Global convergence of adaptive learning in models of pure exchange[1]

Jan Wenzelburger
Fakultät für Wirtschaftswissenschaften, Universität Bielefeld,
Postfach 100 131, D-33501 Bielefeld, Germany (E-mail: jwenzelb@wiwi.uni-bielefeld.de)

Received: revised:

Summary. This paper develops an adaptive learning scheme for a standard version of the OLG model with pure exchange. Perfect forecasting rules which generate perfect-foresight orbits are approximated by cubic spline functions. These approximations are successively constructed using historical data only. Trajectories generated by this scheme converge to perfect-foresight orbits globally for all initial conditions. This result holds for all parameterizations guaranteeing the existence of a monetary steady state and hence is independent of consumers' savings behavior. It generalizes to all one-dimensional models of the Cobweb type.

Keywords and Prases: Learning Dynamics, rational expectations, bounded rationality.

JEL Classification Numbers: D83, D84.

Forthcoming in *Economic Theory*

[1] This research was supported by the Deutsche Forschungsgemeinschaft under contracts No. Bo 635/9-1 and No. We 2287/2-1 and carried out primarily during a visit to Cornell University. I am indebted to Tapan Mitra and Karl Shell for their generous hospitality. I am grateful to Volker Böhm for inspiring this research and I would like to thank Jochen Jungeilges and Thorsten Pampel for valuable discussions.

1 Introduction

The recent economic literature has used the notion of an adaptive learning scheme to justify the rational expectations hypothesis as a long-run concept of agents who succeeded in eliminating all systematic forecast errors. This interpretation was first supported by Bray (1982), Fourgeaud, Gourieroux & Pradel (1986), Marcet & Sargent (1989a), and others, who provide conditions for linear models under which adaptive learning schemes converge to an equilibrium under rational expectations. Unfortunately, however, in economies whose evolution is driven by non-linear maps, the question which learning schemes generate forecasts converging to such rational expectations equilibria remains to a large extent unresolved. This seems to hold equally true for all adaptive schemes such as Bayesian learning, where agents update beliefs according to Bayes's rule (cf. Blume & Easley (1993, 1998) and references therein), or the ordinary least squares scheme (OLS). The latter is popular among many theorists in the field (cf. Evans & Honkapohja 1999).

For the standard OLG model of pure exchange, Bullard (1994) and Schönhofer (1999, 2001) showed that ordinary least squares learning may generate forecasts which do not converge to perfect-foresight orbits of the system. Moreover, the long-run behavior of the system may be irregular and chaotic and for this reason may differ substantially from the dynamics under perfect foresight. As shown in Schönhofer (2001), in some cases the forecast errors associated with such attractors, also called *learning equilibria*, may have vanishing sample means and confirm the hypothesized autocorrelation structure. Hommes & Sorger (1998) and Hommes (1998) call such learning equilibria *consistent*. They argue that agents with limited statistical tools do not identify systematic forecast errors in these equilibria. Hence, agents would see no reason to revise their learning scheme.

This paper argues that the conclusions from these findings are misleading. Apart from the problem of existence of perfect foresight in a particular environment, it is commonplace that the quality of forecasts hinges essentially on the 'rationality' of an economic agent, that is, on both his or her theoretical knowledge of possible economic scenarios and on available statistical tools. On the one hand, it is not surprising that economic agents in deterministic non-linear models may fail in finding orbits with perfect foresight when relying on techniques originally designed for linear stochastic models such as OLS. This observation is supported by Jungeilges (2000) who shows that the underlying linearity hypothesis using OLS in exchange economies can already be rejected within finite time by means of *linear* econometric techniques. On the other hand, the available structural information on a particular market is used insufficiently, when applying linear consistency notions only. For example, it is well-known that the market mechanism determining the actual inflation factors in a stationary OLG model under market clearing is of the Cobweb type, that is, actual interest factors of the economy depend on expected inflation factors only and not on previous interest factors, cf. Böhm & Wenzelburger (1999, 2000a). For these reasons, agents using OLS in nonlinear deterministic environments simply do not use the appropriate statistical tools.

In contrast to the findings of Bullard (1994), Schönhofer (1999), or Marcet & Sargent (1989b), the present paper supports the original intention of the learning literature (see e.g. Blume & Easley 1982) in showing that a forecasting agency in a stationary OLG exchange economy may indeed find the perfect-foresight orbits of the economy. The forecasting agency modeled in the present paper is boundedly rational in the sense of Sargent (1993). It uses well-known theoretical results and, in addition, is aware of the basic market mechanism underlying an OLG exchange economy without knowing consumers' savings behavior. This structural knowledge is comprised in the concept of an *error function* which arises naturally when carefully distinguishing between an *economic law* describing the basic market mechanism of an economy and a *forecasting rule*. The zero contour of an error function determines the so-called perfect forecasting rules which generate perfect-foresight orbits (Böhm & Wenzelburger 1999, 2000a).

The main innovations of this paper are the following. First, by exploiting the Cobweb structure of the economy, it is shown that it is always possible to find the monetary steady state even in cases in which that steady state is unstable under the perfect foresight dynamics. The induced learning dynamics is either globally stable, because the perfect-foresight dynamics is stable or stabilizes an otherwise unstable steady state. The generalization of this learning scheme to one-dimensional models of the Cobweb type extends earlier results (see e.g. Chatterji & Chattopadhyay 2000), because global convergence is obtained under much weaker conditions.

Second, using the structural knowledge provided by an error function this paper provides an algorithm which allows an agency to construct an arbitrarily precise approximation of a (locally) perfect forecasting rule using observed historical data only. It is shown that this algorithm converges globally for all initial conditions. Contrary to Kelly & Shorish (2000), who present a local convergence result, we neither assume the existence of an invariant set where the dynamics under rational expectations or perfect foresight takes place nor its somewhat unrealistic a-priori knowledge to a forecasting agency. This demonstrates that an agency endowed with the appropriate tools can generate forecasts which are more precise than consistent expectations equilibria, since forecast errors vanish pointwise in the long run. As a consequence, an agency which suspects a Cobweb type environment should replace any scheme which generates non-vanishing forecast errors.

The paper is organized as follows. Section 2 briefly reviews the stationary OLG exchange model. The notion of an error function is introduced in Section 3. Section 4 provides a discussion of the learning dynamics. Informational constraints are discussed in Section 5, our adaptive learning schemes are investigated in Sections 6 and 7, conclusions are given in Section 8. An appendix includes an outline of possible generalizations of the approach presented in the main text.

2 The Model

Consider a standard version of the overlapping generations model with one non-storable commodity per period and fiat money as the only store of value between periods. There will be neither growth of the population nor production. Given the usual assumption of price taking behavior of all generations, young agents need to transfer purchasing power from the first to the second period of their lifes. To avoid the problem of heterogenous beliefs, let θ^e_{t+1} denote the common expected gross inflation rate for period $t+1$ on which all members of the young generation in t base their decisions. We assume that a forecasting agency is in charge of issuing these forecasts and that this agency knows the basic market mechanism. However, the agency has no information concerning the households' savings behavior. Under the standard two-period optimizing behavior of a consumer h born in period t, her optimal consumption plan given an initial endowment of goods w^h_1 and w^h_2, respectively, is defined by a savings function $s^h : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$. $s^h(\theta^e_{t+1})$ is the amount saved (and supplied to the market) in period t and $w^h_2 + s^h(\theta^e_{t+1})/\theta^e_{t+1}$ is the amount expected to be consumed in period $t+1$.

A government is infinitely lived and consumes g_t at time t. At time $t = 1$ the government issues $M_1 > 0$ units of currency and agents save by holding this currency. Given a non-negative price level p_t, the government finances its consumption by creating additional currency through the process $M_t = \gamma M_{t-1}$, where $\gamma > 1$ is the gross rate of currency growth. The government's budget constraint is given by $M_t - M_{t-1} = g_t p_t$. The policy rule γ is chosen by the monetary authorities such that the government expenditures g_t become endogenous.

Market clearing on the goods market in any period t requires that the real savings of young consumers, which defines the amount supplied, has to be equal to the demand of the old generation and the government, which is equal to the real purchasing power of the money acquired in the previous period. In other words, aggregate savings S has to be equal to real money balances

$$\frac{M_t}{p_t} = S(\theta^e_{t+1}) := \sum_{h=1}^{H} s^h(\theta^e_{t+1})$$

for all times t. The actual inflation rate $\theta_t = p_t/p_{t-1}$ is then determined by

$$\theta_t = F(\theta^e_t, \theta^e_{t+1}) := \frac{\gamma S(\theta^e_t)}{S(\theta^e_{t+1})}, \tag{1}$$

as long as the expected inflation factors are such that aggregate savings is different from zero.[2] Let $0 < \gamma < \theta_{aut}$, where we assume that the autarky factor θ_{aut}, given by

[2]Notice that inflation factors depend in a deterministic manner on endowments even though these endowments could be governed by a stochastic process, as in Kelly & Shorish (2000). In this sense, inflation factors remain deterministic from the household's perspective. In order to keep the exposition as simple as possible, we abstract from stochastic endowments.

$S(\theta_{aut}) = 0$, is uniquely determined. Thus the monetary steady state γ exists. Observe that from a sequential view point, the forecast θ^e_{t+1} for the inflation rate θ_{t+1} has to be picked before the actual trading takes place and thus prior to the realization of θ_t. Moreover, M_t denotes the final money balance of generation t after trading. The map $F : \mathbb{R}^2_+ \longrightarrow \mathbb{R}_+$ given in (1) includes a two-step-ahead forecast and for this reason defines an economic law with an expectational lead, see Böhm & Wenzelburger (2000a). Since only expected inflation factors enter F as an argument, the economic law (1) has the structure of a Cobweb model. The goods price of period t generated by (1) is given by

$$p_t = \frac{\gamma S(\theta^e_t)}{S(\theta^e_{t+1})} p_{t-1}.$$

3 Perfect forecasting rules

In order to address the question which forecasting rules generate perfect foresight along orbits of the economy, we first consider the error function associated with the economic law (1). For an arbitrary period, let θ^e_{old} and θ^e_{you} denote the forecasts of the old and the young generation, respectively. The forecast error for the old generation is then given by the error function

$$e_F : \mathbb{R}^2_+ \longrightarrow \mathbb{R}, \quad (\theta^e_{old}, \theta^e_{you}) \longmapsto \frac{\gamma S(\theta^e_{old})}{S(\theta^e_{you})} - \theta^e_{old}. \tag{2}$$

The function e_F describes all possible forecast errors, independently of which forecasting rule or learning scheme has been used to obtain the forecasts. Geometrically, the graph of e_F is a surface over the $(\theta^e_{old}, \theta^e_{you})$-plane. The zero-contour of e_F describes the loci of all forecasts with vanishing forecast errors.[3] It follows from $S(\gamma) > 0$ that $e_F(\gamma, \gamma) = 0$ which means that the monetary steady state (γ, γ) belongs to the zero contour of e_F. The error function (2) for an OLG exchange model has some general features which are stated in the next proposition. Let $E_S(\theta) := S'(\theta)\theta/S(\theta)$ denote the elasticity of the savings function.

Proposition 3.1 *Let S be continuously differentiable and $S(\gamma) > 0$. Then the error function e_F for an OLG exchange model has the following properties:*

(i) e_F is linear along the 45°-degree line,

$$e_F(\theta^e, \theta^e) = \gamma - \theta^e \quad \textit{for all} \quad \theta^e \in \mathbb{R}_+ \quad \textit{with} \quad S(\theta^e) > 0;$$

[3] In the case of varying endowments the graph of the error function is no longer a surface. It is a higher dimensional geometric object for which all notions introduced in this section have corresponding analogues, see Böhm & Wenzelburger (1999, 2000a).

(ii) Its gradient along the zero contour satisfies

$$\operatorname{grad} e_F(\theta^e_{old}, \theta^e_{you})_{|e_F = 0} = \left(E_S(\theta^e_{old}) - 1, \frac{-S'(\theta^e_{you})(\theta^e_{old})^2}{\gamma S(\theta^e_{old})} \right),$$

where $E_S(\theta^e_{old}) < 1$ for $\theta^e_{old} \neq \theta_{aut}$;

(iii) The zero contour intersects the $45°$-degree line of $\mathbb{R}^2_+$ transversally in (γ, γ) or, equivalently,

$$\operatorname{grad} e_F(\gamma, \gamma) \neq \lambda(1, -1) \quad \textit{for all} \quad \lambda \in \mathbb{R}.$$

Proof. The first statement is obvious from the definition of e_F. The second statement follows from the definition of a gradient and the fact that $\gamma S(\theta^e_{old}) = \theta^e_{old} S(\theta^e_{you})$ along the zero contour. The Slutsky matrix for an individual savings function s^h is negative semi-definite if and only if

$$s^h(\theta^e) - s^{h\prime}(\theta^e)\theta^e \geq s^h(\theta^e)^2.$$

Therefore

$$S(\theta^e) - S'(\theta^e)\theta^e \geq \sum_{h=1}^{H} s^h(\theta^e)^2$$

and the first component of the gradient in *(ii)* is negative as long as $\theta^e \neq \theta_{aut}$. The third statement follows directly from *(ii)*. *Q.E.D.*

Proposition 3.1 implies that the zero contour of the error function has a unique intersection point with the diagonal of the $(\theta^e_{old}, \theta^e_{you})$-plane which is the monetary steady state. The surface has a singularity at a possible autarky inflation factor where aggregate savings is zero. The gradient of the zero contour shows in the direction of increasingly positive forecast errors. Conversely, the negative gradient shows in the direction of increasingly negative forecast errors. The second statement in Proposition 3.1 says that the gradient of the zero contour always points in the direction of the θ^e_{you}–axis.

It follows from the very definition of the zero contour of the error function (2) that in order to obtain perfect foresight for any previously determined forecast θ^e_{old}, a new forecast θ^e_{you} has to be chosen such that the forecast error $e_F(\theta^e_{old}, \theta^e_{you})$ vanishes. This implies in particular that only forecasting rules of the functional form

$$\psi : \mathbb{R}_+ \longrightarrow \mathbb{R}_+, \quad \theta^e_{you} = \psi(\theta^e_{old}) \tag{3}$$

have a chance of generating perfect foresight along orbits of the system. More precisely, let $\epsilon \geq 0$ be an arbitrary positive number. Following Böhm & Wenzelburger (2000a), forecasting rules of the form (3) which satisfy $|e_F(\theta^e, \psi(\theta^e))| \leq \epsilon$ for all θ^e in an open subset U of $\mathbb{R}_+$ are called *locally ϵ–perfect*. If $U = \mathbb{R}_+$, then ψ is called *globally ϵ–perfect*. To simplify language, locally (globally) 0-perfect forecasting rules are called locally (globally) perfect.

The following proposition shows that locally perfect forecasting rules for the OLG exchange model always exist, whereas globally perfect forecasting rules exist only under very restrictive assumptions.[4]

Proposition 3.2 *Let S be continuously differentiable and $0 < S(\gamma) < w_1$, where $w_1 := \sum_{h=1}^{H} w_1^h$ is aggregate endowment of the young generation and $S'(\gamma) \neq 0$. Then there exists an open neighborhood U of γ and a forecasting rule $\psi_\star$ given by*

$$\psi_\star(\theta^e) = S^{-1}(\gamma S(\theta^e)/\theta^e), \quad \theta^e \in U \tag{4}$$

which is locally perfect on U. If in addition, S is strictly monotonically increasing in θ^e and $\gamma S(\theta^e)/\theta^e < w_1$ for all $\theta^e \in \mathbb{R}_+$, then $\psi_\star$ is globally perfect on $\mathbb{R}_+$.

Proof. As noticed above $e_F(\gamma, \gamma) = 0$. Since $S'(\gamma) \neq 0$, we have $\partial_2 e_F(\gamma, \gamma) \neq 0$ and the equation $e_F(\theta^e_{old}, \theta^e_{you}) = 0$ satisfies the conditions of the implicit function theorem. This shows that $\psi_\star$ given in (4) is well-defined in a neighborhood U of γ. If, in addition, S is strictly monotonically increasing and $\gamma S(\theta^e)/\theta^e < w_1$ for all $\theta^e \in \mathbb{R}_+$, then $\psi_\star$ is well-defined globally on $U = \mathbb{R}_+$ and hence globally perfect on $\mathbb{R}_+$.

Q.E.D.

Corollary 3.3 $\psi'_\star > 0$ *on* U *if and only if* $S' < 0$ *on* U.

The corollary follows from the Slutsky conditions for the individual savings functions. Proposition 3.2 states that perfect forecasting rules depend exclusively on previous forecasts and not on observed states of the economy. Geometrically, the graph of a locally perfect forecasting rule $\psi_\star$ is contained in the zero contour of the error function. If youthful and old-age consumption are normal goods or if the aggregate savings function is non-increasing, then the term $S(\theta^e)/\theta^e$ will become unbounded for sufficiently small θ^e. In these cases globally perfect forecasting rules will not exist. Most preferences with intertemporal substitution properties usually assumed in the OLG literature will therefore not allow for global perfect foresight, see Böhm & Wenzelburger (2000a).

We close this section by illustrating the concept of an error function and the concept of a locally perfect forecasting rule with two examples taken from Bullard (1994) and Schönhofer (1999). An example with multiple forecasting rules which are locally perfect is found in Appendix B.

Example 3.1 *The Cobb-Douglas case. Let each consumer be characterized by the same Cobb-Douglas utility function*

$$u(c_1, c_2) = \ln c_1 + \delta \ln c_2, \quad 0 < \delta \leq 1,$$

[4]Notice that the following conditions depend on endowments. Thus, locally ϵ-perfect forecasting rules depend (pointwise) on endowments.

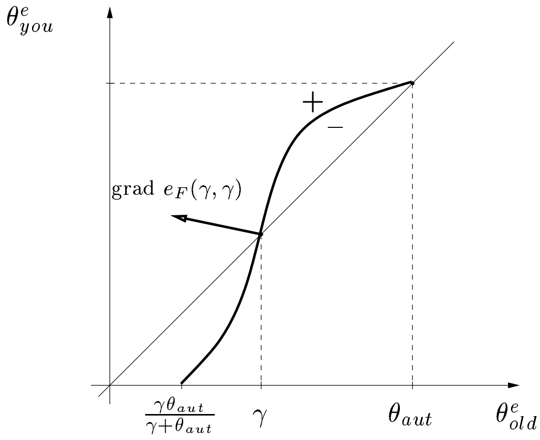

Figure 1: Zero contour for the Cobb-Douglas case.

and the same endowments $w_i^h = w_i^{h'}$, $i = 1, 2$ *for all* $h, h' \in H$. *The savings function is*

$$s^h(\theta^e) = \max\left\{0, \tfrac{1}{1+\delta}[\delta w_1^h - \theta^e w_2^h]\right\}, \quad \theta^e \in \mathbb{R}_+.$$

The common autarky inflation factor of all households is $\theta_{aut} = \frac{\delta w_1^h}{w_2^h}$, *such that aggregate savings becomes*

$$S(\theta^e) = \begin{cases} \frac{1}{1+\delta}[\delta w_1 - \theta^e w_2] & \textit{if } \theta^e \in [0, \theta_{aut}) \\ 0 & \textit{otherwise} \end{cases}$$

with w_1 *and* w_2 *denoting aggregate endowments when young and when old, respectively. Let* $\gamma < \theta_{aut}$ *as before. In this case there exists a locally perfect forecasting rule* $\psi_\star$, *defined by*

$$\psi_\star(\theta^e) = \theta_{aut} - \gamma\left[\frac{\theta_{aut}}{\theta^e} - 1\right] \quad \textit{for} \quad \theta^e \in \left[\frac{\theta_{aut}}{1 + \theta_{aut}/\gamma}, \theta_{aut}\right]. \tag{5}$$

$\psi_\star$ *is uniquely determined and the domain of definition in (5) is maximal. Hence, a globally perfect forecasting rule does not exist, see Fig. 1.*

Example 3.2 *The CES case. Let each consumer be characterized by a CES utility function of the form*

$$u(c_1, c_2) = [c_1^\rho + (\delta\, c_2)^\rho]^{\frac{1}{\rho}}, \quad 0 < \delta \le 1,\ \rho < 1,\ \rho \ne 0.$$

Then aggregate savings becomes

$$S(\theta^e) = w_1 - \frac{w_1 + w_2\,\theta^e}{1 + (\delta^{-1}\theta^e)^{\frac{\rho}{\rho-1}}} \quad \theta^e \in \mathbb{R}_+,$$

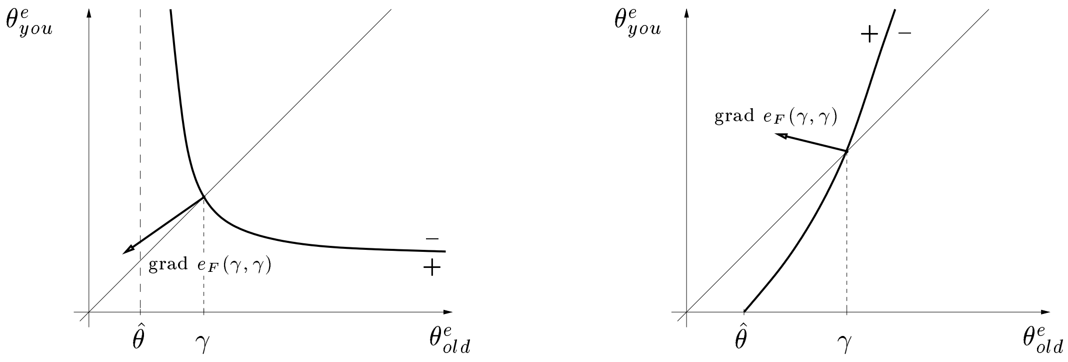

Figure 2: Zero contour for the CES case, $\rho < 0$ (left) and $\rho > 0$ (right).

where w_1 and w_2 denotes aggregate endowment when young and when old, respectively. Suppose $w_2 = 0$. Then there exists a locally perfect forecasting rule $\psi_\star$, defined by

$$\psi_\star(\theta) = \delta \left(\frac{\gamma \theta^{\frac{\rho}{\rho-1}}}{\delta^{\frac{\rho}{\rho-1}}\theta + \theta^{\frac{\rho}{\rho-1}}[\theta - \gamma]} \right)^{\frac{\rho-1}{\rho}} \qquad \theta^e \in (\hat\theta, \infty), \tag{6}$$

where $\hat\theta$ is defined by $\delta^{\frac{\rho}{\rho-1}}\hat\theta^{\frac{1}{1-\rho}} + \hat\theta - \gamma = 0$. $\psi_\star$ is uniquely determined and the interval $(\hat\theta, \infty)$ is its maximal domain. Therefore, a globally perfect forecasting rule does not exist. This case is illustrated in Fig. 2 (left) for $\rho < 0$ and in Fig. 2 (right) for $\rho > 0$.

4 Forecasting dynamics

To discuss the dynamics of the inflation factors under a particular learning scheme, consider for a moment forecasting rules ψ of the simple form (3). Given the economic law F defined in (1), the evolution of the economy is then governed by the two-dimensional map

$$F_\psi : \mathbb{R}^2_+ \longrightarrow \mathbb{R}^2_+, \quad (\theta, \theta^e_{old}) \longmapsto \big(F(\theta^e_{old}, \psi(\theta^e_{old})), \psi(\theta^e_{old})\big),$$

such that $(\theta_t, \theta^e_{t+1}) = F_\psi(\theta_{t-1}, \theta^e_{t+1})$, $t \in \mathbb{N}$. Observe that this evolution is exclusively driven by the forecasting rule ψ implying that the dynamics of the economy is one-dimensional. In mathematical terms, the dynamics governed by F_ψ is topological conjugate to the dynamics of the forecasting rule alone which is given by

$$\psi : \mathbb{R}_+ \longrightarrow \mathbb{R}_+, \quad \theta^e_{old} \longmapsto \psi(\theta^e_{old}).$$

Hence, in the absence of exogenous noise, the dynamics in an OLG exchange economy stems from the expectations formation alone. This fact remains valid for any more complex forecasting rule and any learning scheme (such as OLS) which may depend on whatever variable an agent (or the modeler) may believe in.

Let $\psi_\star : U \to \mathbb{R}_+$ be a locally perfect forecasting rule which is defined on some open subset $U \subset \mathbb{R}_+$ and $V \subset U$ be some (forward-)invariant subset of $\psi_\star$, i.e. $\psi_\star(V) \subset V$. Then $\psi_\star$ when restricted to V defines a perfect-foresight dynamics for all initial conditions $\theta_0 \in V$, given by

$$F_{\psi_\star} : V \times V \longrightarrow V \times V, \quad (\theta, \theta^e) \longmapsto \big(\theta^e, \psi_\star(\theta^e)\big). \tag{7}$$

It is natural to assume $\gamma \in V$ and it follows from Proposition 3.1 that the unique positive fixed point of $F_{\psi_\star}$ with the perfect-foresight property is the monetary steady state $(\theta_\star, \theta^e_\star) = (\gamma, \gamma)$. Clearly, under perfect foresight $\theta^e_t = \theta_t$ for all times t and the dynamics of (7) is equivalent to the dynamics generated by the perfect forecasting rule $\psi_\star$ alone when restricted to V. For this reason, $\psi_\star$ when restricted to an invariant set V satisfies

$$F\big(\psi_\star(\theta_t), \psi_\star(\psi_\star(\theta_t))\big) = \psi_\star(\theta_t), \quad t \in \mathbb{N},\ \theta_0 \in V$$

and hence defines a *functional rational expectations equilibrium* in the sense of Spear (1988).

Observe that by Proposition 3.2 the domain U of a locally perfect forecasting rule $\psi_\star$ a priori need not be invariant under $\psi_\star$. So V will most likely be a proper subset of U. If an invariant set V does not exist, then the perfect-foresight dynamics will eventually leave the region U where perfect foresight is possible. Figs. 1 and 2 (right) illustrate the well-known fact that the monetary steady state in the Cobb-Douglas and in the CES case is unstable. Moreover, the geometry of both figures shows that a non-trivial invariant set V does not exist. Hence for exchange economies, locally perfect forecasting rules are a more general object than functional rational expectations equilibria. Since perfect forecasting rules are completely determined by the economic fundamentals, so is the perfect-foresight dynamics.[5]

Let $\psi_\star : V \longrightarrow V$ define a structurally stable perfect-foresight dynamics. More precisely, according to Devaney (1989, p. 54) $\psi_\star$ is *structurally stable*, if there exists $\delta > 0$ such that whenever

$$|\psi_\star(\theta) - \psi(\theta)| < \delta, \quad \theta \in V$$

for some $\psi : V \longrightarrow V$, then $\psi_\star$ is topologically conjugate to ψ. The conjugacy of the two maps $\psi_\star$ and ψ essentially means that both maps induce the same qualitative dynamics, see Devaney (1989, p. 53). The major economic implication of this observation is that a structurally stable perfect-foresight dynamics can be approximated with ϵ-perfect forecasting rules.

[5]Similar observations remain valid for the case of stochastic endowments. Since a young household must know its income when deciding on the amount to save, perfect foresight is possible even though endowments are stochastic.

5 Informational constraints

Consider a forecasting agency which is in charge of forecasting the future evolution of the economy and hence has to decide on θ^e. We assume here that *all* young consumers in period t share the belief in θ^e_{t+1} issued by the forecasting agency prior to their savings decisions. This excludes heterogenous beliefs and strategic behavior of consumers. This reflects the fact that the effect of strategic behavior of a single consumer can be neglected, if the number of households is large. In order to be credible for households, we assume that the forecasting agency itself has no strategic interest other than issuing good forecasts. The informational constraints faced by our forecasting agency can be described as follows.

Suppose that the agency in some arbitrary period τ has observed past prices and therefore knows all past real rates of return $\{\theta_t\}_{t=0}^{\tau-1}$ and all corresponding forecasts $\{\theta^e_t\}_{t=0}^{\tau}$ including the forecast for the current period. Recall that θ_τ is not available prior to the decision on $\theta^e_{\tau+1}$. Let the agency be aware of the basic market mechanism of the economy, that is, the basic structure of the economic law given by (1) but not the specific parameterization of (1). Hence, neither the preferences nor the savings behavior of young consumers are known to the agency. In other words, the forecasting agency has the concept of the error function without its concrete functional specification and for this reason is boundedly rational in the sense of Sargent (1993). The assumptions on the structural knowledge of a forecasting agency are summarized as follows.

Assumption 5.1 *The information of the forecasting agency encompasses the following:*

(i) the relevant variables of the economy are $(\theta^e_{old}, \theta^e_{you})$ and the market mechanism is of the form $(\theta^e_{old}, \theta^e_{you}) \mapsto \tilde{F}(\theta^e_{old}, \theta^e_{you})$;

(ii) there exists a monetary steady state which lies in some uncertainty interval $[\underline{\theta}, \overline{\theta}]$ *with $e_F(\underline{\theta}, \underline{\theta}) > 0$ and $e_F(\overline{\theta}, \overline{\theta}) < 0$.*

Proposition 3.1 *(i)* ensures the existence of $\underline{\theta}$ and $\overline{\theta}$ with the properties stated in Assumption 5.1, *(ii)*. This assumption can easily be removed such that initially no uncertainty interval is known. Since the error function is linear along the diagonal, it is easy to obtain bounds $\underline{\theta}$ and $\overline{\theta}$. Hence, a forecasting agency needs to know only the qualitative geometry of the error surface graph e_F given in Proposition 3.1. Roughly speaking, this amounts to knowing that the left hand side of the zero-contour line is the positive level and the right hand side the negative level.

For simplicity of exposition, we may therefore assume that the errors $e_F(\underline{\theta}, \underline{\theta})$ and $e_F(\overline{\theta}, \overline{\theta})$ at the boundaries of the uncertainty interval have been observed previously. Notice, however, that unlike Kelly & Shorish (2000), we do not assume the set $[\underline{\theta}, \overline{\theta}]$ to be forward-invariant under the perfect-foresight dynamics.

The key idea now is that the geometric shape of the error function becomes disclosed through the course of time. As shown below, this will allow an agency to improve estimates of the zero contour and hence to improve the quality of the forecasts as long as the market mechanism is time invariant and there are no structural breaks. If $\epsilon_t = \theta_t - \theta_t^e$ denotes the forecast error made in period t, then the point $(\theta_t^e, \theta_{t+1}^e, \epsilon_t)$ belongs to the graph of the error function

$$\operatorname{graph} e_F := \{(\theta_{old}^e, \theta_{you}^e, \epsilon) \in \mathbb{R}_+^2 \times \mathbb{R} \mid e_F(\theta_{old}^e, \theta_{you}^e) = \epsilon\}.$$

At the beginning of an arbitrary period, say τ, the sequence of points $\{(\theta_t^e, \theta_{t+1}^e, \epsilon_t)\}_{t=0}^{\tau-1}$ has been observed and may be visualized as points on the unknown graph of the error function. These points reveal information on the shape of the error function and thus on the location of its zero contour. In Appendix B we show that under Assumption 5.1, a learning scheme which generates complex attractors with non-vanishing forecast errors such as consistent expectations equilibria in the sense of Hommes & Sorger (1998) can be abandoned.

6 Adaptive learning of steady states

In this section we introduce two adaptive learning schemes which find the perfect-foresight monetary steady state. Schemes which approximate a locally perfect forecasting rule will be considered in the next section. There are many reasons for considering the former class. First of all, the perfect-foresight dynamics will most likely fail to exist globally on $\mathbb{R}_+$ as pointed out above. This implies that any successful learning scheme will have to localize a set on which perfect foresight is possible. This holds true in particular for cases in which the perfect-foresight dynamics may become complex as in Grandmont (1985). Such a set may most likely contain a steady state, which in the case of an exchange economy could be the monetary steady state. If this steady state is asymptotically stable under the perfect-foresight dynamics, then all successful learning will in long run end up in that steady state. In this case, a method searching directly for the steady state might be more efficient. On the contrary, if the monetary steady state is unstable under perfect foresight, then orbits with perfect foresight other than the trivial one may fail to exist or may have other unfavorable properties such as convergence to the non-monetary steady state. In either case it therefore is a good strategy to first localize the monetary steady state and from there on approximate a perfect forecasting rule in a second step.

A learning scheme which searches for the monetary steady state will have to search close to the diagonal of the $(\theta_{old}^e, \theta_{you}^e)$-plane. Consider linear forecasting rules of the form

$$\psi_L(\theta^e; \alpha) = \theta^e + \alpha, \tag{8}$$

where $\alpha \in \mathbb{R}$ is some parameter. $\alpha > 0$ will induce a sequence $\theta_t^e = \psi_L(\theta_{t-1}^e; \alpha)$, $t \in \mathbb{N}$ of increasing forecasts, whereas $\alpha < 0$ induces a sequence of decreasing forecasts. Let θ_0^e denote the expected inflation factors of the last generation which is assumed to lie in the uncertainty interval. Using Assumption 5.1, we introduce the following naive learning scheme.

Algorithm 6.1 *Let $\theta_0^e \in [\underline{\theta}, \overline{\theta}]$ be arbitrary, $k \geq 2$ be an integer and $\epsilon > 0$ be a given tolerance level.*

1. Orientation stage.
   (a) *If $|e_F(\theta_0^e, \theta_0^e)| \leq \epsilon$, then stop.*
   (b) *If $e_F(\theta_0^e, \theta_0^e) > \epsilon$, then update $\underline{\theta} = \theta_0^e$ and set $\alpha = [\overline{\theta} - \underline{\theta}]/k$.*
   (c) *If $e_F(\theta_0^e, \theta_0^e) < -\epsilon$, then update $\overline{\theta} = \theta_0^e$ and set $\alpha = -[\overline{\theta} - \underline{\theta}]/k$.*
2. Iteration stage. *Set $\theta_t^e = \psi_L^t(\theta_0^e; \alpha) = \theta_0^e + t\alpha$, with $\psi_L^t(\cdot; \alpha)$ denoting the t-th iterate of $\psi_L(\cdot; \alpha)$, until $\alpha\, e_F(\theta_{t-1}^e, \theta_t^e) \leq 0$ or $t = k - 1$. Let τ denote the first time for which this condition is satisfied.*
3. Updating stage. *Set $\theta_0^e = \theta_\tau^e$ and continue with stage 1.*

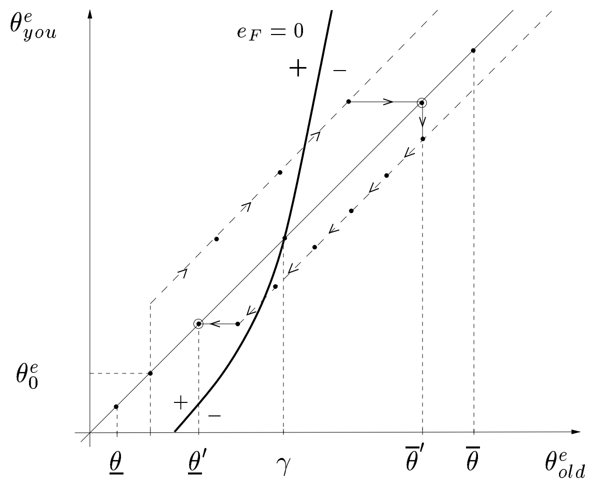

Figure 3: Adaptive learning scheme.

Fig. 3 conveys the economic intuition for this learning scheme. As long as there are positive (negative) forecast errors, increase (reduce) the current forecast by some quantity α. A positive (negative) forecast error means that the expected inflation factor of the old generation was too low (high). As soon as a negative (positive) forecast error is obtained, check how close the current forecast is to the monetary steady state. Geometrically, the resulting learning dynamics takes place along straight lines which are parallel to the 45°-degree line. Since any pair of forecasts $(\theta^e_{old}, \theta^e_{you})$ lying on the diagonal reduces the uncertainty interval, where $e_F(\theta^e, \theta^e) = 0$ if and only if $\theta^e = \gamma$, an initial uncertainty interval $[\underline{\theta}, \overline{\theta}]$ can easily be obtained by repeatedly applying the forecasting rule (8) and checking out the diagonal until the error bounds are obtained.

The next theorem shows that the forecasts generated by repeated application of Algorithm 6.1 converge to the monetary steady state.

Theorem 6.1 *Let Assumption 5.1 be satisfied and $\theta^e_0 \in [\underline{\theta}, \overline{\theta}]$ be an arbitrary initial forecast. Then the iteration stage of Algorithm 6.1 ends in at most k time steps and reduces the initial uncertainty interval $[\underline{\theta}, \overline{\theta}]$ to a smaller uncertainty interval whose length is less than or equal to $\frac{k-1}{k}(\overline{\theta} - \underline{\theta})$. Since the tolerance level ϵ may be arbitrarily small, Algorithm 6.1 yields a sequence of forecasts $\{\theta^e_t\}_{t\in\mathbb{N}}$ which converges to the monetary steady state.*

Proof. By construction, the Algorithm 6.1 picks at least one additional point on the diagonal whose coordinates lie in $(\underline{\theta}, \overline{\theta})$. This yields a new uncertainty interval $[\underline{\theta}', \overline{\theta}']$ which is contained in $[\underline{\theta}, \overline{\theta}]$. Now clearly $\overline{\theta}' - \underline{\theta}' \leq \frac{k-1}{k}(\overline{\theta} - \underline{\theta})$. Since $(k-1)/k < 1$, this yields the theorem. *Q.E.D.*

Since an initial uncertainty interval can be obtained from a slight modification of Algorithm 6.1, we arrived at a learning scheme which in fact converges globally on $\mathbb{R}^2_+$. Algorithm 6.1 stabilizes the monetary steady state in particular for cases in which that steady state is unstable under the perfect-foresight dynamics. It is inspired by a minimization scheme introduced by Berman (1966) and can easily be generalized to practically all one-dimensional models of the Cobweb type (see Appendix A).

Having introduced a surprisingly naive but successful learning scheme, we show next how Newton-type methods can be incorporated in an adaptive scheme searching for the monetary steady state. Since the derivatives of the error function e_F are unknown to the forecasting agency, we apply Newton's secant method (see Ortega & Rheinboldt 1970) to the map $x \mapsto e_F(x, x)$. A Newton step is now given by

$$x_{\tau+1} = x_\tau - \left[\frac{e_F(x_{\tau-1}, x_{\tau-1}) - e_F(x_\tau, x_\tau)}{x_{\tau-1} - x_\tau}\right]^{-1} e_F(x_\tau, x_\tau), \quad \tau \in \mathbb{N}. \tag{9}$$

Since all forecasts known and forecast errors are observable, the only remaining difficulty is that no scheme can directly move on the 45°–degree line of the $(\theta^e_{old}, \theta^e_{you})$-plane.

A learning scheme using a Newton step (9) can proceed as follows. Let $(\theta^e_{-1}, \theta^e_0) \in \mathbb{R}^2_+$ be arbitrary initial forecasts. First, set $\theta^e_1 = \theta^e_0$ and choose some small number ϵ with

$\theta_2^e = \theta_1^e + \epsilon > 0$. Second, set $\theta_3^e = \theta_2^e$. and apply a Newton step (9) yielding θ_4^e. Third, set $\theta_5^e = \theta_4^e$. Now since by Proposition 3.1 e_F is linear along the diagonal, a routine calculation shows that the fourth forecast already is the monetary steady state, i.e. $\theta_4^e = \gamma$. In summary, the monetary steady state (γ, γ) is reached within 5 steps from arbitrary initial conditions in $\mathbb{R}_+^2$.

7 Adaptive learning of perfect forecasting rules

The concept of learning scheme introduced in this section is to construct a locally ϵ-perfect forecasting rule as an approximation of a locally perfect forecasting rule from past observations.[6] Since the specific functional form of a perfect forecasting rule is unknown to a forecasting agency our approximations will be chosen from the class of cubic spline functions which are well known for their good approximation properties, see e.g. Watson (1980).[7] To fix notation, let

$$\Delta = \{\underline{\theta} = a_0 < a_1 < \cdots < a_n = \overline{\theta}\} \tag{10}$$

denote a partition of the compact interval $[\underline{\theta}, \overline{\theta}]$.

Definition 7.1 *A (cubic) spline function ψ_Δ associated with Δ is a real-valued function $\psi_\Delta : [\underline{\theta}, \overline{\theta}] \to \mathbb{R}$ with the following properties:*

(i) ψ_Δ is two times differentiable with continuous second derivatives, $\psi_\Delta \in C^2[\underline{\theta}, \overline{\theta}]$;

(ii) ψ_Δ is a polynomial of degree 3 on each interval $[a_i, a_{i+1}]$, $i = 0, \ldots, n$.

The original idea of spline interpolation was to construct a smooth curve through a prescribed set of points. The following result is standard, see e.g., Watson (1980) or Stoer (1979).

Theorem 7.2 *Let Δ be given. Then for any prescribed set of values b_i, $i = 0, \ldots, n$, there exists a unique cubic spline function ψ_Δ, such that*

$$(i) \quad \psi_\Delta(a_i) = b_i, \; i = 0, \ldots, n \quad and \quad (ii) \quad D^2\psi_\Delta(\underline{\theta}) = D^2\psi_\Delta(\overline{\theta}) = 0.$$

[6] Perfect forecasting rules are directly linked to an inverse function of the induced error function, in this case given by $(\theta^e_{old}, \theta^e_{you}) \mapsto (\theta^e_{old}, e_F(\theta^e_{old}, \theta^e_{you}))$, see Böhm & Wenzelburger (2000a). An alternative approach is to approximate this inverse first and then obtain an ϵ-perfect forecasting rule. This is outlined in Appendix A. The strategy proposed in this section should have less search costs.

[7] Spline functions could as well be replaced here by other classes of approximating functions such as wavelets.

Spline functions can also be used as approximations of continuous functions. To this end, let $\|g\|_\infty$ denote the supremum norm of a real-valued continuous function g, $b_i = g(a_i)$, $i = 0, \ldots, n$ be a prescribed set of values at the knots a_i, $i = 0, \ldots, n$, and replace Condition *(ii)* by $D\psi_\Delta(a_i) = Dg(a_i)$, $i = 0, n$. The following Theorem is due to Carlson & Hall (1973).

Theorem 7.3 *Let $g \in C^m[\underline{\theta}, \overline{\theta}]$ with $m = 1, 2, 3,$ or 4, $\|\cdot\|_\infty$ denote the supremum norm on $[\underline{\theta}, \overline{\theta}]$, and ψ_Δ be the unique cubic spline approximation of g such that*

$$(i) \quad \psi_\Delta(a_i) = g(a_i), \ i = 0, \ldots, n \quad \text{and} \quad (ii) \quad D\psi_\Delta(a_i) = Dg(a_i), \ i = 0, n.$$

Then there exist constants $C_{m,r}$, such that

$$\|D^r\psi_\Delta - D^r g\|_\infty \leq C_{m,r}\,\|D^m g\|_\infty\,\|\Delta\|^{m-r}, \quad 0 \leq r \leq \min\{m, 3\},$$

where $\|\Delta\| := \max_i(a_{i+1} - a_i)$. For $r < 3$, the constants $C_{m,r}$ are independent of the partition Δ.

The precise values of the constants $C_{m,r}$ are found in Carlson & Hall (1973).[8] While it is well-known how to compute a spline approximation for a known function $\psi_\star$, a forecasting agency in our model faces two basic problems. First, the domain of a locally perfect forecasting rule $\psi_\star$ is a priori unknown to the agency. Second, since $\psi_\star$ is unknown, so are the knots (a_i, b_i) with $b_i = \psi_\star(a_i)$ associated with a partition Δ which are needed to construct the spline approximation. The learning scheme which approximates a locally perfect forecasting rule will therefore involve the following steps.

Algorithm 7.1 *The basic steps of the learning scheme to approximate locally perfect forecasting rules are the following:*

1. *Choose an interval $[\underline{\theta}, \overline{\theta}]$;*
2. *Determine knots $(a_i, \psi_\star(a_i))$ for a chosen partition Δ;*
3. *Compute spline approximation ψ_Δ of $\psi_\star$ associated with Δ.*

1. Choosing an interval $[\underline{\theta}, \overline{\theta}]$. In the first step the agency can try a sequence of forecasts of the form

$$(\underline{\theta}, \underline{\theta}), \ (\underline{\theta}, \overline{\theta}), \ (\overline{\theta}, \overline{\theta}), \ (\overline{\theta}, \underline{\theta}).$$

In view of Proposition 3.1, there basically arise two initial situation when appropriately choosing the above sequence (i.e., $\underline{\theta}$ sufficiently small and $\overline{\theta}$ sufficiently large).

[8] It is also relatively straightforward to see that Condition *(ii)* in Theorem 7.3 can be replaced by Condition *(ii)* of Theorem 7.2, see Stoer (1979).

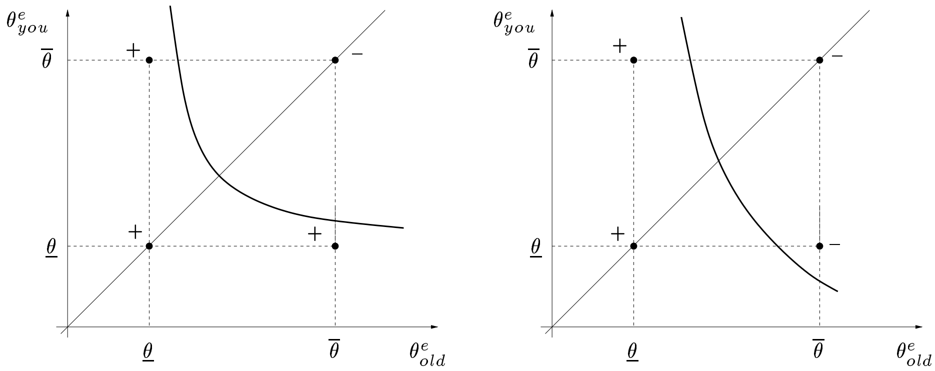

Figure 4: Two initial situations: case 1 (left) and case 2 (right).

Case 1: $-e_F(\underline{\theta},\underline{\theta}) \cdot e_F(\underline{\theta},\overline{\theta}) > 0$ or $-e_F(\overline{\theta},\underline{\theta}) \cdot e_F(\overline{\theta},\overline{\theta}) > 0$;

Case 2: $e_F(\underline{\theta},\underline{\theta}) \cdot e_F(\underline{\theta},\overline{\theta}) > 0$ and $e_F(\overline{\theta},\underline{\theta}) \cdot e_F(\overline{\theta},\overline{\theta}) > 0$.

The two initial cases are qualitatively depicted in Fig. 4. In the first case, the error function has a zero either in between the points $(\underline{\theta},\underline{\theta})$ and $(\underline{\theta},\overline{\theta})$ or in between $(\overline{\theta},\underline{\theta})$ and $(\overline{\theta},\overline{\theta})$. In the second case, the error function has zeros between the points $(\underline{\theta},\underline{\theta})$ and $(\overline{\theta},\underline{\theta})$ as well as in between $(\underline{\theta},\overline{\theta})$ and $(\overline{\theta},\overline{\theta})$.

For what follows, we focus on Case 1 with $e_F(\overline{\theta},\underline{\theta}) > 0$ and $e_F(\overline{\theta},\overline{\theta}) < 0$. All other cases can be treated in a similar manner. A forecasting agency can now infer from Assumption 5.1, that there exists a non-empty subinterval of $[\underline{\theta},\overline{\theta}]$ on which a locally perfect forecasting rule is well defined. Moreover, the graph of this locally perfect forecasting rule will qualitatively look like the left hand picture in Fig. 4.

2. Computing knots. Let a_i be an arbitrary point lying in $[\underline{\theta},\overline{\theta}]$. Then the value $b_i = \psi_\star(a_i)$ at a_i is only defined implicitly through $e_F(a_i, b_i) = 0$. Keeping a_i fixed, b_i can, in principle, be computed by applying any algorithm to compute zeros of the function $\theta^e \mapsto e_F(a_i, \theta^e)$. As the derivatives of the error function e_F are unknown to the forecasting agency, we need derivative-free methods such as Berman's algorithm (Berman 1966) including the bisection method or Newton's secant method (Ortega & Rheinboldt 1970). Applying Newton's secant method to the function $y \mapsto e_F(a_i, y)$, gives

$$y_{\tau+1} = y_\tau - \left[\frac{e_F(a_i, y_\tau) - e_F(a_i, y_{\tau-1})}{y_{\tau-1} - y_\tau}\right]^{-1} e_F(a_i, y_\tau), \tag{11}$$

where y_τ and $y_{\tau-1}$ are previously determined values.

Since the dynamics is one-dimensional, the forecasts cannot be adjusted directly in the way proposed by (11). However, due to the Cobweb nature of our model, this problem can be solved by using a linear forecasting rule of the form

$$\psi_L(\theta^e; a_i, \theta^e_{old}) := a_i - \theta^e + \theta^e_{old}, \tag{12}$$

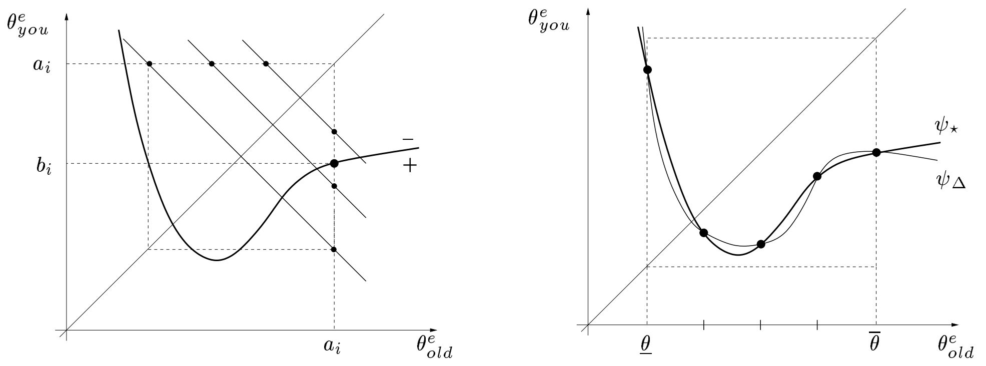

Figure 5: (a) Computing a knot (b) A spline approximation

where $\theta^e_{old} \in \mathbb{R}_+$ is some previously determined forecast. Since $a_i = \psi_L(\theta^e_{old}; a_i, \theta^e_{old})$ for all θ^e_{old}, an application of (12) can be used to direct the system back to the forecast a_i.

The zero (a_i, b_i) of e_F can now be computed through the alternating application of a Newton adjustment step (11) followed by an application of the linear forecasting rule (12), as illustrated in Fig. 5 (a). Suppose that the last two forecasts are (θ^e_{old}, a_i) which is a point located on the horizontal line through $(0, a_i)$–(a_i, a_i). Applying (11), yields an update θ^e_{new} such that the point (a_i, θ^e_{new}) is located on the vertical line $(a_i, 0)$–(a_i, a_i). The linear forecasting rule $\psi_L(\theta^e_{new}; a_i, \theta^e_{new})$ brings the system back to the point (θ^e_{new}, a_i) on the horizontal line $(0, a_i)$–(a_i, a_i). From there on a new Newton adjustment step can be applied. These two steps can be repeated until a sufficiently good approximation of b_i is found.

Observe that the Newton steps can be replaced by either their refinements or by other methods such as Berman's algorithm.[9] A partition Δ of the form (10) can now successively be constructed from the right end to the left end of the interval $[\underline{\theta}, \overline{\theta}]$. For each chosen a_i, $i = 1, \ldots, n$, the corresponding knot (a_i, b_i) is successively computed through an alternating application of the Newton adjustment step (11) and the linear forecasting rule (12). It should be clear that any previously obtained point on the graph of the error function provides information on the location of the zero contour and can therefore be used to apply a Newton step.

3. Computing the spline function. Having determined all knots (a_i, b_i), $i = 1, \ldots, n$, the computation of the spline function ψ_Δ with $\psi_\Delta(a_i) = b_i$ is a routine calculation in

[9] It is well known that Newton's method may not converge and even lead to chaotic behavior if initial guesses are not sufficiently close to the solution of the problem. On the other hand, Newton's method converges quadratically, provided that initial guesses are sufficiently good. It might therefore be advantageous to start out with a global method such as Berman's algorithm, to get good error bounds and then switch to Newton's method. Refinements of Newton's method are found in Ortega & Rheinboldt (1970). These are applicable as well as numerical continuation methods (see Allgower & Georg 1990).

numerical mathematics, see e.g. Watson (1980). If the monetary steady state γ lies in $[\underline{\theta}, \overline{\theta}]$, the stability of γ can be seen from the slope of ψ_Δ, provided the partition was fine enough. If ψ_Δ is monotonically increasing locally around its fixed point, then the perfect-foresight dynamics has no non-trivial forward invariant set. A forecasting agency which realizes this can then direct the system to the monetary steady state γ (by possibly improving the approximation of γ using the methods presented in Section 6).

Assuming that ψ_Δ is decreasing locally around its fixed point, the question now is whether the interval $[\underline{\theta}, \overline{\theta}]$ is invariant under ψ_Δ. This is the case, if

$$\underline{\theta} \leq \min\{\psi_\Delta(\theta) \mid \underline{\theta} \leq \theta \leq \overline{\theta}\} \leq \max\{\psi_\Delta(\theta) \mid \underline{\theta} \leq \theta \leq \overline{\theta}\} \leq \overline{\theta}.$$

If $[\underline{\theta}, \overline{\theta}]$ is invariant under ψ_Δ, a forecasting agency can use and apply the spline approximation ψ_Δ as a forecasting rule. This is illustrated in Fig. 5 (b). By choosing a prescribed tolerance level $\epsilon > 0$, the agency can check whether the forecast errors stay below this level and, if necessary, improve ψ_Δ by adding knots with satisfactory forecast errors to the original partition Δ.

If $[\underline{\theta}, \overline{\theta}]$ is not invariant under ψ_Δ, then there are two strategies. First, one could add knots to the partition Δ to improve the accuracy of the spline approximation at places where ψ_Δ leaves the interval $[\underline{\theta}, \overline{\theta}]$, hoping that the more accurate approximation leaves $[\underline{\theta}, \overline{\theta}]$ invariant. Second, $[\underline{\theta}, \overline{\theta}]$ could be enlarged such that the minimum and maximum of ψ_Δ are both included. This defines a new interval for which new knots can be computed according to step 2. Under the assumption that an invariant set for the perfect-foresight dynamics exists, the following proposition shows that a combination of these two strategies ensures the existence of a an ϵ-perfect forecasting rule together with an invariant set.

Theorem 7.4 *Let $\psi_\star : U \to \mathbb{R}_+$ be a continuous locally perfect forecasting rule and $V \subset U$ a non-trivial compact interval which is forward-invariant under $\psi_\star$. Then there exists $\delta > 0$ such that for each $0 < \epsilon < \delta$, there exists a spline function ψ_ϵ which is ϵ-perfect and for which $\psi_\epsilon(V) \subset V$.*

Proof. Let $V = [\underline{\theta}, \overline{\theta}]$. By Theorem 7.3, for each $\epsilon > 0$ a spline approximation ψ_ϵ on V exists such that $|\psi_\star(\theta) - \psi_\epsilon(\theta)| < \epsilon$ for all $\theta \in V$. If the minimum and the maximum of $\psi_\star$ lie both within the open interval $(\underline{\theta}, \overline{\theta})$, then V is invariant under ψ_ϵ for all sufficiently small $\epsilon > 0$.

Suppose now that either the minimum or the maximum of $\psi_\star$ are equal to $\underline{\theta}$ or $\overline{\theta}$, respectively. Since $\psi_\star$ is continuous, there exists a continuous function $\psi_{\star\star}$ satisfying

$$|\psi_\star(\theta) - \psi_{\star\star}(\theta)| < 2\epsilon/3, \quad \theta \in V$$

such that $\psi_{\star\star}(\theta) \in [\underline{\theta} + \epsilon/3, \overline{\theta} - \epsilon/3]$ for all $\theta \in V$. By Theorem 7.3, there exists a spline approximation ψ_ϵ such that $|\psi_{\star\star}(\theta) - \psi_\epsilon(\theta)| < \epsilon/3$ for $\theta \in V$. Using the triangle inequality, ψ_ϵ is an ϵ-perfect forecasting rule with $\psi_\epsilon(V) \subset V$. *Q.E.D.*

Corollary 7.5 *If, in addition, $\psi_\star$ is differentiable, then there exists a constant $C > 0$ such that*

$$\|\psi_\epsilon - \psi_\star\|_\infty \leq C\,\|D\psi_\star\|_\infty \|\Delta\|_\infty$$

for a partition Δ of V.

Corollary 7.5 follows from Theorem 7.3. Notice that by Proposition 3.2, the derivative of $\psi_\star$ and hence a priori error bounds for ϵ-perfect forecasting rules depend essentially on the savings behavior of households. In summary, Theorem 7.4 states that a sufficiently accurate spline approximation on $[\underline{\theta}, \overline{\theta}] \subset U$ leaves an invariant set V of $\psi_\star$ which is contained in $[\underline{\theta}, \overline{\theta}]$ invariant. Hence, such a spline can be used as an ϵ-perfect forecasting rule ψ_ϵ which generates a dynamics with pointwise bounded forecast errors. If $\psi_\star$ is structurally stable, then the dynamics on V generated by ψ_ϵ is qualitatively the same as the perfect-foresight dynamics. This approximation can be improved on V, in particular along any attractor generated by ψ_ϵ, although this attractor will change with any update of ψ_ϵ. It is evident that the spline approximation provides a useful forecasting rule only, if there exists an invariant set V, since otherwise the dynamics cannot be confined to a compact interval. For cases in which non-trivial invariant sets do not exist, the only chance is to search for a steady state, as carried out in Section 6.

8 Conclusions

The analysis of adaptive learning in a standard OLG economy shows that the correct use of the structural information about the market mechanism enables a forecasting agency to learn those forecasting rules which generate perfect foresight. For the OLG economy it was the fact that only the previous and the current expected inflation factors matter. This structural information is encoded in the error function associated with the economy. This function depends exclusively on the economic fundamentals encoded in the economic law and captures the true nature of the expectations feedback. The graph of the error function is the essential time-invariant object about which information is to be obtained along the evolution of the system. A successful learning scheme must therefore locate the zero contour of the error function through successive approximations.

It was shown that along an orbit of the system, a forecasting agency receives more and more information about the shape of the error function, the location of its zero contour, and hence the location of locally perfect forecasting rules. In the case of the pure exchange economy, it was shown how to compute an arbitrarily precise approximation of such locally perfect forecasting rules. Moreover, exploiting the correct Cobweb type structure of the pure exchange economy, it is always possible to generate forecasts which converge to the monetary steady state, regardless of whether or not this steady state is stable under the perfect foresight dynamics.

The learning scheme proposed in the present paper was based on a simple geometric intuition. It converges globally for all initial conditions and all parameterizations which guarantee the existence of a monetary steady state. Successful performance has been demonstrated for all one-dimensional models of the Cobweb type. Thus, learning schemes which induce diverging or complex behavior in these types of models ignore their essential structural features which may lead to their ultimate failure to obtain perfect foresight. The question *'Can agents learn their way out of chaos?'* in Schönhofer (2001) can clearly be answered with *'yes'* in the case of Cobweb-type models.

The potential of this strand of research is outlined as follows. First, locally perfect forecasting rules for multivariate models of the Cobweb type can, in principle, be approximated using the strategies presented in the paper. Although search algorithms for zeros in the multivariate case are more involved, all techniques used in the present paper have their multivariate analogues. Second, as the notion of an error function can be generalized to models with exogenous stochastic perturbations, the approximation techniques of the paper should appropriately be altered to take stochastic noise into account. By associating a mean error function (Böhm & Wenzelburger 2000b) with a stochastic model, approximations of forecasting rules which generate rational expectations could, in principle, be computed by means of stochastic minimization techniques and approximation theory. While all these possible generalizations involve more technicalities, the basic intuition for successful learning schemes in Cobweb type environments seems to be already contained in the one-dimensional case.

A Learning in models of the Cobweb type

The purpose of this appendix is to outline some generalizations of the learning schemes presented in the main part of the paper. The main idea of the adaptive learning scheme in Section 7 was to interpolate the inverse of an error function associated with an economic law from past observed data. This concept seems to be particularly simple in models of the Cobweb type, where only forecasts drive the dynamics of the system. To illustrate the basic idea, consider the one-dimensional case. Let $F : \mathbb{R}_+ \to \mathbb{R}_+$ define an economic law of the Cobweb type such that

$$x_t = F(x^e_{t+1}), \tag{13}$$

where x^e_{t+1} is the forecast for x_{t+1} based on information available prior to the determination of x_t. This type of model has been studied extensively in the literature (see e.g. Chatterji & Chattopadhyay (2000) and references therein). The corresponding error function is defined by $e_F(x^e_{new}, x^e_{old}) := F(x^e_{new}) - x^e_{old}$. Assume that there exists a (perfect foresight) steady state $\bar{x} = F(\bar{x})$ of F and that F is locally invertible on an open set $W \subset \mathbb{R}$ of $\bar{x}$. Let $U \subset F(W)$. Then a locally perfect forecasting rule $\psi_\star : U \to \mathbb{R}$ is

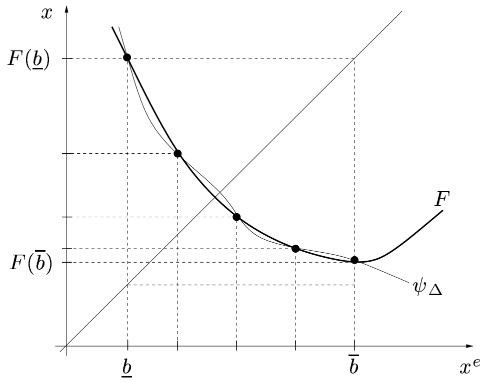

Figure 6: The Cobweb case.

given by the (local) inverse of the function F when restricted to W, such that

$$x^e_{t+1} = \psi_\star(x^e_t) := F^{-1}(x^e_t), \quad x^e_t \in U.$$

The basic idea of learning $\psi_\star$ now is to construct an approximation of $\psi_\star$ using a spline interpolation of past observed data $\{x_t, x^e_{t+1}\}$. This can be done directly by selecting some interval $[\underline{b}, \overline{b}]$ together with a partition $\underline{b} = b_n < b_{n-1} < \ldots < b_0 = \overline{b}$. Choosing the b_i as forecasts, one obtains a sequence of realizations $a_i = F(b_i)$, $i = 0, \ldots, n$ which is monotonically increasing (decreasing), if F is decreasing (increasing). Suppose for a moment that F is decreasing on $[\underline{b}, \overline{b}]$. This implies that the sequence $a_i = F(b_i)$, $i = 0, \ldots, n$ is monotonically increasing such that the original partition of $[\underline{b}, \overline{b}]$ induces a partition $\Delta = \{F(\overline{b}) = a_0 < a_1 < \cdots < a_n = F(\underline{b})\}$ of the interval $[F(\overline{b}), F(\underline{b})]$. By Theorem 7.3, there exists a unique spline approximation ψ_Δ of $\psi_\star$ with $b_i = \psi_\Delta(a_i)$, $i = 0, \ldots, n$. ψ_Δ is an ϵ-perfect forecasting rule, see Fig. 6.

The precision of ψ_Δ depends clearly on the coarseness of Δ. But since this partition can be made arbitrarily fine, ψ_Δ becomes arbitrarily close to $\psi_\star$ in the supremum norm. In this univariate case, it is also easy to separate regions in which F is monotonically decreasing from regions where F is monotonically increasing from past observations. Hence, a forecasting agency can always construct locally perfect forecasting rules, including possible multiplicities. The concept of interpolating an inverse function clearly generalizes to the multivariate case using multivariate approximation theory, although this is technically more cumbersome.

B The geometry of forecast errors

The purpose of this appendix is illustrate that a forecasting agency with the structural knowledge of Assumption 5.1 should abandon any learning scheme which generates

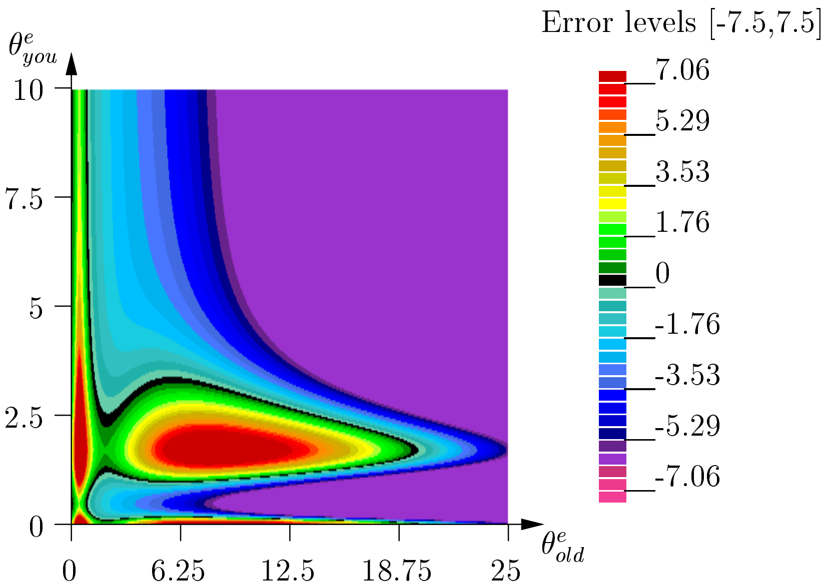

Figure 7: Error levels of an 'irregular' error function with $\gamma = 2.7,\ \rho = 0.5$.

forecasts with non-vanishing forecast errors. To this end, consider an example of an 'irregular' savings function (Bullard 1994, p. 480)

$$S(\theta^e) = \exp\left\{\cos\left[\frac{10}{1+(\theta^e)^{\frac{\rho}{\rho-1}}}\right]\right\}^{-1}, \quad \theta^e \in \mathbb{R}_+.$$

In this case multiple locally perfect forecasting rules exist, as the zero contour of the corresponding error function folds back. Fig. 7 depicts points in the $(\theta^e_{old}, \theta^e_{you})$-plane which lie in the same error range are colored according to the adjacent color code. The contour lines of the corresponding error function are given by the boundaries between to colored regions. (These could, in principle, be computed analytically.)

Consider now the recursive ordinary least squares scheme (see e.g. Chen & Guo 1991) which will be viewed as a *forecasting rule* depending in a well-specified manner on current states and certain other auxiliary variables. In the case of an OLG exchange economy this OLS scheme is given by the maps

$$\begin{aligned} \psi : \mathbb{R}^2_+ \times [0,1] &\longrightarrow \mathbb{R}_+, & \psi(\theta_{t-1}, \theta^e_t, g_{t-1}) &:= \theta^e_t + g_{t-1}[\theta_{t-1} - \theta^e_t] \\ \varphi : \mathbb{R}_+ \times [0,1] &\longrightarrow [0,1], & \varphi(\theta_{t-1}, g_{t-1}) &:= \frac{g_{t-1}\,\theta^2_{t-1}}{1 + g_{t-1}\,\theta^2_{t-1}}, \end{aligned} \tag{14}$$

such that $\theta^e_{t+1} = \psi(\theta_{t-1}, \theta^e_t, g_{t-1})$ is the OLS-based forecast for θ_{t+1} and $g_t = \varphi(\theta_{t-1}, g_{t-1})$ the auxiliary variable.[10] The evolution of the inflation factors is then governed by the

[10]The expression for ψ in (14) is counterintuitive from the econometric point of view. There seem

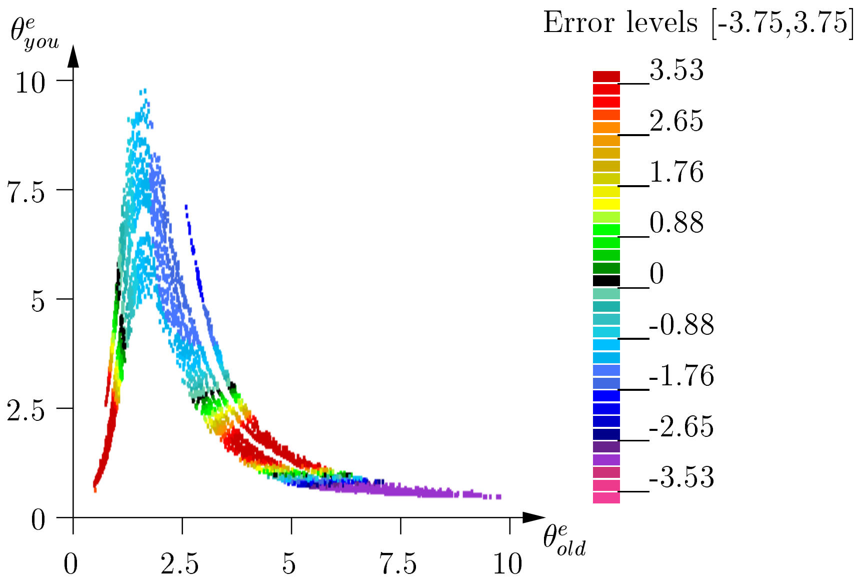

Figure 8: 5000 points of an attractor with corresponding error levels in the 'irregular' case with OLS; parameters $\gamma = 2.7$, $\rho = 0.5$, initial conditions $\theta^e_0 = 0.5$, $\theta^e_{-1} = 1.0$, $g_0 = 0.5$.

map $F_{\psi,\varphi} : \mathbb{R}^2_+ \times [0,1] \longrightarrow \mathbb{R}^2_+ \times [0,1]$, given by

$$(\theta_{t-1}, \theta^e_t, g_{t-1}) \longmapsto \left(F(\theta^e_t, \psi\big(\theta_{t-1}, \theta^e_t, g_{t-1})\big), \psi(\theta_{t-1}, \theta^e_t, g_{t-1}), \varphi\big(\theta_{t-1}, g_{t-1})\big)\right).$$

Replacing θ_{t-1} by $F(\theta^e_{t-1}, \theta^e_t)$, it is seen that the dynamics of $F_{\psi,\varphi}$ is essentially induced by the pair of maps (ψ, φ) and hence by the OLS scheme (14) alone.

The dynamics generated by (14) can now be analyzed in the $(\theta^e_{old}, \theta^e_{you})$-plane as done in Bullard (1994) and Schönhofer (1999, 2001). Fig. 8 shows a complex attractor[11] in the $(\theta^e_{old}, \theta^e_{you})$-plane generated by the OLS scheme. Points $(\theta^e_t, \theta^e_{t+1})$ on the attractor which lie in the same error range are colored according to the adjacent color code.

As is apparent from the figure, the contour lines and hence the shape of the error function becomes to a large extent revealed through the course of time, if the learning dynamics is complex.[12] In particular, the location of the zero contour in this figure is disclosed, giving a clear indication of the location of a locally perfect forecasting rule including the

to appear the wrong differences, namely $\theta_{t-1} - \theta^e_t$ instead of $\theta_t - \theta^e_t$. This is because the regression is done on prices rather than on inflation factors. In fact, θ^e_{t+1} obtained from (14) should be a forecast for θ_t rather than for θ_{t+1} and hence be replaced by $(\theta^e_{t+1})^2$, cf. Wenzelburger (1999, Appendix B).

[11] To be precise, Fig. 8 shows a projection of an attractor onto the $(\theta^e_{old}, \theta^e_{you})$-plane, since the actual system is three-dimensional.

[12] All simulations were carried out using the program package MACRODYN , cf. Böhm & Schenk-Hoppé (1998). We used similar parameterizations to those in Bullard (1994) and Schönhofer (1999).

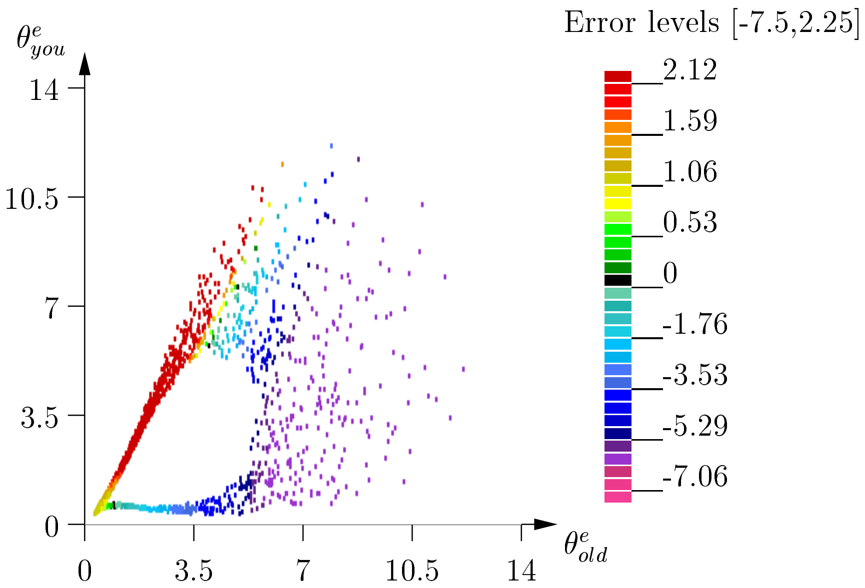

Figure 9: 5000 points of an attractor with corresponding error levels in the CES case with OLS; parameters $\gamma = 1.5$, $\rho = 0.735$, $w_1 = 2$, $w_2 = 0$, $\delta = 1$, initial conditions $\theta^e_0 = 1.2$, $\theta^e_{-1} = 0.75$, $g_0 = 0.5$.

location of the monetary steady state. A similar observation is made for the case of a CES savings function in Fig. 9. In other words, the geometry of errors in Figs. 8 and 9 provide a clear incentive to abandon the employed OLS scheme.

References

Allgower, E. L. & K. Georg (1990): *Numerical Continuation Methods*, Springer Series in Computational Mathematics. Springer, Berlin.

Berman, G. (1966): "Minimization by successive approximation", *SIAM J. Numer. Anal.*, 3(1), 123–133.

Blume, L. & D. Easley (1982): "Learning to Be Rational", *Journal of Economic Theory*, 26, 340–351.

——— (1993): "What Has the Rational Learning Literature Taught Us?", in *Essays in Learning and Rationality in Economics*, ed. by A. Kirman & M. Salmon. Blackwell, Oxford.

——— (1998): "Rational Expectations and Rational Learning", in *Organizations with Incomplete Information: Essays in Economic Analysis: A Tribute to Roy Radner*, ed. by M. Majumdar. Cambridge University Press, Cambridge.

Böhm, V. & K. R. Schenk-Hoppé (1998): "MACRODYN – A User's Guide", Discussion Paper 400, Department of Economics, University of Bielefeld.

Böhm, V. & J. Wenzelburger (1999): "Expectations, Forecasting and Perfect Foresight - A Dynamical Systems Approach", *Macroeconomic Dynamics*, 3(2), 167–186.

——— (2000a): "Expectational Leads in Economic Dynamical Systems", Revised Version October 2000, Discussion Paper No. 373, University of Bielefeld.

——— (2000b): "Perfect Predictions in Economic Dynamical Systems with Random Perturbations", Revised Version September 2000, Discussion Paper No. 356, University of Bielefeld.

Bray, M. (1982): "Learning, Estimation, and the Stability of Rational Expectations Equilibria", *Journal of Economic Theory*, 26, 318–339.

Bullard, J. (1994): "Learning Equilibria", *Journal of Economic Theory*, 64(2), 468–485.

Carlson, R. & C. Hall (1973): "Error Bounds for Bicubic Spline Interpolation", *Journal of Approximation Theory*, 7, 41–47.

Chatterji, S. & S. Chattopadhyay (2000): "Global Stability in Spite of "Local Instability" with Learning", *Journal of Mathematical Economics*, 33, 155–165.

Chen, H.-F. & L. Guo (1991): *Identification and Stochastic Adaptive Control.* Birkhäuser, Boston.

Devaney, R. L. (1989): *An Introduction to Chaotic Dynamical Systems.* Addison-Wesley, Reading, 2nd edn.

Evans, G. & S. Honkapohja (1999): "Learning Dynamics", in *Handbook of Macroeconomics*, ed. by J. Taylor & M. Woodford, vol. 1A, chap. 7, pp. 449–533. Elsevier Science B.V., Amsterdam a.o.

Fourgeaud, C., C. Gourieroux & J. Pradel (1986): "Learning Procedures and Convergence to Rationality", *Econometrica*, 54, 845–869.

Grandmont, J.-M. (1985): "On Endogenous Competitive Business Cycles", *Econometrica*, 53(5), 995–1045.

Hommes, C. (1998): "On The Consistency of Backward Looking Expectations: The Case of The Cobweb", *Journal of Economic Behavior & Organisation*, 33, 333–362.

Hommes, C. H. & G. Sorger (1998): "Consistent Expectations Equilibria", *Macroeconomic Dynamics*, 2, 287–321.

Jungeilges, J. (2000): "On the Temporary Nature of Chaotic Consistent Expectations Equilibria", Mimeo.

Kelly, D. L. & J. Shorish (2000): "Stability of Functional Rational Expectations Equilibria", *Journal of Economic Theory*, 95(2), 215–250.

Marcet, A. & T. J. Sargent (1989a): "Convergence of Least Squares Learning Mechanisms in Self-Referential Linear Stochastic Models", *Journal of Economic Theory*, 48, 337–368.

——— (1989b): "Least Square Learning and the Dynamics of Hyperinflation", in *Chaos, Complexity, and Nonlinearity, Proceedings of the Fourth International Symposium in Economic Theory and Econometrics*, ed. by W. Barnett, J. Geweke & K. Shell. Cambridge University Press, Cambridge (Mass.) a.o.

Ortega, J. & W. Rheinboldt (1970): *Iterative Solution of Nonlinear Equations in Several Variables.* Academic Press, New York a.o.

Sargent, T. J. (1993): *Bounded Rationality in Macroeconomics.* Clarendon Press, Oxford.

Schönhofer, M. (1999): "Chaotic Learning Equilibria", *Journal of Economic Theory*, 89(1), 1–20.

——— (2001): "Can Agents Learn Their Way Out of Chaos?", *Journal of Economic Behavior and Organization*, 44(1), 71–83.

Spear, S. (1988): "Existence and Local Uniqueness of Functional Rational Expectations Equilibria in Dynamic Economic System Models", *Journal of Economic Theory*, 44(1), 124–155.

Stoer, J. (1979): *Einführung in die Numerische Mathematik I*, Heidelberger Taschenbücher. Springer-Verlag, Berlin a.o., 3rd edn.

Watson, G. A. (1980): *Approximation Theory and Numerical Methods.* Wiley, New York a.o.

Wenzelburger, J. (1999): "Convergence of Adaptive Learning in Models of Pure Exchange", Discussion Paper No. 432, University of Bielefeld.