues, that is, the demand function of agent i, prior to time t in terms of the unknown price p_t is

$$D_{i,t} = b(p_{t+1}^e(\theta_{i,t-1}) + \bar{y} - (1+r)p_t),$$

where b is a positive constant related to the absolute risk aversion and perceived risk (which were assumed constant across agents). Once p_t will be quoted, the trading volume required to match agent i's net demand will be equal to

$$V_{i,t} = |D_{i,t} - D_{i,t-1}|.$$

Notice that summing up these individual volumes over the agents would lead to a "double" volume as each transaction is counted twice, as well from the buyer's side as the seller's side. Using the predictor function, which gives $p_{t+1}^e(\theta_{i,t-1}) = \theta_{i,t-1}p_{t-1}$, and the price equation $(1+r)p_t = \bar{\theta}_{t-1}p_{t-1} + \bar{y}$, we can express the demand alternatively as

$$D_{i,t} = b(\theta_{i,t-1} - \bar{\theta}_{t-1})p_{t-1},$$

which clearly shows that individual demands depend only on the previous price through the agent's deviation from consensus. The expected trading volume to meet agent i's demand is

$$E_t[V_{i,t}] = bE_t[|(\theta_{i,t-1} - \bar{\theta}_{t-1})p_{t-1} - (\theta_{i,t-2} - \bar{\theta}_{t-2})p_{t-2}|].$$

where, under the assumption of temporally independent idiosyncrasies, each of the terms have mean zero. The expression for the expected trading volume per agent is obtained under the condition that all agents update their portfolio every time step. That is, all traders are assumed to be active in the market at all times. In practice, the volume will be determined only by the fraction of agents that are actually trading at any moment in time.

Following Brock and LeBaron (1996), the total trading volume can now be calculated as an average of the absolute value of a normally distributed random variable. The expected trading volume per agent can be expressed as

$$E_t[V_{i,t}] = E_t[|Z_t|],$$

where Z_t is a random variable with conditional mean $E_t[Z_t] = 0$ and variance

$$\text{Var}_t[Z_t] = \frac{n-1}{n} b^2 \left(S_t^2 p_{t-1}^2 + S_{t-1}^2 p_{t-2}^2 \right).$$

The factor $(n-1)/n$ takes into account the covariance of $\theta_{i,t}$ with $\bar{\theta}_t$, that is, the fact that $\text{Var}_t[\theta_{i,t} - \bar{\theta}_t] = \frac{n-1}{n} \text{Var}_t[\theta_{i,t}]$. In the case of a utility function based on squared prediction errors, the beliefs $\theta_{i,t}$ are distributed normally around their mean M_t, and we have

$$E_t[V_{i,t}] \propto \left(S_t^2 p_{t-1}^2 + S_{t-1}^2 p_{t-2}^2 \right)^{\frac{1}{2}}.$$

The conditional variance of p_t is equal to $S_t^2 p_{t-1}^2 / n$. For n sufficiently small the stochastic part of the dynamics dominates the deterministic part and leads to positive correlations between volume and volatility. A small effective number of agents can arise when there is dependence between the choices of agents, as we will describe in section 4.3. In this case similar strong dependence between volume and volatility is to be expected as for a small number of agents.

4.2 Market impact

In this subsection we briefly examine the effect of market impact. Although interesting, we consider the study of endogenous wealth effects with continuous choice dynamics beyond the scope of this paper. Endogenous wealth effects have for example been studied by Cabrales and Hoshi (1996) in a discrete choice framework with two types of agents.

If the market weight of agent i is denoted by w_i, such that $\sum_i w_i = 1$, the market aggregate becomes

$$\bar{G}_t = \sum_{i=1}^{n} w_i G_{i,t}$$

which has a conditional mean equal to

$$E_t \left[\sum_{i=1}^{n} w_i G_{i,t} \right] = E_t[G_{i,t}]$$

and, assuming conditional independence of $\theta_{i,t-1}$ given $\mathcal{F}_t$, conditional variance

$$\text{Var}_t \left[\sum_{i=1}^{n} w_i G_{i,t} \right] = \left(\sum_{i=1}^{n} w_i^2 \right) \text{Var}_t[G_{i,t}].$$

The term between brackets on the right hand side of this equation, is also known as the Herfindahl index of concentration. A large concentration thus implies a small effective number of market participants and vice versa. Later, in the empirical section, we will estimate the number of market participants from empirical time series. This is done by introducing an effective number of market parties through the variance equation

$$n_{\text{eff}} = \left(\sum_{i=1}^{n} w_i^2 \right)^{-1}. \tag{15}$$

4.3 Dependence among agents

The next possible source of randomness we consider here is dependence among agents. For our purposes it is useful to think of a continuum of agents, from which we select n agents $i = 1, \ldots, n$, at random. The following lemma states that the covariance between predictors of a pair of agents selected in this way, is non-negative.

Lemma 1 *The covariance $Cov_t[G_{i,t}, G_{j,t}]$ is independent of i and j and nonnegative.*

Proof: The independence of the covariance of i and j follows directly from the permutation symmetry of the mechanism by which agents are selected. Now, concerning the covariance, suppose for the moment that is would be possible to have $\text{Var}_t(G_{i,t}) = v_t > 0$ and $\text{Cov}_t[G_{i,t}, G_{j,t}] = \lambda_t < 0$, and consider n randomly selected agents from the continuum. The variance of $\sum_{i=1}^{n} G_{i,t}$ would then be $nv_t + n(n-1)\lambda_t$. But if $\lambda_t < 0$ this would become negative for n sufficiently large, so that the assumption that $\lambda_t < 0$ cannot be correct. □

For the remaining part of this section we use the conventions introduced in the proof of the lemma, i.e. $v_t = \text{Var}_t[G_{i,t}] > 0$ and $\lambda_t = \text{Cov}_t[G_{i,t}, G_{j,t}] \geq 0$. We consider a CBS with n randomly selected agents and study the limit in which n tends to infinity. First we focus on the case where market impacts are equal ($w_{i,t} = \frac{1}{n}$). The variance of the market expectation then is given by

$$\text{Var}\left(\frac{1}{n} \sum_{i=1}^{n} G_{i,t} \right) = \frac{v_t}{n} + \frac{n(n-1)}{n^2} \lambda_t.$$

The first term, proportional to v_t, tends to zero as $n \to \infty$. However, the second term converges to λ_t, so that whenever $\lambda_t > 0$ stochasticity will be persistent in the limit of an infinite number of agents. This result indicates that the remark made by Muth (1961) that "allowing for cross-sectional differences in expectations is a simple matter because their aggregate effect is negligible as long as the forecasts are not strongly correlated" only applies if correlations between the expectations of a random pair of agents is negligible, that is, if $\lambda_t \ll \frac{v_t}{n-1}$, or $\rho_t \ll \frac{1}{n-1}$, where $\rho_t = \lambda_t / v_t$ is the corerlation between agents' predictors.

More generally, for any given number n of market participants n we can consider the weights, $w_{i,n}$, for $i = 1, \ldots n$. We obtain

$$\text{Var}\left(\sum_{i=1}^{n} w_i G_{i,t} \right) = v_t \sum_{i=1}^{n} w_i^2 + \lambda_t \sum_{i} \sum_{j \neq i} w_i w_j.$$

Both terms in this expression are non-negative, and we will show that they cannot both tend to zero in the limit. If n tends to infinity, the first term vanishes as before, provided that the concentration index $\sum_i w_i^2$ tends to zero. For the second term, we have $\sum_i \sum_{j \neq i} w_i w_j = 1 - \sum w_i^2$. Since $0 < \sum w_i^2 \leq 1$, the second term only vanishes if the concentration index tends to 1. This implies that stochastic dynamics are unavoidable in the limit, regardless of how the market weights are chosen while taking the limit.

Finally, in the presence of dependence, the effective number of agents can be obtained by relating the variance of the aggregate expectation to v_t, the conditional variance of a single expectation $G_{i,t}$. This gives the following generalization of Eq. (15):

$$n_{\text{eff},t} = \left(\sum_{i=1}^{n} w_i^2 + \rho_t \sum_i \sum_{j \neq i} w_i w_j \right)^{-1},$$

where $\rho_t = \lambda_t / v_t$ is the correlation between agent's choices. Clearly, larger values of the correlation coefficient ρ_t lead to a smaller effective number of market parties.

Dependence among agents can be also incorporated directly in the utility function by adding an interaction term, as shown in Brock and Durlauf (2001). In this way it is possible to incorporate several scenarios in the CBS. For example, one might consider a situation where agents coordinate partly on a source of information, being, for example a public exogenous variable. Such a source of information might, but need not necessarily, be related to economic fundamentals. In case the 'signal' on which agents coordinate is an independent exogenous noise source, this can be expected to increase correlation of the agent's predictors, and hence lead to a smaller effective number of agents. Alternatively, interaction terms might be used to model agents coordinating on each other. This type of coordination might also affect the size of endogenous fluctuations. For example, if there is one agent who announces his strategy publically at an early stage, after which others coordinate on this (noisy) signal, the endogenous fluctuations will increase. Interaction scenarios in which the the variance of aggegates decreases appear to be less natural. At first sight, herding to the mean might seem a possible scenario which might reduce the variance. However, at least when the beliefs are linear in the beliefs parameter and the aggregate is also linear, herding to the mean without additional information will only amount to herding towards the sample mean. This might reduce the dispersion of individual beliefs around the sample mean, but will not help reducing the fluctuations of the sample mean itself (which eventually determines the observable) around the average belief μ_t.

4.4 Inherent Randomness

As a final source of randomness we consider an example of the dynamics in a case where $p^e_{t+1}(\theta_{i,t-1})$ does not have a finite conditional mean. Due to the failure of this moment condition, the aggregate expectation $\frac{1}{n}\sum_{i=1}^{n} p^e_{t+k}(\theta_{i,t-1})$ does not tend to a constant, but tends in distribution to a well-defined random variable when n tends to infinity. The starting point is an asset pricing model with agents that choose among constant predictors:

$p^e_{t+1}(\theta) = \theta$. For simplicity we put $\alpha = 0$ (no memory). For the performance measure we take

$$\pi_t(\theta) = -\log(1 + (\theta - p_t)^2)$$

which, like the squared prediction error, is maximal for predictions θ that are equal to realized prices. For the beliefs distribution this gives

$$\phi_t(\theta) = \frac{\Gamma(\beta)}{\Gamma(\frac{1}{2})\Gamma(\beta - \frac{1}{2})}(1 + (\theta - p_t)^2)^{-\beta},$$

from which it follows that

$$\sqrt{2\beta - 1}\,(\theta_{i,t-1} - p_{t-1}) \sim t(2\beta - 1),$$

a t-distribution with $2\beta - 1$ degrees of freedom. Thus, the $\theta_{i,t-1}$ are distributed symmetrically around p_{t-1}, and for $\beta < \frac{3}{2}$ the mean does not exist.

For $\beta = 1$, $\theta_{i,t} - p_t$ given $\mathcal{F}_t$ is Cauchy(0,1) distributed. Since this distribution is closed under averaging, this gives $\bar{p}^e_{n,t+1} = \frac{1}{n}\sum_{i=1}^n \theta_{i,t-1} - p_{t-1} \sim$ Cauchy(0,1). The price equation becomes

$$(1 + r)p_t = p_{t-1} + \bar{y} + \eta_t,$$

where η_t is a Cauchy(0,1) distributed random variable, the pdf of which is $f_{\eta_t}(x) = (\pi(1 + x^2))^{-1}$. The result is a first order autoregressive process with a fat-tailed noise term. The price dynamics is stochastic, and the distribution of the noise term is independent of the number of agents.

The non-vanishing randomness due to a moment condition failure in this example can only occur with a continuum of strategies. If the number of strategies is finite, provided that each of them predicts a finite price, the mean expectation always exists, in which case, by the law of large numbers, the dynamics become deterministic when the number of traders tends to infinity.

Before we move to the empirical section, notice that, so far, we have not allowed for the individual tastes of the agents to change on much longer time scales than that of the price formation process. This might be modelled by a small probability that agents reconsider their strategies per time step, giving rise to temporal dependence of the strategies employed by them. Brock and LeBaron (1996), in a discrete choice setting with dependent agents, argue that such a mechanism tends to increase the persistence in heterogeneity. In the asset pricing model described in our examples this can be expected to lead to an increase in the persistence of volatility.

5 Empirical example

In the previous sections we have developed a framework which is mathematically tractable and provides insights into the feedback between observables and the distribution of beliefs. In this section we present, as an illustration, a highly stylized model of asset returns, to discuss potential empirical applications of a CBS.

The discussion in the previous sections suggest at least three possible types of empirical applications. First of all, after deriving the equations for the aggregate observables, one might use maximum likelihood methods to estimate the macroscopic model from the data. The model can then be used for several standard purposes, such as prediction and/or risk evaluation. A second possibility is to use the estimated model parameters for classification and/or ranking based on the estimated model parameters. The fact that the model parameters can be interpreted directly in terms of economic quantities can be very convenient there. A third possibility is to use the model to estimate the beliefs distribution, and to visualize its evolution over time. Based on the extremely simple model presented here, we were able to address the first two issues in full. The third, being visualization of beliefs over time, turns out to be difficult for the simple model considered here, but not impossible in principle.

Once more, we state that this empirical section is only meant to serve as an illustration, and to provide a background for our discussion on potential empirical applications of a CBS. By no means do we intend to claim that the model discussed here is sufficiently sophisticated to capture asset price dynamics in full detail.

For convenience we consider a model based on returns rather than prices. Again, the framework is one in which agents can invest in an asset with uncertain payoff. Agents are assumed to expect that today's return $r_t = \ln(p_t/p_{t-1})$ can be expressed in terms of yesterday's return r_{t-1} through

$$r_t^e(\theta) = \theta r_{t-1}. \tag{16}$$

A performance measure based on squared prediction errors is used,

$$\pi_t(\theta) = -(r_t^e(\theta) - r_t)^2,$$

while the utility function again involves memory as in Eq. (2). The realized return is assumed to be the market expectation of returns:

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

Empirical returns often exhibit small but significant positive autocorrelation at small lags, which can be captured by taking an opportunity function which is a pdf with nonzero mean and finite variance. This would then reflect a certain aversion of agents against using unrealistically large negative or positive parameter values. Upon taking the opportunity function to be equal to the pdf of a $N(\mu_{\text{op}}, \sigma_{\text{op}}^2)$ random variable, the following dynamical system is obtained:

$$r_t = \mu_{t-1} r_{t-1} + h_t^{\frac{1}{2}} \epsilon_t, \tag{17}$$

where μ_t is updated according to

$$\begin{aligned} \frac{\mu_t}{\sigma_t^2} &= \alpha \frac{\mu_{t-1}}{\sigma_{t-1}^2} + (1-\alpha)\left(r_t r_{t-1} + \frac{\mu_{\text{op}}}{\sigma_{\text{op}}^2} \right), \\ \frac{1}{\sigma_t^2} &= \alpha \frac{1}{\sigma_{t-1}^2} + (1-\alpha)\left(r_{t-1}^2 + \frac{1}{\sigma_{\text{op}}^2} \right). \end{aligned} \tag{18}$$

and $\{\epsilon_t\}$ represent an IID standard normal process, while the conditional variance h_t is given by

$$h_t = \sigma_u^2 + \frac{\sigma_{t-1}^2}{n} r_{t-1}^2.$$

Here σ_u^2 represents the variance of exogenous noise, which we assume independent of the endogenous noise, and IID normally distributed. The model can thus be interpreted as a time varying ARCH type model with a time-varying AR coefficient

$$\mu_{t-1} = \frac{2\beta(1-\alpha)\sum_{i=0}^{\infty} \alpha^i r_{t-i-1} r_{t-i-2} + \frac{\mu_{\text{op}}}{\sigma_{\text{op}}^2}}{2\beta(1-\alpha)\sum_{i=0}^{\infty} \alpha^i r_{t-i-1}^2 + \frac{1}{\sigma_{\text{op}}^2}} \tag{19}$$

and a time-varying ARCH parameter σ_{t-1}^2/n, in which

$$\sigma_{t-1}^2 = \left(2\beta(1-\alpha)\sum_{i=0}^{\infty} \alpha^i r_{t-1-i}^2 + \frac{1}{\sigma_{\text{op}}^2}\right)^{-1}. \tag{20}$$

The parameters of the model given in Eqs (17) and (18) are the memory parameter α, the intensity of choice β, the effective number of traders, n, and the parameters describing the opportunity distribution: μ_{op} and σ_{op}^2. However, the following lemma states that the model as such is over-parameterized:

Observation 1 *The parameters β and n enter the dynamics only through their product βn.*

For a derivation, see appendix A.1. □

As a result, without loss of generality, we may fix β at any positive value and use n only. For convenience, we choose to take $\beta = \frac{1}{2}$. Under the assumption that β is constant across markets, n then provides a proxy for the number of market participants. If this assumption is considered to be too strong, one might alternatively interpret βn as a single indicator of market efficiency, which takes into account the (effective) number of market parties as well as their intensity of choice.

If the memory parameter α is smaller than 1, it is possible, in principle, to use maximum likelihood methods to estimate the full model from empirical returns time series. However, in most applications we found that the memory parameter is often close to 1 for empirical data when daily data are used. This indicates a potential practical difficulty in recovering μ_t and σ_t^2, since these depend on a long history of past returns, including those before the period of observation. For α close to 1 large data sets would be required for estimation based on daily returns data. Indeed, a memory half-time as short as one year (approximately 250 trading days) already implies a value of α of $0.5^{1/250} \simeq 0.9972$, while downsampling to obtain reasonably small values of α seriously reduces the number of data available (monthly data still imply a value of α of $0.5^{1/12} \simeq 0.944$).

For model estimation purposes (as opposed to estimation of the beliefs distribution, which will be discussed below) we have first examined the model theoretically in the limit

$\alpha \to 1$. It turns out that the time-varying ARCH model in this limit becomes a standard ARCH model. Obviously, the time variations in the beliefs distribution can no longer be estimated in the ARCH limit (in this limit the beliefs distribution is constant). However, the the model parameters still contain meaningful information regarding the underlying micro model. For example, the effective number of traders can be estimated from the ARCH parameter, as we will show next.

In the limit $\alpha \to 1$ the CBS reduces to an ARCH model with an AR(1) coefficient (see Engle, 1982). This will be shown here assuming stationarity and ergodicity of the model for any given $\alpha < 1$. The stationarity of $\{r_t\}$ requires some restrictions on the model parameters, such as those imposed in the following lemma.

Lemma 2 *The variables* μ_t *and* $\frac{1}{\sigma_t^2}$*, in the limit* $\alpha \to 1$*, for all* $-1 < \mu_{\text{op}} < 1$ *and* $0 < \sigma_{\text{op}}^2 < \infty$*, converge to*

$$\begin{aligned} \mu_t &\to \mu = \frac{E[r_t r_{t-1}]}{E[r_t^2]} = \mu_{\text{op}} \\ \frac{1}{\sigma_t^2} &\to \frac{1}{\sigma^2} = E[r_t^2] + \frac{1}{\sigma_{\text{op}}^2}. \end{aligned}$$

Proof: See appendix A.2. □

Clearly, the model is non-stationary for $|\mu_{\text{op}}| > 1$, while for $-1 < \mu_{\text{op}} < 1$ it is stable. This is not surprising, as the agents, on average, then have a tendency to extrapolate trends away from equilibria.

Notice that the model can be considered a random coefficient AR model,

$$r_t = \left(\mu_{t-1} + \frac{\sigma}{\sqrt{n}}\epsilon_t\right) r_{t-1} + \sigma_u \epsilon_t',$$

where $\{\epsilon_t\}$ and $\{\epsilon_t'\}$ are independent IID standard normal processes. The term between brackets exactly corresponds with the average $\bar{\theta}_{t-1} = \frac{1}{n}\sum_{i=1}^{n}\theta_{i,t-1}$ of the predictors over the agents. It is known that random coefficient AR models include ARCH type models as special cases (see e.g. Tsay, 1987; Bera *et al.*, 1992). The dynamics in the limit $\alpha \to 1$ can indeed be shown to converge to an ARCH model:

$$\begin{aligned} r_t &= \gamma r_{t-1} + h_t^{\frac{1}{2}}\epsilon_t \\ h_t &= c + a r_{t-1}^2. \end{aligned} \tag{21}$$

This model can be estimated straightforwardly by standard maximum likelihood methods.

The relations between the agent model and the ARCH model are as follows. The AR coefficient γ is equal to the mean μ_{op} of the opportunity distribution, c can be interpreted as the variance of the exogenous noise, σ_u^2, and the ARCH coefficient a satisfies

$$a = \frac{1}{n(v + \frac{1}{\sigma_{\text{op}}^2})},$$

where n is the (effective) number of agents, v the unconditional variance of r_t and $\sigma^2_{\rm op}$ the variance of the opportunity function, which we had taken to be the pdf of an $N(\mu_{\rm op}, \sigma^2_{\rm op})$ distribution.The parameter of our interest here is the (effective) number n of market participants. To reduce the number of parameters, we assume that agents have little aversion against using extreme strategies if the market provides evidence in favor of their use. It is assumed that $\sigma^2_{\rm op}$ is large enough to ensure that $1/\sigma^2_{\rm op}$ is negligible with respect to v, i.e. $1/\sigma^2_{\rm op} \ll v$.[1] We might alternatively have defined a ψ through $1/\sigma^2_{\rm op} = \psi v$ to replace $\sigma^2_{\rm op}$. In that case, fixing ψ to any contant would lead to similar results as those presented here (ψ fixed at zero) in that the relative ordering of the estimates of n are independent of the value at which ψ is held constant.

Let us consider some applications where the effective number of market parties is estimated. If we denote the estimated ARCH coefficient from the model given in Eq. (21) by $\widehat{a}$, the effective number of agents $n_{\rm eff}$ can be estimated through

$$n_{\rm eff} = \frac{1}{\widehat{a}\widehat{v}},$$

where $\widehat{v}$ is the sample variance of r_t. Figure 1 shows the estimated effective number of agents in a moving time window of 1000 trading days, for daily returns of 5 large stock indices, the SP500, FTSE, AEX, Hang-Seng and Nikkei indices in the period between January 1, 1988 and October 24, 2002 (3864 trading days). At first sight the plot appears to provide evidence that the effective number of market parties across markets tend to move toward eachother over time. However, the results displayed here should be interpreted with care for several reasons. The first reason is that there is a quite large overlap in data points between consecutive points plotted. The time window required for estimation of the ARCH coefficient with a reasonably small standard error turned out ot be as high as 1000. However, the lag between the points in the plot only amounts to 400 trading days. Therefore, the estimates can be expected to be fairly dependent. This will smooth the figure and might introduce apparant trends, even when in fact none are present. Some results which should suffer from dependence to a lesser extent are shown in Table 1. The estimates in this table correspond to two non-overlapping parts of the same data as used for Figure 1. The table clearly shows that apart from the Nikkei index, the effective number of agents for the other four markets has reduced from the first period to the second. The relative ordering of the these four markets is in line with what one would expect: the SP500 and FTSE are large in terms of the effective number of market participants, followed by the AEX and the Hang Seng index. The recent increase in the effective number of traders for the Nikkei might be related to a structural change in the market. From the results presented so far, it becomes clear that the typical time scale on which the effective number of agents changes can be very long. The following example demonstrates this.

The time evolution of the effective number of agents for daily SP500 returns over a total of 113 years, is shown in Figure 2. Here the window is moved by 500 trading days before

[1] In order to do so, one should verify that this assumption is consistent with the model. See the appendix for details. There we also discuss the advantage of our approach to starting with a constant opportunity function.

a new estimate is calculated. During this long period we estimated an effective number of agents as low as 3000 in the thirties and as high as 3 million in the late ninetees. It becomes clear from this figure that the effective number of agents changes on a time scale of decades rather than years or months. Therefore, the time series might look non-stationary when only a small portion (e.g. 2 decades) is considered. Taking this into account the behavior observed in a period as 'short' as that of Figure 1 could by no means be representative for a much longer historic sequence of estimates.

Finally we briefly wish to discuss the estimation of the beliefs parameters over time in a continuous beliefs system. (For comparison, in the dynamic programming literature methods for estimating discrete choice models have been described in a dynamic setting. See e.g. Hotz and Miller, 1993; Keane and Wolpin, 1994). As noted above, the hidden variables are difficult to identify for α close to 1, since their values depend on a long price history as described by Eqs. (20) and (19). For α close to 1 the conditional AR and ARCH parameters both tend to constants. The small deviations from these limiting values, that remain when α is close to, but not equal to, one, are difficult to identify from empirical data. Possibly these identification problems might be overcome partly by Kalman filtering techniques (see e.g. Harvey, 1990). However, one of the difficulties that may arise is that h_t is a nonlinear function of the hidden variables, so that one would have to use a nonlinear Kalman filter, the theory of which is not nearly developed as well as that of linear Kalman filters. For te moment, we leave the study of optimal estimation methods for the hidden variables for future research. Instead we focus on some qualitative properties of the response of the belief parameters to observables, that are relatively insensitive to the exact model parameters. For example, when α is close to one, both μ_t/σ_t^2 and $1/\sigma_t^2$ are close to a random walk driven by $r_t r_{t-1}$ and r_{t-1}^2 respectively. Therefore, the level of σ_t^2 can be expected to decrease (increase) in periods of low (high) volatility. The mean μ_t can be expected to increase (decrease) in periods where returns are positively (negatively) correlated. Figure 3 shows the 'simulated response' of the belief parameters over time to empirical returns of the SP500 index, for a period of 1000 consecutive trading days. In these simulations α is fixed at 0.95 and $\frac{1}{\sigma_{\text{op}}^2}$ is taken to be equal to v, the unconditional variance of the returns. While the exact level and smoothness of the simulated responses will depend on the parameter values, the qualitative properties mentioned above are clearly visible.

6 Concluding Remarks

Many will agree that most observed economic aggregates, such as prices, are to a large extent determined by the expectations of those who are active in the market. In turn, these individual expectations are typically shaped by the observed past, as agents will try to incorporate the underlying economic laws at force in their expectations. New observations will provide agents with an incentive to update their beliefs whenever their prior appears inconsistent with what is observed. Therefore, one would expect that individual beliefs will coevolve over time with realized prices. Being a hidden variable, it is the distribution of individual beliefs that has been of our key interest in this paper. More precisely, it is the feedback between observed prices and the unobserved beliefs distribution that is at the heart of our concept – the continuous beliefs system (CBS).

We have defined the distribution of beliefs as a probability density function on a space of possible strategies, which we denoted the beliefs space. The continuous choice model is employed to update the beliefs distribution, while the incentives for the agents to switch are provided by the past performances of strategies. Coevolution of unobserved beliefs with observed prices thus emerges due to the ongoing evaluation of predictors.

This approach to modeling the evolution of the beliefs distribution over time provides several insights into the nature of the feedback between economic observables and the dispersion of beliefs. The results and implications of our concept are at least three-fold. The first two reflect, we believe, a certain theoretical elegance, while the third potentially might have important empirical consequences.

Firstly, while being applicable in a wide range of theoretical models where expectations feedback plays a key role, the concept often allows one to model the coevolution of aggregate economic observables (such as prices) and the distribution of beliefs explicitly. The beliefs distribution becomes a state variable of the dynamical system. Moreover, the distribution is endogenously shaped by the assumptions about the functional form of predictors, and the type of performance measure (squared prediction error or profits etc.) employed. In this paper, we have provided some stylized examples that lead to a normal- and a log-normal distribution for individual beliefs. In these cases, the distribution of beliefs can thus completely be described by the average belief and the dispersion of beliefs, the latter of which can be related to the degree of heterogeneity. The CBS then prescribes explicitly how the dispersion of beliefs affects prices, and in turn how prices affect the dispersion of beliefs.

Secondly, our concept provides what we refer to as 'natural sources of randomness'. In many economic models, noise, which is required when matching stylized facts is the objective, is often included as additive exogenous shocks. Popular justifications include: exogenous news shocks, model approximation error, and noise induced by trading. Our natural source of randomness introduces a type of endogenous uncertainty that can not be associated with exogenous shocks but rather with the dispersion of beliefs. Moreover, the endogenous noise term was shown to inherit its statistical properties directly from the beliefs distribution. To understand where the stochasticity comes from, note that in a CBS the beliefs distribution describes the likelihood, through the eyes of the econometrician, according to which individual agents select their beliefs. When the number of agents is finite, the aggregate belief by definition is a random variable, marking our first natural source of randomness. From simulation studies (Lux and Marchesi, 1999) it is known that the stochastic nature of the dynamics in simulated markets vanishes when the number of agents becomes large. Our results indicate that there might be several machanisms preventing endogenous noise from becoming insignificant in a large market. This occurs if, for example: (i) wealth is disproportionately distributed (ii) agents' choices exhibit dependence(herding); and (iii) when the combination of the performance measure and the functional form of beliefs induces nonexiststence of the average perceived price. We refer to the latter as inherently random dynamics.

Thirdly, we believe that the two theoretic implications just described might open the door to a number of potential empirical applications. To start with probably the most ambitious, it would be desirable to have an implementation of the concept that allows one to visualize the evolution of beliefs directly from empirical data. Note that if the latter is feasible, one can derive from that (an estimate for) a measure of the degree of heterogeneity and its time evolution. As a second empirical application we coin estimation of the number of effective market participants.

In section 5 we have provided a stylized example that can at best be considered as a first step in the empirical directions just described. All agents were assumed to employ AR(1) beliefs, and consider past squared prediction errors as a measure for performance. Dispersion of beliefs emerges whenever agents disagree on the AR(1) parameter. When agents have finite memory, the model gives rise to a time-varying AR(1)-ARCH(1) process for the returns, where the evolving ARCH parameter is proportional to the dispersion of beliefs. The number of effective market participants enters as a scale parameter. The traditional ARCH model (with constant parameters) is obtained in the limit where memory tends to infinity. In that case, the beliefs distribution is static i.e. constant over time.

In our stylized empirical example, we have only been able to address the latter issue – estimating the number of effective agents. Since the beliefs distribution is hidden, one would generally need to consider Kalman filters if visualization by means of estimation is the objective. In a CBS, these Kalman filters easily become of the nonlinear type, so that estimation of the hidden variables is not straightforward. For this reason, we have confined ourselves here to the static case where the beliefs distribution is constant over time. Doing so, the approach partly overlaps with that of a study by Lewbel (1994) in which the key objective was to estimate a static distribution of AR(1) parameters from empirical data. However, our example allows us additionally to relate the ARCH parameters to what we define as the number of effective heterogenous agents in the market. Also the latter might be considered a measure for the degree of heterogeneity. In general, if a CBS leads to a feasible and perhaps well-known econometric model, one has as an immediate result an interpretation of the parameters in terms of economically relevant or behavioral quantities.

Our data set contains the log returns of the indices of six leading stock markets, namely the S&P 500, FTSE 100, Hang Seng, NIKKEI, DAX, and the AEX. Estimation of the number of agents in different periods of time indicated that the largest stock market in terms of the number of active agents is that of the S&P 500, followed by the stock market in London, the FTSE 100. As these results are in line with what one would expect, we feel that our first step might indeed be a step in the right direction. Regarding the second issue – visualizing the evolution of beliefs – we have yet only 'scratched' the surface of what we believe is potentially 'large'. Our first step has indicated that visualizing hidden beliefs from empirical data, although difficult, is not impossible in principle. We hope that future smarter empirical applications of CBS will prove us right.

One can think of many directions for further research. To continue with empirics, it might be worthwhile to consider nested models that allows one to test hypothesizes w.r.t. the beliefs of agents. Questions that might be addressed include: (i) Is the degree of heterogeneity constant over time, and if not, does it increase/decrease in periods of high volatility?; (ii) Are the beliefs of agents best described by linear or non-linear rules?; and (iii) Is it more realistic to agents as updating their beliefs based upon past profits or prediction errors?

From a theoretical point of view, we believe that a number of extensions and generalizations might also prove to provide new insights. For example building a bridge between CBS and that of Rational Beliefs (RB), the latter being a concept recently developed by Kurz (2001) in which another notion of a beliefs distribution is introduced. In the work of Kurz, a beliefs distribution describes the individual belief of an agent, and reflects the uncertainty each individual holds regarding future economic variables. This concept is designed to generalize Rational Expectations (RE) theory where all agents have exact knowledge about the future. Although Kurz's approach is different from ours, in particular since a beliefs distribution represent a different object, Kurz's definition of 'Endogenous Uncertainty' as 'that component of the volatility of quantities and prices in the economy which is generated by the distributions of beliefs' has similar implications. Also he is able to relate his notion of endogenous uncertainty to common financial market "anomalies" such as excess volatility, and ARCH-type structure. When reading the work of Kurz, a logical extension of the CBS would be to endow agents with notions of uncertainty about future values rather than having point predictors. Then each agent's belief is described by a pdf, so that agents will generally have more dimensions to disagree upon. The distribution of beliefs, as defined in our CBS, then would become a pdf on a beliefs space of pdf's.

A Appendix

A.1 Derivation of Observation 1

Defining $\tilde{\sigma}_t^2 = 2\beta\sigma_t^2$ and $\tilde{\sigma}_{\text{op}}^2 = 2\beta\sigma_{\text{op}}^2$ leads to

$$\begin{array}{rcl} r_t & = & \mu_{t-1} r_{t-1} + h_t^{\frac{1}{2}} \epsilon_t, \\ \frac{\mu_t}{\tilde{\sigma}_t^2} & = & \alpha \frac{\mu_{t-1}}{\tilde{\sigma}_{t-1}^2} + (1-\alpha) r_t r_{t-1} + (1-\alpha) \frac{\mu_{\text{op}}}{\tilde{\sigma}_{\text{op}}^2}, \\ \frac{1}{\tilde{\sigma}_t^2} & = & \alpha \frac{1}{\tilde{\sigma}_{t-1}^2} + (1-\alpha) r_{t-1}^2 + (1-\alpha) \frac{1}{\tilde{\sigma}_{\text{op}}^2}, \end{array} \tag{22}$$

where

$$h_t = \frac{\tilde{\sigma}_{t-1}^2 r_{t-1}^2}{2\beta n} + \sigma_u^2$$

which shows that β and n enter the dynamics only through their product. □

A.2 Proof of Lemma 3

The assumed ergodicity, together with the invariance of the dynamics under the simultaneous inversion of the sign of all of the r_s's: $r_s \leftrightarrow -r_s \; \forall s$, implies $E[r_t] = 0$, or $\text{Var}[r_t] = E[r_t^2]$

whenever the mean and variance are finite. From Eq. (22) it can be seen that $\frac{1}{\sigma_t^2} - \frac{1}{\sigma_{\text{op}}^2}$ tends to $E[r_t^2]$, provided that $E[r_t^2]$ has a well-defined finite limit for $\alpha \to 1$. Similarly, $\frac{\mu_t}{\sigma_t^2} - \frac{\mu_{\text{op}}}{\sigma_{\text{op}}^2}$ converges to the first order auto-covariance $E[r_t r_{t-1}]$. If we define $\lim_{\alpha\to 1} \mu_t = \mu$ and $\lim_{\alpha \to 1} \sigma_t^2 = \sigma^2$, we obtain the following set of equations for μ and σ^2:

$$\begin{aligned} \frac{\mu}{\sigma^2} - \frac{\mu_{\text{op}}}{\sigma_{\text{op}}^2} &= E[r_t r_{t-1}] \\ \frac{1}{\sigma^2} - \frac{1}{\sigma_{\text{op}}^2} &= E[r_t^2]. \end{aligned}$$

The solutions to these equations are

$$\begin{aligned} \mu &= \frac{\mu_{\text{op}} + \sigma_{\text{op}}^2 E[r_t r_{t-1}]}{1 + \sigma_{\text{op}}^2 E[r_t^2]} \\ \sigma^2 &= (E[r_t^2] + 1/\sigma_{\text{op}}^2)^{-1}. \end{aligned} \tag{23}$$

This implies that the dynamics in the limit $\alpha \to 1$ can be described by

$$r_t = \mu r_{t-1} + h_t^{\frac{1}{2}} \epsilon_t, \tag{24}$$

where $h_t = \frac{\sigma^2}{2\beta n} r_{t-1}^2 + \sigma_u^2$. A simple consistency requirement on the first order autocorrelation coefficient and the above time varying AR model for r_t shows that in fact μ should be equal to the first order autocorrelation $\frac{E[r_t r_{t-1}]}{E[r_t^2]}$. By inserting this result into the top line of Eq. (23) it can be seen that this implies $\mu = \mu_{\text{op}}$, for all $0 < \sigma_{\text{op}}^2 < \infty$. □

A.3 Model restriction consistency

We consider the consistency of the assumption made in section 5 that $1/\sigma_{\text{op}}^2 \ll v = \text{Var}[r_t]$, which we used to reduce the number of parameters to be estimated from the model (apart from putting $\beta = \frac{1}{2}$). Since v is a function of the model parameters, it is not immediately obvious that $1/\sigma_{\text{op}}^2$ can become small with respect to v. The relation between v and the parameters can be found by taking the variance on both sides of Eq. (21). Upon substitution of $\gamma = \mu_0$, $c = \sigma_u^2$ and $a = (n(v + 1/\sigma_{\text{op}}^2))^{-1}$ we find

$$v = \mu_{\text{op}}^2 v + \frac{v}{n(v + \frac{1}{\sigma_{\text{op}}^2})} + \sigma_u^2,$$

which clearly shows that the model indeed allows $1/\sigma_{\text{op}}^2 \ll v$. Under this assumption, we have $n \simeq \frac{1}{av}$ so that the ARCH parameter can be used, together with the unconditional variance v, to estimate n, the effective number of market parties.

As an alternative to assuming a model with a normal opportunity function, taking the limit $\alpha \to 1$, and then considering large values of σ_{op}^2, one might alternatively consider starting with a constant opportunity function $\varphi(\theta) = 1$. Intuitively, taking a constant opportunity function is equivalent to fixing σ^2_{op} at infinity. Indeed, the dynamics in the limit $\alpha \to 1$ for any fixed value of $\mu_0 \in (-1, 1)$ obtained with a normal opportunity function is, in the limit $\sigma^2_{\text{op}} \to \infty$, interpretable as a model with constant opportunity function. However, problems in taking the limit $\alpha \to 1$ would arise if one would take a model with a constant opportunity function as a starting point. One obvious, but not severe disadvantage would be that, even if convergence of the time varying AR coefficient μ_t to a constant μ could be shown, the value of μ would not be interpretable in terms of the mean of the pdf associated with the opportunity function. More seriously, the convergence of μ_t to a constant μ would be hard to justify in the first place, since any value of $\mu \in (-1, 1)$ is compatible with a model based on a constant opportunity function.

References

Ben-Akiva, M. and Watanatada, T. (1981). Application of a continuous spatial choice logit model. In *Structural Analysis of Discrete Data and Econometric Applications* (eds C. F. Manski and D. L. McFadden), pp. 320–343. MIT Press, Cambridge, MA.

Bera, A. K., Higgins, M. L. and Lee, S. (1992). Interaction between autocorrelation and conditional heteroskedasticity: A random-coefficient approach. *Journal of Business and Economic Statistics*, **10,** 133–142.

Brock, W. A. and Durlauf, S. N. (2001). Discrete choice with social interactions. *Review of Economic Studies*, **68,** 235–260.

Brock, W. A. and Hommes, C. H. (1997). Rational routes to randomness. *Econometrica*, **65,** 1059–1095.

Brock, W. A., Hommes, C. H. and Wagener, F. O. O. (2001). Evolutionary dynamics in financial markets with many trader types. Technical Report. CeNDEF Working paper 01-01, University of Amsterdam.

Brock, W. A. and LeBaron, B. (1996). A dynamic structural model for stock return volatility and trading volume. *Review of Economics and Statistics*, **78,** 94–110.

Cabrales, A. and Hoshi, T. (1996). Heterogeneous beliefs, wealth accumulation, and asset price dynamics. *Journal of Economic Dynamics and Control*, **20,** 1073–1100.

Dagsvik, J. K. (1994). Discrete and continuous choice, max-stable processes, and independence from irrelevant attributes. *Econometrica*, **62,** number 5, 1179–1205.

Dagsvik, J. K. (2002). Discrete choice in continuous time: Implications of an intertemporal version of the IIA property. *Econometrica*, **70,** 817–831.

Diether, K. B., Malloy, C. J. and Scherbina, A. (2002). Differences of opinion and the cross section of stock returns. *Journal of Finance*, **57,** 2113–2141.

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, **50,** 987–1008.

Frankel, J. A. and Froot, K. A. (1990). Chartists, fundamentalists and trading in the foreign-exchange market. *American Economic Review*, **80,** 181–185.

Gallant, A. R., Rossi, P. E. and Tauchen, G. (1992). Stock prices and volume. *Review of Financial Studies*, **5,** 199–242.

Guesnerie, R. (2002). Anchoring economic prediction in common knowledge. *Econometrica*, **70,** 439–480.

Hanemann, W. M. (1984). Discrete/continuous models of consumer demands. *Econometrica*, **52,** number 3, 541–561.

Harvey, A. C. (1990). *The Econometric Analysis of Time Series*, 2nd edn. MIT Press, Cambridge, MA.

Hotz, V. J. and Miller, R. A. (1993). Conditional choice probabilities and the estimation of dynamic models. *Review of Economic Studies*, **60,** 497–529.

Kandel, E. and Pearson, N. D. (1995). Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy*, **4,** 831–872.

Keane, M. P. and Wolpin, K. I. (1994). The solution and estimation of discrete choice dynamic-programming models by simulation and interpolation – monte-carlo evidence. *Review of Economics and Statistics*, **76,** 648–672.

Kurz, M. (2001). Endogenous uncertainty and market volatility. *Economic Theory*, **17,** 497–544.

LeBaron, B. (2000). Agent-based computational finance: Suggested readings and early research. *Journal of Economic Dynamics and Control*, **24,** 679–702.

Lewbel, A. (1994). Aggregation and simple dynamics. *American Economic Review*, **84,** 905–918.

Lux, T. and Marchesi, M. (1999). Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature*, **397,** 498–500.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics*, *New York* (ed. P. Zarembka), pp. 105–142. Academic Press, NY.

Michaely, R. and Vila, J.-L. (1996). Trading volume with price valuation: Evidence from the ex-dividend day. *Review of Financial Studies*, **9,** 471–509.

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica*, **29,** number 3, 315–335.

Resnick, S. I. (1998). *A Probability Path.* Birkhäuser, Boston.

Resnick, S. I. and Roy, R. (1994). Superextremal processes, max-stability and dynamic continuous choice. *Annals of Applied Probability*, **4,** 791–811.

Shalen, C. T. (1993). Volume, volatility, and the dispersion of beliefs. *Review of Financial Studies*, **6,** 405–434.

Shiller, R. J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, **71,** 421–436.

Tsay, R. (1987). Conditional heteroskedastic time series models. *Journal of the American Statistical Association*, **82,** 590–604.

Ziegler, A. (2002). State-price densities under heterogeneous beliefs, the smile effect, and implied risk aversion. *European Economic Review*, **46,** 1539–1557.

First subsample, 1–2500				
Index	$\log n_{\text{eff}}$	AR	ARCH	$\text{Var}[r_t]$
SP500	5.39	0.049	0.067	0.008
FTSE	5.32	0.061	0.080	0.008
AEX	4.99	0.033	0.126	0.009
Hang Seng	4.30	0.112	0.264	0.014
Nikkei	4.26	0.126	0.435	0.011
Second subsample, 2501–3864				
S&P500	4.57	0.006	0.148	0.014
FTSE	4.44	0.049	0.202	0.013
AEX	4.01	0.066	0.327	0.017
Hang Seng	4.05	0.073	0.216	0.020
Nikkei	4.96	0.112	0.054	0.014

Table 1: *Separate estimates for the first period, from January 1, 1988 until August 1, 1997* (2500 *returns) and the remaining* 1364 *returns, from August 2, 1997, until October 24, 2002. (approx. the first 10 and the 5 most recent years).*

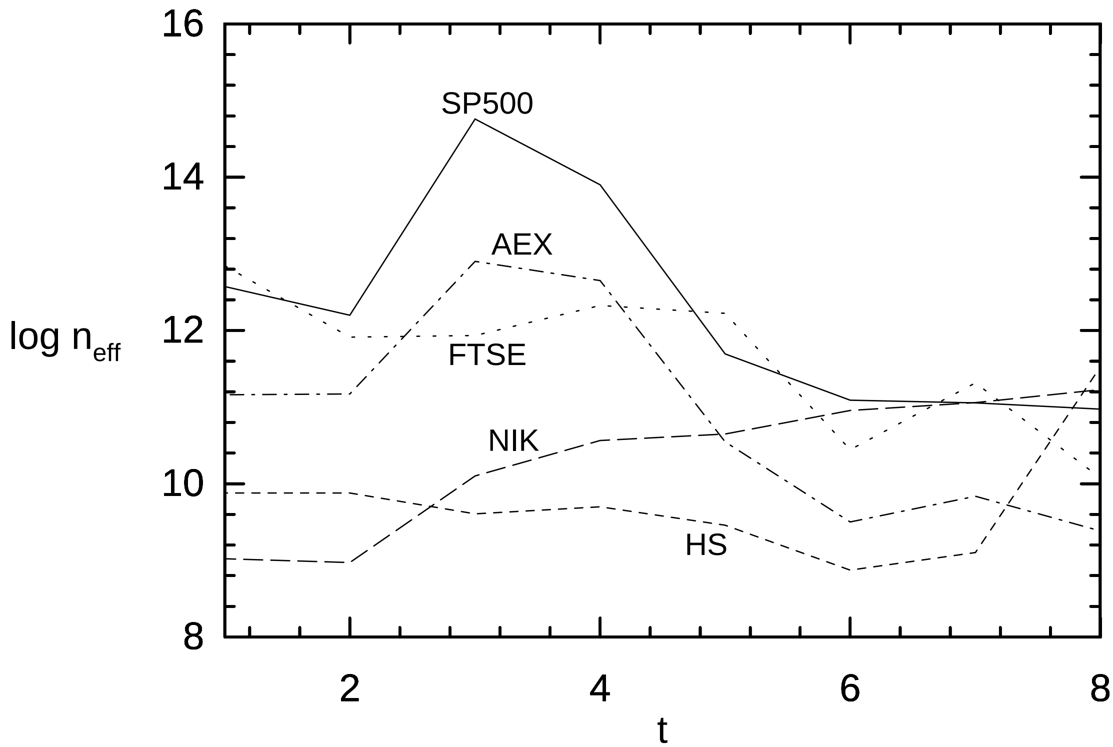

Figure 1: *Estimated effective number of agents, $n_{\rm eff}$, for five big stock indices: the S&P500, FTSE, AEX, Nikkei (NIK), Hang-Seng (HS). Estimation is based on estimated ARCH parameters in a moving time window (size* 1000 *trading days). The window is moved by* 400 *trading days for each consecutive estimate. The graph is based on daily returns from January 1st, 1988 until October 24, 2002.*

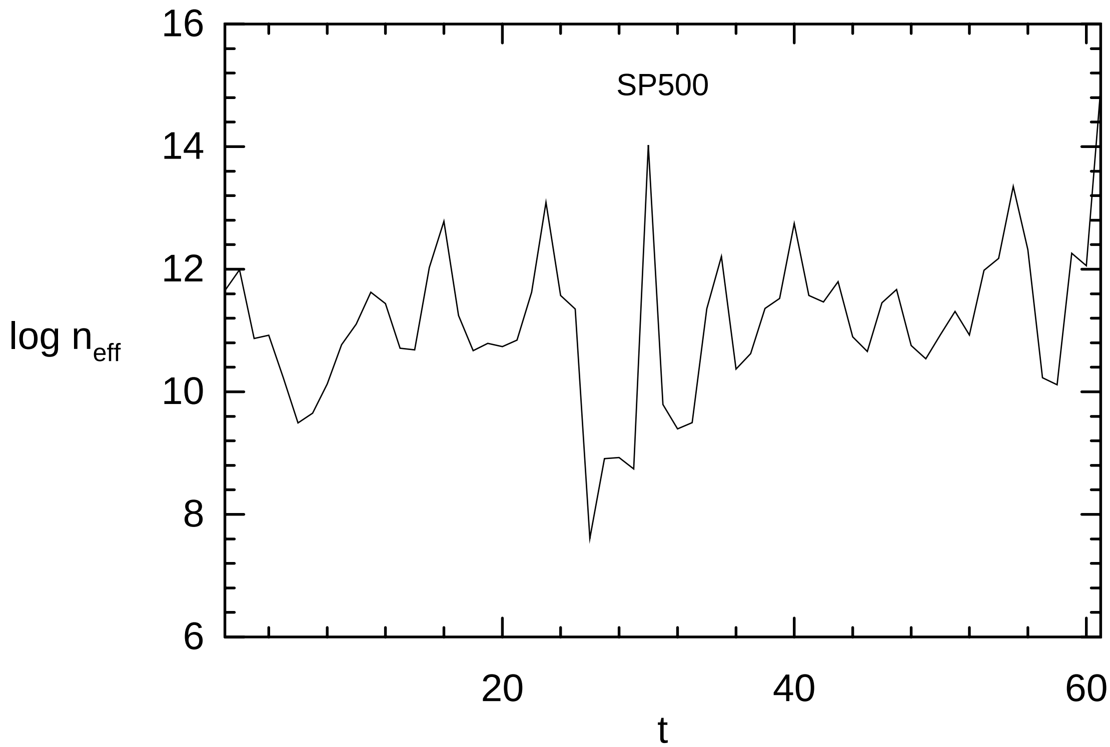

Figure 2: *Estimated effective number of agents,* $n_{\rm eff}$*, for the SP500 only, over a period of* 113 *years. Estimation is based on the ARCH parameter in a moving time window (size* 1000 *trading days). For this plot the window is moved by* 500 *trading days for each consecutive estimate. The data are daily returns from February 17, 1885 until October 6, 1998. The windows for the points labelled* $t = 20$, 40 *and* 60 *are roughly centered around the years 1918, 1952 and 1991*

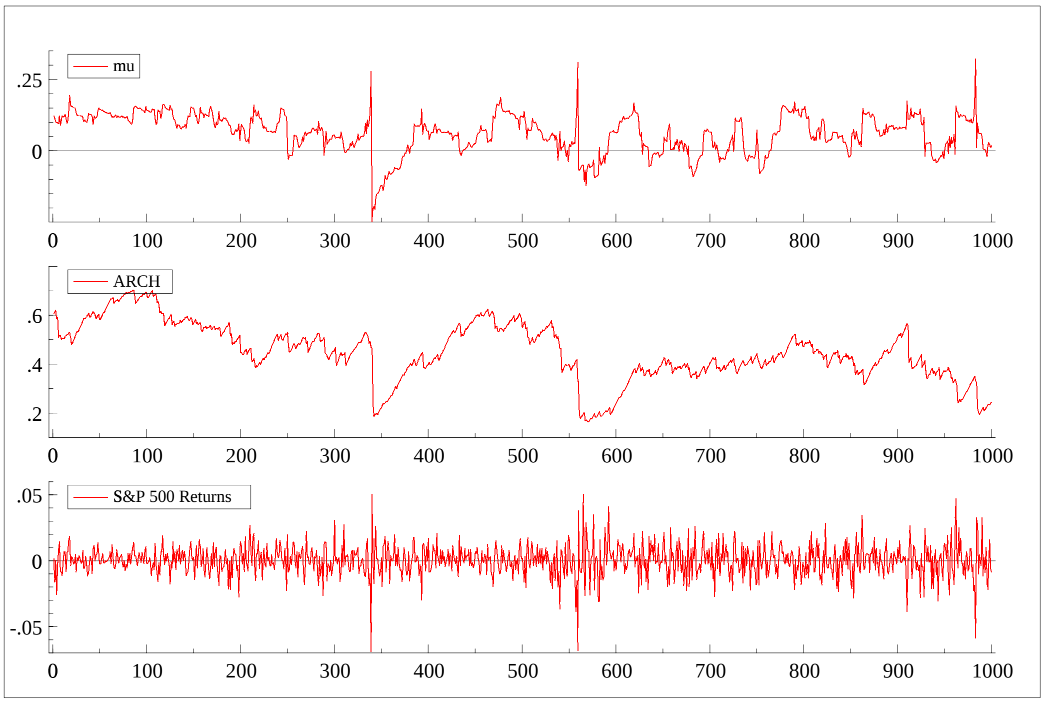

Figure 3: *Simulated response of the beliefs parameters to daily SP500 returns. The parameter values used are* $\alpha = 0.95$ *and* $\psi = 1$.