UvA-DARE (Digital Academic Repository)

Consistency of linear forecasts in a nonlinear stochastic economy

Hommes, C.; Sorger, G.; Wagener, F.

DOI

[10.1007/978-3-642-29503-4_10](https://doi.org/10.1007/978-3-642-29503-4_10)

Publication date

2013

Document Version

Final published version

Published in

Global analysis of dynamic models in economic and finance: essays in honour of Laura Gardini

[Link to publication](#)

Citation for published version (APA):

Hommes, C., Sorger, G., & Wagener, F. (2013). Consistency of linear forecasts in a nonlinear stochastic economy. In G. I. Bischi, C. Chiarella, & I. Sushko (Eds.), *Global analysis of dynamic models in economic and finance: essays in honour of Laura Gardini* (pp. 229-287). Springer. https://doi.org/10.1007/978-3-642-29503-4_10

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)

Consistency of Linear Forecasts in a Nonlinear Stochastic Economy

Cars Hommes, Gerhard Sorger, and Florian Wagener

Abstract The notion of consistent expectations equilibrium is extended to economies that are described by a nonlinear stochastic system. Agents in the model do not know the nonlinear law of motion and use a simple linear forecasting rule to form their expectations. Along a stochastic consistent expectations equilibrium (SCEE), these expectations are correct in a linear statistical sense, i.e., the unconditional mean and autocovariances of the actual (but unknown) nonlinear stochastic process coincide with those of the linear stochastic process on which the agents base their beliefs. In general, the linear forecasts do not coincide with the true conditional expectation, but an SCEE is an ‘approximate rational expectations equilibrium’ in the sense that forecasting errors are unbiased and uncorrelated. Adaptive learning of SCEE is studied in an overlapping generations framework.

Keywords Bounded rationality • Expectation formation • Consistent expectations • Sample autocorrelation learning • Path dependence

1 Introduction

Modeling the formation of expectations is a crucial step in every dynamic economic theory. Agents base their consumption, production, and investment decisions upon expectations or beliefs about future prices and quantities. These decisions affect

C. Hommes (✉)
CeNDEF, University of Amsterdam,
Amsterdam, The Netherlands
e-mail: C.H.Hommes@uva.nl

G. Sorger
University of Vienna, Vienna, Austria
e-mail: gerhard.sorger@univie.ac.at

F. Wagener
CeNDEF, University of Amsterdam, Amsterdam, The Netherlands
e-mail: F.O.O.Wagener@uva.nl

G.I. Bischi et al. (eds.), *Global Analysis of Dynamic Models in Economics and Finance*,
DOI 10.1007/978-3-642-29503-4_10, © Springer-Verlag Berlin Heidelberg 2013

the current and future market equilibrium outcomes which, in turn, are taken into account when new expectations are formed. A dynamic market economy is thus an expectations feedback system.

In the 1930s, simple habitual rule of thumb forecasting rules such as naive or adaptive expectations were popular in modeling the formation of expectations. An important problem with using such simple, mechanical forecasting rules is that the agents make *systematic* mistakes, i.e., their forecasting errors are biased and/or autocorrelated. This means that the use of these simple forecasting rules can only be justified if the agents are assumed to act irrationally, because any rational agent would learn from systematic forecasting errors, adapt his behavior, and change his forecasting rule accordingly. This type of argument against habitual rule of thumb forecasting rules was used, e.g., by Muth (1961) in his pioneering paper introducing the rational expectations hypothesis. Since its application to macroeconomics by Lucas (1971) and others, this hypothesis has become the predominant paradigm in expectation formation in economics. It postulates that an agent's subjective forecast of a future variable equals the true (mathematical) expectation of that variable conditional upon all available information (including economic theory). Thus, a rational expectations equilibrium (REE) can be considered as a fixed point of the expectations feedback system. Forecasting errors in an REE do not have any exploitable structure.

This paper argues that, in an unknown nonlinear stochastic environment, simple (linear) forecasting rules need not be irrational but may in fact be consistent with observations *in a linear statistical sense*. We extend the notion of consistent expectations equilibrium, developed by Hommes and Sorger (1998), to nonlinear stochastic dynamic economic models. The key feature of a *stochastic consistent expectations equilibrium* (SCEE) is that agents' expectations about endogenous variables are consistent with the actual realizations of these variables in the sense that the unconditional mean and auto-covariances of the unknown nonlinear stochastic process, which describes the actual behavior of the economy, coincide with the unconditional mean and auto-covariances of the linear stochastic process agents believe in. Along an SCEE, the endogenous variable may be autocorrelated (with the same autocorrelation structure as the linear forecasting rule), but forecasting errors are uncorrelated. In other words, along an SCEE, agents do make mistakes but these mistakes are not systematic. If agents were using linear statistical tests, they would not be able to distinguish between the (true) nonlinear stochastic law of motion and their (perceived) linear stochastic model. As a consequence, agents would have no reason to deviate from their simple linear forecasting rule and, therefore, the situation qualifies as an equilibrium. Stated differently, an SCEE is an 'approximate rational expectations equilibrium' because forecasting errors are unbiased and uncorrelated.

There are situations in which an SCEE qualifies also as an REE. This happens for example in a so-called steady state SCEE, in which the endogenous variable follows an i.i.d. random process. In general, however, an SCEE does not coincide with an REE, since the true conditional expectation of the unknown nonlinear stochastic

law of motion is not equal to the linear perceived law of motion (i.e., the linear forecasting rule employed by the agents). An important difference between the notion of SCEE and REE is that, in the former, the agents do not need to have any knowledge about the underlying market equilibrium equations. In other words, an SCEE is a fixed point of the expectations feedback system in terms of the *observable* sample average and sample autocorrelations.

In an earlier paper, [Hommes and Sorger \(1998\)](#), we introduced the notion of consistent expectations equilibrium (CEE) in a deterministic framework and studied three types of CEE: steady state CEE, 2-cycle CEE, and chaotic CEE. The concept of CEE may be seen as a formalization of the notion of a *self-fulfilling mistake* introduced by [Grandmont \(1998\)](#), where agents incorrectly believe that the economy follows a stochastic process whereas the actual dynamics is generated by a deterministic chaotic process which is indistinguishable from the former (stochastic) process by linear statistical tests; see also [Sorger \(1998\)](#) and [Hommes \(1998\)](#). [Sögner and Mitlöhner \(2002\)](#) applied the CEE concept to a standard asset pricing model. [Hommes and Rosser \(2001\)](#) investigate CEE in an optimal fishery management model and use numerical methods to study adaptive learning of CEE in the presence of dynamic noise.

The present paper may be viewed as an extension of the concept of CEE to a stochastic framework. It studies SCEE for models of the form $p_t = F(p_{t+1}^e, \eta_t)$, where p_t represents the endogenous variable (henceforth we shall refer to that variable as the price) at date t , p_{t+1}^e is the expected price for period $t + 1$, $(\eta_t)_{t=0}^{+\infty}$ is an i.i.d. random process with mean zero, and F is a continuous function. The overlapping generations model (OG-model) and the standard present value asset pricing model are well known examples. In a pioneering paper, [Grandmont \(1985\)](#) showed that in a standard version of the OG-model, when the offer curve is sufficiently backward bending, infinitely many periodic and even chaotic REE exist.¹ We study sample autocorrelation learning (SAC-learning) in this OG-model, that is, an adaptive learning process in which agents update the parameters of their perceived law of motion – a linear AR(1) model – according to the observations of the sample mean and the sample autocorrelations. We find that, in Grandmont's specification of the OG-model, only steady state SCEE and 'noisy 2-cycle' SCEE occur as long run outcomes under SAC-learning. Thus, even in an economy in which infinitely many (chaotic) perfect foresight equilibria exist, boundedly rational agents, who search for an optimal linear forecasting rule, learn to coordinate on one of these simple equilibria.

[Branch and McGough \(2005\)](#) show the existence of (first order) SCEE for a class of nonlinear stochastic processes and numerically investigate their stability under real time adaptive learning. In a recent paper, [Hommes and Zhu \(2011\)](#) study existence of SCEE and their stability under adaptive learning in linear stochastic

¹The model in [Grandmont \(1985\)](#) is deterministic such that REE are actually perfect foresight equilibria.

models driven by autocorrelated processes, with agent using a simple univariate linear forecasting rule. They also emphasize a behavioral interpretation of SCEE and learning.

Our work fits well into the recent literature on bounded rationality, expectations formation, and learning, as discussed extensively in [Sargent \(1993, 1999\)](#) and [Evans and Honkapohja \(2001\)](#). In contrast to fully rational agents, it is assumed that boundedly rational agents do not know market equilibrium equations, but that they base their beliefs only upon time series observations. For example, the agents might compute their expectations from actual time series observations in the past by ordinary least squares regressions; see, e.g., [Bray \(1982\)](#), [Bray and Savin \(1986\)](#), and [Marcet and Sargent \(1989\)](#). Under this assumption it may or may not be the case that the asymptotic outcome of such a learning process is an REE. Our SCEE may be seen as an example of a misspecification equilibrium, where agents try to find an optimal linear forecast in an unknown nonlinear world. It may also be viewed as an example of a Restricted Perceptions Equilibrium (RPE), as defined by [Evans and Honkapohja \(2001\)](#), formalizing the idea that agents have misspecified beliefs, but within the context of their forecasting model they are unable to detect their misspecification. See Branch (2006) for an excellent survey, arguing that the RPE is a natural alternative to rational expectation equilibrium.

The concepts of CEE and SCEE are also related to quasi rational expectations introduced in [Nerlove, Grether, and Carvalho \(1979, Chap. XIII\)](#), where the expectations about both exogenous and endogenous variables are given by those predictors which minimize the mean squared prediction errors in an ARIMA model. However, our focus is on the expectations feedback and on the consistency between linear expectations and a nonlinear actual law of motion. Other recent work related to bounded rationality and expectation formation includes the rational belief equilibria in [Kurz \(1994\)](#), the pseudo rational learning in [Marcet and Nicolini \(2003\)](#), the expectational stability and adaptive learning rules in [Evans and Honkapohja \(1994, 1995\)](#), the perfect predictors in [Böhm and Wenzelburger \(1999\)](#), and the adaptive rational equilibrium dynamics in [Brock and Hommes \(1997, 1998\)](#). Instability of adaptive learning processes has been investigated, e.g., by [Bullard \(1994\)](#), [Schönhöfer \(1999\)](#), and [Tuinstra and Wagener \(2007\)](#).

The paper is organized as follows. Section 2 introduces the main concepts, i.e., CEE, SCEE, and SAC-learning. Section 3 presents examples of (deterministic) CEE. Section 4 discusses the relation between SCEE and REE. This section also states the main theorem, that along an SCEE forecasting errors are always unbiased and uncorrelated. Moreover, it is shown that first order non-steady state SCEE exist under fairly general conditions. Section 5 investigates adaptive learning of CEE and SCEE in two different specifications of the OG-model with and without noise. An intuitive graphical analysis of the SAC-learning process is given, showing that the learning process may exhibit path dependence that may lead to (two) different long run outcomes. Finally, Sect. 6 concludes.

2 Conceptual Framework

This section introduces the main concepts of the paper. We start by reviewing some background material and by presenting a motivating example. Then we define deterministic CEE, stochastic CEE, and SAC-learning.

2.1 An Example

Consider a continuous map of a compact interval into itself, say, $f : [a, b] \mapsto [a, b]$, where a and b are real numbers with $a < b$. Such a map defines a deterministic dynamical system through the difference equation

$$p_t = f(p_{t-1}). \quad (1)$$

Given any initial state $p_0 \in [a, b]$, (1) determines a unique trajectory $(p_t)_{t=0}^{+\infty}$. Deterministic dynamical systems and their possible complicated, chaotic behavior have been explored in great detail; see, e.g., [Devaney \(1989\)](#). For our purpose, the so-called ergodic approach to dynamical system will be the most relevant one because it highlights the statistical, or probabilistic properties of the trajectories; see, e.g., [Lasota and Mackey \(1985, 1994\)](#). To get started, we shall therefore introduce some of the basic concepts of the ergodic approach and illustrate them by means of a simple example.

Suppose that a probability measure μ on the set $[a, b]$ is given, and that the initial state p_0 is drawn randomly according to that measure. In that case, (1) defines a stochastic process with the peculiar feature that, conditional on p_0 , the process is deterministic. The probability measure μ is said to be invariant under f , if $\mu(f^{-1}(B)) = \mu(B)$ holds for all measurable sets $B \subseteq [a, b]$, where $f^{-1}(B) = \{p \in [a, b] \mid f(p) \in B\}$. If μ is invariant under f and p_0 is drawn according to μ , then it follows that for every $t \geq 0$ the unconditional distribution of p_t is given by μ , and that the stochastic process defined by (1) is stationary. It has been shown that, for every dynamical system defined by a continuous map from a compact interval to itself, there exists at least one invariant probability measure. In general, there will exist many such measures.

Of particular interest are so-called ergodic measures. The measure μ is ergodic under f if, for every measurable set $B \subseteq [a, b]$ which satisfies $f^{-1}(B) = B$, it holds that either $\mu(B) = 0$ or $\mu(B) = 1$. To explain the importance of ergodic invariant measures let us denote the j -th iterate of f by $f^{(j)}$, that is, $f^{(j)}$ is the map f composed with itself j times. The ergodic theorem says that, for every probability measure μ which is ergodic and invariant under f and for every integrable function g , the equation

$$\lim_{T \rightarrow +\infty} \frac{1}{T+1} \sum_{j=0}^T g(f^{(j)}(p_0)) = \int_{[a,b]} g(p) d\mu(p)$$

holds for μ -almost all initial states p_0 . In other words, the time average of the function g along the trajectory of (1) starting in p_0 coincides with the space average of g with respect to the measure μ . In particular, the time averages are constant (i.e., independent of the initial state p_0) on a set of full μ -measure. This property is especially useful if the probability measure μ is absolutely continuous with respect to Lebesgue measure because in that case it follows that any property that holds on a set of μ -measure 1 automatically holds for Lebesgue-almost all initial values p_0 .

Finally, let us define the sample average and the sample autocorrelation coefficients of a trajectory generated by the dynamical system (1).² The sample average along the trajectory starting in p_0 is defined as

$$\bar{p}(p_0) = \lim_{T \rightarrow \infty} \frac{1}{T+1} \sum_{t=0}^T p_t = \lim_{T \rightarrow \infty} \frac{1}{T+1} \sum_{t=0}^T f^{(t)}(p_0), \quad (2)$$

and the sample autocorrelation coefficient at lag $j \geq 0$ along this trajectory is defined by

$$\rho_j(p_0) = \lim_{T \rightarrow \infty} \frac{c_{j,T}(p_0)}{c_{0,T}(p_0)}, \quad (3)$$

where

$$\begin{aligned} c_{j,T}(p_0) &= \frac{1}{T+1} \sum_{t=0}^{T-j} [p_t - \bar{p}(p_0)][p_{t+j} - \bar{p}(p_0)] \\ &= \frac{1}{T+1} \sum_{t=0}^{T-j} [f^{(t)}(p_0) - \bar{p}(p_0)][f^{(t+j)}(p_0) - \bar{p}(p_0)]. \end{aligned}$$

We are now ready to present an example that will help us explain certain features of CEE and SCEE. To this end, let us define the (asymmetric) tent map $T_{\beta,[a,b]} : [a,b] \mapsto [a,b]$ by

$$T_{\beta,[a,b]}(x) = \begin{cases} \frac{2}{1+\beta}(x-a) + a & \text{if } a \leq x \leq a + \frac{1+\beta}{2}(b-a), \\ \frac{2}{1-\beta}(b-x) + a & \text{if } a + \frac{1+\beta}{2}(b-a) < x \leq b, \end{cases}$$

where a, b , and β are real numbers satisfying $a < b$ and $\beta \in (-1, 1)$. Figure 1 shows the graph of the tent map for different values of β . Note that $T_{\beta,[a,b]}$ is

²See, e.g., Box, Jenkins, and Reinsel (1994) for a discussion of these definitions.

Fig. 1 Graphs of $T_{\beta, [a, b]}$ for $\beta = -0.7$ (top left), $\beta = 0$ (top right), and $\beta = 0.7$ (bottom). The geometric construction of trajectories starting at a given initial state is also shown

piecewise linear and continuous, and that it is uniformly expanding, that is, the absolute value of its slope has a lower bound that is greater than 1. The figure also shows trajectories generated by the corresponding dynamical system

$$p_t = T_{\beta, [a, b]}(p_{t-1}) \quad (4)$$

for different initial values p_0 .

It is well known that the dynamical system (4) has a periodic orbit for every period $j \geq 1$. If $(p_1, p_2, \dots, p_j)$ is such a periodic orbit, then we can define a probability measure μ by $\mu(B) = \#\{p \in B \mid p = p_i \text{ for some } i = 1, 2, \dots, j\}/j$. It is straightforward to see that μ defined in this way is invariant and ergodic under $T_{\beta, [a, b]}$. Thus, there exist infinitely many invariant and ergodic probability measures for (4). It is also straightforward to prove that the uniform distribution on $[a, b]$ is an invariant and ergodic probability measure for (4). As a matter of fact, the uniform distribution on $[a, b]$ is the unique invariant probability measure for (4) which is both ergodic and absolutely continuous with respect to Lebesgue measure.³ Denoting

³Dynamical systems which are uniformly expanding – such as (4) – typically have ‘nice’ invariant probability measures; see, e.g., Lasota and Mackey (1985).

the uniform distribution on $[a, b]$ by μ , it has been shown by Sakai and Tokumaru (1980) that

$$\int_{[a,b]} p \, d\mu(p) = (a + b)/2 \quad (5)$$

and

$$\int_{[a,b]} [T_{\beta,[a,b]}^{(j)}(p) - (a + b)/2][p - (a + b)/2] \, d\mu(p) = \frac{(b - a)^2}{12} \beta^j \quad (6)$$

for all $j \geq 0$. Equation 5 shows that the unconditional mean of the stochastic process defined by the difference equation (4) and the assumption of a uniformly distributed initial state p_0 is given by $(a + b)/2$. Analogously, (6) shows that the autocorrelation coefficient at lag j of this process is given by β^j . Using these results and the absolute continuity of μ with respect to Lebesgue measure, the ergodic theorem implies that the following two properties hold.

- T1 For Lebesgue-almost all initial states $p_0 \in [0, 1]$, the sample average of the trajectory of (4) starting in p_0 satisfies $\bar{p}(p_0) = (a + b)/2$.
- T2 For Lebesgue-almost all initial states $p_0 \in [a, b]$, the sample autocorrelation coefficient at lag j of the trajectory of (4) starting in p_0 satisfies $\rho_j(p_0) = \beta^j$.

Now consider the stochastic AR(1) process

$$p_t = \alpha + \beta(p_{t-1} - \alpha) + \delta_t \quad (7)$$

where α and β are real numbers, $\beta \in (-1, 1)$, and where $(\delta_t)_{t=0}^{+\infty}$ is an i.i.d. stochastic process with $\mathbb{E}\delta_t = 0$ and $\mathbb{E}\delta_t^2 = \sigma_\delta^2 > 0$. Furthermore, assume that p_0 is a random variable with $\mathbb{E}p_0 = \alpha$ and $\mathbb{E}(p_0 - \alpha)^2 = \sigma^2 := \sigma_\delta^2 / (1 - \beta^2)$. Under these assumptions it follows that the unconditional first and second moments of p_t exist for all t and that they are given by $\mathbb{E}p_t = \alpha$ and $\mathbb{E}(p_t - \alpha)(p_{t+j} - \alpha) = \sigma^2 \beta^j$. In other words, the unconditional first and second moments of the AR(1) process are stationary, the mean is α , and the autocorrelation coefficient at lag j is β^j .⁴ Comparing these observations with the results stated in (5), (6), (T1) and (T2) above, we see that the stochastic process (4) with uniformly distributed initial state has exactly the same unconditional first and second moments as the AR(1) process (7) with $\alpha = (a + b)/2$. Note that the latter process involves stochastic shocks δ_t in every period t , whereas the former evolves completely deterministically once the initial state has been selected.

The above results are statements about the means and autocorrelation coefficients of two stochastic processes. Because of the ergodic theorem, these statements carry over to the sample averages and sample autocorrelations. More specifically, using

⁴Note that stationarity of the first two moments does not necessarily imply stationarity of the process itself.

properties T1 and T2 from above it follows that, for Lebesgue-almost all initial states p_0 , the sample average of a trajectory generated by the tent map dynamics (4) coincides with the unconditional mean of the AR(1) process (7). Analogously, for Lebesgue-almost all initial states p_0 , the sample autocorrelation coefficients of a trajectory generated by the tent map dynamics (4) coincide with the unconditional autocorrelation coefficients of the AR(1) process (7). In other words, *for an econometrician who observes only sample means and sample autocorrelations, the two processes (4) and (7) are indistinguishable*. These findings are illustrated in Fig. 2, which shows 2,000 points of a chaotic time series (top left) generated by the asymmetric tent map $T_{0.7,[0,1]}$. The sample autocorrelation plot (top right) looks like the autocorrelation function of an AR(1) process with its exponentially decreasing autocorrelation pattern. Now assume that we use the linear forecasting rule $p_t^e = \alpha + \beta(p_{t-1} - \alpha)$ with $\alpha = (a + b)/2 = 0.5$ and $\beta = 0.7$ to predict the future evolution of the tent map dynamics. The resulting forecast error in period t is given by $\epsilon_t = p_t - p_t^e = p_t - [\alpha + \beta(p_{t-1} - \alpha)]$. Figure 2 (bottom left) shows the sample autocorrelation coefficients of these forecast errors. All coefficients are close to 0, which illustrates the fact that the forecast errors are uncorrelated. We also fitted an AR(1) model to the chaotic time series and obtained parameter estimates $\alpha = 0.483$ and $\beta = 0.699$. The sample autocorrelation plot (bottom right) of the residuals of the estimated AR(1) model is close to 0, showing that the residuals are uncorrelated, too. In fact, the AR(1) model cannot be rejected at the 5% level. To an observer using linear statistical techniques, deterministic chaos generated by a tent map is therefore practically indistinguishable from the dynamics of a stochastic AR(1) process.

Hommes and Sorger (1998) define a CEE for a deterministic economy as a price sequences for which sample average and sample autocorrelations coincide with the unconditional mean and autocorrelations of the agents' perceived law of motion; see Sect. 2.2 below. Note that the actual dynamics generating the sequence of prices is linked to the perceived law of motion via the expectations feedback. Considering the material presented above, a natural way to generalize the CEE concept is to replace sample average and sample autocorrelation coefficients, which are defined as time averages, by corresponding values that can be computed as space averages with respect to an invariant measure. This approach is especially useful in a stochastic model, because it takes the probabilistic nature of equilibria better into account and shifts the emphasis away from single time series to the entire process. Thus, for an SCEE we will require that the unconditional mean and autocorrelation coefficients under an invariant measure of the law of motion of the economy coincide with the unconditional mean and autocorrelation coefficients of the agents' perceived law of motion. As in the case of CEE, the situation is complicated by the fact that the two processes (the actual law of motion and the perceived law of motion) are connected to each other by the expectations feedback.

Fig. 2 Two-thousand observations of a chaotic time series generated by the asymmetric tent map with $\beta = 0.7$ (top left), the corresponding sample autocorrelation function (top right), the sample autocorrelation function of one-period ahead forecasting errors (bottom left), and the sample autocorrelation function of the residuals of an estimated AR(1) model (bottom right)

2.2 Deterministic CEE

CEE have been introduced in [Hommes and Sorger \(1998\)](#) for the ‘cobweb-type’ model $p_t = G(p_t^e)$. In the present paper, on the other hand, we study the model

$$p_t = G(p_{t+1}^e). \quad (8)$$

According to (8), today’s market equilibrium price p_t depends on the expected price p_{t+1}^e for tomorrow. Throughout the paper we assume that the map G is continuous. Typically, it will be a nonlinear function. Equation (8) arises for example in the study of temporary equilibria in [Grandmont \(1998\)](#), in overlapping generations models (OG-models), and in standard asset pricing models. In Sect. 5 we will discuss the OG-model as the main application and, therefore, we refer to (8) as an OG-type model. In Sect. 2.3 we will generalize equation (8) to a stochastic setting.

We assume that the agents in the model do not know G and, thus, they are not able to use knowledge about G in forming their expectations. Instead, they form expectations based only upon time series observations. More specifically, the agents know all past prices $p_0, p_1, \dots, p_{t-1}$ and use this information in generating the forecast p_{t+1}^e . Notice that the current equilibrium price p_t is a function of the next period's forecast p_{t+1}^e and, consequently, p_t is not known to the agents when they make that forecast.

Suppose the agents believe that the price sequence follows a linear stochastic process. More specifically, suppose that all agents believe that prices are generated by the AR(1) process

$$p_t = \alpha + \beta(p_{t-1} - \alpha) + \delta_t. \quad (9)$$

Here, α and β are real numbers, $\beta \in (-1, 1)$, and $(\delta_t)_{t=0}^{+\infty}$ is a white noise process, i.e., a sequence of i.i.d. random variables with mean zero and $\mathbb{E}\delta_t^2 = \sigma_\delta^2 > 0$. We assume that the belief process in (9) has stationary first and second moments which is the case if and only if $\mathbb{E}p_0 = \alpha$ and $\mathbb{E}(p_0 - \alpha)^2 = \sigma^2 := \sigma_\delta^2 / (1 - \beta^2)$. As has been mentioned in Sect. 2.1, this implies that the unconditional mean of p_t is given by α and the autocorrelation coefficient at lag j is given by β^j .

We will refer to (9) as the *perceived law of motion*. Given the perceived law of motion and knowledge of all prices observed up to period $t - 1$, the unique 2-period ahead forecasting rule for p_{t+1} that minimizes the mean squared forecasting error is

$$p_{t+1}^e = \alpha + \beta^2(p_{t-1} - \alpha). \quad (10)$$

The 2-period ahead forecast is thus the unconditional mean price α plus the second order autocorrelation coefficient β^2 times the latest observed deviation of the price from its unconditional mean. Given that agents use the linear predictor (10), the *implied actual law of motion* becomes

$$p_t = G_{\alpha, \beta}(p_{t-1}) := G(\alpha + \beta^2(p_{t-1} - \alpha)). \quad (11)$$

Recalling the definitions of the sample average and the sample autocorrelations of a time series $(p_t)_{t=0}^{+\infty}$ from (2)–(3), we are now ready to define a CEE.

Definition 1. A triple $\{(p_t)_{t=0}^{+\infty}; \alpha, \beta\}$, where $(p_t)_{t=0}^{+\infty}$ is a sequence of prices and α and β are real numbers, $\beta \in [-1, 1]$, is called a (deterministic) consistent expectations equilibrium (CEE) if

- C1 the sequence $(p_t)_{t=0}^{+\infty}$ satisfies the implied actual law of motion (11),
- C2 the sample average $\bar{p}(p_0)$ of the sequence $(p_t)_{t=0}^{+\infty}$ is equal to α , and
- C3 for the sample autocorrelation coefficients $\rho_j(p_0)$ of the sequence $(p_t)_{t=0}^{+\infty}$ the following is true:

(a) if $(p_t)_{t=0}^{+\infty}$ is a convergent sequence, then $\text{sgn}[\rho_j(p_0)] = \text{sgn}(\beta^j)$ for all $j \geq 1$,

(b) if $(p_t)_{t=0}^{+\infty}$ is not convergent, then $\rho_j(p_0) = \beta^j$ for all $j \geq 1$.

Property C1 simply states that the sequence $(p_t)_{t=0}^{+\infty}$ satisfies the equilibrium equations of the economy, (8), given that the agents use the linear forecasting rule (10). Condition C2 requires that the sample average of the observed time series equals the unconditional mean, α , of the stochastic AR(1) belief (9). Condition C3 tries to capture the requirement that the sample autocorrelation coefficients of the observed time series equal the autocorrelation coefficients of the stochastic AR(1) belief. A few remarks are in order concerning this condition. First of all, we allow the parameter β to take values in the closed interval $[-1, 1]$, that is, the endpoints -1 and 1 are not excluded. Note that for $\beta = -1$ or $\beta = 1$, the autocorrelation coefficients of (9) are not given by β^j because the second moments are not even stationary. Therefore, the interpretation of a CEE as a fixed point of the expectations feedback in terms of first and second moments is not strictly true for the boundary points $\beta = -1$ and $\beta = 1$. On the other hand, what actually matters for the implied actual law of motion is the linear forecasting rule (10) and not the underlying stochastic process (9) itself. It will be seen in Sect. 3 that the case $\beta = -1$ is of considerable interest as it occurs quite naturally in so-called 2-cycle CEE. The second remark concerns the distinction made between convergent time series and nonconvergent ones. If $(p_t)_{t=0}^{+\infty}$ does not converge, then condition C3b requires that the sample autocorrelation coefficients of the actual prices are exactly the same as the autocorrelation coefficients of the perceived law of motion. If the price sequence converges, however, we make the weaker requirement that autocorrelation coefficients of observations and beliefs have the same sign but not necessarily the same value. A CEE for which the sequence $(p_t)_{t=0}^{+\infty}$ converges will be called a steady state CEE; see Sect. 3 below. The reason for the weaker consistency requirement for steady state CEE is related to the singularity of the sample autocorrelation coefficient $\rho_j(p_0)$ at any fixed point of $G_{\alpha, \beta}$. If p_0 is such a fixed point, the trajectory starting at p_0 must be constant and the autocorrelation coefficients $\rho_j(p_0)$ are not defined.⁵ For SCEE defined in Sect. 2.3 below, this problem will disappear because the implied actual law of motion will be a nontrivial stochastic process for which all autocorrelation coefficients are well defined.

Summarizing, a CEE is a price sequence and an AR(1) belief process such that expectations are self-fulfilling in terms of the observable sample average and sample autocorrelations.

For some of our results it will be convenient to define also a weaker equilibrium concept, which we call k -th order CEE. The only difference to Definition 1 is that for a k -th order CEE, condition C3 is required to hold only for all $j \leq k$. Formally, we have the following definition.

Definition 2. Let k be a positive integer. A triple $\{(p_t)_{t=0}^{+\infty}; \alpha, \beta\}$, where $(p_t)_{t=0}^{+\infty}$ is a sequence of prices and α and β are real numbers, $\beta \in [-1, 1]$, is called a k -th

⁵See also Hommes and Sorger (1998, footnote 2).

order CEE if conditions C1 and C2 from Definition 1 hold and if for the sample autocorrelation coefficients $\rho_j(p_0)$ of the sequence $(p_t)_{t=0}^{+\infty}$ the following is true:

(a) if $(p_t)_{t=0}^{+\infty}$ is a convergent sequence, then

$$\text{sgn}[\rho_j(p_0)] = \text{sgn}(\beta^j)$$

for all $j \in \{1, 2, \dots, k\}$,

(b) if $(p_t)_{t=0}^{+\infty}$ is not convergent, then $\rho_j(p_0) = \beta^j$ for all $j \in \{1, 2, \dots, k\}$.

2.3 SCEE

Now suppose that the law of motion is given by the stochastic difference equation

$$p_t = F(p_{t+1}^e, \eta_t), \quad (12)$$

where F is a continuous function (usually a nonlinear one) and $(\eta_t)_{t=0}^{+\infty}$ is a white noise process. Given that the perceived law of motion is still the stochastic AR(1) process (9) and that agents use the linear 2-period ahead forecasting rule (10), the implied actual law of motion becomes

$$p_t = F_{\alpha,\beta}(p_{t-1}, \eta_t) := F(\alpha + \beta^2(p_{t-1} - \alpha), \eta_t). \quad (13)$$

An equivalent way of defining the stochastic process in (13) uses the so-called transition function $Q_{\alpha,\beta}$. More precisely, for every real number p and every measurable set $A \subseteq \mathbb{R}$, we denote by $Q_{\alpha,\beta}(p, A)$ the conditional probability that $p_t \in A$ given that $p_{t-1} = p$ and that (13) holds. Thus, $Q_{\alpha,\beta}(p, A)$ is the probability of moving from price p to the set A in a single period. Analogously, let $Q_{\alpha,\beta}^{(j)}(p, A)$ be the probability of moving from price p to the set A along a trajectory of (13) in exactly $j \geq 0$ periods. Note furthermore that the stochastic difference equation (13) defines an operator $\mathcal{T}_{\alpha,\beta}^*$ on the set of probability measures in the following way. If p_{t-1} is a random variable with distribution μ_{t-1} , then the distribution of p_t generated by (13) is $\mu_t = \mathcal{T}_{\alpha,\beta}^* \mu_{t-1}$. Under certain assumptions on the function $F_{\alpha,\beta}$ and the distribution of η_t , there exists a unique invariant probability measure for this stochastic difference equation, that is, a probability measure μ such that $\mathcal{T}_{\alpha,\beta}^* \mu = \mu$; see, e.g., Theorem 11.10 in [Stokey and Lucas \(1989\)](#). Due to the expectations feedback, this measure depends in general on the belief parameters α and β . Under the assumptions referred to above, it can be shown that the time average of any continuous function along a trajectory of (13) converges with probability one to the corresponding space average with respect to the unique invariant probability measure μ ; see, e.g., Theorem 14.7 in [Stokey and Lucas](#). In particular, for every initial prices p_0 , the condition

$$\bar{p}(p_0) = \lim_{T \rightarrow +\infty} \frac{1}{T+1} \sum_{t=0}^T p_t = \int p \, d\mu(p)$$

holds with probability 1. In other words, the sample average of the price sequence generated by (13) coincides with the expected value of the invariant probability measure μ . Since this result implies that $\bar{p}(p_0)$ is independent of p_0 , we shall henceforth simplify the notation by writing $\bar{p}$ instead of $\bar{p}(p_0)$. Analogously to the above equation, for every initial price p_0 and every $j \geq 0$, the condition

$$\lim_{T \rightarrow +\infty} \frac{1}{T+1} \sum_{t=0}^{T-j} (p_t - \bar{p})(p_{t+j} - \bar{p}) = \iint (p' - \bar{p})(p - \bar{p}) Q_{\alpha, \beta}^{(j)}(p, dp') \, d\mu(p)$$

holds with probability 1. Using these results, one can see that the following definition is a direct translation of Definition 1 of CEE into a stochastic framework.

Definition 3. A triple $\{\mu; \alpha, \beta\}$, where μ is a probability measure and α and β are real numbers with $\beta \in [-1, 1]$, is called a *stochastic consistent expectations equilibrium* (SCEE) if

- S1 the probability measure μ is a non-degenerate invariant measure for the stochastic difference equation (13),⁶
- S2 the stationary stochastic process defined by (13) and the invariant measure μ has unconditional mean α , that is, $\mathbb{E}_\mu(p) = \int p \, d\mu(p) = \alpha$, and
- S3 the stationary stochastic process defined by (13) and the invariant measure μ has unconditional autocorrelation coefficients β^j , that is

$$\frac{\iint (p' - \alpha)(p - \alpha) Q_{\alpha, \beta}^{(j)}(p, dp') \, d\mu(p)}{\int (p - \alpha)^2 \, d\mu(p)} = \beta^j$$

holds for all $j \geq 1$.

If conditions S1 and S2 are satisfied and the equation in S3 holds for all $j \in \{1, 2, \dots, k\}$, then we call $\{\mu; \alpha, \beta\}$ a *k-th order SCEE*.

According to Definition 3, an SCEE is characterized by the fact that both the unconditional mean and the autocorrelation coefficients generated by the unknown nonlinear stochastic process (13) and the invariant measure μ coincide with the corresponding values for the perceived linear AR(1) process. Stated differently, an SCEE is self-fulfilling in terms of the unconditional mean and the autocorrelation coefficients. The SCEE concept is a natural generalization of the deterministic CEE concept, obtained by replacing the sample mean and sample autocorrelation

⁶By ‘non-degenerate’ we mean that it has positive variance.

coefficients by the corresponding mean and autocorrelation coefficients of the stochastic difference equation (13) under the invariant measure μ .

Note that in analogy to the case of deterministic CEE we allow β to take the values -1 and 1 although the interpretation of an SCEE as a fixed point of the expectations feedback is not correct for these cases. As for condition S3 of Definition 3, it is worth pointing out that there is no need to distinguish between convergent and nonconvergent sequences (as it was the case in the corresponding condition C3 of Definition 1), because of the non-degeneracy of μ assumed in S1. In other words, in an SCEE, the implied actual law of motion is a stationary stochastic process with nonzero variance and, therefore, autocorrelation coefficients are well defined.

2.4 SAC-Learning

In the definitions of CEE and SCEE, agents' beliefs are described by the forecasting rule (10) with *fixed* parameters α and β . In this subsection we consider a natural extension where the agents change their forecasting rule by updating their belief parameters as additional observations become available. In other words, agents' beliefs change over time and agents use an adaptive learning algorithm. Note that, although the belief parameters change over time, the agents' forecasting rule always belongs to the class of linear rules derived from a perceived law of motion given by an AR(1) model. The adaptive learning algorithm tries to find an optimal or consistent linear forecasting rule. A natural learning scheme fitting the SCEE framework is based upon sample averages and sample autocorrelation coefficients.

For any finite set of observations $\{p_0, p_1, \dots, p_t\}$, the sample average is given by

$$\alpha_t = \frac{1}{t+1} \sum_{i=0}^t p_i, \quad (14)$$

and the first order sample autocorrelation coefficient is given by

$$\beta_t = \frac{\sum_{i=0}^{t-1} (p_i - \alpha_t)(p_{i+1} - \alpha_t)}{\sum_{i=0}^t (p_i - \alpha_t)^2}. \quad (15)$$

When the agents update the belief parameters in every period according to the most recent observations of the sample average and the first order sample autocorrelation coefficient, the *temporary law of motion* for the OG-type model (12) becomes

$$p_t = F_{\alpha_{t-1}, \beta_{t-1}}(p_{t-1}, \eta_t) = F(\alpha_{t-1} + \beta_{t-1}^2(p_{t-1} - \alpha_{t-1}), \eta_t). \quad (16)$$

Notice that the 2-period ahead SAC-forecast $p_{t+1}^e = \alpha_{t-1} + \beta_{t-1}^2(p_{t-1} - \alpha_{t-1})$ uses only price observations for periods $i \leq t-1$. We will refer to the dynamical

system (14)–(16) as *sample autocorrelation learning* (SAC-learning). The initial state for the system (14)–(16) can be any triple (p_0, α_0, β_0) with $\beta_0 \in [-1, 1]$. It is well known (and easy to check) that, independently of the choice of these initial values, it always holds that $\beta_1 = -1/2$, and that the first order sample autocorrelation $\beta_t \in [-1, 1]$ for all $t \geq 1$.

The SAC-learning process is closely related to, but not identical to OLS-learning.⁷ If the SAC-estimate for β_t and the OLS-estimate for β_t are close in the initial phase of the learning process then, assuming that p_t is bounded, they are likely to remain close to each other and converge to the same limit in the long run. However, in the initial phase of the learning process, the SAC-estimate and the OLS-estimate may differ from each other and, in general, these differences lead to different outcomes in the long run. SAC-learning has an important advantage over OLS-learning. The SAC-estimate (15) satisfies $\beta_t \in [-1, 1]$ for all $t \geq 1$. In contrast, the OLS-estimate for β_t may be outside the interval $[-1, 1]$, which can destabilize the learning process in its initial stage. Because of this reason, [Marcet and Sargent \(1989\)](#) imposed a so-called ‘projection facility’, that is, an interval $[-\mu_1, \mu_2]$ of allowable values for the OLS-estimate β_t . There is no need for such an artificial device in the SAC-learning process.

The SAC-learning scheme naturally fits our framework since traders believe in an AR(1) process with stationary first and second moments. In the analysis of adaptive learning in the OG-type model in Sect. 5, we shall therefore focus on SAC-learning. Simulations with OLS-learning often (but not always) yield similar results in the long run.

The SAC-learning dynamics as formulated in (14)–(16) is a high-dimensional, nonlinear, and non-autonomous system. In the appendix it is shown that one can rewrite the system in a recursive form as a 4-dimensional non-autonomous stochastic difference equation. But there is a much simpler and quite intuitive way to think about the SAC-learning dynamics. Since p_t in (16) is a stochastic variable, the belief parameters α_t and β_t are also stochastic variables. The role of the state variable p_t , however, is quite different from the role of the belief parameters α_t and β_t , both mathematically and economically. In particular, if the price p_t remains bounded, then both the changes in α_t and the changes in β_t become small as time t approaches $+\infty$. When prices are bounded, the belief parameters α_t and β_t are thus ‘slow’ variables in the SAC-learning dynamics. After some (short) transient phase, the SAC-learning process may thus be viewed as a price generating system with a slowly changing temporary law of motion. In the long run, the time averages α_t and β_t will usually converge and the temporary law of motion will therefore settle down to some stationary limiting law of motion. If the limits $\alpha = \lim_{t \rightarrow +\infty} \alpha_t$ and $\beta = \lim_{t \rightarrow +\infty} \beta_t$ are such that there exists a probability measure μ such that $(\mu; \alpha, \beta)$ is an SCCE, then this would mean that the agents *learn to believe in an*

⁷The OLS-estimate for α is identical to (14). The OLS-estimate for β is slightly different from (15), namely $\beta_{t-1} = [\sum_{i=0}^{t-2} (p_i - \bar{p}_{t-1}^-)(p_{i+1} - \bar{p}_{t-1}^+)] / [\sum_{i=0}^{t-2} (p_i - \bar{p}_{t-1}^-)^2]$ for $t \geq 3$, where $\bar{p}_{t-1}^- = [1/(t-1)] \sum_{i=0}^{t-2} p_i$ and $\bar{p}_{t-1}^+ = [1/(t-1)] \sum_{i=1}^{t-1} p_i$.

SCEE. Note that the (nonlinear) limiting law of motion $F_{\alpha,\beta}$ remains unknown to the agents and that, in general, it does not coincide with the limiting perceived law of motion or its corresponding linear forecasting rule $p_{t+1}^e = \alpha + \beta^2(p_{t-1} - \alpha)$. However, the observable sample mean and sample auto-covariances of the unknown limiting law of motion and the limiting perceived law of motion coincide. It will be proved in the following section (Theorem 6) that the limiting forecasting errors are unbiased and uncorrelated. According to this scenario, *agents are thus learning to believe in linearity in an unknown and nonlinear stochastic world*.

In Sect. 5 we will study SAC-learning in a stochastic OG-model and we will see that the limiting law of motion is in general not unique, but that the SAC-learning process exhibits *path dependence* and may converge to different limiting laws of motion depending on the initial values (p_0, α_0, β_0) and/or the realizations of the stochastic process $(\eta_t)_{t=0}^{+\infty}$.

3 Existence and Properties of CEE

This section presents analytical results concerning deterministic CEE. We start by discussing existence of simple (steady state and 2-cycle) CEE, thereafter consider existence of chaotic CEE and finally discuss the relation between CEE and REE.

3.1 Simple Examples of CEE

In this subsection we investigate existence of steady state CEE and 2-cycle CEE. We will argue that for monotonic maps, either increasing or decreasing, these will be the only deterministic CEE that can arise.

Let G be the law of motion (8). Recall that a fixed point of G is a number p satisfying $G(p)=p$, whereas a point with (minimum) period 2 satisfies $G(G(p)) = p \neq G(p)$. Note that whenever p_1^* is a point of period 2 then so is $p_2^* = G(p_1^*)$. Thus, points of period 2 always come in pairs $\{p_1^*, p_2^*\}$, such that $p_1^* \neq p_2^*, p_2^* = G(p_1^*)$, and $p_1^* = G(p_2^*)$. The pair $\{p_1^*, p_2^*\}$ is also called a 2-cycle of G . We are now ready to define the special types of CEE that we are going to discuss:

- *Steady state CEE* in which the price sequence $(p_t)_{t=0}^{+\infty}$ converges to a fixed point of G .
- *2-cycle CEE* in which the price sequence $(p_t)_{t=0}^{+\infty}$ converges to a 2-cycle of G .⁸
- *Chaotic CEE* in which the price sequence $(p_t)_{t=0}^{+\infty}$ is chaotic.

⁸Here we mean orbital convergence, that is, the existence of a 2-cycle $\{p_1^*, p_2^*\}$ such that $\lim_{t \rightarrow +\infty} p_{2t} = p_1^*$ and $\lim_{t \rightarrow +\infty} p_{2t+1} = p_2^*$, or vice versa. Since $p_1^* \neq p_2^*$ the sequence $(p_t)_{t=0}^{+\infty}$ is not convergent in the usual sense.

The following two theorems clarify the relation between steady state CEE and 2-cycle CEE and fixed points and 2-cycles, respectively, of the function G .

- Theorem 1.** 1. If $\{(p_t)_{t=0}^{+\infty}; \alpha, \beta\}$ is a first order steady state CEE converging to p^* , then $\alpha = p^*$ and p^* is a fixed point of G .
 2. If $\{(p_t)_{t=0}^{+\infty}; \alpha, \beta\}$ is a first order 2-cycle CEE converging to $\{p_1^*, p_2^*\}$, then $\alpha = (p_1^* + p_2^*)/2$ and $\beta = -1$. Furthermore, $\{p_1^*, p_2^*\}$ is a 2-cycle of G .

Proof. 1. Whenever $\lim_{t \rightarrow +\infty} p_t$ exists, the sample average $\bar{p}(p_0)$ must be equal to this limit. Together with the consistency requirement C2 of Definition 1 this implies $\alpha = p^*$. The implied actual price dynamics (11) is therefore $p_t = G(p^* + \beta^2(p_{t-1} - p^*))$. Since p^* must be a steady state of this difference equation it follows that p^* is a fixed point of G .

2. Analogously to case 1 it can be seen that $\alpha = \bar{p}(p_0) = (p_1^* + p_2^*)/2$. Using this it is also straightforward to show that $\lim_{T \rightarrow +\infty} c_{0,T}(p_0) = -\lim_{T \rightarrow +\infty} c_{1,T}(p_0) = (p_1^* - p_2^*)^2/4$. This implies $\rho_1(p_0) = -1$ and it follows from the consistency requirement C3(b) of Definition 1 that $\beta = -1$. The implied actual law of motion is therefore given by $p_t = G(p_{t-1})$, and $\{p_1^*, p_2^*\}$ must be a 2-cycle of G . This concludes the proof of the theorem. $\square$

The above theorem shows that, along a first order steady state CEE, prices will converge to a fixed point of G whereas, along a first order 2-cycle CEE, prices will converge to a 2-cycle of G . The latter result is in contrast to a similar theorem in Hommes and Sorger (1998), where it was shown that 2-cycle CEE converge to a pair of fixed points of G . The difference between these results is due to the differences between the cobweb-type model studied in Hommes and Sorger (1998) and the OG-type model of the present paper. Note furthermore that the results in Theorem 1 are stated for first order CEE. Consequently, they are also true for the stronger equilibrium concepts of CEE and k -th order CEE with $k > 1$.

The following result is sort of a converse to Theorem 1. It shows that one can always construct CEE when fixed points or 2-cycles of G are known. Note that the theorem yields CEE and not only k -th order CEE.

- Theorem 2.** 1. Suppose that p^* is a fixed point of the map G and define $p_t = p^*$, for all t , $\alpha = p^*$, and $\beta = 0$. Then $\{(p_t)_{t=0}^{+\infty}; \alpha, \beta\}$ is a steady state CEE.
 2. Suppose that $\{p_1^*, p_2^*\}$ is a 2-cycle of the map G and define $p_{2t-1} = p_1^*$ and $p_{2t} = p_2^*$ for all $t \geq 0$, $\alpha = (p_1^* + p_2^*)/2$, and $\beta = -1$. Then $\{(p_t)_{t=0}^{+\infty}; \alpha, \beta\}$ is a 2-cycle CEE.

Proof. 1. If $\alpha = p^*$ is a fixed point of G then the constant sequence $(p^*, p^*, p^*, \dots)$ satisfies the implied actual law of motion (11) so that condition C1 in Definition 1 is satisfied. Conditions C2 and C3 hold trivially.

2. If $\alpha = (p_1^* + p_2^*)/2$ and $\beta = -1$, where $\{p_1^*, p_2^*\}$ is a 2-cycle of G , then the periodic sequence $(p_1^*, p_2^*, p_1^*, p_2^*, \dots)$ satisfies the implied actual law of motion (11) so that condition C1 in Definition 1 is satisfied. Conditions C2 and C3 are again easily verified. $\square$

We call a (k -th order) CEE $\{(p_t)_{t=0}^{+\infty}; \alpha, \beta\}$ *bounded* if the sequence $(p_t)_{t=0}^{+\infty}$ is bounded. Our next theorem proves that in the case where G is a monotonic function, all bounded CEE of model (8) are either steady state CEE or 2-cycle CEE.

Theorem 3. *Let $\{(p_t)_{t=0}^{+\infty}; \alpha, \beta\}$ be a bounded CEE or a bounded k -th order CEE, where k is a positive integer.*

1. *If the map G is nondecreasing then it follows that $(p_t)_{t=0}^{+\infty}$ converges to a steady state and that β is nonnegative.*
2. *If the map G is non-increasing then it follows that $(p_t)_{t=0}^{+\infty}$ converges either to a steady state or to a 2-cycle. In the case of convergence to a steady state the parameter β is non-positive, and in the case of convergence to a 2-cycle CEE it holds that $\beta = -1$.*

Proof. 1. If G is nondecreasing the actual law of motion $G_{\alpha, \beta}$ in (11) must be nondecreasing as well. It follows therefore that the sequence $(p_t)_{t=0}^{+\infty}$ must be monotonic and, since it is bounded, it must converge to a steady state p^* . Moreover, because $(p_t)_{t=0}^{+\infty}$ is monotonic the first order sample autocorrelation coefficient has to be nonnegative which implies that β must be nonnegative. This concludes the proof of part 1.

2. Now suppose that G is non-increasing. In that case the implied actual law of motion $G_{\alpha, \beta}$ is non-increasing as well which, in turn, implies that the sequence $(p_t)_{t=0}^{+\infty}$ either converges to a steady state or to a 2-cycle. In the case of convergence to a steady state, convergence takes place in an oscillatory way with prices being below the steady state in one period and above it in the next period. This implies that the first order sample autocorrelation coefficient is non-positive which, in turn, shows that β must be non-positive. In the case of convergence to a 2-cycle CEE, consistency requirement C3(b) implies that $\beta = -1$. $\square$

3.2 Chaotic CEE

As long as the function G is monotonic, Theorem 3 shows that only steady state CEE or 2-cycle CEE can arise. If G is non-monotonic, on the other hand, much more complicated CEE can arise. We will now demonstrate this by using examples related to the tent map introduced in Sect. 2.1. The following lemma from [Hommes and Sorger \(1998\)](#) will be useful for the construction of chaotic CEE.

Lemma 1. *Let real numbers a, b, γ , and δ be given such that $a < b$, $\gamma \neq 0$, and $\delta \in (-1, 1)$. Furthermore, define $\alpha = (a + b)/2$. Then there exist real numbers A, B, C, D with $A > 0$ and $C > 0$ such that the continuous and piecewise linear map $G : \mathbb{R} \rightarrow \mathbb{R}$ defined by*

$$G(x) = \begin{cases} Ax - B & \text{if } x \leq (B + D)/(A + C) \\ -Cx + D & \text{if } x \geq (B + D)/(A + C) \end{cases}$$

satisfies $G(\alpha + \gamma(x - \alpha)) = T_{\delta, [a, b]}(x)$ for all $x \in [a, b]$.

As an immediate consequence of this lemma we obtain the following result.

Theorem 4. *Let real numbers α and β be given such that $\beta \in (-1, 1)$ and $\beta \neq 0$.⁹ There exists a continuous and piecewise linear function G such that the OG-type model (8) has an infinite set (of positive Lebesgue measure) of chaotic CEE $\{(p_t)_{t=0}^{+\infty}; \alpha, \beta\}$.*

Proof. Choose any real numbers a and b such that $a < b$ and $(a + b)/2 = \alpha$. Applying Lemma 1 with $\gamma = \beta^2$ and $\delta = \beta$ it follows that there exists a piecewise linear function G such that $G_{\alpha, \beta}(p) = G(\alpha + \beta^2(p - \alpha)) = T_{\beta, [a, b]}(p)$ holds for all $p \in [a, b]$. The implied actual law of motion (11) therefore coincides with the tent map dynamics (4). According to Sect. 2.1, the sample average and sample autocorrelation coefficients of the trajectory $(p_t)_{t=0}^{+\infty}$ are given by $\bar{p}(p_0) = (a + b)/2 = \alpha$ and $\rho_j(p_0) = \beta^j$ for Lebesgue-almost all initial states $p_0 \in [a, b]$. Thus, all conditions of Definition 1 are satisfied. Moreover, it is well known that the trajectories of (4) are chaotic for Lebesgue-almost all initial states $p_0 \in [a, b]$. This concludes the proof of the theorem. $\square$

3.3 Relation Between CEE and REE

An important question is what is the relation between CEE and REE? Let us start with the following observation for the deterministic model (8): if p^* is a fixed point of G and $p_t = p^*$ for all t , then $(p_t)_{t=0}^{+\infty}$ is an REE.¹⁰ This is quite obvious because in that case $(p_t)_{t=0}^{+\infty}$ satisfies equation (8) with $p_t^e = p_t$ for all t . An analogous argument shows that 2-cycles of G also qualify as perfect foresight equilibria. More specifically, if $\{p_1^*, p_2^*\}$ is a 2-cycle for G and $p_{2t-1} = p_1^*$ and $p_{2t} = p_2^*$ for all t , then $(p_t)_{t=0}^{+\infty}$ satisfies (8) with $p_t^e = p_t$ for all t . From the results presented in the previous subsection we can therefore conclude that every first order steady state CEE or first order 2-cycle CEE has the property that it converges to a perfect foresight equilibrium. Interestingly, these perfect foresight equilibria can be supported by the simple linear forecasting rule (10) provided the belief parameters α and β are chosen appropriately.

In contrast to the steady state CEE and 2-cycle CEE, chaotic CEE do not approach perfect foresight equilibria. To see this just note that $p_{t+1}^e = p_{t+1}$ and (10) implies that $p_{t+1} = \alpha + \beta^2(p_{t-1} - \alpha)$. Since this is a linear difference equation with nondecreasing right-hand side, the two sequences $(p_{2t-1})_{t=0}^{+\infty}$ and $(p_{2t})_{t=0}^{+\infty}$ must be monotonic and convergent. This, in turn, rules out that $(p_t)_{t=0}^{+\infty}$ is a chaotic time series.

⁹The case $\beta = 0$ must be excluded since for $\beta = 0$ there is no dynamics in (11). Sorger (1998) presents an example of an OG-model of a more general form than (8) for which a chaotic CEE with $\beta = 0$ exists.

¹⁰Because the model is deterministic, REE are equivalent to perfect foresight equilibria.

The existence of chaotic CEE is caused by the similarity of chaotic processes to stochastic processes – in the present example the similarity between the dynamics generated by asymmetric tent maps and stochastic AR(1) processes. Agents using only linear statistical techniques cannot distinguish between these processes, and there is no way they could ever detect that they are constantly making forecasting mistakes. Those mistakes are self-fulfilling because they are fully consistent with the agents' model of the world; see [Grandmont \(1998\)](#). In fact, forecasting errors along a chaotic CEE are unbiased and uncorrelated. We will not give an independent proof of this result, because it essentially follows from a similar result for SCEE in the next section.

4 Existence and Properties of SCEE

In this section we return to the stochastic case, with the economy described by the nonlinear stochastic law of motion (12). We first discuss the relation between an SCEE and an REE. We will argue that the simplest form of an SCEE, a steady state SCEE, coincides with an REE. In contrast, a non-steady state SCEE is generally *not* an REE. However, our main theorem states that forecasting errors along an SCEE are always *unbiased* and *uncorrelated*. In Sect. 4.2 we discuss first order non-steady state SCEE and prove their existence under fairly general conditions.

4.1 The Relation Between SCEE and REE

In this subsection, we restrict ourselves to the important case of steady state SCEE for the stochastic law of motion (12).

Theorem 5. *Let p^* be a fixed point of the mapping $p \mapsto \mathbb{E}_\eta F(p, \eta_t)$, where the expectation is taken with respect to the (time-invariant) distribution of the shocks η_t . Let μ be the distribution of $F(p^*, \eta_t)$ and define $\alpha = p^*$ and $\beta = 0$. Then it follows that $(\mu; \alpha, \beta)$ is an SCEE.*

Proof. If $\alpha = p^*$ and $\beta = 0$, the implied actual law of motion (13) becomes

$$p_t = F(p^*, \eta_t). \quad (17)$$

Note that the right-hand side of this equation does not depend on p_{t-1} such that the corresponding transition function $Q_{p^*, 0}(p, A)$ is independent of p . Obviously, it holds that $\mu(A) = Q_{p^*, 0}(p, A)$ for all measurable sets A . It is straightforward to see that μ is the unique invariant probability measure for (17). The definition of p^* together with (17) implies that $\mathbb{E}_\mu(p_t) = \mathbb{E}_\eta F(p^*, \eta_t) = p^*$. Finally, since η_t and η_s are independent random variables whenever $t \neq s$ it follows from (17) that p_t and p_s are independent as well and therefore $\rho_j = 0$ for all $j \geq 1$. These properties imply that the triple $(\mu; p^*, 0)$ satisfies the conditions of an SCEE. $\square$

We will refer to an SCEE of the form described in the above theorem as a *steady state SCEE*. Notice that in the case of additive noise, i.e. $p_t = F(p_{t+1}^e) + \eta_t$, with η_t IID noise with zero mean, a fixed point of the map $p \mapsto \mathbb{E}_\eta F(p, \eta_t)$ coincides with a fixed point of the map F . In the next subsection we will discuss existence of non-steady state SCEE. But first we will argue that, along any SCEE, prediction errors of the linear forecasting rules are unbiased and uncorrelated.

The actual law of motion (13) implies that

$$E_{t-1}(p_{t+1}) = E_{t-1}[F_{\alpha,\beta}(F_{\alpha,\beta}(p_{t-1}, \eta_t), \eta_{t+1})] = \iint F_{\alpha,\beta}(F_{\alpha,\beta}(p_{t-1}, \eta), \eta') d\eta d\eta'. \quad (18)$$

In general, there do not exist real numbers α and $\beta \in [-1, 1]$ such that this conditional expectation coincides with the linear forecasting rule $p_{t+1}^e = \alpha + \beta^2(p_{t-1} - \alpha)$. This suggests that an SCEE is in general not an REE. It is interesting to point out, however, that a steady state SCEE always qualifies as an REE. To see this, remember that in a steady state SCEE $(\mu; \alpha, \beta)$ we have $\beta = 0$ and $\alpha = p^*$, where p^* is a fixed point of the mapping $p \mapsto \mathbb{E}_\eta F(p, \eta_t)$. It follows therefore that $F_{\alpha,\beta}(p, \eta) = F(p^*, \eta)$ independently of p . Using this result in (18) we get

$$E_{t-1}(p_{t+1}) = \iint F(p^*, \eta') d\eta d\eta' = \int F(p^*, \eta') d\eta' = p^*.$$

Since $\alpha + \beta^2(p_{t-1} - \alpha) = p^*$, it follows that the agents' forecast of p_{t+1} coincides with the conditional expectation of p_{t+1} and, hence, that a steady state SCEE qualifies as an REE.

Our next step is to show that, along any SCEE, forecasting errors are unbiased and uncorrelated. To this end, recall that the agents in the model use the 2-period ahead forecast $p_{t+1}^e = \alpha + \beta^2(p_{t-1} - \alpha)$; see (10). The forecast error in period t is therefore given by $\epsilon_t = p_t - \alpha - \beta^2(p_{t-2} - \alpha)$. We have the following theorem.

Theorem 6 (unbiased and uncorrelated forecasting errors). *Let $(\mu; \alpha, \beta)$ be an SCEE for the stochastic expectations feedback system (12) and let $(\epsilon_t)_{t=2}^{+\infty}$ be the corresponding process of forecasting errors. Then it holds that $\mathbb{E}_\mu(\epsilon_t) = 0$ for all $t \geq 2$ and $\mathbb{E}_\mu(\epsilon_t \epsilon_{t+j}) = 0$ for all $t \geq 2$ and all $j \geq 2$. Moreover, there exists a constant $\sigma^2 > 0$ such that $\mathbb{E}_\mu(\epsilon_t \epsilon_{t+1}) = \beta(1 - \beta^2)\sigma^2$ for all $t \geq 2$.*

Proof. From conditions S1 and S2 in Definition 3 we have $\mathbb{E}_\mu(p_t) = \mathbb{E}_\mu(p_{t-2}) = \alpha$ and therefore $\mathbb{E}_\mu(\epsilon_t) = 0$. To derive the uncorrelatedness of forecast errors, first note that for all $j \geq 2$

$$\begin{aligned} \epsilon_t \epsilon_{t+j} &= [p_t - \alpha - \beta^2(p_{t-2} - \alpha)][p_{t+j} - \alpha - \beta^2(p_{t+j-2} - \alpha)] \\ &= (p_t - \alpha)(p_{t+j} - \alpha) - \beta^2(p_{t-2} - \alpha)(p_{t+j} - \alpha) - \beta^2(p_t - \alpha)(p_{t+j-2} - \alpha) \\ &\quad + \beta^4(p_{t-2} - \alpha)(p_{t+j-2} - \alpha). \end{aligned} \quad (19)$$

Let us define σ^2 by $\sigma^2 = \int (p - \alpha)^2 d\mu(p)$. Because of condition S1 we have $\sigma^2 > 0$. According to condition S3 of Definition 3 we have

$$\begin{aligned}
\mathbb{E}_\mu(p_t - \alpha)(p_{t+j} - \alpha) &= \mathbb{E}_\mu(p_{t-2} - \alpha)(p_{t+j-2} - \alpha) \\
&= \iint (p' - \alpha)(p - \alpha) Q_{\alpha, \beta}^{(j)}(p, dp') d\mu(p) = \beta^j \sigma^2, \\
\mathbb{E}_\mu(p_{t-2} - \alpha)(p_{t+j} - \alpha) &= \iint (p' - \alpha)(p - \alpha) Q_{\alpha, \beta}^{(j+2)}(p, dp') d\mu(p) = \beta^{j+2} \sigma^2, \\
\mathbb{E}_\mu(p_t - \alpha)(p_{t+j-2} - \alpha) &= \iint (p' - \alpha)(p - \alpha) Q_{\alpha, \beta}^{(j-2)}(p, dp') d\mu(p) = \beta^{j-2} \sigma^2.
\end{aligned}$$

Together with (19) this shows that

$$\mathbb{E}_\mu(\epsilon_t \epsilon_{t+j}) = (\beta^j - \beta^2 \beta^{j+2} - \beta^2 \beta^{j-2} + \beta^4 \beta^j) \sigma^2 = 0.$$

For $j = 1$, computations analogous to those above yield

$$\begin{aligned}
\epsilon_t \epsilon_{t+1} &= [p_t - \alpha - \beta^2(p_{t-2} - \alpha)][p_{t+1} - \alpha - \beta^2(p_{t-1} - \alpha)] \\
&= (p_t - \alpha)(p_{t+1} - \alpha) - \beta^2(p_{t-2} - \alpha)(p_{t+1} - \alpha) - \beta^2(p_t - \alpha)(p_{t-1} - \alpha) \\
&\quad + \beta^4(p_{t-2} - \alpha)(p_{t-1} - \alpha)
\end{aligned}$$

and, therefore,

$$\mathbb{E}_\mu(\epsilon_t \epsilon_{t+1}) = (\beta - \beta^2 \beta^3 - \beta^2 \beta + \beta^4 \beta) \sigma^2 = \beta(1 - \beta^2) \sigma^2.$$

This completes the proof of the theorem. $\square$

The above theorem shows that, along an SCEE, forecast errors have the mean 0 and all autocorrelation coefficients of lags $j \geq 2$ are 0. In other words, along an SCEE, agents do not make any *systematic* mistakes. This is true despite the fact that the linear forecasting rule is misspecified. According to the last statement of the theorem, the *first order* autocorrelation of forecast errors is 0 if and only if $\beta \in \{-1, 0, 1\}$. In general, it can be nonzero due to the assumed information structure in the OG-type model. More precisely, the autocorrelation between the forecasting errors ϵ_t and ϵ_{t+1} cannot be exploited by the agents, because neither the price p_t nor the forecast error ϵ_t are known when the forecast for p_{t+1} has to be made. However, the first order autocorrelation coefficient of forecasting errors is 0 in a steady state SCEE, because in that case $\beta = 0$.

Theorem 6 has been stated for SCEE (and holds also for nonconvergent CEE). There is also an analogous result for k -th order SCEE. As a matter of fact, following the same proof as in Theorem 6, one can show that, along a k -th order SCEE, it holds that $\mathbb{E}_\mu(\epsilon_t \epsilon_{t+j}) = 0$ for all $t \geq 2$ and all $j \in \{2, 3, \dots, k-2\}$.

In an REE, agents have the correct conditional expectation, which requires perfect knowledge of the underlying nonlinear function F or G , respectively. In an SCEE, on the other hand, agents do not know the true conditional expectation, but they do have knowledge about the mean and the autococorrelation coefficients

of the unknown data generating mechanism. It is important to note that the mean and the autocorrelation coefficients can be inferred from time series observations without any knowledge of the true underlying law of motion, as has been discussed in Sect. 2.4. An SCEE may therefore be seen as an approximate rational expectations equilibrium, i.e., a situation in which expectations are selffulfilling in terms of the first two moments of the linear perceived law of motion and the unknown and nonlinear actual law of motion.

4.2 First Order SCEE

In this subsection we prove existence of non-steady state first order SCEE. These simple types of first order SCEE are of special interest because, if the SAC-learning process converges, it must converge to an SCEE at least of first order. Our examples will all be of the form “deterministic dynamics plus additive noise”, that is, of the form

$$p_t = F(p_{t+1}^e) + \eta_t, \quad (20)$$

where η_t is a white noise process as before. Branch and McGough (2005) have proven existence of non-REE first order SCEE for the case where the map F is increasing and symmetric with respect to the origin (such that F has a fixed point at the origin), and where some other regularity conditions are met. Branch and McGough apply the stochastic framework described in Stokey and Lucas (1989) and point out that this approach does not work for the case where F is decreasing. We will give a direct proof, both for the increasing and the decreasing case, and our conditions will also be more general.

Our motivating example is the family of processes

$$p_t = \tanh(\nu p_{t-1}) + \eta_t \quad (21)$$

for $\nu \in \mathbb{R}$. Note that the function $p \mapsto t_\nu(p) = \tanh(\nu p)$ is monotonically increasing for $\nu > 0$, and decreasing for $\nu < 0$; the cases correspond to positive and negative feedback respectively. Though our results will be formulated for general processes, we shall indicate the consequences for this ‘motivating family’.

In analogy to (13) we write $F_{\alpha,\beta}(p) = F(\alpha + \beta^2(p - \alpha))$. Let $Q_{\alpha,\beta}$ denote the transition function of the process $p_t = F_{\alpha,\beta}(p_t) + \eta_t$. In terms of the density function φ of the noise process $(\eta_t)_{t=0}^{+\infty}$, the transition function can be written as

$$Q_{\alpha,\beta}(p, A) = \int_A \varphi(p' - F_{\alpha,\beta}(p)) \, dp'.$$

We call a function monotone if either $x \leq y$ implies $f(x) \leq f(y)$ for all $x, y \in \mathbb{R}$, or if $x \leq y$ implies $f(x) \geq f(y)$ for all $x, y \in \mathbb{R}$. We make the following assumptions:

A1 (assumptions on the noise process)

The density function $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ of the white noise process $(\eta_t)_{t=0}^{+\infty}$ is continuous and satisfies $\varphi(-p) = \varphi(p)$ for all p . Moreover, φ is monotone decreasing on the positive real axis.

A2 (assumptions on the map F)

- (i) The map $F: \mathbb{R} \rightarrow \mathbb{R}$ is a monotone function.
- (ii) The function F is odd, that is $F(-p) = -F(p)$ for all $p \in \mathbb{R}$.
- (iii) There exist constants a and b satisfying $0 < a < 1$ and $b \geq 0$ such that $|F(p)| \leq a|p| + b$ for all p .

A3 (ergodicity assumption)

The process $(p_t)_{t=0}^{+\infty}$ generated by equation

$$p_t = F_{0,\beta}(p_{t-1}) + \eta_t$$

has a unique invariant probability distribution for every β .

The symmetry assumptions in A1 and A2 simplify the proofs considerably. Assumption A2(ii) implies that F has a fixed point at 0.¹¹ Let $\sigma^2 = \mathbb{V}\text{ar } \eta_t$. From the stated properties of η_t it follows that $\mathbb{E}\eta_t = 0$ and $\varphi(p) \leq \frac{3}{2}\sigma^2/p^3$; the latter property is a consequence of

$$\sigma^2 = \int t^2 \varphi(t) dt \geq \int_{-p}^p t^2 \varphi(p) dt = \frac{2}{3} p^3 \varphi(p). \quad (22)$$

Assumption A2(iii) essentially says that F is contracting for large arguments. Assumption A3 avoids certain technical issues. Note that it is satisfied if the support of the density function φ is $\mathbb{R}$. Thus, it holds for instance if η_t is Gaussian.

We now want to study stochastic consistent expectations equilibria for the model (20). Because of the assumed symmetry of F , we can fix the constant α in the linear AR(1) forecasting rule to $\alpha = 0$, so that the implied actual law of motion becomes

$$p_t = F_{0,\beta}(p_{t-1}) + \eta_t = F(\beta^2 p_{t-1}) + \eta_t. \quad (23)$$

Using the invariant measure μ , the existence of which was assumed in A3, the unconditional first-order autocorrelation coefficient of (23) can be computed. Let us denote this coefficient by $\rho_1(F_{0,\beta})$. If $(\mu; 0, \beta)$ is a first order SCEE, then it must hold that $\beta = \rho_1(F_{0,\beta})$. The first order SCEE is said to be *nontrivial* if $\beta \neq 0$.

¹¹We have chosen F to be symmetric around 0 but, without loss of generality, we could have chosen F to be symmetric around any fixed point α .

4.2.1 Impossibility of Nontrivial First Order SCEE

The following theorem says that if the map F has a fixed point at 0 and if it is contracting then there does not exist a nontrivial (first order) SCEE for the process (20).

Theorem 7. *Assume that there exists a real number λ such that $0 < \lambda < 1$ and $|F(p)| \leq \lambda|p|$ for all p . Then it follows that the process (20) has no nontrivial (first order) SCEE.*

Note that, according to this theorem, our motivating family $p_t = \tanh(\nu p_{t-1}) + \eta_t$ has no non-trivial SCEE if $|\nu| < 1$.

Proof. Suppose that there exists a non-trivial SCEE $(\mu; 0, \beta)$ and let $Q_{0,\beta}$ be the transition function corresponding to (23). Under the stated assumptions it holds that

$$\left| \iint p p' Q_{0,\beta}(p, dp') d\mu(p) \right| = \left| \int p F(\beta^2 p) d\mu(p) \right| \leq \lambda \beta^2 \int p^2 d\mu(p)$$

and consequently that $|\rho_1(F_{0,\beta})| \leq \lambda \beta^2$. In order to satisfy $\rho_1(F_{0,\beta}) = \beta$, it must necessarily hold that $|\beta| \leq 1$. Therefore, it follows that $|\lambda \beta| < 1$ and, consequently,

$$|\rho_1(F_{0,\beta})| \leq \lambda \beta^2 = |\lambda \beta| |\beta| < |\beta|$$

whenever $\beta \neq 0$. Hence, the equation $\rho_1(F_{0,\beta}) = \beta$ does not have a solution $\beta \neq 0$. This is a contradiction to the assumption that there exists a non-trivial SCEE. $\square$

4.2.2 First Order SCEE for Arbitrary Noise Levels

Next we investigate conditions for which a nontrivial first order SCEE exists. The first result is that model (20) has a nontrivial first order SCEE if the function F is sufficiently expanding at the origin.

Theorem 8. *Suppose assumptions A1–A3 hold. There exists $\lambda_0 > 1$ such that the following is true. If $c > 0$ and $\lambda > \lambda_0$ are such that $|F(p)| > \lambda|p|$ holds for all p satisfying $0 < |p| < c$, then there exists a non-trivial first order SCEE.*

In case of the motivating family (21), the theorem yields some $\lambda_0 > 1$. If $|\nu| > \lambda_0$, then also $|t'_\nu(0)| > \lambda_0$. Choose λ such that $\lambda_0 < \lambda < |\nu|$. Then, by continuity, there is a $c > 0$ such that $|t'_\nu(p)| > \lambda$ for all $|p| < c$; since $t_\nu(0) = 0$, it then follows that $|t_\nu(p)| > \lambda|p|$ for $|p| < c$, and we may apply the theorem.

Note that for $|\nu| > 1$, the origin $p=0$ is a hyperbolic repeller of the deterministic map $p \mapsto \tanh \nu p$. For $\nu > 1$ this map has furthermore two attracting fixed points; for $\nu < -1$ it has an attracting period-2 cycle.

The theorem now says that there is some $\lambda_0 > 1$ such that for every $|v| > \lambda_0$ these deterministic attractors are echoed by nontrivial SCEE of the stochastic process.

4.2.3 First Order SCEE Close to Bifurcation

Finally, we show the existence of first order SCEE for systems close to a bifurcation in which the origin has just lost its stability; for this result, the fourth moment of the noise has to be assumed sufficiently small.

Theorem 9. *Let the noise η_t satisfy assumption A1 and let F satisfy assumptions A2 and A3. Assume moreover that $F''(p)F(1) < 0$ for all $p > 0$. Set $\lambda = |F'(0)|$ and $c = F'''(0)/6$. If $c \neq 0$, and if $\lambda > 1$ is sufficiently close to 1, there exists $\bar{\tau}$ such that if $\mathbb{E}\eta_t^4 \leq \bar{\tau}^4$, then F has a first order SCEE.*

In the case of our motivating example $F(p) = \tanh(vp)$, we have $\lambda = |F'(0)| = |v|$,

$$F''(p)F(1) = -2 \frac{\tanh vp}{\cosh^2 vp} \tanh v < 0 \quad \text{for all } p > 0,$$

and $c = F'''(0)/6 = -\frac{v^3}{3}$. Note that for $|v| > 1$ we have that $c \neq 0$. The theorem says then that if $|v|$ is larger than 1, but both $|v| - 1$ and $\mathbb{E}\eta_t^4$ are sufficiently small, then the motivating family has a nontrivial SCEE.

Now, since $v = 1$ and $v = -1$ are bifurcation values of the deterministic system $p \mapsto \tanh(vp)$, we can formulate this theorem as follows: if $0 < |v| - 1 \ll 1$, the deterministic skeleton $p_t = \tanh(vp_{t-1})$ of the motivating family has a nontrivial attractor. If the fourth order moment of the noise term η_t is sufficiently small, then the family itself has a nontrivial SCEE.

5 Adaptive Learning in the OG-Model

In this section we investigate adaptive learning of SCEE in overlapping generations economies with fiat money. Our main application is the OG-model as specified in [Grandmont \(1985\)](#), which has served as a benchmark in the learning literature. In Sect. 4.1 we recall the model assumptions and its dynamics under perfect foresight and under naive expectations. Section 4.2 studies the same OG-model under SAC-learning. It is shown that for Grandmont's specification of the OG-model the SAC-learning dynamics has only two possible long run outcomes: a steady state SCEE and a 'noisy 2-cycle SCEE'. Adaptive learning exhibits path dependence in the sense that, even for the same initial states, the long run outcome may depend upon the realizations of the random shocks to the economy. Section 4.3 presents a graphical analysis of the SAC-learning dynamics providing an explanation why

the adaptive learning process exhibits path dependence with positive probability for both long run outcomes. Finally, Sect. 4.4 presents another specification of an OG-model, for which, in the deterministic case, the SAC-learning process converges to a first order chaotic CEE. The persistence of chaotic price fluctuations under random shocks is also studied.

5.1 Model Formulation

The presentation of the OG-model follows [Grandmont \(1985\)](#). In each period, a continuum (of measure 1) of identical, 2-period lived agents is born who work when young and consume only when old. Agents have access to a constant returns to scale production technology which transfers one unit of labor into one unit of output. Each agent is endowed with l^* units of labor in his first period of life. There exists a fixed amount M of fiat money in the economy which is initially held by the members of the first old generation. Fiat money is the only store of value.

Denoting by l_t , m_t , and c_{t+1} the labor supply, money holding, and consumption of a member of generation t , and by p_t and p_{t+1}^e the price of output in period t and the expected price in period $t + 1$, we obtain the following budget constraints for individuals of generation t :

$$m_t = p_t l_t, \quad p_{t+1}^e c_{t+1} = m_t. \quad (24)$$

Each agent maximizes an additively separable utility function of the form

$$V_1(l^* - l_t) + V_2(c_{t+1}), \quad (25)$$

where the first term describes the utility of labor (or the utility of leisure) and the second term the utility derived from consumption. We assume that both V_1 and V_2 are increasing and concave functions. The market clearing conditions for the money market and the goods market are

$$m_t = M, \quad c_t = l_t \quad (26)$$

for all t .

The first order condition for the utility maximization is

$$V_1'(l^* - l_t)/p_t = V_2'(c_{t+1})/p_{t+1}^e. \quad (27)$$

From (24) and (26) it follows that in equilibrium one has $l_t = M/p_t$ and $c_{t+1} = M/p_{t+1}^e$. Substituting this into the first order condition (27) we obtain $v_1(M/p_t) = v_2(M/p_{t+1}^e)$, where

$$v_1(x) = x V_1'(l^* - x), \quad v_2(x) = x V_2'(x). \quad (28)$$

It is straightforward to see that v_1 is a strictly increasing function from $[0, l^*)$ into $\mathbb{R}$, so that v_1^{-1} exists and the law of motion generating the equilibrium dynamics can therefore be written in the form

$$p_t = G(p_{t+1}^e) = \frac{M}{\chi(M/p_{t+1}^e)}, \quad (29)$$

where $\chi = v_1^{-1} \circ v_2$. The graph of χ is called the offer curve and describes the optimal pairs of labor and consumption. [Grandmont \(1985\)](#) has shown that the function χ is increasing if substitution effects are dominating. In the presence of a sufficiently strong income effect, however, the offer curve χ and hence the function G becomes non-monotonic.

Before we embark on simulations of SAC-learning, it will be useful to discuss equilibria under perfect foresight and under naive expectations. Under perfect foresight we must have $p_{t+1}^e = p_{t+1}$, and the dynamics (29) becomes

$$p_t = G(p_{t+1}) = \frac{M}{\chi(M/p_{t+1})}. \quad (30)$$

[Grandmont \(1985\)](#) refers to (30) as the backward perfect foresight dynamics. If the offer curve χ is non-monotonic the forward perfect foresight dynamics may not be uniquely defined and can be studied only locally. [Grandmont](#) showed that the forward perfect foresight dynamics may have infinitely many period cycles and chaotic perfect foresight equilibria.¹²

Here we study the dynamics under backward looking expectations to address the question which of the infinitely many perfect foresight equilibria may arise under adaptive learning. The dynamics under naive expectations, i.e. when the agents use the forecast $p_{t+1}^e = p_{t-1}$, is given by

$$p_t = G(p_{t-1}) = \frac{M}{\chi(M/p_{t-1})}.$$

[Grandmont \(1985\)](#) presents a detailed investigation of the backward perfect foresight dynamics in the constant elasticity case, that is, with utility functions

$$V_1(l^* - l) = \frac{(l^* - l)^{(1-\gamma_1)}}{1 - \gamma_1}, \quad V_2(c) = \frac{(c + c^*)^{(1-\gamma_2)}}{1 - \gamma_2}. \quad (31)$$

Following [Grandmont \(1985\)](#) we fix the parameters $l^* = 2$, $c^* = 1/2$, and $\gamma_1 = 1/2$. For $\gamma_2 < 1$, current leisure and future consumption are gross substitutes and the

¹²See also [Medio and Raines \(2007\)](#) and [Gardini, Hommes, Tramontana, and de Vilder \(2009\)](#) for an extensive discussion and characterization of the forward perfect foresight dynamics in the case of a non-monotonic offer curve.

offer curve is increasing. In that case, the dynamics under naive expectations always converges to a unique positive (monetary) steady state. When γ_2 becomes larger, the income effect becomes stronger and the offer curve becomes non-monotonic for $\gamma_2 > 1$. [Grandmont](#) showed that as γ_2 increases, the dynamics under naive expectations (or equivalently, the backward perfect foresight dynamics) becomes increasingly more complicated and exhibits a period doubling bifurcation route to chaos. This implies that, for γ_2 large, infinitely many periodic as well as chaotic perfect foresight equilibria exist.

Here we are interested in the question whether there is any linear, detectable structure in the forecasting errors of simple expectations schemes such as naive expectations. The answer depends on the (long run) outcome of the price dynamics. It may be that the forecasting errors vanish in the long run, as it is the case for example when the dynamics under naive expectations converge to a steady state or to a 2-cycle. In these cases, naive expectations lead to perfect foresight in the long run. In general, however, the dynamics under naive expectations does not approach a perfect foresight equilibrium and agents make systematic forecasting errors forever. For example, if the dynamics converges to a (stable) cycle of period $k > 2$, naive agents make systematic forecasting errors forever. The same is true if the resulting time path is chaotic. Figure 3 shows 500 observations of a chaotic price series under naive expectations, as well as the corresponding forecasting errors $\epsilon_t = p_t - p_t^e = p_t - p_{t-2}$. In this simulation, the parameter γ_2 measuring the concavity of the old traders' utility function, has been fixed at $\gamma_2 = 12$. This value is in the chaotic region of the parameter space as can be seen in the bifurcation diagram in [Grandmont \(1985, p. 1030\)](#).

The time series in Fig. 3 show that naive agents make large forecasting errors, often of the order of plus or minus the amplitude of the price oscillations. Furthermore, the forecasting errors seem to be systematic. In particular, there appears to be clear linear structure in the observed forecasting errors as can be seen from the sample autocorrelation plots of the forecasting errors, whose first 10 lags are strongly significant. Thus, we may say that in this chaotic economy the assumption of naive expectations leads to systematic forecasting errors.

5.2 SAC-Learning

We have seen in the previous subsection that naive agents make systematic forecasting errors. By using time series observations only and by employing simple, linear statistical techniques the agents would be able to detect these errors. We now study price fluctuations when the agents adapt their forecasting rule according to the SAC-learning dynamics. Let us ignore noise for the moment and focus on the deterministic OG-model first. In this situation the prices evolve according to the difference equation

Fig. 3 OG-model under naive expectations. *Top left*: chaotic price fluctuations; *bottom left*: chaotic forecasting errors under naive expectations; *top right*: sample autocorrelations of prices; *bottom right*: sample autocorrelations of forecasting errors. *Dotted lines* indicate the approximate two standard errors significance bands $\pm 2/\sqrt{N}$, $N = 500$

$$p_t = G(p_{t+1}^e) = G_{\alpha_{t-1}, \beta_{t-1}}(p_{t-1}) = \frac{M}{\chi(M/[\alpha_{t-1} + \beta_{t-1}^2(p_{t-1} - \alpha_{t-1})])}, \quad (32)$$

where the learning parameters α_{t-1} and β_{t-1} are given by (14) and (15), respectively.

As long as the map G is monotonically increasing, it follows immediately from Theorem 3 that the only bounded CEE is a steady state CEE. Hence, if current leisure and future consumption are gross substitutes, the only CEE is a steady state CEE.

For the utility function (31) with $\gamma_2 > 1$, that is, when the concavity of the old traders' utility function is large enough, the function G in (32) is non-monotonic. For all $\gamma_2 > 1$ only two possible outcomes of the SAC-learning process have been observed in all our numerical simulations: fast convergence to the monetary steady state CEE as illustrated in Fig. 4 and slow convergence to a 2-cycle CEE as illustrated in Fig. 5. In Figs. 4 and 5, the parameters in the OG-model are exactly the same as those used in the simulations depicted in Fig. 3. In the case of the steady state CEE in Fig. 4, SAC-learning leads to convergence to a stationary

Fig. 4 SAC-learning in the deterministic (chaotic) OG-model (32) with initial states $p_0 = \alpha_0 = 1.99$ and $\beta_0 = -1$

perfect foresight equilibrium and forecasting errors vanish quickly. Analogously, the forecasting errors in Fig. 5 (convergence to a 2-cycle) are roughly 10,000 times smaller than under naive expectations. We may therefore conclude that, although the adapted linear forecasting rules are not perfect, they lead to a large improvement in forecasting performance and to a dampening of price fluctuations and forecasting errors. In a highly nonlinear, chaotic environment, SAC-learning thus seems to have a stabilizing effect on the price dynamics. Moreover, *even in the presence of infinitely many chaotic perfect foresight equilibria, the SAC-learning dynamics converges either to the steady state CEE or to a 2-cycle CEE, that is, the SAC-learning process selects the simplest types of equilibria among many complicated perfect foresight equilibria.*

As a next step, we study how dynamic noise affects the SAC-learning process. We are particularly interested in whether convergence to steady state CEE or 2-cycle CEE is robust with respect to dynamic noise. In order to address these issues, we investigate SAC-learning for the stochastic model with the same parameter values as before. To avoid negative prices due to exceptionally large stochastic shocks, we add a nonnegativity constraint in the numerical simulations with noise. Under these assumptions, the SAC-learning is given by

$$p_t = G_{\alpha_{t-1}, \beta_{t-1}}(p_{t-1}, \eta_t) = \max \left\{ \frac{M}{\chi(M/[\alpha_{t-1} + \beta_{t-1}^2(p_{t-1} - \alpha_{t-1})])} + \eta_t, 0 \right\},$$

with α_{t-1} and β_{t-1} given by (14) and (15), respectively.

In all our numerical simulations of this version of the model, we only observed two different outcomes: convergence to a steady state SCEE or to a ‘noisy 2-cycle’

Fig. 5 SAC-learning in the deterministic (chaotic) OG-model (32) with initial states $p_0 = \alpha_0 = 2.25$ and $\beta_0 = -1$

SCEE. Convergence to a steady state SCEE is illustrated in Fig. 6, where η_t is uniformly distributed on the interval $[-2.5, 2.5]$. Prices seem to fluctuate randomly around their mean, whereas the belief parameters seem to converge. The belief parameter α_t converges to the mean price whereas the first order autocorrelation coefficient β_t converges slowly to zero. The bottom panel in Fig. 6 shows sample autocorrelations of prices, forecasting errors, and residuals from a regression of prices on a constant. Prices and forecasting errors are uncorrelated. Estimation of an AR(1) model on prices shows that the estimated coefficient for lagged prices is not significant. Agents therefore cannot reject the null hypothesis that prices fluctuate randomly around their mean. This example shows that even in a noisy chaotic economy the SAC-learning dynamics may, and in fact often does, enforce convergence to a steady state SCEE.

Figures 7 and 8 show simulations with the same parameter values as in Fig. 6, but with different initial values $p_0 = \alpha_0 = 2.25$ and $\beta_0 = -1$. The figures suggest that the SAC-learning process converges to a ‘noisy 2-cycle’ SCEE.¹³ The only difference between these two simulations is the noise level, which is twice as large in Fig. 8 than it is in Fig. 7. In both simulations, the belief parameter α_t converges to the mean price, whereas the first order sample autocorrelation coefficient β_t converges to a negative value around -0.6 . Also in both cases, the sample autocorrelation

¹³The name ‘noisy 2-cycle’ SCEE captures the feature that price fluctuations look like a noisy 2-cycle. In Sect. 5.3 we will see that the underlying (deterministic) limiting law of motion of this economy has indeed a stable 2-cycle.

Fig. 6 SAC-learning in the stochastic OG-economy with initial values $p_0 = \alpha_0 = 1.99$ and $\beta_0 = -1$. Prices (*top left*), expected prices (*top middle*), forecast errors (*top right*), and the belief parameters α_t (*middle left*) and β_t (*middle right*). The *bottom panel* shows sample autocorrelations of prices, forecasting errors, and residuals from a regression of prices on a constant

plot of the price series looks like the autocorrelation function of an AR(1) process with negative first order autocorrelation coefficient. A more careful look, however, reveals that, in the case of small noise, the sample autocorrelation plot of the price series does not decay exponentially fast as it must be the case in an AR(1) process (compare the bottom left plots in Figs. 7 and 8). In the case of small noise, forecasting errors are correlated and the residuals from an estimated AR(1) model on prices still exhibit some significant autocorrelations. In fact, in Fig. 7 the SAC-learning process has settled down only to a *first order SCEE*: β_t has converged to the first order autocorrelation coefficient, but higher order autocorrelation coefficients ρ_j are not exactly equal (but close) to β^j . A careful boundedly rational agent would conclude that his AR(1) forecasting rule is not correct and that his perceived law of motion is misspecified. However, when the noise level is larger as in Fig. 8, both forecasting errors and residuals from an estimated AR(1) model are uncorrelated, so that the hypothesis that prices follow an AR(1) model cannot be rejected. The

Fig. 7 SAC-learning in the stochastic OG-economy with initial states $p_0 = \alpha_0 = 2.25$ and $\beta_0 = -1$ and with η_t uniformly distributed on the interval $[-1.25, 1.25]$. Prices (top left), expected prices (top middle), forecast errors (top right), and the belief parameters α_t (middle left) and β_t (middle right). The bottom panel shows sample autocorrelations of prices, forecasting errors, and residuals from an estimated AR(1) model

SAC-learning process has converged to an SCEE, because a sufficiently high noise level prevents the agents from discovering that their perceived linear law of motion is misspecified.

It is useful to discuss the volatility of price fluctuations in these three examples of SAC-learning. Although the noise level in the example of the steady state SCEE in Fig. 6 is larger than or equal to the noise level in the two examples of ‘noisy 2-cycle’ SCEE, the amplitude of the price fluctuations of the steady state SCEE is about half the amplitude of the price fluctuations in the ‘noisy 2-cycle’ SCEE. This shows that a boundedly rational SAC-learning equilibrium – such as a ‘noisy 2-cycle’ SCEE – may cause excess volatility in price fluctuations compared to the benchmark of an REE.¹⁴ Bullard and Duffy (2001) have also discussed the importance of excess volatility due to learning processes.

¹⁴Note that a steady state SCEE is also an REE.

Fig. 8 SAC-learning in the stochastic OG-economy with initial states $p_0 = \alpha_0 = 2.25$ and $\beta_0 = -1$ and with η_t uniformly distributed on the interval $[-2.5, 2.5]$. Prices (*top left*), expected prices (*top middle*), forecast errors (*top right*), and the belief parameters α_t (*middle left*) and β_t (*middle right*). The *bottom panel* shows sample autocorrelations of prices, forecasting errors, and residuals from an estimated AR(1) model

5.3 Graphical Analysis of the Temporary Law of Motion

We have seen two possible outcomes of the SAC-learning process in the [Grandmont \(1985\)](#) specification of the OG-model with noise: convergence to a steady state SCEE or convergence to a (first order) ‘noisy 2-cycle’ SCEE. The SAC-learning dynamics (14)–(16) is a high-dimensional, nonlinear system, which is difficult to analyze analytically. A simple and intuitive graphical analysis however can explain why, depending on the initial states of the learning process and the realizations of the noise, both possibilities occur with positive probability. To this end consider the temporary law of motion of the deterministic model, G_{α_t, β_t} , as shown in [Fig. 9](#), and recall that the equilibrium price p_{t+1} in period $t + 1$ is given by $G_{\alpha_t, \beta_t}(p_t)$. The SAC-learning process may be viewed as a 1-dimensional dynamical system with a *slowly changing temporary law of motion* G_{α_t, β_t} . The initial state (p_0, α_0, β_0) and

Fig. 9 Graphs of the temporary actual laws of motion G_{α_t, β_t} (top) and its second iterate $G_{\alpha_t, \beta_t}^{(2)}$ (bottom) for $\alpha_t = 2.2$ and $\beta_t = -0.4$, $\beta_t = -0.5$, and $\beta_t = -0.6$

the initial phase of the SAC-learning process are critical since, in the early stage, the temporary law of motion is most sensitive to new observations. Figures 9 illustrates the graphs of the function $G_{\alpha, \beta}$ and its second iterate $G_{\alpha, \beta}^{(2)}$ for $\alpha = 2.2$ and three different values of β : $\beta = -0.4$, $\beta = -0.5$, and $\beta = -0.6$. Notice that these values for α and β have been chosen close to the values of the initial phase of the SAC-learning process in Figs. 6–8. For $\beta = -0.4$ and $\beta = -0.5$, the temporary law of motion has a (globally) stable steady state around $p = 1.7$, and prices thus tend to converge. However, for $\beta = -0.6$, the temporary law of motion has an unstable steady state and a stable 2-cycle with values close to 0.5 and 4.5. Hence, for $\alpha_t \approx 2.2$ and $\beta_t \leq -0.6$ prices will oscillate up and down, which will lead to a further decrease of β_t .¹⁵ If, in the initial phase of the learning process, β_t becomes small enough, say $\beta_t \leq -0.6$, then the SAC-learning process is likely to lock into a ‘noisy 2-cycle’ SCEE. If on the other hand, $\beta_t \geq -0.5$ in the initial phase, price fluctuations are stable and the SAC-learning process converges to a steady state SCEE. Both the steady state SCEE and the ‘noisy 2-cycle CEE’ can therefore occur with positive probability.

¹⁵The reader should have a look at the time series of the learning parameters α_t and β_t in Figs. 6–8 again. In these simulations, in the initial phase of the learning process $1.8 \leq \alpha_t \leq 3$, whereas $-0.6 \leq \beta_t \leq -0.2$.

It is interesting to note that Bullard and Duffy (1998) investigate the same OG-model with a population of artificial, heterogeneous, adaptive agents whose forecasting rules are updated by a genetic algorithm. They find that for high values of γ_2 , for which many chaotic perfect foresight equilibria exist, in most of the cases the agents learn to coordinate on two simple equilibria: either on the monetary steady state or on the (noisy) perfect foresight 2-cycle. Also in the experimental heterogeneous agents OG-model of Marimon, Spear, and Sunder (1993) coordination on the steady state and a (noisy) 2-cycle are the observed outcomes. Our results for SCEE provide a possible theoretical explanation of these results. In none of our simulations we observed convergence to one of the many complicated, chaotic perfect foresight trajectories that exist. *Simple AR(1) forecasting rules and SCEE may thus explain why, in a world of adaptive agents, coordination on a simple steady state or a noisy 2-cycle are more likely to occur.*

Another interesting feature of the SAC-learning process, suggested by the graphical analysis above, is that the adaptive learning process exhibits *path dependence*. If both long run outcomes, steady state SCEE and ‘noisy 2-cycle’ SCEE, occur with positive probability, the initial realizations of the random shocks η_t may be critical in determining to which of the two long run outcomes the process will settle down. Indeed we have observed in our numerical simulations that for the same initial state (p_0, α_0, β_0) , different realizations of the random shocks may lead to different long run outcomes.¹⁶ This path dependence of SAC-learning has similar features as *nonlinear urn process*, proposed by Arthur, Ermoliev, and Kaniovski (1987) to model ‘lock in’ effects due to technological change. In these models of technological change, the two competing attractors are both steady states and the process locks into one of the two different steady states, one superior and one inferior technology. In contrast, for the SAC-learning process the competing attractors are of a different kind, a steady state SCEE versus a ‘noisy 2-cycle’ SCEE, and the adaptive learning process locks into one of them depending upon the initial state and the random shocks.

5.4 Example of a First Order Chaotic CEE and its Robustness to Noise

In the previous subsection we have seen that, for constant elasticity utility functions, chaotic CEE do not occur as the long run outcome of the SAC-learning process. The purpose of this subsection is to construct a utility function for which the long run outcome of the SAC-learning process in the OG economy is a first order chaotic

¹⁶For example, for the initial state $(p_0, \alpha_0, \beta_0) = (2.25, 2.25, -1)$ as in the simulation of the ‘noisy 2-cycle’ SCEE in Fig. 8, for different realizations of the random shocks η_t we have indeed also observed converge to the steady state SCEE of Fig. 6. OLS-learning in this model exhibits the same type of path dependence.

CEE and to study the sensitivity of this first order chaotic CEE under dynamic noise. Except for the utility function, the OG-model is exactly the same as in the previous subsection.

Each agent maximizes his separable utility function (25), with $V_1(x) = 2\sqrt{x}$ and $V_2(x) = W(x)/\sqrt{l^*}$, where

$$W(x) = \begin{cases} -\int_x^{M/d} (M/d)e^{-\mu d} e^{\mu M/z}/(\lambda z) dz & \text{if } x \in (0, M/d], \\ (x - M/d)/\lambda & \text{if } x \in (M/d, +\infty) \end{cases}$$

and where λ , μ , and d are parameters. As before, we denote the money supply be M and the labor endowment by l^* . Note that V_1 is defined as in (31) with $\gamma_1 = 1/2$. Both V_1 and V_2 are increasing, continuously differentiable, and concave for all $x \in (0, +\infty)$. We will demonstrate that for this choice of the utility function, the function G in (29) has a tent shape when l^* is sufficiently large, as illustrated in Fig. 10. To this end first observe that (28) and $V_1(x) = 2\sqrt{x}$ imply $v_1(x) = x/\sqrt{l^*} - x$ and therefore

$$v_1^{-1}(y) = (y/2) \left[\sqrt{y^2 + 4l^*} - y \right].$$

Furthermore, $v_2(x) = xV_2'(x) = xW'(x)/\sqrt{l^*}$. The function G in (29) therefore becomes

$$\begin{aligned} G(p) &= \frac{M}{v_1^{-1} \left(M W'(M/p) / [p\sqrt{l^*}] \right)} \\ &= \frac{2p\sqrt{l^*}}{W'(M/p) \left[\sqrt{(M/p)^2 [W'(M/p)]^2 / l^* + 4l^*} - (M/p) W'(M/p) / \sqrt{l^*} \right]}. \end{aligned}$$

Since the function W does not depend on l^* it follows from the above equation that

$$\lim_{l^* \rightarrow +\infty} G(p) = \frac{p}{W'(M/p)}.$$

Moreover, using the definition of W we obtain

$$W'(M/p) = \begin{cases} 1/\lambda & \text{if } p < d, \\ p e^{\mu(p-d)}/(\lambda d) & \text{if } p \geq d. \end{cases}$$

Combining these results we obtain

$$\lim_{l^* \rightarrow +\infty} G(p) = \begin{cases} \lambda p & \text{if } p < d, \\ (\lambda d) e^{-\mu(p-d)} & \text{if } p \geq d. \end{cases}$$

Fig. 10 The graph of the function G for $M = 0.25, \lambda = 2.5$, $d = 2, \mu = 2$, and several values of l^* . The graph labelled 'limit' is the limit as l^* approaches $+\infty$

The graph of the function G is shown in Fig. 10 for several values of l^* . In the limit, the slope of the graph to the left of the kink is equal to λ , whereas the slope to the right depends on the parameter μ and the critical point is given by the parameter d . Since no restrictions are placed on these parameters one can generate tent-shaped maps that are tilted to either side of the critical point.

We will consider a typical example of the SAC-learning process in this OG-model, with the parameters fixed at $\mu = 2, d = 2, l^* = 10, \lambda = 2.5$, and $M = 0.25$, with and without noise. Figure 11 presents a numerical simulation of the SAC-learning dynamics without noise, showing 2,000 observations of the time series of realized prices p_t , forecasts p_t^e , forecasting errors ϵ_t , the sample average α_t , and the sample autocorrelation coefficient β_t . Prices fluctuate irregularly, with seemingly unpredictable forecasting errors. The sample mean α_t converges to $\alpha^* \approx 2.24$ and the sample autocorrelation β_t converges to $\beta^* \approx -0.77$. Figure 11 also shows the sample autocorrelation plots of prices, forecasting errors, and residuals from an estimated AR(1) model. The sample autocorrelation plot of prices exhibits a regular up and down oscillatory pattern, that is, the sample autocorrelation coefficients do not decrease exponentially as the lag increases but the amplitude of up and down oscillations stays more or less constant. In particular, we do not have $\rho_j(p_0) = (\beta^*)^j$ for all j . This shows that the SAC-learning process does not converge to a CEE. This finding is corroborated by the observation that forecasting errors also exhibit significant autocorrelations. Based on linear statistical techniques, agents would thus be able to discover regularities in forecasting errors and conclude that their model is misspecified. On the other hand, the first order sample autocorrelation coefficient converges to its correct value β^* , so that the situation qualifies as a first order chaotic CEE.¹⁷

¹⁷See also Jungeilges (2007) for a similar example in the cobweb framework, where under SAC-learning the first order sample autocorrelation coefficient may converge to its correct value β while higher order sample autocorrelation coefficients need not converge to the correct values β^j . In a related paper Tuinstra (2003) shows that, under OLS-learning in an OG-model with money growth and inflation, 'beliefs equilibria' may arise where the first order autocorrelation coefficient converges while prices fluctuate on a quasi-periodic or chaotic attractor. A beliefs equilibrium is

Fig. 11 Convergence of SAC-learning process to a chaotic first order CEE. *Top panel:* time series of realized prices, forecasts, and forecasting errors; *middle panel:* time series of the sample average α_t and the sample autocorrelation β_t ; *bottom panel:* sample autocorrelation plots of prices, forecasting errors, and residuals from an estimated AR(1) model

Figure 12 presents a graphical explanation of the observed long run outcome of the SAC-learning dynamics, showing the graphs of the limiting actual law of motion $G_{\alpha,\beta}$, its second iterate $G_{\alpha,\beta}^{(2)}$, its fourth iterate $G_{\alpha,\beta}^{(4)}$, and its 16-th iterate $G_{\alpha,\beta}^{(16)}$ for the limiting values of the SAC-learning process $\alpha = 2.24$ and $\beta = -0.77$. The graphs of $G_{\alpha,\beta}$ and $G_{\alpha,\beta}^{(2)}$ show that prices jump back and forth between two disjoint intervals which, in turn, implies that the price process has long memory. The fact that the 16-th iterate $G_{\alpha,\beta}^{(16)}$ is expanding (i.e. its derivative always exceeds one in absolute value) implies that the dynamics is chaotic and exhibits sensitive dependence on initial conditions. Chaotic dynamics jumping back and forth between two disjoint

in fact a first order CEE, where agents have fitted the correct regression line to a quasi-periodic or chaotic attractor.

Fig. 12 Graphs of the implied actual law of motion $G_{\alpha,\beta}$ (top left), its second iterate $G_{\alpha,\beta}^{(2)}$ (top right), its fourth iterate $G_{\alpha,\beta}^{(4)}$ (bottom left), and its 16-th iterate $G_{\alpha,\beta}^{(16)}$, for the limiting values of the SAC-learning process $\alpha = 2.24$ and $\beta = -0.77$

intervals explains the oscillatory pattern in sample autocorrelations of prices with long range autocorrelations.

As a next step we add dynamic noise to the model. Figure 13 shows numerical simulation of the same OG-model buffeted with dynamic noise, normally distributed with standard deviation $\sigma=2$. Prices fluctuate irregularly, with seemingly unpredictable forecasting errors. The sample mean α_t converges to $\alpha^* \approx 2.6$ and the sample autocorrelation β_t converges to $\beta^* \approx -0.6$. The sample autocorrelation plot of prices exhibits a regular up and down oscillatory pattern but long memory has disappeared and the autocorrelation plot seems to decrease exponentially fast very much like the autocorrelation function of an AR(1) process. Forecasting errors exhibit no significant autocorrelations.¹⁸ Estimation of an AR(1) model on past prices yields estimates $\alpha = 2.65$ and $\beta = -0.6$ with uncorrelated residuals. Hence, in the presence of noise the SAC-learning process seems to converge to an SCEE and the null hypothesis of an AR(1) model cannot be rejected.

Figure 14 shows the graph of the deterministic part of the limiting law of motion $G_{\alpha,\beta}$ and its second iterate $G_{\alpha,\beta}^{(2)}$ for the limiting values of the SAC-learning process

¹⁸The first order autocorrelation coefficient is significant, but recall that first order autocorrelation can not be exploited, since agents have to make a 2-period ahead forecast.

Fig. 13 Convergence of SAC-learning process to a ‘noisy 2-cycle CEE’. Parameters and initial states are as in Fig. 12, but dynamic noise has been added. *Top panel*: time series of realized prices, forecasts, and forecasting errors; *middle panel*: time series of the sample average α_t and the sample autocorrelation β_t ; *bottom panel*: sample autocorrelation plots of prices, forecasting errors, and residuals from an estimated AR(1) model

with noise, i.e., $\alpha = 2.65$ and $\beta = -0.6$. The reader should compare these graphs to the implied actual law of motion in the case without noise as shown in Fig. 12. Dynamic noise clearly has affected the limiting law of motion. In the presence of noise, the limiting law of motion $G_{\alpha,\beta}$ is less steep and the second iterate $G_{\alpha,\beta}^{(2)}$ has two stable fixed points, indicating that the long run SAC-learning dynamics with noise has converged to a ‘noisy 2-cycle’ SCEE.

To summarize, in the deterministic case the SAC-learning converges to a chaotic *first order* CEE. Although the first order autocorrelation coefficient β_t converges, forecasting errors still exhibit significant autocorrelations at higher order lags and the null hypothesis that the time series is generated by an AR(1) model is rejected. Careful boundedly rational agents are therefore able to discover that their perceived law of motion is misspecified. In the presence of noise however, the misspecification becomes harder to detect. With a fair amount of noise, the SAC-learning process

Fig. 14 Graphs of the implied actual law of motion $G_{\alpha,\beta}$ and its second iterate $G_{\alpha,\beta}^{(2)}$, for the limiting values of the SAC-learning process with noise $\alpha = 2.65$ and $\beta = -0.6$.

converges to a ‘noisy 2-cycle’ SCEE. Consequently, by using linear statistical tools only, agents would not be able to reject the null hypothesis that prices are generated by a stochastic AR(1) process.

6 Concluding Remarks

There seems to be a growing consensus among economists that the rational expectations hypothesis implies extremely strong rationality of economic agents, and that deviations from full rationality need to be explored. The merit of studying REE is that they form a natural benchmark and discipline research, as nicely expressed in the saying that there is only one way (or perhaps a few ways) one can be right. Stated more formally, in an REE beliefs have to coincide with the implied actual law of motion which puts strong restrictions upon agents’ beliefs. In contrast, researchers using the assumption of bounded rationality are usually forced to make ad hoc assumptions because there are many ways one can be wrong. In an equilibrium under bounded rationality beliefs and their implied laws of motion typically do not coincide leaving room for a “wilderness of bounded rationality”. Recent work on bounded rationality may be seen as, what [Sargent \(1999\)](#) calls a search for approximate REE, where differences between a belief and its implied actual law of motion are in some sense hard to detect from time series observations. A good theory of bounded rationality then should be able to say which of the approximate REE are more likely than others.

SCEE is an equilibrium concept where the perceived law of motion is linear and the implied actual law of motion is a nonlinear stochastic process with the same linear statistical properties, i.e., with the same unconditional mean and autocovariance structure. A nice feature of our SCEE framework is that the class of perceived laws of motion is simple and seems to be natural, at least as a starting point. SCEE puts discipline on the forecasting rules, since there are only a few linear forecasting rules which are right. The main theorem in this paper shows that *SCEE is always an approximate REE in the sense that forecasting errors are unbiased and*

uncorrelated. In an example OG-model with infinitely many periodic and chaotic REE, we have seen that SAC-learning selects simple equilibria, namely either a steady state SCEE or a noisy 2-cycle SCEE. These simple equilibria may be the most plausible approximate REE, on which a large group of agents can coordinate.

An attractive feature of the SCEE concept is the simplicity of the linear forecasting rules. Simplicity of a forecasting strategy should make coordination by a large population of agents on that particular rule more likely. It may be useful to compare SCEE with sunspot equilibria, as surveyed, e.g., by [Guesnerie \(2001\)](#). Self-fulfilling sunspot equilibria occur if all agents coordinate, or learn to coordinate, their beliefs on the same stochastic sunspot variable and if such a belief becomes self-fulfilling. In his elegant paper “learning to believe in sunspots” [Woodford \(1990\)](#) has shown that for an open set of initial states adaptive learning processes such as OLS-learning may converge to sunspot equilibria. Due to randomness, self-fulfilling sunspot equilibria may be very complicated however, and coordination on an irregular sunspot variable, even if possible with positive probability, seems unlikely in a heterogeneous world with many boundedly rational agents. SCEE are based on simple linear rules and are therefore much simpler equilibria than most sunspot equilibria, and coordination on a simple SCEE thus seems more likely than coordination on a complicated sunspot equilibrium. Stated differently, *learning to believe in linearity seems more likely than learning to believe in sunspots*.

In this paper we have focussed on SCEE with a linear AR(1) forecasting rule. It would be interesting to investigate a generalized form of SCEE in which agents use a simple linear forecasting rules derived from an AR(k) model with $k \geq 2$. A particularly interesting question is whether SCEE based on an AR(2) model exist in which the actual implied law of motion has a stable period- q cycle with $q > 2$, a quasi-periodic attractor, or a strange attractor. Stochastic linear AR(2) models often yield a good fit for (detrended) macroeconomic data. An SCEE with a linear AR(2) rule thus may be a boundedly rational equilibrium concept explaining this stylized fact of macroeconomics.

An important caveat concerning SCEE is that we have assumed that agents cannot distinguish between any two processes with the same mean and autocorrelation coefficients. One might argue that there exist simple *nonlinear* time series techniques which could easily distinguish between the realizations of the unknown nonlinear actual law of motion and the linear stochastic perceived law of motion. In fact, for the simple OG-models presented in this paper, a simple phase space plot (p_{t-1}, p_t) would reveal immediately that prices are generated by a 1-dimensional (chaotic) map and not by a stochastic AR(1) model. Figure 15 shows scatter plots of (p_{t-1}, p_t) of the SAC-learning process in the OG-model of Sect. 5.4, both without noise (left figure) and with noise (right figure), with an estimated regression line fitted to the observed data. When there is no noise, the simple scatter plot clearly reveals the tent-shaped graph of the limiting law of motion, and boundedly rational agents employing such a plot could improve their forecast. We emphasize though that these techniques are sensitive to the introduction of noise as illustrated in

Fig. 15 Scatter plots of the time series (p_{t-1}, p_t) generated by the SAC-learning process in the OG-model of Sect. 5, without noise (*left*) and with dynamic noise (*right*), with estimated *regression line* to observed data

the right plot in Fig. 15.¹⁹ In the presence of (small) dynamic noise, the graph of the limiting law of motion becomes invisible. Hence, in the presence of (small) dynamic noise the implied limiting law of motion becomes hard to detect and an estimated linear regression line may yield the most reasonable, boundedly rational forecast. There may be other, more sophisticated nonlinear time series techniques that could improve forecasting performance, but we would like to emphasize that it seems unlikely that many different agents would coordinate on the same nonlinear technique, especially when there is a simple linear rule that performs reasonably well.

Acknowledgements Earlier versions of this paper have been presented at the IFAC symposium Computational Economics, Finance and Engineering, Barcelona, July 6–8, 2000, the CeNDEF workshops on Economic Dynamics, Amsterdam, January 4–6, 2001 and Leiden, June 17–28, 2002, the 8th Viennese workshop on Optimal Control, Dynamic Games and Nonlinear Dynamics, Vienna, May 14–16, 2003 as well as various department seminars. Stimulating discussions with participants are gratefully acknowledged. In particular, we would like to thank Buz Brock, Jim Bullard, George Evans, Jean-Michel Grandmont and Seppo Honkapohja for stimulating discussions. This research has been supported by the Netherlands Organization for Scientific Research (NWO).

¹⁹The scatter plot technique works very well for 1-dimensional systems but becomes less informative for higher-dimensional systems. Brock, Hsieh, and LeBaron (1991), Brock and Dechert (1991) and Barnett et al. (1998) contain extensive discussions of the sensitivity to increasing dimension and the sensitivity to noise of nonlinear time series embedding methods.

Appendix

Appendix 1: Recursive form of SAC-Learning

Define $n_t = \sum_{i=0}^t (p_i - \alpha_t)^2$ and $z_t = \sum_{i=0}^{t-1} (p_i - \alpha_t)(p_{i+1} - \alpha_t)$. Obviously, we have $\beta_t = z_t/n_t$ for all t . The actual dynamics (16) can therefore be written as

$$\begin{aligned} p_t &= \Phi_1(p_{t-1}, \alpha_{t-1}, n_{t-1}, z_{t-1}, \eta_t) \\ &:= F(\alpha_{t-1} + (z_{t-1}/n_{t-1})^2(p_{t-1} - \alpha_{t-1}), \eta_t). \end{aligned}$$

Together with (14) this implies

$$\begin{aligned} \alpha_t &= \Phi_2(p_{t-1}, \alpha_{t-1}, n_{t-1}, z_{t-1}, \eta_t, t) \\ &:= \frac{t}{t+1}\alpha_{t-1} + \frac{1}{t+1}\Phi_1(p_{t-1}, \alpha_{t-1}, n_{t-1}, z_{t-1}, \eta_t). \end{aligned}$$

Finally, it is straightforward to verify that

$$\begin{aligned} n_t &= \Phi_3(p_{t-1}, \alpha_{t-1}, n_{t-1}, z_{t-1}, \eta_t, t) \\ &:= n_{t-1} + \frac{t}{t+1} [\alpha_{t-1} - \Phi_1(p_{t-1}, \alpha_{t-1}, n_{t-1}, z_{t-1}, \eta_t)]^2 \end{aligned}$$

and

$$\begin{aligned} z_t &= \Phi_4(p_{t-1}, \alpha_{t-1}, n_{t-1}, z_{t-1}, \eta_t, t, p_0) \\ &:= z_{t-1} + [\Phi_1(p_{t-1}, \alpha_{t-1}, n_{t-1}, z_{t-1}, \eta_t) - \alpha_{t-1}] \Phi_4(p_{t-1}, \alpha_{t-1}, n_{t-1}, z_{t-1}, \eta_t, t, p_0), \end{aligned}$$

where

$$\begin{aligned} &\Phi_4(p_{t-1}, \alpha_{t-1}, n_{t-1}, z_{t-1}, \eta_t, t, p_0) \\ &= p_{t-1} + \frac{p_0}{t+1} - \frac{t^2 + 3t + 1}{(t+1)^2} \alpha_{t-1} - \frac{1}{(t+1)^2} \Phi_1(p_{t-1}, \alpha_{t-1}, n_{t-1}, z_{t-1}, \eta_t). \end{aligned}$$

The above equations provide a recursive form of the SAC-learning dynamics.

Appendix 2: Existence of First Order SCEE

Before we can prove Theorems 8 and 9, we need a few auxiliary results.

First Order Coefficient

Let assumptions A1–A3 be satisfied and denote the invariant probability distribution mentioned in A3 by μ . As has been discussed in the main text, the transition function for the stochastic difference equation $p_t = F_{0,\beta}(p_{t-1}) + \eta_t$ can be expressed in the form

$$Q(p, A) = \int_A \varphi(p' - F_{0,\beta}(p)) \, dp',$$

where φ is the density function of η_t . Furthermore, the unconditional first order autocorrelation coefficient of the stochastic process is given by

$$\rho_1 = \frac{\iint p' p \, Q(p, dp') \, d\mu(p)}{\int p^2 \, d\mu(p)} = \frac{\iint p' p \, \varphi(p' - F_{0,\beta}(p)) \, dp' \, d\mu(p)}{\int p^2 \, d\mu(p)}.$$

The following lemma gives an lower bound on the first order autocorrelation coefficient of $(p_t)_{t=0}^{+\infty}$ in terms of its variance and the probability that p_t is far away from 0.

Lemma 2. *For every $A > 0$ it holds that*

$$\frac{\rho_1}{F_{0,\beta}(A)} \geq \frac{A}{\int p^2 \, d\mu(p)} \mathbb{P}(|p_t| > A).$$

Proof. Define $g(p) = pF_{0,\beta}(p)/F_{0,\beta}(A)$. Since $\int \varphi(p) \, dp = 1$ and $\int p \varphi(p) \, dp = 0$, we have

$$\begin{aligned} \iint p' p \, \varphi(p' - F_{0,\beta}(p)) \, dp' \, d\mu(p) &= \iint p(p' + F_{0,\beta}(p)) \varphi(p') \, dp' \, d\mu(p) \\ &= F_{0,\beta}(A) \int g(p) \, d\mu(p). \end{aligned}$$

Note that $g(p)$ is even, positive for all $p \neq 0$ and equal to 0 at $p = 0$. Moreover, it is nondecreasing on the positive real axis. We have therefore

$$\begin{aligned} F_{0,\beta}(A)^{-1} \iint p' p \, \varphi(p' - F_{0,\beta}(p)) \, dp' \, d\mu(p) &= 2 \int_0^A g(p) \, d\mu(p) + 2 \int_A^{+\infty} g(p) \, d\mu(p) \\ &\geq 2 \int_A^{+\infty} g(p) \, d\mu(p) \\ &\geq 2A \int_A^{+\infty} d\mu(p). \end{aligned} \tag{33}$$

The first inequality follows from $g(p) > 0$ for all $p > 0$, and the second one from $g(p) \geq g(A) = A$ for all $p > A > 0$. Because of the definition of ρ_1 and because of $2 \int_A^{+\infty} d\mu(p) = \mathbb{P}(|p_t| > A)$, the proof of the lemma is complete. $\square$

Moments

The following lemma gives estimates for all unconditional even moments of the process $(p_t)_{t=0}^{+\infty}$.

Lemma 3. *Suppose that $\mathbb{E}\eta_t^n$ exists and that a and b are as specified in A2. Then it follows that*

$$\int p^n d\mu(p) \leq \left(\frac{b + \sqrt[n]{\mathbb{E}\eta_t^n}}{1-a} \right)^n.$$

Proof. Note that $\mathbb{E}\eta_t^{2m+1} = 0$ for all $m \in \mathbb{N}$. Let β be fixed, and set $f = F_{0,\beta}$.

For n even, we have

$$\begin{aligned} \int p^n d\mu(p) &= \int [f(p) + \eta]^n \varphi(\eta) d\eta d\mu(p) = \sum_{k=0}^n \binom{n}{k} \mathbb{E}\eta_t^{n-k} \int f(p)^k d\mu(p) \\ &\leq \sum_{k=0}^n \sum_{j=0}^k \binom{n}{k} \binom{k}{j} a^j b^{k-j} \mathbb{E}\eta_t^{n-k} \int |p|^j d\mu(p) \end{aligned}$$

Recall that for $0 \leq j \leq n$, it follows from Jensen's inequality that $(\mathbb{E}|X|^j)^{\frac{n}{j}} \leq \mathbb{E}|X|^n$. In particular we have for n even that

$$\mathbb{E}|X|^j \leq (\mathbb{E}X^n)^{j/n}. \quad (34)$$

Putting

$$\xi = \left(\int p^n d\mu(p) \right)^{1/n}, \quad \tau = (\mathbb{E}\eta_t^n)^{1/n}$$

and using (34) yields that

$$\xi^n \leq \sum_{k=0}^n \sum_{j=0}^k \binom{n}{k} \binom{k}{j} a^j b^{k-j} \xi^j \tau^{n-k} = (a\xi + b + \tau)^n.$$

This implies obviously that $\xi \leq a\xi + b + \tau$ and, consequently, that

$$\xi \leq \frac{b + \tau}{1-a}.$$

Thus, the lemma is proved. □

The next lemma provides an estimate for the quantity $\mathbb{P}(|\eta_t| > A)$.

Lemma 4. *Assume that the fourth moment $\mathbb{E}\eta_t^4$ is finite. If $A^2 \leq \mathbb{E}\eta_t^2$, then*

$$\mathbb{P}(|\eta_t| > A) \geq \left(\frac{\mathbb{E}\eta_t^2 - A^2}{\sqrt{\mathbb{E}\eta_t^4}} \right)^2.$$

Proof. Let us denote the moments $\mathbb{E}\eta_t^2$ and $\mathbb{E}\eta_t^4$ by σ^2 and τ^4 , respectively, and the probability $\mathbb{P}(|\eta_t| > A)$ by ζ^2 . Furthermore, we denote the interval $[-A, A]$ by U and its complement by U^c . Then, using the Cauchy–Schwarz inequality, we find

$$\begin{aligned} \sigma^2 &= \int_U \eta^2 \varphi(\eta) \, d\eta + \int_{U^c} \eta^2 \varphi(\eta) \, d\eta \\ &\leq A^2 \int_U \varphi(\eta) \, d\eta + \sqrt{\int_{U^c} \varphi(\eta) \, d\eta} \sqrt{\int_{U^c} \eta^4 \varphi(\eta) \, d\eta} = A^2(1 - \zeta^2) + \zeta \tau^2. \end{aligned}$$

Rewrite this inequality as

$$\zeta^2 - \frac{\tau^2}{A^2} \zeta + \frac{\sigma^2 - A^2}{A^2} \leq 0,$$

and note that it implies

$$\begin{aligned} \zeta &\geq \frac{1}{2} \frac{\tau^2}{A^2} - \sqrt{\frac{1}{4} \frac{\tau^4}{A^4} - \frac{\sigma^2 - A^2}{A^2}} = \frac{1}{2} \frac{\tau^2}{A^2} \left(1 - \sqrt{1 - \frac{4A^2}{\tau^4} (\sigma^2 - A^2)} \right) \\ &\geq \frac{\sigma^2 - A^2}{\tau^2}; \end{aligned}$$

in the last line, the familiar inequality $\sqrt{1+x} \leq 1 + \frac{1}{2}x$ has been used. This proves the lemma. $\square$

Invariant Density

From here on, we restrict attention to the case that the invariant probability distribution μ is absolutely continuous with respect to the Lebesgue density dp , and moreover that the invariant probability density (or Jacobian) $\psi = d\mu/dp$ is continuously differentiable with respect to p . This will always be the case if both $F_{0,\beta}$ and the probability density φ of η_t are smooth functions.

If the origin $p = 0$ is a hyperbolic repeller for the dynamical system $p \mapsto F_{0,\beta}(p)$, that is, if $|F'_{0,\beta}(0)| > 1$, and if the variance of the noise term η_t is sufficiently small, it seems reasonable to expect that the invariant probability distribution of the process $p_t = F_{0,\beta}(p_{t-1}) + \eta_t$ has little mass around the origin $p = 0$. This is illustrated in Fig. 16: there will be only very little probability mass in the interval $[-A, A]$.

Fig. 16 Illustration of $g(p)$

The next, rather technical, lemma makes this precise. We shall restrict here our attention to the case that $F_{0,\beta}$ is differentiable and strictly monotonic. In order to be able to deal with the two cases $F_{0,\beta}$ increasing or decreasing simultaneously, we consider the second iterate $g(p) = F_{0,\beta}(F_{0,\beta}(p))$, which is always an increasing function. In fact, we shall require that g satisfies the following definition.

Definition 4. Let $\lambda > 1, \delta > 0$ and $\kappa > 0$. A function $f : \mathbb{R} \rightarrow \mathbb{R}$ is said to be of type $(\lambda, \delta, \kappa)$ with first fixed point $p_* > 0$, if

1. f is differentiable, strictly monotonic and odd;
2. $f(f(p_*)) = p_*$.
3. $|f'(p)| > \lambda$ if $|p| < \delta$ and $|f'(p)| < \lambda^{-1}$ if $|p - p_*| < \delta$.
4. $|f(p) - p| > \kappa$ for all $p \in [\delta, p_* - \delta]$.
5. $|f'(p)| > \kappa$ for all $p \in (0, p_*)$.

Note that the following lemma is ‘uniform’ with respect to the precise shape of $F_{0,\beta}$; this will be used later on.

Lemma 5. For every $\lambda > 1, \delta > 0, \kappa > 0, A > 0$ and $\varepsilon > 0$, there is

$$\sigma_0 = \sigma_0(\lambda, \delta, \kappa, A, \varepsilon) > 0,$$

such that the following holds.

Let $p_* > A$ and let $F_{0,\beta}$ be of type $(\lambda, \delta, \kappa)$ with first fixed point p_* . If μ is the invariant measure of the process $p_t = F_{0,\beta}(p_{t-1}) + \eta_t$ and if $\mathbb{E}\eta_t^2 < \sigma_0^2$, then

$$\mu([-A, A]) < \varepsilon.$$

Proof. For notational simplicity we write f instead of $F_{0,\beta}$ throughout this proof.

We shall construct increasing sequences of intervals U_n, V_n , such that $V_n \subset U_n \subset V_{n+1}$, and we will recursively estimate $\max |\varphi(p)|$ on these intervals using the integral equation representation of the stochastic process.

Let $U_0 = (-\delta, \delta)$. Then $|f'(p)| \geq \lambda > 1$ for all $p \in U_0$. Define intervals U_n recursively by $U_n = f(U_{n-1})$; note that $U_{n-1} \subset U_n$ and that their union satisfies

$$\bigcup_{n=0}^{\infty} U_n = (-p_*, p_*).$$

In particular $U_1 = (-|f(\delta)|, |f(\delta)|)$.

Set $\Delta = (|f(\delta)| - \delta)/2$ and define first

$$V_1 = (-\delta - \Delta, \delta + \Delta),$$

and then, recursively, $V_n = f(V_{n-1})$. We have that

$$V_n \subset U_n \subset V_{n+1}.$$

Denote by $d(U_n, \mathbb{C}V_{n+1})$ the distance between U_n and the complement $\mathbb{C}V_{n+1}$ of V_{n+1} ; particularly

$$d(U_0, \mathbb{C}V_1) = \Delta.$$

Recall that we assumed that the invariant probability density $\psi = d\mu/dp$ of p_t is continuously differentiable. Let moreover $\zeta(p)$ be the probability density of the stochastic variable $f(p_{t-1})$. The transformation formula for probabilities yields

$$\zeta(p) = \psi(f^{-1}(p)) |(f^{-1})'(p)| = \frac{\psi(f^{-1}(p))}{|f'(f^{-1}(p))|}.$$

From this it follows that

$$\max_{p \in V_n} \zeta(p) \leq \frac{\max_{p \in V_{n-1}} \psi(p)}{\min_{p \in V_{n-1}} |f'(p)|}.$$

The equality $p_t = f(p_{t-1}) + \eta_t$ is equivalent to the integral equation

$$\psi(p) = \int \varphi(p - \tilde{p}) \zeta(\tilde{p}) d\tilde{p}.$$

We find consequently for $n \geq 0$ that

$$\begin{aligned} \max_{p \in U_n} \psi(p) &= \max_{p \in U_n} \int \varphi(p - \tilde{p}) \zeta(\tilde{p}) d\tilde{p} \\ &= \max_{p \in U_n} \int_{V_{n+1}} \varphi(p - \tilde{p}) \zeta(\tilde{p}) d\tilde{p} + \max_{p \in U_n} \int_{\mathbb{C}V_{n+1}} \varphi(p - \tilde{p}) \zeta(\tilde{p}) d\tilde{p} \\ &\leq \frac{\max_{p \in V_n} \psi(p)}{\min_{p \in V_n} |f'(p)|} + \varphi(d(U_n, \mathbb{C}V_{n+1})). \end{aligned} \quad (35)$$

Analogously, we find for $n \geq 0$ that

$$\max_{p \in V_{n+1}} \psi(p) \leq \frac{\max_{p \in U_n} \psi(p)}{\min_{p \in U_n} |f'(p)|} + \varphi(d(V_{n+1}, \mathbb{C}U_{n+1})). \quad (36)$$

From these inequalities, we already obtain a result for $n = 0$. For if we recall that

$$\min_{p \in U_0} |f'(p)| \geq \lambda,$$

inequality (36) implies that

$$\max_{p \in U_0} \psi(p) \leq \max_{p \in V_1} \psi(p) \leq \frac{\max_{p \in U_0} \psi(p)}{\lambda} + \varphi(\Delta),$$

and consequently that

$$\max_{p \in U_0} \psi(p) \leq \frac{\lambda}{\lambda - 1} \varphi(\Delta).$$

Let $N > 0$ be such that $U_{N-1} \subset [-A, A] \subset U_N$. Note that N an upper bound of N can be given in terms of $\delta, \lambda, p_*, \kappa$ and A . Let moreover $d_* = \min_{0 \leq n \leq N} d(\mathbb{C}V_n, U_n)$ and recall that $\kappa < \min_{p \in (-p_*, p_*)} |f'(p)|$. By recursively applying estimates (35) and (36), it follows that

$$\begin{aligned} \max_{p \in U_N} \psi(p) &\leq \kappa^{-2N} \max_{p \in U_0} \psi(p) + \varphi(d_*) \sum_{j=0}^{2N} \kappa^{-j} \\ &\leq \kappa^{-2N} \frac{\lambda}{\lambda - 1} \varphi(\Delta) + \varphi(d_*) \sum_{j=0}^{2N} \kappa^{-j} \\ &\leq \kappa^{-2N} \left(\frac{\lambda}{\lambda - 1} + \frac{1}{1 - \kappa} \right) \varphi(\min\{d_*, \Delta\}), \end{aligned}$$

recalling from (22) the inequality $\varphi(d) \leq \frac{2}{3} \sigma^2 / d^3$. Choosing $\sigma_0 > 0$ sufficiently small then yields the lemma. $\square$

Lemma 6. *For any given constants $C, \bar{p} > 0$, there is a $\lambda_0 > 0$ such that the following holds. If the map $f: \mathbb{R} \rightarrow \mathbb{R}$ is differentiable, odd, strictly increasing, such that $f'(0) > \lambda_0$ and $f''(p) < 0$ for all $p > 0$, and if*

$$C^{-1} \leq -f^{(3)}(p) \leq C \quad \text{for all } |p| < \bar{p},$$

then the process

$$p_t = f(\beta^2 p_{t-1}) + \eta_t$$

has a nontrivial SCEE.

Proof. Introduce the stochastic variables $q_t = \beta^2 p_t$. Note that the process (q_t) has the same autocorrelation coefficients as (p_t) . The q_t satisfy

$$q_t = \beta^2 f(q_{t-1}) + \beta^2 \eta_t \quad (37)$$

for all t . Consider first the deterministic system $q \mapsto \beta^2 f(q)$. Note that $q = 0$ is a fixed point of this system, and putting $f'(0) = \lambda$ and $f^{(3)}(0) = -6c$, we have that

$$\beta^2 f(q) = \lambda \beta^2 q - c \beta^2 q^3 + \beta^2 q^5 r(q^2, \beta^2);$$

here r is a smooth function. For $0 < \beta^2 \lambda - 1 \ll 1$, there will be two other fixed points $\pm q_*$, solutions of the equation

$$\frac{\lambda \beta^2 - 1}{c \beta^2} = q^2 - \frac{q^4}{c} r.$$

Invoking the implicit function theorem, we obtain

$$q_* = \sqrt{\frac{\lambda \beta^2 - 1}{c \beta^2}} + \dots$$

Lemma 7. *Let assumptions A2 and A3 hold. We have for arbitrary $A > 0$ that*

$$\mathbb{P}(|p_t| > A) \geq \mathbb{P}(|\eta_t| > A).$$

Proof. Note that the invariant measure μ satisfies the integral equation

$$\mathrm{d}\mu(p) = \int Q(\tilde{p}, \mathrm{d}p) \mathrm{d}\mu(\tilde{p}) = \mathrm{d}p \int \varphi(p - F_{0,\beta}(\tilde{p})) \mathrm{d}\mu(\tilde{p})$$

and that hence

$$\mu([-A, A]) = \int_{-A}^A \left[\int_{-A}^{-A} + \int_{-A}^A + \int_A^\infty \varphi(p - F_{0,\beta}(\tilde{p})) \mathrm{d}\mu(\tilde{p}) \right] \mathrm{d}p. \quad (38)$$

To estimate the right hand side of this expression, the following lemma is needed.

Lemma 8. *Let $h(s) = \int_{-A}^A \varphi(p + s) \mathrm{d}p$. Then $h(0) \geq h(s)$ for every $s \geq 0$.*

Proof. By definition

$$h(s) - h(0) = \int_{-A+s}^{A+s} \varphi(p) \mathrm{d}p - \int_{-A}^A \varphi(p) \mathrm{d}p.$$

There are two cases to be distinguished: $s \leq 2A$ and $s > 2A$. For the latter case, since φ is nonincreasing on the positive real axis, we have $\varphi(p) \leq \varphi(-A + s)$ for all $p \in [-A + s, A + s]$, while $\varphi(p) \geq \varphi(A)$ for all $p \in [-A, A]$; hence

$$h(s) - h(0) \leq \varphi(-A + s) \int_{-A+s}^{A+s} dp - \varphi(A) \int_{-A}^A dp,$$

and since φ is nonincreasing, this is smaller than or equal to 0; this proves the lemma in this case.

In the case that $s \leq 2A$, note that

$$\begin{aligned} h(s) - h(0) &= \int_A^{A+s} \varphi(p) dp - \int_{-A}^{-A+s} \varphi(p) dp = \int_A^{A+s} \varphi(p) dp - \int_{A-s}^A \varphi(p) dp \\ &\leq \varphi(A) \int_A^{A+s} dp - \varphi(A) \int_{A-s}^A dp = 0, \end{aligned}$$

where the inequality follows from the fact that $\varphi(p) \geq \varphi(A)$ for $p \in [A - s, A]$ if $s \leq 2A$, and that φ is nonincreasing on the positive real axis. $\square$

Returning to the proof of Lemma 7, consider the third term of the right hand side of (38). Changing the order of integration (note that all arguments are positive functions), and using Lemma 8, it can be estimated as follows:

$$\int_A^\infty \left(\int_{-A}^A \varphi(p - F_{0,\beta}(\tilde{p})) dp \right) d\mu(\tilde{p}) \leq \int_A^\infty \left(\int_{-A}^A \varphi(p) dp \right) d\mu(\tilde{p}).$$

The other terms are treated similarly. Setting $q_A = \int_A^\infty \psi(p) dp$, it follows that

$$\begin{aligned} \int_{-A}^A d\mu(p) &\leq q_A \int_{-A}^A \varphi(p) dp + (1 - 2q_A) \int_{-A}^A \varphi(p) dp + q_A \int_{-A}^A \varphi(p) dp \\ &\leq \int_{-A}^A \varphi(p) dp \end{aligned}$$

This proves the lemma. $\square$

We are now ready to prove Theorems 8 and 9.

Proof of Theorem 8

Recall from Lemma 2 that the first autocorrelation coefficient ρ_1 satisfies

$$\frac{\rho_1}{F_{0,\beta}(A)} \geq \frac{A}{\mathbb{E} p_t^2} \mathbb{P}(|p_t| \geq A),$$

for any $A > 0$. Using Lemmas 3, 4 and 7, as well as the fact that $\sqrt{\mathbb{E}p_t^2} \leq \sqrt[4]{\mathbb{E}p_t^4}$, this is seen to imply that

$$\frac{\rho_1}{F(\beta^2 A)} \geq A \left(\frac{1-a}{\frac{b}{A} + \frac{\tau}{A}} \right)^2 \left(1 - \frac{A^2}{\tau^2} \right)^2 = A \left(\frac{1-a}{1 + \frac{b}{\tau}} \right)^2 \frac{A^2}{\tau^2} \left(1 - \frac{A^2}{\tau^2} \right)^2,$$

for $A^2 < \mathbb{E}\eta_t^2$. Note that for $\beta_0^2 < c/A$, we have $|F(\beta_0^2 A)| > \lambda\beta_0^2 A$. Choose $A^2 = \tau^2/2$.

We have to consider the cases of increasing and decreasing F separately; for decreasing F , we obtain

$$\rho_1 \leq -\frac{\lambda\beta_0^2}{4} \left(\frac{1-a}{1 + \frac{b}{\tau}} \right)^2.$$

This is smaller than β_0 for λ sufficiently large. On the other hand, since p_{t+1} and p_t are not perfectly negatively correlated, we have $\rho_1(F_{0,-1}) > -1$; hence there is a $\beta_* < 0$ such that $\rho_1(F_{0,\beta_*}) = \beta_*$. For increasing F the argument is similar, with only the signs reversed.

The theorem is proved. $\square$

Proof of Theorem 9

Consider first a parameterised family f_μ of the form

$$f_\mu(p) = -p - \mu p + cp^3 + \mathcal{O}(p^5),$$

which has the properties that $f_\mu(-p) = -f_\mu(p)$ and $f_\mu''(p) > 0$ for all $p > 0$. Let $p_*(\mu)$ solve $f_\mu(p_*) = p_*$; we have that

$$p_*(\mu) = \sqrt{\mu/c} + \mathcal{O}(\mu^{3/2}).$$

The tangent to the graph of f at $p = p_*$ is the graph of $\ell(p) = -ap - b$, where $a = 1 - 2\mu + \mathcal{O}(\mu^2)$ and $b = 2\mu(\mu/c)^{1/2} + \mathcal{O}(\mu^{5/2})$. As f was assumed to be convex on the positive real axis, it follows that there is a constant $C > 0$ such that

$$|f(p)| \leq (1 - 2\mu + C\mu^2)p + 2\mu \left(\frac{\mu}{c} \right)^{1/2} + C\mu^{5/2}.$$

Setting $\tau = \sqrt[4]{\mathbb{E}p_t^4}$ and assuming that $\tau < \tau_0 = C\mu^{5/2}$, it follows from Lemma 3 that

$$\mathbb{E}p_t^2 \leq \left(\frac{2\mu(\mu/c)^{1/2} + \tau + \mathcal{O}(\mu^{5/2})}{2\mu + \mathcal{O}(\mu^2)} \right)^2 \leq \frac{\mu}{c} + \mathcal{O}(\mu^2).$$

Choose $\delta > 0$ arbitrarily, and set $A = p_*(1 - \delta)$; note that $f(A) < -A$. By Lemma 5, there is $\sigma_0 > 0$ such that for $\sigma < \sigma_0$, we have $\mathbb{P}(|p_t| < A) < \delta$. Taking $\bar{\tau} = \min\{\sigma_0, \tau_0\}$, Lemma 2 now yields that

$$\rho_1 \leq -\frac{A^2}{\mu/c + \mathcal{O}(\mu^2)}(1 - \delta) \leq -(1 - \delta)^3 + \mathcal{O}(\mu).$$

Let $F(p) = -\lambda p + \tilde{c}p^3 + \mathcal{O}(p^5)$, and put $\beta = -1 + \eta$. Then $F_{0,\beta}(p) = F(\beta^2 p)$ and

$$F_{0,\beta}(p) = -\lambda(1 - \eta)^2 p + \tilde{c}(1 - \eta)^3 + \mathcal{O}(p^5).$$

Choose $\eta > 0$ such that $\lambda(1 - \eta)^2 = 1 + \mu = 1 + \eta^2$. Then, for $\delta > 0$ sufficiently small, and for $\lambda > 1$ sufficiently close to 1:

$$\rho_1(F_{0,\beta}) \leq -(1 - \delta)^3 + \mathcal{O}(\eta^2) < -1 + \eta = \beta.$$

Since $\rho_1(F_{0,-1}) > -1$, this proves the existence of a first order SCEE also in this case.

The argument for a family of the form

$$f_\mu(p) = p + \mu p - cp^3 + \mathcal{O}(p^5)$$

is similar and is therefore omitted. □

References

- Arthur, W. B., Ermoliev, Y. M., & Kaniovski, Y. (1987). Path-dependent processes and the emergence of macro-structure. *European Journal of Operational Research*, 30, 294–303.
- Barnett, W. A., Gallant, A. R., Hinich, M. J., Jungeilges, J. A., Kaplan, D. T., & Jensen, M. J. (1998). A single-blind controlled competition among tests for nonlinearity and chaos. *Journal of Econometrics*, 82, 157–192.
- Böhm, V., & Wenzelburger, J. (1999). Expectations, forecasting, and perfect foresight – A dynamical systems approach. *Macroeconomic Dynamics*, 3, 167–186.
- Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). *Time series analysis. Forecasting and control* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.
- Branch, W., & McGough, B. (2005). Consistent expectations and misspecification in stochastic non-linear economies. *Journal of Economic Dynamics and Control*, 29, 659–676.
- Branch, W. A. (2006). Restricted perceptions equilibria and learning in macroeconomics. In D. Colander (Ed.), *PostWalrasian macroeconomics: Beyond the dynamic stochastic general equilibrium model*. New York: Cambridge University Press, pp. 135–160.
- Bray, M. M. (1982). Learning, estimation, and the stability of rational expectations. *Journal of Economic Theory*, 26, 318–339.
- Bray, M. M., & Savin, N. E. (1986). Rational expectations equilibria, learning, and model specification. *Econometrica*, 54, 1129–1160.
- Brock, W. A., & Dechert, W. D. (1991). Nonlinear dynamical systems: Instability and chaos in economics. In W. Hildenbrand & H. Sonnenschein (Eds.), *Handbook of mathematical economics IV*. Amsterdam: North-Holland.

- Brock, W. A., & Hommes, C. H. (1997). A rational route to randomness. *Econometrica*, 65, 1059–1095.
- Brock, W. A., & Hommes, C. H. (1998). Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control*, 22, 1235–1274.
- Brock, W. A., Hsieh, D. A., & LeBaron, B. (1991). *Nonlinear dynamics, chaos and instability: Statistical theory and economic evidence*. Cambridge, MA: MIT Press.
- Bullard, J. (1994). Learning equilibria. *Journal of Economic Theory*, 64, 468–485.
- Bullard, J., & Duffy, J. (1998). Learning and the stability of cycles. *Macroeconomic Dynamics*, 2, 22–48.
- Bullard, J., & Duffy, J. (2001). Learning and excess volatility. *Macroeconomic Dynamics*, 5(2), 272–302.
- Devaney, R. L. (1989). *An introduction to chaotic dynamical systems* (2nd ed.). Redwood City, NJ: Addison-Wesley.
- Evans, G. W., & Honkapohja, S. (1994). On the local stability of sunspot equilibria under adaptive learning rules. *Journal of Economic Theory*, 64, 142–161.
- Evans, G. W., & Honkapohja, S. (1995). Local convergence of recursive learning to steady states and cycles in stochastic nonlinear models. *Econometrica*, 63, 195–206.
- Evans, G. W., & Honkapohja, S. (2001). *Learning in macroeconomics*. Princeton, NJ: Princeton University Press.
- Gardini, L., Hommes, C., Tramontana, F., & de Vilder, R. (2009). Forward and backward dynamics in implicitly defined overlapping generations models. *Journal of Economic Behavior and Organization*, 71, 110–129.
- Grandmont, J.-M. (1985). On endogenous competitive business cycles. *Econometrica*, 53, 995–1045.
- Grandmont, J.-M. (1998). Expectations formation and stability in large socio-economic systems. *Econometrica*, 66, 741–781.
- Guesnerie, R. (2001). *Assessing rational expectations: Sunspot multiplicity and economic fluctuations*. Cambridge, MA: MIT Press.
- Hommes, C. H. (1998). On the consistency of backward-looking expectations: The case of the cobweb. *Journal of Economic Behavior and Organization*, 33, 333–362.
- Hommes, C. H., & Rosser, J. B. (2001). Consistent expectations equilibria and complex dynamics in renewable resource markets. *Macroeconomic Dynamics*, 5, 180–203.
- Hommes, C. H., & Sorger, G. (1998). Consistent expectations equilibria. *Macroeconomic Dynamics*, 2, 287–321.
- Hommes, C. H., & Zhu, M. (2011). *Behavioral learning equilibria* (Technical Report, CeNDEF Working Paper). University of Amsterdam.
- Jungeilges, J. (2007). On chaotic consistent expectations equilibria. *Analyse und Kritik*, 29, 269–289.
- Kurz, M. (1994). On rational belief equilibria. *Economic Theory*, 4, 859–876.
- Lasota, A., & Mackey, M. C. (1985). *Probabilistic properties of deterministic systems*. Cambridge, UK: Cambridge University Press.
- Lasota, A., & Mackey, M. C. (1994). *Applied Mathematical Sciences: Vol. 97. Chaos, fractals and noise. Stochastic aspects of dynamics* (2nd ed.). New York: Springer.
- Lucas, R. E. (1971). Econometric testing of the natural rate hypothesis. In O. Eckstein (Ed.), *The econometrics of price determination Conference*. Washington DC: Board of Governors of the Federal Reserve System and Social Science Research Council.
- Marcet, A., & Nicolini, J. P. (2003). Recurrent hyperinflations and learning. *American Economic Review*, 93, 1476–1498.
- Marcet, A., & Sargent, T. J. (1989). Convergence of least squares learning in mechanisms in self referential linear stochastic models. *Journal of Economic Theory*, 48, 337–368.
- Marimon, R., Spear, S. E., & Sunder, S. (1993). Expectationally driven market volatility: An experimental study. *Journal of Economic Theory*, 61, 74–103.
- Medio, A., & Raines, B. (2007). Backward dynamics in economics. the inverse limit approach. *Journal of Economic Dynamics and Control*, 31, 1633–1671.

- Muth, J. (1961). Rational expectations and the theory of price movements. *Econometrica*, 29, 315–335.
- Nerlove, M., Grether, D. M., & Carvalho, J. L. (1979). *Analysis of economic time series: A synthesis*. New York: Academic Press.
- Sakai, H., & Tokumaru, H. (1980). Autocorrelations of a certain chaos. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 28, 588–590.
- Sargent, T. J. (1993). *Bounded rationality in macroeconomics*. Oxford, UK: Clarendon Press.
- Sargent, T. J. (1999). *The conquest of American inflation*. Princeton, NJ: Princeton University Press.
- Schönhofer, M. (1999). Chaotic learning equilibria. *Journal of Economic Theory*, 89, 1–20.
- Sögner, L., & Mitlöhner, H. (2002). Consistent expectations equilibria and learning in a stock market. *Journal of Economic Dynamics and Control*, 26, 171–185.
- Sorger, G. (1998). Imperfect foresight and chaos: An example of a self-fulfilling mistake. *Journal of Economic Behavior and Organization*, 33, 363–383.
- Stokey, N. L., & Lucas, R. E. (1989). *Recursive methods in economic dynamics*. Cambridge, MA: Harvard University Press.
- Tuinstra, J. (2003). Beliefs equilibria in an overlapping generations model. *Journal of Economic Behavior and Organization*, 50, 145–164.
- Tuinstra, J., & Wagener, F. (2007). On learning equilibria. *Economic Theory*, 30, 493–513.
- Woodford, M. (1990). Learning to believe in sunspots. *Econometrica*, 58, 277–307.