UvA-DARE (Digital Academic Repository)

Behavioral learning equilibria

Hommes, C.; Zhu, M.

Link to publication

Behavioral Learning Equilibria

Cars Hommes[a][*], Mei Zhu[b,a]

[a] CeNDEF, School of Economics, University of Amsterdam

and Tinbergen Institute

[b] Institute for Advanced Research, Shanghai University of Finance and Economics,

and the Key Laboratory of Mathematical Economics(SUFE), Ministry of Education, Shanghai 200433, China

August 16, 2013

Abstract

We propose behavioral learning equilibria, where boundedly rational agents learn to use a simple univariate linear forecasting rule with correctly specified unconditional mean and first-order autocorrelation. In the long run, agents learn the best univariate linear forecasting rule, without fully recognizing the more complex structure of the economy. An important feature of behavioral learning equilibria is simplicity and parsimony, which makes coordination of individual expectations on such an aggregate outcome more likely. In a first application, an asset pricing model driven by AR(1) dividends, a unique behavioral learning equilibrium exists characterized by high persistence and excess volatility, and it is stable under learning. In a second application, the New Keynesian Phillips curve, multiple equilibria co-exist, learning exhibits path dependence and inflation may switch between low and high persistence regimes.

Keywords: Bounded rationality; Stochastic consistent expectations equilibrium; Adaptive learning; Excess volatility; Inflation persistence

JEL classification: E30; C62; D83; D84

[*] *Corresponding Author:* e-mail: *C.H.Hommes@uva.nl*; phone *+31-20-525 4246*; address: *Valckenierstraat 65-67, 1018 XE Amsterdam, Netherlands*; e-mail co-author: *M.Zhu@uva.nl*

1

1 Introduction

Since the 1970's the Rational Expectations Hypothesis (REH), introduced in Muth (1961) and applied in macroeconomics by Lucas (1972) and others, has become the dominant paradigm in macroeconomics. An Rational Expectations Equilibrium (REE) requires that economic agents' subjective probability distributions coincide with the objective distribution that is determined, in part, by their subjective beliefs. There is a vast literature that studies the drawbacks of REE. Some of these drawbacks include the fact that REE requires an unrealistic degree of computational power and information on the part of agents. Alternatively, the adaptive learning literature (see, e.g., Evans and Honkapohja (2001, 2011) and Bullard (2006) for extensive surveys and references) replaces rational expectations with beliefs that come from an econometric forecasting model with parameters updated using observed time series. A large part of this literature involves studying under which conditions learning will converge to the rational expectations equilibrium. When the perceived law of motion (PLM) of agents is correctly specified, convergence of adaptive learning to an REE can occur. However, generally REE is not the only fixed point in self-referential systems and one should not expect learning to always converge to an REE. Whenever agents have *misspecified* PLMs a reasonable learning process may settle down to some sort of misspecification equilibrium. In the existing literature, different types of misspecification equilibria have been proposed: a Restricted Perceptions Equilibrium (RPE) where the forecasting model is underparameterized (Sargent, 1991; Evans and Honkapohja, 2001); a self-confirming equilibrium where beliefs are only correctly specified on the equilibrium path (Sargent, 1999); and a Stochastic Consistent Expectations Equilibrium (SCEE) (Hommes and Sorger, 1998; Hommes et al., 2013)[1].

A SCEE is a very natural misspecification equilibrium, where agents in the economy do not know the actual law of motion or even recognize all of the explanatory variables, but prefer a parsimonious forecasting model. The economy is too complex to fully understand and therefore, as a first order approximation, agents forecast the state of the economy by simple autoregressive models. In the simplest model applying this idea, agents run an univariate AR(1) regression to generate out-of-sample forecasts of the state of the economy. The idea was first introduced as Consistent Expectations Equilibrium (CEE) in Hommes and Sorger (1998), following Grandmont's (1998) idea of a self-fulfilling mistake,

---

[1]Branch (2006) provides a stimulating survey discussing the connection between these types of misspecification equilibria.

where agents believe that prices follow a linear AR(1) stochastic process, whereas the implied actual law of motion is a deterministic chaotic nonlinear process with exactly the same autocorrelation structure. Hommes et al. (2013) generalize the notion of CEE to *nonlinear stochastic* dynamic economic models, introducing the concept of Stochastic Consistent Expectations Equilibrium (SCEE). A SCEE arises when agents' perceptions about endogenous variables are consistent with the actual realizations of these variables in the sense that the unconditional mean and autocorrelations of the unknown nonlinear stochastic process, which describes the actual behavior of the economy, coincide with the unconditional mean and autocorrelations of the AR(1) process agents believe in. In a SCEE agents use the optimal (univariate) linear forecasting rule in an unknown nonlinear stochastic economy. Although an SCEE is *not* an REE, because generally the linear forecast does not coincide with the true conditional expectation, along a SCEE forecasting errors are unbiased and uncorrelated. A SCEE may be seen as an "approximate rational expectations equilibrium", in which the misspecified perceived law of motion is the *best univariate linear approximation* of the actual (unknown) nonlinear law of motion.

In nonlinear stochastic models, unfortunately, SCEE are very hard to compute analytically[2]. In this paper we study *first-order SCEE* in the context of *linear* stochastic self-referential models. Suppose the actual law of motion (ALM) of the economy is a high dimensional linear stochastic system. But agents might fail to recognize all explanatory variables because of cognitive limitations or because they simply prefer a parsimonious univariate time series prediction rule instead of figuring out exactly how another explanatory variable might affect the state of the economy. As argued e.g. in Fuster et al. (2011), whatever the mix of reasons pragmatic, psychological/suboptimal, and statistical economic agents usually do use simple models to understand economic dynamics. Hence, we assume that agents use a simple univariate AR(1) rule to forecast the economy. A first-order SCEE arises when the two parameters of the AR(1) forecasting rule are pinned down by two consistency requirements: the mean and the first-order autocorrelation of the perceived model (the AR(1) process) are identical to these same moments of the true higher dimensional model, which is generated by these beliefs through the self-referential feature of the model.

First-order SCEE have a simple, *intuitive behavioral* interpretation and therefore, we refer to them as a *Behavioral Learning Equilibrium* (BLE). Although it is possible for some

---

[2]Some existence results on first-order SCEE in 1-D nonlinear stochastic models have been obtained by Branch and McGough (2005) and Hommes et al. (2013).

agents to use more sophisticated models, one may argue that these practices are neither straightforward nor widespread. A simple BLE seems more likely as a description of aggregate behavior, because a large population of individual agents may coordinate their expectations more easily to learn a simple, parsimonious behavioral equilibrium. The simple adaptive expectations or error-correction rule, used in the New Keynesian Philips Curve in Lansing (2009) and the concept of *natural expectations* in Fuster et al. (2010) and Fuster et al. (2011, 2012) –where agents use simple, misspecified models, e.g., linear autoregressive models, for their perceived law of motion– are closely related to our BLE. Natural expectations, however, do not pin down the parameters of the forecasting model through consistency requirements as for a restricted perceptions equilibrium nor do they allow the agents to learn an optimal misspecified model through empirical observations.

We formalize the concept of BLE in the simplest class of models one can think of: a one-dimensional linear stochastic model driven by an exogenous linear stochastic AR(1) process. Agents do not recognize, however, that the economy is driven by an exogenous AR(1) process y_t, but simply forecast the state of the economy x_t using an univariate AR(1) rule. Within this simple, but general, class of models we are able to fully characterize the existence of BLE. While this class of models is strikingly simple, it yields surprisingly rich dynamical behavior, including high persistence, excess volatility and multiple equilibria. We study the stability of BLE under a simple adaptive learning scheme –Sample Autocorrelation Leaning (SAC-learning)– and present the first proof that the first-order SCEE (i.e. BLE) can be attainable via SAC-learning and provide simple and intuitive stability conditions[3]. SAC-learning also has a simple behavioral interpretation. Agents simply estimate or "guestimate" the sample average and first-order persistence (i.e. autocorrelation) from observed time series.

Although the class of models we consider is simple, it contains two important standard applications: an asset pricing model and the New Keynesian Philips curve. In the asset pricing model driven by an exogenous stochastic dividend process, the BLE is unique and SAC-learning always converges to it. The BLE is characterized by highly persistent prices (near unit root) and excess volatility with asset price volatility more than doubled compared to REE. In the second application –the New Keynesian Philips curve (NKPC)– driven by an exogenous AR(1) process for the output gap and an independent and identically distributed (i.i.d.) stochastic shock to inflation - multiple stable BLE may co-exist.

---

[3]In earlier work, learning of SCEE has been studied only by numerical simulations, see, for example, Hommes and Rosser (2001) and Tuinstra (2003).

4

In particular, for empirically plausible parameter values a BLE with highly persistent inflation exists, matching the stylized facts of US-inflation data.

There is a large literature showing that learning models can generate stylized facts such as excess volatility in stock prices and high persistence in inflation. Excess volatility in asset prices through learning has e.g. been shown in Timmerman (1993, 1996), Bullard and Duffy (2001), Bullard et al. (2010), Adam and Marcet (2011, 2012), Lansing (2010) and Branch and Evans (2010, 2011). High persistence in inflation through learning has been studied e.g. in Milani (2007), Bullard et al. (2008) and Lansing (2009). With many alternative learning approaches, simplicity and parsimony become an important issue, which makes our BLE approach attractive as a possible explanation for coordination of individual expectations in a large population. A distinguishing feature of our BLE concept is that the parameters of the AR(1) forecasting rule are pinned down by consistency requirements for the two most important time series statistics, the mean and the first-order autocorrelation, both of which can be guestimated from observed realizations.

The paper is organized as follows. Section 2 introduces the main concepts, the first-order SCEE, sample autocorrelation learning and their interpretation as a behavioral learning equilibrium. Section 3 focusses on existence of first-order SCEE and stability under SAC-learning within a simple linear class of one-dimensional models driven by an exogenous AR(1) process. Section 4 discusses two applications, an asset pricing model and a New Keynesian Philips curve, illustrating the empirical relevance of BLE in explaining excess volatility, inflation persistence and regime switching. Finally, Section 5 concludes.

2   Main concepts

Suppose that the law of motion of an economic system is given by the stochastic difference equation

$$x_t = f(x_{t+1}^e, \ y_t, \ u_t), \tag{2.1}$$

where x_t is the state of the system (e.g. an asset price or inflation) at date t and x_{t+1}^e is the expected value of x at date $t + 1$. This denotation highlights that expectations may not be rational. Here f is a continuous function, $\{u_t\}$ is an i.i.d. noise process with mean zero and finite absolute moments[4], where the variance is denoted by σ_u^2, and y_t is an

---

[4]The condition on finite absolute moments is required to obtain convergence results under SAC-learning.

exogenous driving variable (e.g. dividends, marginal costs, or the output gap), assumed to follow a stochastic AR(1) process

$$y_t = a + \rho y_{t-1} + \varepsilon_t, \qquad 0 \le \rho < 1, \tag{2.2}$$

where $\{\varepsilon_t\}$ is another i.i.d. noise process with mean zero and finite absolute moments, with variance σ_ε^2, and uncorrelated with $\{u_t\}$. The mean of the stationary process y_t is $\bar{y} = \frac{a}{1-\rho}$, the variance is $\sigma_y^2 = \frac{\sigma_\varepsilon^2}{1-\rho^2}$ and the kth-order autocorrelation coefficient of y_t is ρ^k, see for example, Hamilton (1994).

Agents are boundedly rational and do not know the exact form of the actual law of motion (2.1). In particular, agents fail to recognize that the state of the economy x_t is driven by an exogenous stochastic process y_t, either because of cognitive limitations or because they simply prefer a parsimonious univariate time series prediction rule instead of figuring out exactly which other explanatory variable might affect the state of the economy. Although agents may observe y_t they do not realize or they do not believe that the state of the economy is partly determined by the exogenous process y_t. Instead agents only use past observations, $x_{t-1}, x_{t-2}, \cdots$, and a simple linear model to forecast the future state of the economy, whose structure and complexity they do not fully understand. We focus on the simplest case where agents' perceived law of motion (PLM) is a simple, parsimonious AR(1) process.[5] Numerous empirical studies show that simple, parsimonious models often empirically outperform more complex models in out-of-sample forecasting (e.g. Nelson, 1972; Stock and Watson, 2007; and Enders, 2010). From a theoretical perspective, we follow the idea of a *self-fulfilling mistake* (Grandmont, 1998) imposing consistency requirements on the parameters of the AR(1) rule, as explained below, so that the mistake becomes self-fulfilling.

Agents' perceived law of motion is an AR(1) process

$$x_t = \alpha + \beta(x_{t-1} - \alpha) + \delta_t, \tag{2.3}$$

where α and β are real numbers with $\beta \in (-1, 1)$ and $\{\delta_t\}$ is a white noise process; α is the unconditional mean of x_t and β is the first-order autocorrelation coefficient. Given the perceived law of motion (2.3), the 2-period ahead forecasting rule for x_{t+1} that minimizes the mean-squared forecasting error is

$$x_{t+1}^e = \alpha + \beta^2(x_{t-1} - \alpha). \tag{2.4}$$

---

[5]Here we focus on a univariate stochastic process for the law of motion of the economy (2.1) and an AR(1) PLM (2.3). More generally one may consider an N-dimensional state vector X_t and a higher-order linear AR(p) or a VAR forecasting model.

Combining the expectations (2.4) and the law of motion of the economy (2.1), we obtain the implied actual law of motion (ALM)

$$x_t = f(\alpha + \beta^2(x_{t-1} - \alpha), \ y_t, \ u_t), \tag{2.5}$$

with y_t an AR(1) process as in (2.2).

Stochastic Consistent Expectations Equilibrium (SCEE)

We are now ready to recall the definition of SCEE. Following Hommes et al. (2013)[6], the concept of first-order SCEE is defined as follows.

Definition 2.1 *A triple (μ, α, β), where μ is a probability measure and α and β are real numbers with $\beta \in (-1, 1)$, is called a first-order stochastic consistent expectations equilibrium (SCEE) if the following three conditions are satisfied:*

- *S1 The probability measure μ is a nondegenerate invariant measure for the stochastic difference equation (2.5);*

- *S2 The stationary stochastic process defined by (2.5) with the invariant measure μ has unconditional mean α, that is, $E_\mu(x) = \int x \, d\mu(x) = \alpha$;*

- *S3 The stationary stochastic process defined by (2.5) with the invariant measure μ has unconditional first-order autocorrelation coefficient β.*

That is to say, a first-order SCEE is characterized by two natural consistency requirements: the unconditional mean and the unconditional first-order autocorrelation coefficient generated by the actual (unknown) stochastic process (2.5) coincide with the corresponding statistics for the perceived linear AR(1) process (2.3), as given by the parameters α and β. This means that in a first-order SCEE agents correctly perceive the mean and the first-order autocorrelation (i.e., the persistence) of the state of the economy, without fully understanding its structure and recognizing all explanatory variables.

Our SCEE concept may be viewed as the simplest example of a RPE, where agents predict an unknown stochastic law of motion by a first-order linear approximation. It should be stressed that the SCEE has an intuitive behavioral interpretation, and therefore we refer to a first-order SCEE as a *behavioral learning equilibrium* (BLE). In a SCEE

---

[6]In Hommes et al. (2013), the actual law of motion is $x_t = f(x_{t+1}^e, \ u_t)$, without the driving variable y_t. However, the definitions of SCEE and SAC-learning can still be applied here.

agents use a linear forecasting rule with two parameters, the mean α and the first-order autocorrelation β. Both can be detected from past observations by inferring the average price (or inflation) level and the (first-order) persistence of the time series. For example, $\beta = 0.5$ means that, on average, prices mean-revert toward their long-run mean by 50%. These observations could be made by "guestimating" the mean and the persistence from an observed time series of aggregate variables. It is interesting to note that in learning-to-forecast laboratory experiments with human subjects, individual forecasting behavior is well described by simple rules, such as a simple AR(1) rule, see for example, Hommes et al. (2005), Adam (2007), Heemeijer et al. (2009) and Hommes (2011).

Finally, we note that a first-order SCEE imposes cross-equation restrictions on the model. In particular, the orthogonality condition imposed by a Restricted Perceptions Equilibrium (RPE)[7]

$$Ex_{t-1}[x_t - \alpha - \beta(x_{t-1} - \alpha)] = E(x_{t-1} - \alpha)[x_t - \alpha - \beta(x_{t-1} - \alpha)] = 0$$

is satisfied. The orthogonality condition shows that agents can not detect the correlation between their forecasting errors and perceived model, see Branch (2006). The first-order SCEE is a RPE where agents have their model incorrect; but within the context of their forecasting model agents are unable to detect their misspecification.

Sample autocorrelation learning

In the above definition of first-order SCEE, agents' beliefs are described by the linear forecasting rule (2.4) with fixed parameters α and β. However, the parameters α and β are usually unknown. In the adaptive learning literature, it is common to assume that agents behave like econometricians using time series observations to estimate the parameters as additional observations become available. Following Hommes and Sorger (1998), we assume that agents use sample autocorrelation learning (SAC-learning) to learn the parameters α and β. That is, for any finite set of observations $\{x_0, x_1, \cdots, x_t\}$, the sample average is given by

$$\alpha_t = \frac{1}{t+1} \sum_{i=0}^{t} x_i, \tag{2.6}$$

and the first-order sample autocorrelation coefficient is given by

$$\beta_t = \frac{\sum_{i=0}^{t-1}(x_i - \alpha_t)(x_{i+1} - \alpha_t)}{\sum_{i=0}^{t}(x_i - \alpha_t)^2}. \tag{2.7}$$

---

[7]Readers are referred to Evans and Honkapohja (2001) and Branch (2006) for further discussion on the orthogonality condition and RPE.

Hence α_t and β_t are updated over time as new information arrives. It is easy to check that, independently of the choice of the initial values (x_0, α_0, β_0), it always holds that $\beta_1 = -\frac{1}{2}$, and that the first-order sample autocorrelation $\beta_t \in [-1, 1]$ for all $t \geq 1$.[8]

Adaptive learning is sometimes referred to as *statistical learning*, because agents act as statisticians or econometricians and use a statistical procedure, such as OLS, to estimate and update parameters over time. SAC-learning may be viewed as another statistical learning procedure. We would like to stress, however, that SAC-learning has a simple *behavioral interpretation* that agents simply infer the sample average and persistence (i.e. first-order autocorrelation) from time series observations. Eqs. (2.6) and (2.7) specify the sample average α_t and first-order sample autocorrelation β_t over the entire time-horizon, but one could also restrict the learning to the last T observations with T relatively small (e.g., $T = 100$ or even smaller). In fact, it is relatively easy for agents to "guestimate" the mean and first-order autocorrelation directly based on an observed time series.

Define

$$R_t = \frac{1}{t+1} \sum_{i=0}^{t} (x_i - \alpha_t)^2,$$

then the SAC-learning is equivalent to the following recursive dynamical system (see Appendix A).

$$
\begin{cases}
\alpha_t = \alpha_{t-1} + \dfrac{1}{t+1}(x_t - \alpha_{t-1}), \\[2mm]
\beta_t = \beta_{t-1} + \dfrac{1}{t+1} R_t^{-1} \Big[(x_t - \alpha_{t-1})\big(x_{t-1} + \dfrac{x_0}{t+1} - \dfrac{t^2 + 3t + 1}{(t+1)^2}\alpha_{t-1} - \dfrac{1}{(t+1)^2}x_t\big) \\[2mm]
\qquad - \dfrac{t}{t+1}\beta_{t-1}(x_t - \alpha_{t-1})^2 \Big], \\[2mm]
R_t = R_{t-1} + \dfrac{1}{t+1}\Big[\dfrac{t}{t+1}(x_t - \alpha_{t-1})^2 - R_{t-1} \Big].
\end{cases}
\tag{2.8}
$$

The actual law of motion under SAC-learning is therefore given by

$$x_t = f(\alpha_{t-1} + \beta_{t-1}^2(x_{t-1} - \alpha_{t-1}), \ y_t, \ u_t), \tag{2.9}$$

---

[8]The definition of the first-order sample autocorrelation coefficient in (2.7) is only slightly different from least-squares learning, where in fact $\beta_t = \big(\sum_{i=0}^{t-1}(x_i - \bar{x}_t^-)(x_{i+1} - \bar{x}_t^+)\big)/\big(\sum_{i=0}^{t-1}(x_i - \bar{x}_t^-)^2\big)$, with $\bar{x}_t^- = \frac{1}{t}\sum_{i=0}^{t-1}x_i$, $\bar{x}_t^+ = \frac{1}{t}\sum_{i=1}^{t}x_i$. However, the sample autocorrelation coefficient in (2.7) always satisfies $|\beta_t| \leq 1$, while the OLS estimate does not. Under SAC-learning agents believe that prices do not explode. This restriction is a natural "projection facility" for the SAC-learning process, which is the terminology used in Evans and Honkapohja (2001) to bound the parameter interval in ordinary least-squares learning to avoid explosive dynamics.

9

with α_t, β_t as in (2.8) and y_t as in (2.2).

In Hommes and Sorger (1998), the map f in (2.9) is a nonlinear deterministic function depending only on $\alpha_{t-1} + \beta_{t-1}^2(x_{t-1} - \alpha_{t-1})$, without the driving variable y_t and the noise u_t. Hommes et al. (2013) extend the CEE framework to SCEE, with f a nonlinear stochastic process (without exogenous driving variable y_t), but existence and stability under learning are hard to obtain in a nonlinear stochastic framework. Here, in order to make the model analytically tractable, the map f is assumed to be a linear function, depending on the forecast $\alpha_{t-1} + \beta_{t-1}^2(x_{t-1} - \alpha_{t-1})$, the noise u_t, and on an exogenous AR(1) process y_t.

3 Main results in a simple linear framework

Assume that the true law of motion of the economy is a one-dimensional linear stochastic process x_t, driven by an exogenous AR(1) process y_t. More precisely, the actual law of motion of the economy is given by

$$x_t = f(x_{t+1}^e, \ y_t, \ u_t) = b_0 + b_1 x_{t+1}^e + b_2 y_t + u_t, \tag{3.1}$$

$$y_t = a + \rho y_{t-1} + \varepsilon_t, \tag{3.2}$$

where $0 < \rho < 1$, as before, and b_1 is in the interval $(-1, 1)$.[9] Before turning to SCEE, consider rational expectations first.

3.1 Rational expectations equilibrium

Under the assumption that agents are rational, a straightforward computation (see Appendix B) shows that the rational expectations equilibrium x_t^* satisfies

$$x_t^* = \frac{b_0}{1 - b_1} + \frac{ab_1 b_2}{(1 - b_1 \rho)(1 - b_1)} + \frac{b_2}{1 - b_1 \rho} y_t + u_t. \tag{3.3}$$

Thus based on the expression of the rational expectations equilibrium x_t^* in (3.3), its unconditional mean and variance are, respectively,

$$\overline{x^*} := E(x_t^*) = \frac{b_0(1 - \rho) + ab_2}{(1 - b_1)(1 - \rho)}, \tag{3.4}$$

$$Var(x_t^*) = E(x_t^* - \overline{x^*})^2 = \frac{b_2^2 \sigma_\varepsilon^2}{(1 - b_1 \rho)^2(1 - \rho^2)} + \sigma_u^2. \tag{3.5}$$

---

[9]This assumption is made to ensure stationarity; for $|b_1| > 1$ the dynamics under learning easily becomes explosive.

Furthermore, the first-order autocovariance and autocorrelation of rational expectations equilibrium x_t^* are, respectively,

$$E(x_t^* - \overline{x^*})(x_{t-1}^* - \overline{x^*}) = \frac{b_2^2 \rho \sigma_\varepsilon^2}{(1 - b_1\rho)^2(1 - \rho^2)},$$

$$Corr(x_t^*, x_{t-1}^*) = \frac{\rho b_2^2}{b_2^2 + (1 - b_1\rho)^2(1 - \rho^2)\frac{\sigma_u^2}{\sigma_\varepsilon^2}}.$$

Note that in the special case $\sigma_u = 0$, the above expression reduces to $Corr(x_t^*, x_{t-1}^*) = \rho$, that is, when there is no exogenous noise u_t in (3.1), the persistence of the REE coincides exactly with the persistence of the exogenous driving force y_t.

3.2 Existence of first-order SCEE

Now assume that agents are boundedly rational and do not recognize that the economy is driven by an exogenous AR(1) process y_t, but use a simple univariate linear rule to forecast the state x_t of the economy. Given that agents' perceived law of motion is an AR(1) process (2.3), the actual law of motion becomes

$$x_t = b_0 + b_1[\alpha + \beta^2(x_{t-1} - \alpha)] + b_2 y_t + u_t, \tag{3.6}$$

with y_t given in (3.2). The mean of x_t in (3.6), denoted by $\bar{x}$, is computed as

$$\bar{x} = \frac{b_0 + b_1\alpha(1 - \beta^2) + b_2 a/(1 - \rho)}{1 - b_1\beta^2}. \tag{3.7}$$

Imposing the first consistency requirement of a SCEE on the mean, i.e. $\bar{x} = \alpha$, and solving for α yields

$$\alpha^* = \frac{b_0(1 - \rho) + ab_2}{(1 - b_1)(1 - \rho)}. \tag{3.8}$$

Comparing with (3.4), we conclude that in a SCEE the unconditional mean α^* coincides with the REE mean. That is to say, in a SCEE the state of the economy x_t fluctuates on average around its RE fundamental value x^*.

Consider the second consistency requirement of a SCEE on the first-order autocorrelation coefficient β of the PLM. A straightforward computation (see Appendix C) shows that the first-order autocorrelation coefficient $Corr(x_t, x_{t-1})$ of the ALM (3.6) is

$$Corr(x_t, x_{t-1}) = b_1\beta^2 + \frac{b_2^2\rho(1 - b_1^2\beta^4)}{b_2^2(b_1\beta^2\rho + 1) + (1 - \rho^2)(1 - b_1\beta^2\rho)\frac{\sigma_u^2}{\sigma_\varepsilon^2}} =: F(\beta). \tag{3.9}$$

The second consistency requirement of first-order autocorrelation coefficient β yields

$$F(\beta) = \beta. \tag{3.10}$$

11

The actual law of motion (3.1-3.2) depends on seven parameters b_0, b_1, b_2, a, ρ, σ_ε^2 and σ_u^2. The constants b_0 and a only affect the level of fluctuations through the mean α^* in (3.8), but not the persistence, i.e. they do not affect $F(\beta)$ in (3.9). Moreover, only the ratio $\sigma_u^2/\sigma_\varepsilon^2$ of noise terms matters for the persistence $F(\beta)$ in (3.9). Hence, the existence of first-order SCEE (α^*, β^*) depends on four parameters b_1, b_2, ρ and $\sigma_u^2/\sigma_\varepsilon^2$

Define $G(\beta) := F(\beta) - \beta$. Since $0 < \rho < 1$ and $|b_1| < 1$,

$$G(0) = \frac{b_2^2 \rho}{b_2^2 + (1 - \rho^2)\frac{\sigma_u^2}{\sigma_\varepsilon^2}} > 0$$

and

$$
\begin{aligned}
G(1) &= \frac{b_2^2(b_1 + \rho) + b_1(1 - \rho^2)(1 - b_1\rho)\frac{\sigma_u^2}{\sigma_\varepsilon^2}}{b_2^2(b_1\rho + 1) + (1 - \rho^2)(1 - b_1\rho)\frac{\sigma_u^2}{\sigma_\varepsilon^2}} - 1 \\
&= \frac{-b_2^2(1 - b_1)(1 - \rho) - (1 - b_1)(1 - \rho^2)(1 - b_1\rho)\frac{\sigma_u^2}{\sigma_\varepsilon^2}}{b_2^2(b_1\rho + 1) + (1 - \rho^2)(1 - b_1\rho)\frac{\sigma_u^2}{\sigma_\varepsilon^2}} < 0.
\end{aligned}
$$

Therefore, by continuity of G, there exists at least one $\beta^* \in (0, 1)$, such that $G(\beta^*) = 0$, i.e. $F(\beta^*) = \beta^*$. That is,

Proposition 1 *In the case that $0 < \rho < 1$ and $|b_1| < 1$, there exists at least one nonzero first-order stochastic consistent expectations equilibrium (SCEE) (α^*, β^*) for the economic system (3.6) with $\alpha^* = \frac{b_0(1-\rho)+ab_2}{(1-b_1)(1-\rho)} = \overline{x^*}$ and $0 < \beta^* < 1$.*

It is useful to discuss the special case without dependence on an exogenous AR(1) driving variable y_t, that is, $b_2 = 0$ (no driving variable), $\rho = 0$ (no autocorrelation in the driving variable), or $\sigma_\varepsilon^2 = 0$ (no stochasticity in the driving variable). In all these cases, (3.9) reduces to $F(\beta) = b_1\beta^2$. Hence, without an exogenous driving AR(1) process, the unique first-order SCEE $\beta^* = 0$ and coincides with the REE.

Since the solutions β^* of the consistency requirement (3.10) depend continuously on the parameters, we conclude that for $b_2 \approx 0$ (a weak driving variable), $\rho \approx 0$ (almost no autocorrelation in the driving variable), or $\sigma_\varepsilon^2 \approx 0$ (weak stochasticity in the driving variable) the system has a unique SCEE $\beta^* \approx 0$. Hence, when the exogenous AR(1) driving force is weak, there is a unique *low persistence* SCEE.

On the other hand, consider the other extreme case with strong dependence on the AR(1) driving variable y_t, i.e. $|b_2| \to \infty$ (strong dependence on the AR(1) driving variable) or $\sigma_u^2 = 0$ (no exogenous shock u_t, but only an AR(1) driving variable y_t). In both cases, (3.9) reduces to

$$F(\beta) = \frac{b_1\beta^2 + \rho}{b_1\beta^2\rho + 1}. \tag{3.11}$$

In this case we have a unique SCEE (see Appendix D). Furthermore, in the case of positive expectations feedback, i.e., $b_1 > 0$, because $F(0) = \rho$ and $F'(\beta) = \frac{2b_1\beta(1-\rho^2)}{(\rho b_1 \beta^2 + 1)^2} > 0$ for $\beta \in (0,1)$, we have $F(\beta) > \rho$. Consequently

$$\beta^* > \rho.$$

In the special case where also $b_1 = 0$, $F(\beta) \equiv \rho$ and hence $\beta^* = \rho$. Based on the above analysis, we have the following proposition.

Proposition 2 *Under the conditions in Proposition 1, if $b_2 \to \infty$ or $\sigma_u^2 \to 0$, then the nonzero first-order stochastic consistent expectations equilibrium (SCEE) (α^*, β^*) is unique. Furthermore in the case $0 \leq b_1 < 1$, the unique SCEE satisfies $\beta^* \geq \rho$.*

The fact that $\beta^* \geq \rho$ means that along the first-order SCEE the persistence of the economy is larger than under REE. Hence, the fact that agents do not recognize that the economy is driven by a relatively strong exogenous AR(1) process leads to *excess volatility*.

To summarize, when the dependence on the AR(1) driving variable is weak, a unique low persistence SCEE exists. If, on the other hand, the dependence is strong, a unique high persistence, excess volatility SCEE exists. Hence, depending on whether the exogenous driving force is weak or strong, the self-fulfilling mistake leads to a low persistence or a high persistence BLE respectively. It turns out that for intermediate values of the parameter b_2, multiple SCEE may coexist. The next proposition states, however, that at most three different SCEE coexist (the proof is given in appendix E).

Proposition 3 *For the economic system (3.1-3.2) with $0 < \rho < 1$ and $|b_1| < 1$, at most three first-order stochastic consistent expectations equilibria (SCEE) (α^*, β^*) coexist.*

In our applications in Section 4 we will see that the asset pricing model has a unique excess volatility SCEE, while the New Keynesian Philips curve can have multiple SCEE.

3.3 Stability under SAC-learning

In this subsection we study the stability of SCEE under SAC-learning. The ALM of the economy under SAC-learning is given by

$$\begin{cases} x_t = b_0 + b_1[\alpha_{t-1} + \beta_{t-1}^2(x_{t-1} - \alpha_{t-1})] + b_2 y_t + u_t, \\ y_t = a + \rho y_{t-1} + \varepsilon_t. \end{cases} \quad (3.12)$$

with α_t, β_t updated based upon realized sample average and sample autocorrelation as in (2.8). Appendix F shows that the E-stability principle applies and that the stability under SAC-learning is determined by the associated ordinary differential equation (ODE)[10]

$$\begin{cases} \dfrac{d\alpha}{d\tau} = \bar{x}(\alpha, \beta) - \alpha = \dfrac{b_0 + \alpha(b_1 - 1) + b_2 a/(1 - \rho)}{1 - b_1 \beta^2}, \\[3mm] \dfrac{d\beta}{d\tau} = F(\beta) - \beta = \dfrac{b_2^2(b_1\beta^2 + \rho) + b_1\beta^2(1 - \rho^2)(1 - b_1\beta^2\rho)\frac{\sigma_u^2}{\sigma_\varepsilon^2}}{b_2^2(b_1\beta^2\rho + 1) + (1 - \rho^2)(1 - b_1\beta^2\rho)\frac{\sigma_u^2}{\sigma_\varepsilon^2}} - \beta, \end{cases} \tag{3.13}$$

where $\bar{x}(\alpha, \beta)$ is the implied mean given by (3.7) and $F(\beta)$ the implied first-order auto-correlation given by (3.9). A first-order SCEE (α^*, β^*) corresponds to a fixed point of the ODE (3.13). Moreover, a SCEE (α^*, β^*) is locally stable under SAC-learning, if it is a stable fixed point of the ODE (3.13).

A straightforward computation shows that the eigenvalues of the Jacobian $JG(\alpha^*, \beta^*)$ of (3.13) are given by $(b_1 - 1)/(1 - b_1(\beta^*)^2)$ (the coefficient of α in the first ODE) and $F'(\beta^*) - 1$ (since the second ODE is independent of α). Since, by assumption, $|b_1| < 1$ the first eigenvalue is always < 0. Hence, the local stability of a first-order SCEE (α^*, β^*) under SAC-learning only depends on the slope $F'(\beta^*)$:

Proposition 4 *A first-order SCEE (α^*, β^*) is locally stable under SAC-learning if*

$$F'(\beta^*) < 1,$$

where $F(\beta)$ is the implied first-order autocorrelation in (3.9).

Proof. See Appendix F.

Recall from Subsection 3.2 that $F(0) > 0$ and $F(1) < 1$, so that at least one first-order SCEE exists. If the SCEE is unique, then by continuity of F and F' it must be that at the unique intersection point $F'(\beta^*) < 1$ and, according to proposition 4, the unique SCEE is (locally) stable under SAC-learning.[11] Numerical simulations suggest that a unique SCEE is even globally stable under SAC-learning. In the case of multiple first-order SCEE, the graph of the map F has multiple fixed points. Since $F(0) > 0$ and $F(1) < 1$, typically F will then have three fixed points, two locally stable first-order SCEE separated by an unstable SCEE. Indeed in the application of the New Keynesian Philips curve in Subsection 4.2 we will encounter exactly this situation.

---

[10]See Evans and Honkapohja (2001) for discussion and a mathematical treatment of E-stability.

[11]The only exception is a hairline case where the graph of F is tangent to the diagonal at its unique fixed point β^* and $F'(\beta^*) = 1$. In such a hairline case, the SCEE may also be locally stable under SAC-learning, but stability does not follow directly from the E-stability principle.

4 Two applications

In this section we discuss two applications: an asset pricing model driven by AR(1) dividends and a New Keynesian Philips curve driven by an exogenous AR(1) process for the output gap. In both applications we study existence of first-order SCEE and stability under SAC-learning with parameters taken from the empirical literature.

4.1 An asset pricing model with AR(1) dividends

A simple example of the general framework (3.1-3.2) is the standard present value asset pricing model with stochastic dividends; see for example Campbell et al. (1997) and Brock and Hommes (1998). Here we consider autocorrelated AR(1) dividends instead of i.i.d. dividends.

4.1.1 The model

Assume that agents can invest in a risk free asset or in a risky asset. The risk-free asset is perfectly elastically supplied at a gross return $R > 1$. p_t denotes the price (ex dividend) of the risky asset and y_t denotes the (random) dividend process. Let $\widetilde{E}_t, \widetilde{V}_t$ denote the subjective beliefs of a representative agent about the conditional expectation and conditional variance of excess return $p_{t+1} + y_{t+1} - Ry_t$. The representative agent is a myopic mean-variance maximizer of next period's wealth. Optimal demand z_t for the risky asset by the representative agent is then given by

$$z_t = \frac{\widetilde{E}_t(p_{t+1} + y_{t+1} - Rp_t)}{\widetilde{a}\widetilde{V}_t(p_{t+1} + y_{t+1} - Rp_t)} = \frac{\widetilde{E}_t(p_{t+1} + y_{t+1} - Rp_t)}{\widetilde{a}\sigma^2},$$

where $\widetilde{a} > 0$ denotes the risk aversion coefficient and the belief about the conditional variance of the excess return is assumed to be constant over time[12], i.e. $\widetilde{V}_t(p_{t+1} + y_{t+1} - Rp_t) \equiv \sigma^2$.

Equilibrium of demand and supply implies

$$\frac{\widetilde{E}_t(p_{t+1} + y_{t+1} - Rp_t)}{\widetilde{a}\sigma^2} = z_s,$$

---

[12]This assumption is consistent with the assumption that agents believe that prices follow an AR(1) process and dividends follow a stochastic AR(1) process with finite variance. Of course, as discussed in Branch and Evans (2010, 2011), agents might also not know $\widetilde{V}_t$ and need to learn it. However in this paper we focus on the theoretical analysis of SCEE and its stability under SAC-learning in a relatively simple framework. We leave the case of learning of the variance for future work.

where z_s denotes the supply of outside shares in the market, assumed to be constant over time. Without loss of generality[13], we assume zero supply of outside shares, i.e. $z_s = 0$. The market clearing price in the standard asset pricing model is then given by

$$p_t = \frac{1}{R}\left[p_{t+1}^e + y_{t+1}^e\right], \tag{4.1}$$

where p_{t+1}^e is the conditional expectation of next period price p_{t+1} and y_{t+1}^e is the conditional expectation of next period dividend y_{t+1}.

Dividend $\{y_t\}$ follows an AR(1) process (2.2) and the process is known. Suppose that the risky asset (share) is traded, after payment of real dividends y_t, at a competitively determined price p_t, so that y_t is known by the agents, and the dividend forecast is[14]

$$y_{t+1}^e = a + \rho y_t. \tag{4.2}$$

The market clearing price in the standard asset pricing model with AR(1) dividends is then given by

$$p_t = \frac{1}{R}\left[p_{t+1}^e + a + \rho y_t\right]. \tag{4.3}$$

Compared to the general framework (3.1), we have $b_0 = \frac{a}{R}$, $b_1 = \frac{1}{R}$, $b_2 = \frac{\rho}{R}$ and $\sigma_u = 0$.

4.1.2 Theoretical results

Following the general results on SCEE in Section 3, the rational expectations equilibrium p_t^* becomes

$$p_t^* = \frac{aR}{(R-1)(R-\rho)} + \frac{\rho}{R-\rho}y_t. \tag{4.4}$$

In particular, if $\{y_t\}$ is i.i.d., i.e., $a = \bar{y}$ and $\rho = 0$, then $p_t^* \equiv \frac{a}{R-1} = \frac{\bar{y}}{R-1}$ is constant. The corresponding mean, variance and first-order autocorrelation coefficient of the rational

---

[13]In the case $z_s > 0$, the difference with the analysis below only lies in the mean of the SCEE $\alpha^* = \frac{\bar{y}-\tilde{a}\sigma^2 z_s}{R-1}$. The analysis on autocorrelations and variances remains the same.

[14]Agents are thus assumed to know the exogenous dividend process and forecast it correctly. An exogenous dividend process is easier to forecast than endogenously determined equilibrium prices. In a homogeneous rational world, agents believe that prices are completely determined by dividends and use the dividend process to compute rational equilibrium prices. Our agents however are boundedly rational and, although they observe the dividends y_t, believe that prices are not completely determined by dividends, but that "other factors" may affect prices in an economy whose structure they do not fully understand. As a first order approximation of capturing these "other factors", our boundedly rational agents simply use a parsimonious univariate AR(1) rule to forecast endogenous asset prices.

expectation price p_t^* are given by, respectively,

$$\overline{p^*} := E(p_t^*) \quad = \quad \frac{a}{(R-1)(1-\rho)} = \frac{\bar{y}}{R-1}, \tag{4.5}$$

$$Var(p_t^*) \quad = \quad E(p_t^* - \overline{p^*})^2 = \frac{\rho^2 \sigma_\varepsilon^2}{(R-\rho)^2(1-\rho^2)} \tag{4.6}$$

$$Corr(p_t^*, p_{t-1}^*) \quad = \quad \rho. \tag{4.7}$$

Under the assumption that agents are boundedly rational and believe that the price p_t follows a univariate AR(1) process, the implied actual law of motion for prices is

$$\begin{cases} p_t = \dfrac{1}{R}\big[\alpha + \beta^2(p_{t-1} - \alpha) + a + \rho y_t\big], \\[2mm] y_t = a + \rho y_{t-1} + \varepsilon_t. \end{cases} \tag{4.8}$$

Applying the theoretical results in Section 3, the corresponding first-order autocorrelation coefficient $F(\beta)$ of the ALM (4.8) satisfies

$$F(\beta) = \frac{\beta^2 + R\rho}{\rho\beta^2 + R}. \tag{4.9}$$

Using Propositions 2 and 4 we have the following property for the asset pricing model[15]:

Corollary 1 *For the asset pricing model (4.8), the first-order SCEE (α^*, β^*) is unique, $\alpha^* = \frac{\bar{y}}{R-1} = \overline{p^*}$ and $\beta^* > \rho$ (excess volatility), and it is stable under SAC-learning.*

4.1.3 Numerical analysis

We illustrate these results by numerical simulations for empirically plausible parameter values. We choose $R = 1.05, \rho = 0.9, a = 0.005, \varepsilon_t \sim$ i.i.d. $U(-0.01, 0.01)$ (i.e. uniform distribution on $[-0.01, 0.01])$[16]. Figure 1a illustrates the existence of a unique first-order SCEE $(\alpha^*, \beta^*) = (1, 0.997)$, stable under SAC-learning. The time series of fundamental prices and market prices along the first-order SCEE, i.e., with $(\alpha, \beta) = (\alpha^*, \beta^*)$, are shown in Figure 1b, illustrating that the market price fluctuates around the fundamental price but has much more persistence and exhibits excess volatility. Recall from Corollary 1, that in a SCEE the mean of the market prices is equal to that of the fundamental prices and the first-order autocorrelation coefficient β^* of the market prices is greater than that

---

[15]Sögner and Mitlöhner (2002) consider the case with IID dividends, i.e. $\rho = 0$, and show that the unique SCEE coincides with the REE.

[16]As shown theoretically above, the numerical results are independent of the selection of the parameter values within plausible ranges, sample paths, initial values and distribution of noise.

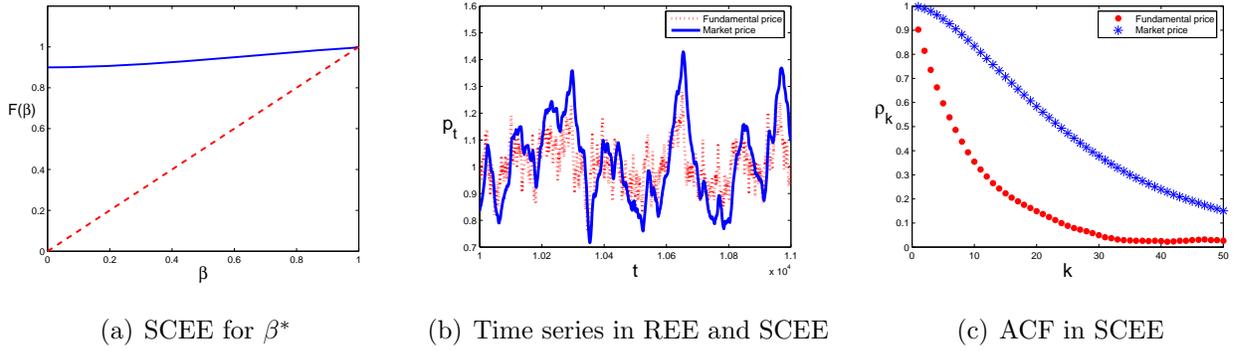(a) SCEE for β^* (b) Time series in REE and SCEE   (c) ACF in SCEE

Figure 1: (a) SCEE $\beta^*(= 0.997)$ is the intersection point of the first-order autocorrelation coefficient $F(\beta) = \frac{\beta^2 + R\rho}{\rho\beta^2 + R}$ (bold curve) with the perceived first-order autocorrelation β (dotted line); (b) RE fundamental prices (dotted curve) and market prices (bold curve) along the SCEE; (c) Autocorrelation Functions (ACF) of RE fundamental prices (lower dots) and market prices (higher stars) along the SCEE.

of the fundamental prices ρ, implying that the market prices have higher persistence. The autocorrelation functions of the market prices and the fundamental prices are shown in Figure 1c. The autocorrelation coefficients of the market prices along a SCEE are substantially higher than those of the fundamental prices and hence the market prices have much higher persistence.

We now investigate how the excess volatility of market prices along a SCEE depends on the autoregressive coefficient of dividends ρ, which is also the first-order autocorrelation of fundamental prices. Consistent with Corollary 1, Figure 2a illustrates that the first-order autocorrelation β^* of market prices is significantly higher than that of fundamental prices, especially for $\rho > 0.4$. For $\rho \geq 0.5$ we have $\beta^* > 0.9$, implying that asset prices are close to a random walk and therefore quite unpredictable. In fact, based on empirical findings, e.g. Timmermann (1996) and Branch and Evans (2010), the autoregressive coefficient of dividends ρ is about 0.9, where the corresponding $\beta^* \approx 0.997$, very close to a random walk. In the case $\rho > 0.4$, the corresponding unconditional variance of market prices is larger than that of fundamental prices. As illustrated in Figure 2b, the ratio of the variance of market prices and the variance of fundamental prices is greater than 1 for $0.4 < \rho < 1$, with a peak around 3.5 for $\rho = 0.7$. For $\rho = 0.9$, $\frac{\sigma_p^2}{\sigma_{p^*}^2} \approx 2.5$, that is, excess volatility by a factor of more than two for empirically relevant parameter values. Given the variance of fundamental prices (4.6) and the variance of market prices (C.8), with
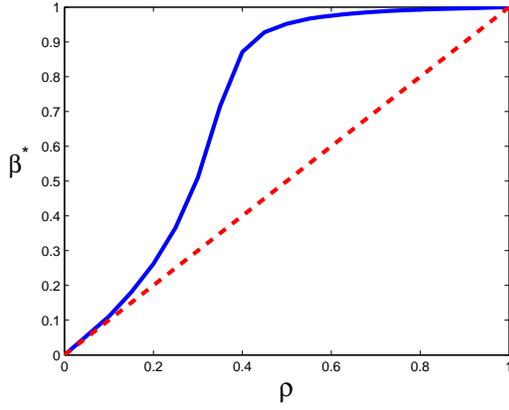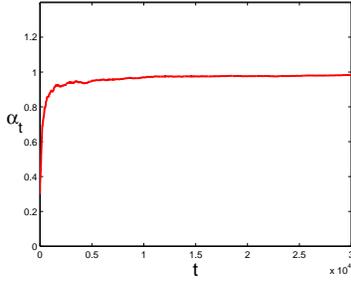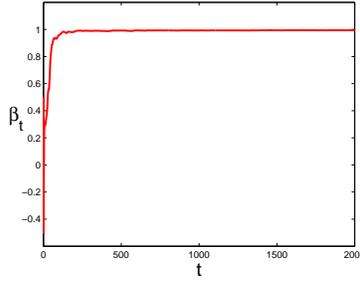
(a) first-order autocorrelation in SCEE
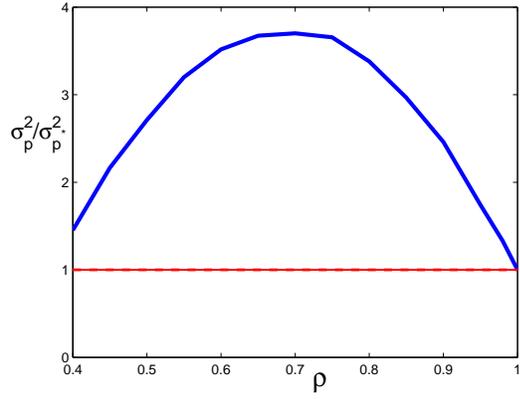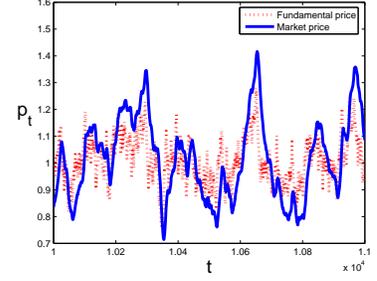
(b) Ratio unconditional variances in SCEE

Figure 2: (a) first-order SCEE β^* with respect to ρ; (b) ratio of unconditional variances of market prices and fundamental prices with respect to ρ, where $R = 1.05$.

(a)

(b)

(c)

Figure 3: (a) Time series $\alpha_t \to \alpha^*(1.0)$; (b) time series $\beta_t \to \beta^*(0.997)$; (c) time series of market prices under SAC-learning and fundamental prices.

$b_1 = \frac{1}{R}$, $b_2 = \frac{\rho}{R}$ and $\sigma_u = 0$,

$$\frac{\sigma_p^2}{\sigma_{p^*}^2} = \left. \frac{(\beta^2 \rho + R)(R - \rho)^2}{(R^2 - \beta^4)(R - \rho\beta^2)} \right|_{\beta = \beta^*(\rho)}.$$

Corollary 1 demonstrates $\rho < \beta^*(\rho) < 1$ for $0 < \rho < 1$, and hence $\beta^*(\rho)$ converges to 1 as ρ tends to 1. Thus as ρ tends to 1, $\frac{\sigma_p^2}{\sigma_{p^*}^2}$ converges to 1, consistent with Figure 2b. For plausible parameter values of ρ, the variance of market prices is significantly greater than that of RE fundamental prices, indicating that market prices exhibit excess volatility along the SCEE.

Figure 3 illustrates that the unique SCEE (α^*, β^*) is stable under SAC-learning. Figure 3a shows that the sample mean of the market prices under SAC-learning, α_t, tends

to the mean $\alpha^* = 1$, while Figure 3b shows that the first-order sample autocorrelation coefficient of the market prices under SAC-learning, β_t, tends to the first-order autocorrelation coefficient $\beta^* = 0.997$. Figure 3c shows the asset price under SAC-learning, using the same sample path of noise, as the time series of the SCEE in Figure 1c. Since the times series are almost the same, SAC-learning converges to the SCEE rather quickly.

In summary, the first-order SCEE and SAC-learning offer an explanation of high persistence, excess volatility and bubbles and crashes in asset prices within a stationary time series framework.[17]

4.2   A New Keynesian Philips curve

Our second application of SCEE and SAC-learning uses the New Keynesian macro model (Woodford, 2003). We consider the New Keynesian Philips curve (NKPC) with inflation driven by an exogenous AR(1) process for the output gap (often measured by detrended real GDP) or the firm's real marginal cost (often measured by labor's share of income), as in Lansing (2009).

There are two approaches in the adaptive learning literature in the NKPC. Bullard and Mitra (2002) and Evans and Honkapohja (2003, 2006) study monetary policy interest rate rules using the Euler Equation (EE) learning approach, where agents base their decisions on Euler equations derived under one-period-ahead subjective expectations. Preston (2005) studies monetary policy in the NKPC under Infinite Horizon (IH) learning, where decisions of individual agents are made under infinite horizon subjective expectations. A number of recent papers address the question under which conditions the EE and IH learning approaches lead to the same learning dynamics and/or the same stability conditions, e.g. Branch and McGough (2009), Honkapohja et al. (2013) and Woodford (2013). Evans and Mitra (2013) recently showed for the stochastic Ramsey model that the rational expectations equilibrium is stable under both EE and IH learning, but there are differences in the transitory learning dynamics. Branch, Evans and McGough (2012) develop a theory of bounded rationality called finite-horizon learning, generalizing EE and

---

[17]We also simulated SAC-learning with a constant gain parameter (see the online Supplementary Material) and, similar to Branch and Evans (2011), obtained persistent near unit root bubble and crash dynamics. When the autocorrelation in the driving process is low these unit root bubble and crash dynamics are transitory and recurrent after a series of shocks; for higher values of $\rho \geq 0.4$ persistent near-unit root bubble and crash dynamics arise, because of the existence of a unique stable high persistence SCEE; cf. Figure 2a, where the SCEE β^* is plotted as a function of ρ.

IH learning, where agents decisions are based on a finite planning horizon. The interplay between optimization, learning and the forecasting horizon of boundedly rational agents is an important (and difficult) topic that needs to be addressed carefully in future work on microfoundations of macroeconomics with boundedly rational agents[18]. Here we study SCEE and SAC-learning in the NKPC with EE learning, which fits exactly in our general linear framework of Section 3[19].

In the New Keynesian Philips curve (NKPC) with inflation driven by an exogenous AR(1) process y_t for the firm's real marginal cost or the output gap, inflation and the real marginal cost (output gap) evolve according to (see e.g. Woodford, 2003)

$$
\begin{cases}
\pi_t = \delta \pi_{t+1}^e + \gamma y_t + u_t, \\
y_t = a + \rho y_{t-1} + \varepsilon_t,
\end{cases}
\tag{4.10}
$$

where π_t is the inflation at time t, π_{t+1}^e is the subjective expected inflation at date $t+1$ and y_t is the output gap or real marginal cost, $\delta \in [0,1)$ is the representative agent's subjective time discount factor, $\gamma > 0$ is related to the degree of price stickiness in the economy and $\rho \in [0,1)$ describes the persistence of the AR(1) driving process. u_t and ε_t are i.i.d. stochastic disturbances with zero mean and finite absolute moments with variances σ_u^2 and σ_ε^2, respectively. The most important difference with the asset pricing model in Subsection 4.1 is that (4.10) includes two stochastic disturbances, not only the noise ε_t of the AR(1) driving variable, but also an additional noise term u_t in the New Keynesian Philips curve. We refer to u_t as a markup shock, which is often motivated by the presence of an uncertain variable tax rate and to ε_t as a demand shock, that is uncorrelated with the markup shock. Compared with our general framework (3.1), the corresponding parameters are $b_0 = 0$, $b_1 = \delta$ and $b_2 = \gamma$.

---

[18]An interesting new approach is Evans and McGough (2013), who derive simple behavioral primitives, called shadow price learning, for agents who are boundedly rational in their individual allocation decisions as well as in forecasting.

[19]In the online Supplementary Material we derive the NKPC with IH learning and an AR(1) forecasting rule from microfoundations and show that IH learning leads to a (somewhat) different system. The NKPC under IH learning is more complicated with additional nonlinearities and does not fit exactly into our simple general framework, but, as shown in the online Supplementary Material, the SCEE and SAC-learning dynamics turn out to be quite similar to the case of EE learning.

4.2.1 Theoretical results

Following the general results in Section 3, the rational expectations equilibrium

$$\pi_t^* = \frac{\gamma \delta a}{(1-\delta)(1-\delta\rho)} + \frac{\gamma}{1-\delta\rho} y_t + u_t. \tag{4.11}$$

The corresponding mean, variance and first-order autocorrelation coefficient of the rational expectations equilibrium π_t^* are given by, respectively,

$$\overline{\pi^*} := E(\pi_t^*) = \frac{\gamma a}{(1-\delta)(1-\rho)}, \tag{4.12}$$

$$Var(\pi_t^*) = E(\pi_t^* - \overline{\pi^*})^2 = \frac{\gamma^2 \sigma_\varepsilon^2}{(1-\delta\rho)^2(1-\rho^2)} + \sigma_u^2, \tag{4.13}$$

$$Corr(\pi_t^*, \pi_{t-1}^*) = \frac{\rho\gamma^2}{\gamma^2 + (1-\delta\rho)^2(1-\rho^2)\frac{\sigma_u^2}{\sigma_\varepsilon^2}}. \tag{4.14}$$

Note that, the larger the noise level σ_u^2 in the markup shock, the smaller the first-order autocorrelation in the fundamental rational equilibrium inflation.

Under the assumption that agents are boundedly rational and believe that inflation π_t follows a univariate AR(1) process, the implied actual law of motion becomes[20]

$$\begin{cases} \pi_t = \delta[\alpha + \beta^2(\pi_{t-1} - \alpha)] + \gamma y_t + u_t, \\ y_t = a + \rho y_{t-1} + \varepsilon_t. \end{cases} \tag{4.15}$$

Following the theoretical results in Section 3, the corresponding first-order autocorrelation coefficient $F(\beta)$ of the implied ALM (4.15) is

$$F(\beta) = \delta\beta^2 + \frac{\gamma^2\rho(1-\delta^2\beta^4)}{\gamma^2(\delta\beta^2\rho+1) + (1-\rho^2)(1-\delta\beta^2\rho)\cdot\frac{\sigma_u^2}{\sigma_\varepsilon^2}}. \tag{4.16}$$

Applying Proposition 1 in Section 3.2 we obtain

Corollary 2 *In the case that $0 < \rho < 1$ and $0 \leq \delta < 1$, there exists at least one nonzero first-order stochastic consistent expectations equilibrium (SCEE) (α^*, β^*) for the New Keynesian Philips curve (4.15) with $\alpha^* = \frac{\gamma a}{(1-\delta)(1-\rho)} = \overline{\pi^*}$.*

For the New Keyensian Philips curve (4.15), however, multiple SCEE may coexist.

---

[20]In line with Subsection 3.2 we assume that boundedly rational agents do not recognize or do not believe that inflation is driven by output or marginal costs, but simply forecast inflation by an univariate AR(1) rule; see also footnote 14 for similar reasoning in the asset pricing model.

4.2.2 Numerical analysis

The NKPC under SAC-learning is given by

$$\begin{cases} \pi_t = \delta[\alpha_t + \beta_t^2(\pi_{t-1} - \alpha_t)] + \gamma y_t + u_t, \\ y_t = a + \rho y_{t-1} + \varepsilon_t, \end{cases} \qquad (4.17)$$

with α_t, β_t updated based upon realized sample average and sample autocorrelation as in (2.8). In this subsection we investigate the multiplicity of SCEE and their stability under learning. Based on empirical findings, e.g., in Lansing (2009), Gali et al. (2001) and Fuhrer (2006, 2009), we examine a range of empirically plausible parameter values.[21] First, fix the parameters $\delta = 0.99$, $\gamma = 0.075$, $a = 0.0004$, $\rho = 0.9$, $\sigma_\varepsilon = 0.01$ $[\varepsilon_t \sim N(0, \sigma_\varepsilon^2)]$, and $\sigma_u = 0.003162$ $[u_t \sim N(0, \sigma_u^2)]$, so that $\frac{\sigma_u^2}{\sigma_\varepsilon^2} = 0.1$.

Figure 4a illustrates an example where $F(\beta)$ has three fixed points $\beta_1^* \approx 0.3066$, $\beta_2^* \approx 0.7417$ and $\beta_3^* \approx 0.9961$. Hence, we have coexistence of three first-order SCEE (α^*, β_j^*), $j = 1, 2, 3$. Figures 4b and 4c illustrate the time series of inflation along the coexisting SCEE. Inflation has low persistence along the SCEE (α^*, β_1^*), but very high persistence along the SCEE (α^*, β_3^*). The time series of inflation along the high persistence SCEE in Figure 4c has in fact similar persistence characteristics and amplitude of fluctuation as in empirical inflation data, e.g., in Tallman (2003). Furthermore, Figure 4c illustrates that inflation in the high persistence SCEE has much stronger persistence than REE inflation, where the first-order autocorrelation coefficient of REE inflation is 0.865, significantly less than $\beta_3^* = 0.9961$.

If multiple SCEE coexist, the convergence under SAC-learning depends on the initial state of the system, as illustrated in Figure 5. Since $0 < F'(\beta_j^*) < 1$, for $j = 1$ and $j = 3$, while $F'(\beta_2^*) > 1$, (see Figure 4a), Proposition 4 implies that the first-order SCEE (α^*, β_1^*) and (α^*, β_3^*) are (locally) stable under SAC-learning, while (α^*, β_2^*) is unstable. For initial state $(\pi_0, y_0) = (0.028, 0.01)$ (Figures 5a and 5b), the SAC-learning dynamics (α_t, β_t) converges to the stable low-persistence SCEE $(\alpha^*, \beta_1^*) = (0.03, 0.3066)$. Figure 5b also illustrates that the convergence of the first-order autocorrelation coefficient β_t to

---

[21] As shown in Lansing (2009), based on regressions using either the output gap or labor's share of income over the period 1949.Q1 to 2004.Q4, $\rho = 0.9, \sigma_\varepsilon = 0.01$. Estimates of the NKPC parameters δ, γ, σ_u are sensitive to the choice of the driving variable, the sample period, and the econometric model, etc., but our choices are within a plausible range. Furthermore, based on the above theoretical results, the constant a only affects the mean of inflation $\bar{\pi}$, and not its autocorrelation coefficient $F(\beta)$. Moreover, $F(\beta)$ only depends on the ratio $\sigma_u/\sigma_\varepsilon$, but not on their absolute values.
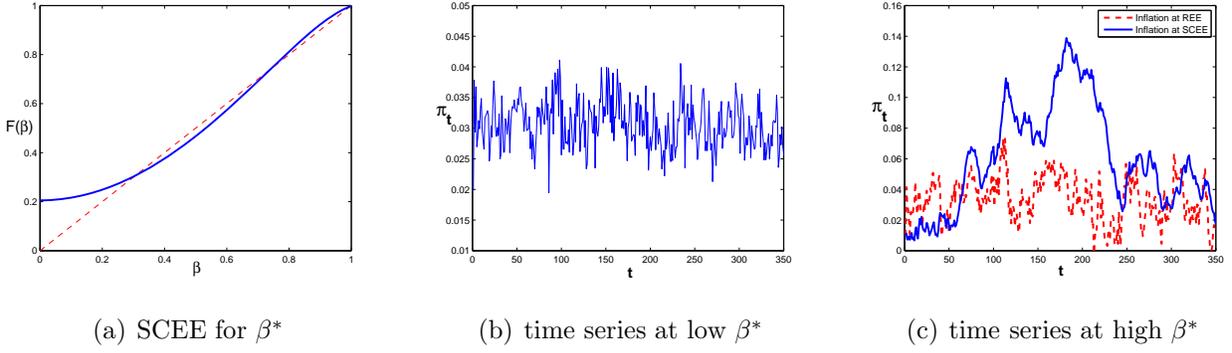
(a) SCEE for β^* (b) time series at low β^* (c) time series at high β^*

Figure 4: (a) The first-order autocorrelation β^* of the SCEE correspond to the three intersection points of $F(\beta)$ in (4.16) (bold curve) with the perceived first-order autocorrelation β (dotted line); (b) time series of inflation in low-persistence SCEE $(\alpha^*, \beta_1^*) = (0.03, 0.3066)$; (c) times series of inflation in high-persistence SCEE $(\alpha^*, \beta_3^*) = (0.03, 0.9961)$ (bold curve) and time series of REE inflation (dotted curve).
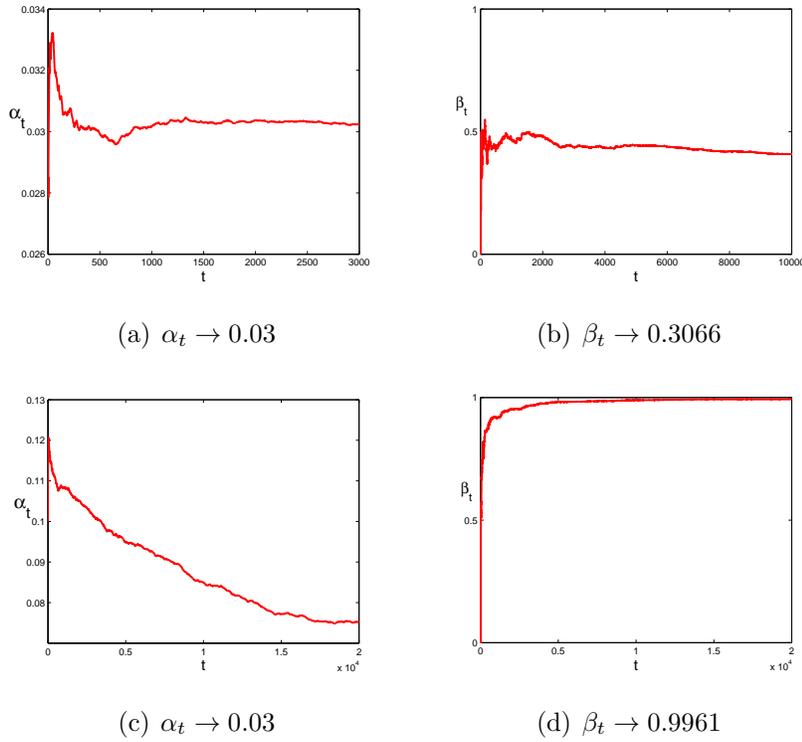

(a) $\alpha_t \to 0.03$ (b) $\beta_t \to 0.3066$

(c) $\alpha_t \to 0.03$ (d) $\beta_t \to 0.9961$

Figure 5: Time series of α_t and β_t under SAC-learning for different initial values. (a-b) For $(\pi_0, y_0) = (0.028, 0.01)$ SAC-learning converges to the low persistence SCEE $(\alpha^*, \beta_1^*) = (0.03, 0.3066)$; (c-d) For $(\pi_0, y_0) = (0.1, 0.15)$ SAC-learning converges to the high persistence SCEE $(\alpha^*, \beta_1^*) = (0.03, 0.9961)$.

24

the low-persistence first-order autocorrelation coefficient $\beta_1^* = 0.3066$ is very slow. For a different initial state, $(\pi_0, y_0) = (0.1, 0.15)$, our numerical simulation shows that the sample mean α_t still tends to $\alpha^* = 0.03$, but only slowly[22] (see Figure 5c), while β_t tends to the high persistence SCEE $\beta_3^* \approx 0.9961$[23] (see Figure 5d).

Numerous simulations (not shown) show that for initial values π_0 of inflation higher than the mean $\alpha^* = 0.03$, the SAC-learning β_t generally enters the high-persistence region. In particular, a large shock to inflation may easily cause a jump of the SAC-learning process into the high-persistence region.[24] In the following we further indicate how high and low persistence SCEE depend on different parameters.

4.2.3 Multiple equilibria and parameter dependence

Figure 6 illustrates how the number of SCEE depends on the parameter γ. For sufficiently small $\gamma(< 0.05)$, there exists only one, low persistence SCEE β^* (Figure 6a). This is similar to the case $\gamma = 0$, where correspondingly $F(\beta) = \delta\beta^2$ and hence the unique SCEE $\beta^* = 0$. Moreover, since

$$\frac{\partial F}{\partial \gamma} = \frac{2\rho(1 - \delta^2\beta^4)(1 - \rho^2)(1 - \delta\beta^2\rho)\frac{\sigma_u^2}{\sigma_\varepsilon^2}}{\gamma^3\left[(\delta\beta^2\rho + 1) + (1 - \rho^2)(1 - \delta\beta^2\rho)\frac{1}{\gamma^2}\frac{\sigma_u^2}{\sigma_\varepsilon^2}\right]^2} > 0,$$

the graph of $F(\beta)$ in (4.16) shifts upward as γ increases. At some critical γ-value, a tangent bifurcation occurs. Immediately thereafter, there exist three SCEE, β_1^*, β_2^* and β_3^* (see Figure 6b). The low persistence SCEE β_1^* and the high persistence SCEE β_3^* are stable under SAC-learning, since $0 < F'(\beta_j^*) < 1$, $j = 1$ and $j = 3$, separated by an unstable SCEE β_2^*, with $F'(\beta_2^*) > 1$. As γ further increases, another tangent bifurcation occurs and the low persistence SCEE disappears. A unique high persistence SCEE then remains, which is stable under SAC-learning (Figure 6c).

The dependence of the number of SCEE and their persistence upon the parameter γ are quite intuitive. Recall that γ in (4.10) measures the relative strength of the driving variable, the output gap or marginal costs, to inflation.[25] When the driving force is

---

[22]The slow convergence is caused by the slope coefficient $\frac{\delta - \delta\beta^2}{1 - \delta\beta^2}$ for α in the expression for the mean $\overline{\pi}$, which is very close to 1 for $\delta = 0.99 \approx 1$.

[23]As shown in Figure 4a, $F'(\beta_3^*)$ is close to 1 and, hence, the convergence of SAC-learning is very slow.

[24]We also simulated the NKPC under SAC-learning with a constant gain parameter (see the online Supplementary Material) and, similar to Branch and Evans (2010), obtained irregular regime switching between phases of very low persistence and phases of high persistence with near unit root behavior.

[25]Note that γ corresponds to the parameter b_2 in the general linear specification (3.1-3.2). See Proposition 2 and the discussion in Subsection 3.2 how the low and high persistence SCEE depend on b_2.
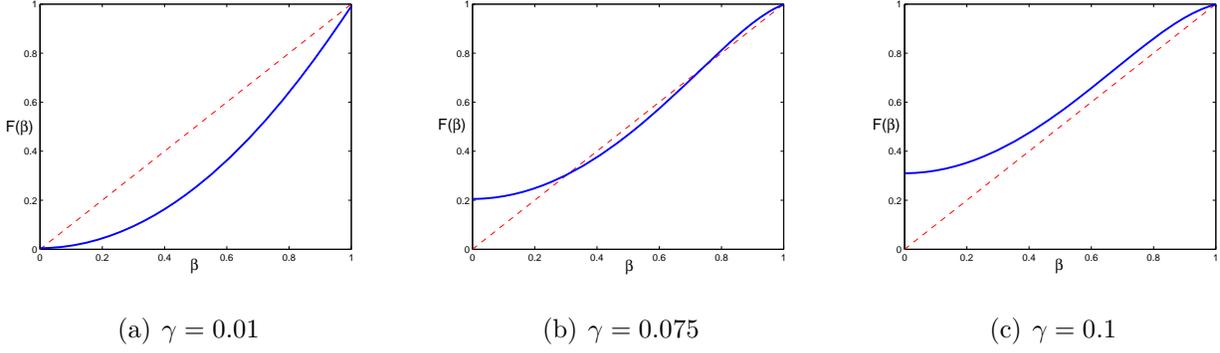
(a) $\gamma = 0.01$ (b) $\gamma = 0.075$ (c) $\gamma = 0.1$

Figure 6: The figure illustrates how the (co-)existence of low and high persistence SCEE β^* depends upon the parameter γ, measuring the relative strength of inflation upon the driving variable, the output gap. (a) $\gamma = 0.01$; (b) $\gamma = 0.075$; (c) $\gamma = 0.1$. Other parameters: $\frac{\sigma_u^2}{\sigma_\varepsilon^2} = 0.1$, $\rho = 0.9$ and $\delta = 0.99$.

relatively weak, a unique, stable low persistence SCEE prevails, with much weaker autocorrelation than in the driving variable. At the other extreme, when the driving force is sufficiently strong, a unique, stable high persistence SCEE prevails, with significantly stronger autocorrelation and higher persistence than in the driving variable. In the intermediate case, multiple SCEE coexist and the system exhibits path dependence, where, depending on initial conditions, inflation converges to a low or a high persistence SCEE.

In a similar way, the dependence of the SCEE upon the noise ratio $\frac{\sigma_u^2}{\sigma_\varepsilon^2}$ can be analyzed. $F(\beta)$ in (4.16) can be rewritten as

$$F(\beta) = \delta\beta^2 + \frac{\rho(1 - \delta^2\beta^4)}{(\delta\beta^2\rho + 1) + (1 - \rho^2)(1 - \delta\beta^2\rho) \cdot \frac{\sigma_u^2}{\sigma_\varepsilon^2}\frac{1}{\gamma^2}}.$$

Consequently, the effect of the noise ratio $\frac{\sigma_u^2}{\sigma_\varepsilon^2}$ is inversely related to the effect of γ. Hence, when the ratio $\frac{\sigma_u^2}{\sigma_\varepsilon^2}$ is high, that is, when the markup shocks to inflation are high compared to the noise of the driving variable, a unique, stable low persistence SCEE prevails. If on the other hand, the markup shocks to inflation are low compared to the noise of the driving variable, a unique, stable high persistence SCEE prevails.

Furthermore, Figure 7 illustrates how the SCEE β^*, together with the first-order autocorrelation coefficient of REE inflation, depends upon the parameter ρ, measuring the persistence in the driving variable. For intermediate values of $\rho(\in [0.84, 0.918])$, two stable SCEE β^* coexist separated by an unstable SCEE. In the high persistence SCEE, β^* is larger than the first-order autocorrelation coefficient of REE inflation, while in the low persistence SCEE β^* is smaller than the first-order autocorrelation coefficient of REE
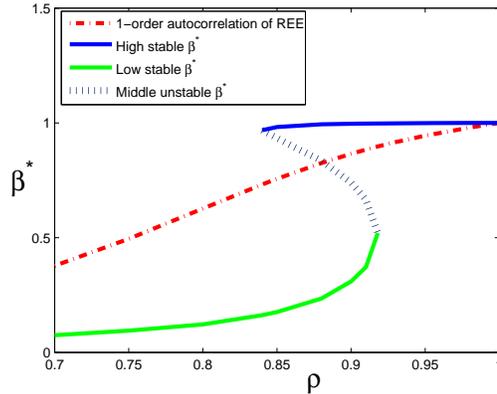
Figure 7: First-order autocorrelation coefficient of REE inflation (dotted real curve), stable SCEE β^* with respect to ρ (bold curves), unstable SCEE β^* (dotted curve), where $\gamma = 0.075, \sigma_u = 0.003162, \sigma_\varepsilon = 0.01, \delta = 0.99$.

inflation. For small values of ρ, $\rho < 0.84$, a unique, stable low persistence SCEE prevails, while for large values of ρ, $\rho > 0.918$, a unique, stable high persistence SCEE prevails.

Simulations show that, for plausible values of ρ around 0.9, for a large range of initial values of inflation, the SAC-learning converges to the stable, high persistence SCEE β^* with very strong persistence in inflation (see e.g. Figure 5d). This result is consistent with the empirical finding in Adam (2007) that the Restricted Receptions Equilibrium (RPE) describes subjects' inflation expectations surprisingly well and provides a better explanation for the observed persistence of inflation than REE.

In summary, the dependence of the number of equilibria and whether their persistence is high or low are quite intuitive. This intuition essentially follows from the signs of the partial derivatives of the first-order autocorrelation coefficient $F(\beta)$ in (4.16) of the implied ALM (4.15) satisfying (see Appendix G):

$$\frac{\partial F}{\partial \gamma} > 0 \qquad \frac{\partial F}{\partial (\frac{\sigma_u^2}{\sigma_\varepsilon^2})} < 0 \qquad \frac{\partial F}{\partial \rho} > 0 \qquad \frac{\partial F}{\partial \delta} > 0. \tag{4.18}$$

Hence, as in Figure 6, the graph of $F(\beta)$ shifts upwards when γ increases, $\frac{\sigma_u^2}{\sigma_\varepsilon^2}$ decreases, ρ increases or δ increases, and consequently, the equilibria shift from low persistence to high persistence equilibria. Depending on the shape of $F(\beta)$ there are then two possibilities. When F is only weakly nonlinear, e.g., as in Figure 1a for the asset pricing model, the equilibrium is unique and only a gradual shift from a low to a high persistence equilibrium arises. When the nonlinearity is stronger and F is S-shaped, e.g., as in Figure 6 for empirically relevant parameter values in the NKPC, both the persistence and the number

27

of equilibria shift, and a transition from a unique stable low persistence SCEE, through coexisting stable low and high persistence equilibria, to a unique stable high persistence equilibrium occurs. Such a transition from a unique low persistence SCEE, through coexisting low and high persistence SCEE, toward a unique high persistence SCEE occurs when the strength of the AR(1) driving force (the parameter γ) increases, when the ratio of the model noise compared to the noise of the driving force (i.e. $\frac{\sigma_u^2}{\sigma_{\bar{\varepsilon}}^2}$) decreases, when the autocorrelation (i.e., the parameter ρ) in the driving force increases, and when the strength of the expectations feedback (i.e., the parameter δ) increases.

5   Concluding remarks

We have introduced the concept of behavioral learning equilibrium, a very simple type of misspecification equilibrium with an intuitive behavioral interpretation and learning process. Boundedly rational agents use a univariate linear forecasting rule and in equilibrium correctly forecast the unconditional mean and first-order autocorrelation. Hence, to a first order approximation the simple linear forecasting rule is consistent with observed market realizations. Sample autocorrelation learning means that agents are gradually updating the two coefficients –sample mean and first-order autocorrelation– of their linear rule. In the long run, agents thus learn the best univariate linear forecasting rule, without fully recognizing the more complex structure of the economy.

We have applied our behavioral learning equilibrium concept to a standard asset pricing model with AR(1) dividends and a New Keynesian Philips curve driven by an AR(1) process for the output gap or marginal costs. In both applications, the law of motion of the economy is linear, but it is driven by an exogenous stochastic AR(1) process. Agents however are not fully aware of the exact linear structure of the economy and all explanatory variables, but use a simple univariate forecasting rule, to predict asset prices or inflation. In the asset pricing model a unique SCEE exists and it is stable under SAC-learning. An important feature of the SCEE is that it is characterized by high-persistence and excess volatility in asset prices, significantly higher than under rational expectations. In the New Keynesian model, multiple SCEE arise and a low and a high-persistence misspecification equilibrium coexist. The SAC-learning exhibits path dependence and it depends on the initial states whether the system converges to the low-persistence or the high-persistence inflation regime. In particular, when there are shocks– e.g. oil shocks– temporarily causing high inflation, SAC-learning may lock into the high-persistence inflation regime.

Are these behavioral learning equilibria empirically relevant or would smart agents recognize their (second order) mistakes and learn to be perfectly rational? This empirical question should be addressed in more detail in future work, but we provide some arguments for the empirical relevance of our equilibrium concept. Firstly, in our applications the SCEE already explain some important stylized facts: (i) high persistence (close to unit root) and excess volatility in asset prices, (ii) high persistence in inflation and (iii) regime switching in inflation dynamics, which could explain a long phase of high US inflation in the 1970s and early 1980s as well as a long phase of low inflation in the 1990s and 2000s. Secondly, we stress simplicity, parsimony and the *behavioral* interpretation of our learning equilibrium concept. The univariate AR(1) rule and the SAC-learning process are examples of simple forecasting heuristics that can be used without any knowledge of statistical techniques, simply by observing a time series and roughly "guestimating" its sample average and its first-order persistence coefficient. Coordination on a behavioral forecasting heuristic that performs reasonably well to a first-order approximation seems more likely than coordination on more complicated learning or sunspot equilibria as, for example, in Woodford (1990). Even though some smart individual agents might be able to improve upon the best linear, univariate forecasting rule, a majority of agents might still stick to their simple univariate rule. It therefore seems relevant to describe aggregate phenomena by simple misspecification equilibria and behavioral learning processes. Our behavioral learning equilibrium concept also relates to the "natural expectations" in Fuster et al. (2010) and Beshears et al. (2013), emphasizing parsimonious forecasting rules giving much weight to recent changes to explain the long-run persistence of economic shocks. Our simple univariate AR(1) rule may be seen as the most parsimonious forecasting rule leading to long-run persistence. There is experimental evidence for the relevance of misspecification equilibria in Adam (2007). More recently Assenza et al. (2012) and Pfajfar and Zakelj (2011) ran learning to forecasting experiments with human subjects in a New Keynesian framework with expectations feedback from individual inflation and output gap forecasts. Coordination on simple linear univariate models explain a substantial part of individual inflation and output gap forecasting behavior.

In future work we plan to consider more general economic settings to study behavioral learning equilibria. An obvious next step is to apply our BLE and SAC-learning framework to higher dimensional linear economic systems, with agents forecasting by univariate linear rules. In particular, the fully specified New Keynesian model of inflation and output dynamics would be an interesting (two-dimensional) application. Including asset prices

in a New Keynesian model, as in Bernanke and Gertler (1999, 2001), provides another interesting (three-dimensional) application. It is also interesting and challenging to study BLE and misspecification under *heterogeneous expectations* and allow for switching between different rules. Branch (2004) and Hommes (2011) provide some empirical and experimental evidence on heterogeneous expectations, while Berardi (2007) and Branch and Evans (2006, 2007) have made some related studies on heterogeneous expectations and learning in similar settings. Future work should focus on the robustness and survival of behavioral forecasting rules, such as AR(1) and SAC-learning, in a heterogeneous expectations environment. In addition to theoretical work, it would be of interest to study coordination and learning of BLE in laboratory settings with multiple restricted perception and/or sunspot equilibria. Last but not least, studying macroeconomic policy under BLE is a promising and important area of future research.

Acknowledgements

Appendix

A   Recursive dynamics of SAC-learning

The sample average is

$$
\begin{aligned}
\alpha_t &= \frac{1}{t+1}[x_0 + x_1 + \cdots + x_t] \\
&= \frac{1}{t+1}[t\alpha_{t-1} + x_t] \\
&= \frac{1}{t+1}[(t+1)\alpha_{t-1} + x_t - \alpha_{t-1}] \\
&= \alpha_{t-1} + \frac{1}{t+1}[x_t - \alpha_{t-1}].
\end{aligned}
$$

Let

$$
\begin{aligned}
z_t &:= (x_0 - \alpha_t)(x_1 - \alpha_t) + \cdots + (x_{t-1} - \alpha_t)(x_t - \alpha_t) \\
&= (x_0 - \alpha_{t-1} - \frac{1}{t+1}(x_t - \alpha_{t-1}))(x_1 - \alpha_{t-1} - \frac{1}{t+1}(x_t - \alpha_{t-1})) + \\
&\quad \cdots + (x_{t-1} - \alpha_{t-1} - \frac{1}{t+1}(x_t - \alpha_{t-1}))(x_t - \alpha_{t-1} - \frac{1}{t+1}(x_t - \alpha_{t-1})) \\
&= (x_0 - \alpha_{t-1})(x_1 - \alpha_{t-1}) + \cdots + (x_{t-2} - \alpha_{t-1})(x_{t-1} - \alpha_{t-1}) \\
&\quad + \frac{x_t - \alpha_{t-1}}{t+1}(2\alpha_{t-1} - x_0 - x_1 + \cdots + 2\alpha_{t-1} - x_{t-2} - x_{t-1}) + \frac{t-1}{(t+1)^2}(x_t - \alpha_{t-1})^2 \\
&\quad + \frac{t}{t+1}(x_{t-1} - \alpha_{t-1})(x_t - \alpha_{t-1}) - \frac{t}{(t+1)^2}(x_t - \alpha_{t-1})^2 \\
&= z_{t-1} + \frac{1}{t+1}(x_t - \alpha_{t-1})[2(t-1)\alpha_{t-1} - x_0 - 2x_1 - \cdots - 2x_{t-2} - x_{t-1} + t(x_{t-1} - \alpha_{t-1})] \\
&\quad - \frac{1}{(t+1)^2}(x_t - \alpha_{t-1})^2, \\
&= z_{t-1} + \frac{1}{t+1}(x_t - \alpha_{t-1})[x_0 + (t+1)x_{t-1} - (t+2)\alpha_{t-1}] - \frac{1}{(t+1)^2}(x_t - \alpha_{t-1})^2 \\
&= z_{t-1} + (x_t - \alpha_{t-1})\left[x_{t-1} + \frac{x_0}{t+1} - \frac{t+2}{t+1}\alpha_{t-1} + \frac{1}{(t+1)^2}\alpha_{t-1} - \frac{1}{(t+1)^2}x_t\right] \\
&= z_{t-1} + (x_t - \alpha_{t-1})\Phi_4,
\end{aligned}
$$

31

where $\Phi_4 = x_{t-1} + \frac{x_0}{t+1} - \frac{t^2+3t+1}{(t+1)^2}\alpha_{t-1} - \frac{1}{(t+1)^2}x_t$.

Write

$$
\begin{aligned}
n_t \;:=\; & (x_0 - \alpha_t)^2 + (x_1 - \alpha_t)^2 + \cdots + (x_t - \alpha_t)^2 \\
=\; & (x_0 - \alpha_{t-1} - \frac{1}{t+1}(x_t - \alpha_{t-1}))^2 + \cdots + (x_t - \alpha_{t-1} - \frac{1}{t+1}(x_t - \alpha_{t-1}))^2 \\
=\; & (x_0 - \alpha_{t-1})^2 + (x_1 - \alpha_{t-1})^2 + \cdots + (x_{t-1} - \alpha_{t-1})^2 + \frac{t+t^2}{(t+1)^2}(x_t - \alpha_{t-1})^2 \\
=\; & n_{t-1} + \frac{t}{t+1}(x_t - \alpha_{t-1})^2.
\end{aligned}
$$

All these results are consistent with those in Appendix 1 of Hommes, Sorger & Wagener (2004). Note that in our paper R_t is different from n_t in Hommes et al. (2004). In fact,

$$
\begin{aligned}
R_t \;=\; & \frac{1}{t+1}n_t \\
=\; & \frac{1}{t+1}[n_{t-1} + \frac{t}{t+1}(x_t - \alpha_{t-1})^2] \\
=\; & \frac{t}{t+1}\frac{n_{t-1}}{t} + \frac{t}{(t+1)^2}(x_t - \alpha_{t-1})^2 \\
=\; & \frac{t}{t+1}R_{t-1} + \frac{t}{(t+1)^2}(x_t - \alpha_{t-1})^2 \\
=\; & R_{t-1} + \frac{1}{t+1}\Big[\frac{t}{t+1}(x_t - \alpha_{t-1})^2 - R_{t-1}\Big].
\end{aligned}
$$

Furthermore,

$$
\begin{aligned}
\beta_t \;=\; & \frac{z_t}{n_t} \\
=\; & \beta_{t-1} + \Big[\frac{z_t}{n_t} - \frac{z_{t-1}}{n_{t-1}}\Big] \\
=\; & \beta_{t-1} + \frac{1}{n_t n_{t-1}}[z_t n_{t-1} - z_{t-1}n_t] \\
=\; & \beta_{t-1} + \frac{1}{n_t n_{t-1}}\Big[\big(z_{t-1} + (x_t - \alpha_{t-1})\Phi_4\big)n_{t-1} - z_{t-1}\big(n_{t-1} + \frac{t}{t+1}(x_t - \alpha_{t-1})^2\big)\Big] \\
=\; & \beta_{t-1} + \frac{1}{n_t n_{t-1}}\Big[(x_t - \alpha_{t-1})\Phi_4 n_{t-1} - z_{t-1}\frac{t}{t+1}(x_t - \alpha_{t-1})^2\Big] \\
=\; & \beta_{t-1} + \frac{1}{n_t}\Big[(x_t - \alpha_{t-1})\Phi_4 - \beta_{t-1}\frac{t}{t+1}(x_t - \alpha_{t-1})^2\Big] \\
=\; & \beta_{t-1} + \frac{R_t^{-1}}{t+1}\Big[(x_t - \alpha_{t-1})\big(x_{t-1} + \frac{x_0}{t+1} - \frac{t^2+3t+1}{(t+1)^2}\alpha_{t-1} - \frac{x_t}{(t+1)^2}\big) - \frac{t}{t+1}\beta_{t-1}(x_t - \alpha_{t-1})^2\Big].
\end{aligned}
$$

B   Rational expectations equilibrium

Under the assumption that the transversality condition $\lim_{k\to\infty} b_1^k E_t(x^*_{t+k}) = 0$ holds, the REE x^*_t can be computed as

$$
\begin{aligned}
x^*_t &= b_0 + b_1 E_t x^*_{t+1} + b_2 y_t + u_t \\
&= b_0 + b_1 E_t[b_0 + b_1 E_{t+1} x^*_{t+2} + b_2 y_{t+1} + u_{t+1}] + b_2 y_t + u_t \\
&= b_0(1 + b_1) + b_1^2 E_t x^*_{t+2} + b_1 b_2 E_t y_{t+1} + b_2 y_t + u_t \\
&= b_0(1 + b_1) + b_1^2 E_t x^*_{t+2} + b_1 b_2 (a + \rho y_t) + b_2 y_t + u_t \\
&= b_0(1 + b_1 + \cdots + b_1^{n-1}) + b_1^n E_t x^*_{t+n} + \sum_{k=1}^{n-1}[b_1^k b_2(a + \rho a + \rho^{k-1} a + \rho^k y_t)] + b_2 y_t + u_t \\
&= b_0 \sum_{k=0}^{n-1} b_1^k + b_1^n E_t x^*_{t+n} + \sum_{k=1}^{n-1} \frac{b_2 a}{\rho - 1} b_1^k(\rho^k - 1) + b_2 y_t \sum_{k=0}^{n-1} b_1^k \rho^k + u_t \\
&= \cdots \\
&= \frac{b_0}{1 - b_1} + \frac{a b_1 b_2}{(1 - b_1 \rho)(1 - b_1)} + \frac{b_2}{1 - b_1 \rho} y_t + u_t.
\end{aligned}
\tag{B.1}
$$

C   First-order autocorrelation coefficient

We rewrite model (3.6) as

$$
\begin{cases}
x_t - \bar{x} = b_1 \beta^2 (x_{t-1} - \bar{x}) + b_2(y_t - \bar{y}) + u_t, \\
y_t - \bar{y} = \rho(y_{t-1} - \bar{y}) + \varepsilon_t.
\end{cases}
\tag{C.1}
$$

That is,

$$
\begin{cases}
x_t - \bar{x} = b_1 \beta^2 (x_{t-1} - \bar{x}) + b_2 \rho(y_{t-1} - \bar{y}) + b_2 \varepsilon_t + u_t, \\
y_t - \bar{y} = \rho(y_{t-1} - \bar{y}) + \varepsilon_t.
\end{cases}
\tag{C.2}
$$

$$
\begin{aligned}
&E[(x_t - \bar{x})(x_{t-1} - \bar{x})] \\
&= E\left[b_1 \beta^2 (x_{t-1} - \bar{x})^2 + b_2 \rho(x_{t-1} - \bar{x})(y_{t-1} - \bar{y}) + b_2(x_{t-1} - \bar{x})\varepsilon_t + (x_{t-1} - \bar{x})u_t\right] \\
&= b_1 \beta^2 Var(x_t) + b_2 \rho E[(x_{t-1} - \bar{x})(y_{t-1} - \bar{y})] + b_2 E[(x_{t-1} - \bar{x})\varepsilon_t] + (x_{t-1} - \bar{x})u_t \\
&= b_1 \beta^2 Var(x_t) + b_2 \rho E[(x_{t-1} - \bar{x})(y_{t-1} - \bar{y})] \\
&= b_1 \beta^2 Var(x_t) + b_2 \rho E[(x_t - \bar{x})(y_t - \bar{y})].
\end{aligned}
\tag{C.3}
$$

$$Var(x_t)$$

$$= E(x_t - \bar{x})^2$$

$$= E\Big[b_1\beta^2(x_t - \bar{x})(x_{t-1} - \bar{x}) + b_2\rho(x_t - \bar{x})(y_{t-1} - \bar{y}) + b_2(x_t - \bar{x})\varepsilon_t + (x_t - \bar{x})u_t\Big]$$

$$= b_1\beta^2 E[(x_t - \bar{x})(x_{t-1} - \bar{x})] + b_2\rho E[(x_t - \bar{x})(y_{t-1} - \bar{y})] + b_2(x_t - \bar{x})\varepsilon_t + (x_t - \bar{x})u_t]$$

$$= b_1\beta^2 E[(x_t - \bar{x})(x_{t-1} - \bar{x})] + b_2\rho E[(x_t - \bar{x})(y_{t-1} - \bar{y})] + b_2^2\sigma_\varepsilon^2 + \sigma_u^2, \qquad (C.4)$$

where the last equation is based on the fact that $E[(x_t - \bar{x})\varepsilon_t] = E\Big[b_1\beta^2(x_{t-1} - \bar{x})\varepsilon_t + b_2\rho(y_{t-1} - \bar{y})\varepsilon_t + b_2\varepsilon_t^2 + u_t\varepsilon_t\Big] = b_2\sigma_\varepsilon^2$ and $E[(x_t - \bar{x})u_t] = E\Big[b_1\beta^2(x_{t-1} - \bar{x})u_t + b_2\rho(y_{t-1} - \bar{y})u_t + b_2\varepsilon_t u_t + u_t^2\Big] = \sigma_u^2$.

Based on (C.3) and (C.4),

$$\begin{aligned}
Var(x_t) &= b_1\beta^2 E[(x_t - \bar{x})(x_{t-1} - \bar{x})] + b_2\rho E[(x_t - \bar{x})(y_{t-1} - \bar{y})] + b_2^2\sigma_\varepsilon^2 + \sigma_u^2 \\
&= b_1\beta^2\Big[b_1\beta^2 Var(x_t) + b_2\rho E[(x_t - \bar{x})(y_t - \bar{y})]\Big] + b_2\rho E[(x_t - \bar{x})(y_{t-1} - \bar{y})] + b_2^2\sigma_\varepsilon^2 + \sigma_u^2 \\
&= b_1^2\beta^4 Var(x_t) + b_1\beta^2 b_2\rho E[(x_t - \bar{x})(y_t - \bar{y})] + b_2\rho E[(x_t - \bar{x})(y_{t-1} - \bar{y})] + b_2^2\sigma_\varepsilon^2 + \sigma_u^2.
\end{aligned}$$

That is,

$$Var(x_t) = \frac{b_1\beta^2 b_2\rho E[(x_t - \bar{x})(y_t - \bar{y})] + b_2\rho E[(x_t - \bar{x})(y_{t-1} - \bar{y})] + b_2^2\sigma_\varepsilon^2 + \sigma_u^2}{1 - b_1^2\beta^4}. \qquad (C.5)$$

Thus, in order to obtain $E[(x_t - \bar{x})(x_{t-1} - \bar{x})]$ and $Var(x_t)$, we need calculate $E[(x_t - \bar{x})(y_t - \bar{y})]$ and $E[(x_t - \bar{x})(y_{t-1} - \bar{y})]$.

$$\begin{aligned}
E[(x_t - \bar{x})(y_t - \bar{y})] &= E\Big[b_1\beta^2(x_{t-1} - \bar{x})(y_t - \bar{y}) + b_2\rho(y_{t-1} - \bar{y})(y_t - \bar{y}) + b_2\varepsilon_t(y_t - \bar{y}) + u_t(y_t - \bar{y})\Big] \\
&= b_1\beta^2 E\{(x_{t-1} - \bar{x})[\rho(y_{t-1} - \bar{y}) + \varepsilon_t]\} + b_2\rho E[(y_{t-1} - \bar{y})(y_t - \bar{y})] \\
&\quad + b_2 E\{\varepsilon_t[\rho(y_{t-1} - \bar{y}) + \varepsilon_t]\} + E[u_t(y_t - \bar{y})] \\
&= b_1\beta^2\rho E[(x_{t-1} - \bar{x})(y_{t-1} - \bar{y})] + 0 + \frac{b_2\rho^2\sigma_\varepsilon^2}{(1 - \rho^2)} + b_2\sigma_\varepsilon^2 + 0.
\end{aligned}$$

Thus

$$E[(x_t - \bar{x})(y_t - \bar{y})] = \frac{b_2\sigma_\varepsilon^2}{(1 - \rho^2)(1 - b_1\beta^2\rho)}. \qquad (C.6)$$

Hence based on (C.6),

$$
\begin{aligned}
&E[(x_t - \bar{x})(y_{t-1} - \bar{y})] \\
&= E\Big[b_1\beta^2(x_{t-1} - \bar{x})(y_{t-1} - \bar{y}) + b_2\rho(y_{t-1} - \bar{y})^2 + b_2\varepsilon_t(y_{t-1} - \bar{y}) + u_t(y_{t-1} - \bar{y})\Big] \\
&= b_1\beta^2 E[(x_{t-1} - \bar{x})(y_{t-1} - \bar{y})] + b_2\rho E(y_{t-1} - \bar{y})^2 + 0 + 0 \\
&= b_1\beta^2 \cdot \frac{b_2\sigma_\varepsilon^2}{(1-\rho^2)(1-b_1\beta^2\rho)} + b_2\rho \cdot \frac{\sigma_\varepsilon^2}{1-\rho^2} \\
&= \frac{b_2\sigma_\varepsilon^2}{(1-\rho^2)}\left[\frac{b_1\beta^2}{1-b_1\beta^2\rho} + \rho\right] \\
&= \frac{b_2\sigma_\varepsilon^2}{(1-\rho^2)} \cdot \frac{b_1\beta^2(1-\rho^2) + \rho}{1-b_1\beta^2\rho} \\
&= \frac{b_2\sigma_\varepsilon^2}{(1-b_1\beta^2\rho)}\left[b_1\beta^2 + \frac{\rho}{1-\rho^2}\right]. \tag{C.7}
\end{aligned}
$$

Therefore, based on (C.5), (C.6) and (C.7),

$$
\begin{aligned}
Var(x_t) &= \frac{1}{1-b_1^2\beta^4}\Big\{b_1\beta^2 b_2\rho E[(x_t - \bar{x})(y_t - \bar{y})] + b_2\rho E[(x_t - \bar{x})(y_{t-1} - \bar{y})] + b_2^2\sigma_\varepsilon^2 + \sigma_u^2\Big\} \\
&= \frac{1}{1-b_1^2\beta^4}\left\{\frac{b_1\beta^2 b_2^2\rho\sigma_\varepsilon^2}{(1-\rho^2)(1-b_1\beta^2\rho)} + \frac{b_2^2\rho\sigma_\varepsilon^2}{(1-b_1\beta^2\rho)}\left[b_1\beta^2 + \frac{\rho}{1-\rho^2}\right] + b_2^2\sigma_\varepsilon^2 + \sigma_u^2\right\} \\
&= \frac{\sigma_\varepsilon^2}{1-b_1^2\beta^4}\left\{\frac{b_2^2\rho\big[b_1\beta^2(2-\rho^2) + \rho\big]}{(1-\rho^2)(1-b_1\beta^2\rho)} + b_2^2 + \frac{\sigma_u^2}{\sigma_\varepsilon^2}\right\} \\
&= \frac{\sigma_\varepsilon^2}{1-b_1^2\beta^4}\left\{\frac{b_2^2(b_1\beta^2\rho + 1)}{(1-\rho^2)(1-b_1\beta^2\rho)} + \frac{\sigma_u^2}{\sigma_\varepsilon^2}\right\}. \tag{C.8}
\end{aligned}
$$

According to (C.3),

$$
\begin{aligned}
E[(x_t - \bar{x})(x_{t-1} - \bar{x})] &= b_1\beta^2 Var(x_t) + b_2\rho E[(x_t - \bar{x})(y_t - \bar{y})] \\
&= b_1\beta^2 Var(x_t) + \frac{b_2^2\rho\sigma_\varepsilon^2}{(1-\rho^2)(1-b_1\beta^2\rho)}. \tag{C.9}
\end{aligned}
$$

Thus, the correlation coefficient $Corr(x_t, x_{t-1})$ satisfies

$$
\begin{aligned}
Corr(x_t, x_{t-1}) &= E[(x_t - \bar{x})(x_{t-1} - \bar{x})]/Var(x_t) \\
&= b_1\beta^2 + \frac{\frac{b_2^2\rho\sigma_\varepsilon^2}{(1-\rho^2)(1-b_1\beta^2\rho)}}{\frac{\sigma_\varepsilon^2}{1-b_1^2\beta^4}\left\{\frac{b_2^2(b_1\beta^2\rho+1)}{(1-\rho^2)(1-b_1\beta^2\rho)} + \frac{\sigma_u^2}{\sigma_\varepsilon^2}\right\}} \\
&= b_1\beta^2 + \frac{b_2^2\rho(1-b_1^2\beta^4)}{b_2^2(b_1\beta^2\rho+1) + (1-\rho^2)(1-b_1\beta^2\rho)\frac{\sigma_u^2}{\sigma_\varepsilon^2}} \\
&= \frac{b_2^2(b_1\beta^2 + \rho) + b_1\beta^2(1-\rho^2)(1-b_1\beta^2\rho)\frac{\sigma_u^2}{\sigma_\varepsilon^2}}{b_2^2(b_1\beta^2\rho+1) + (1-\rho^2)(1-b_1\beta^2\rho)\frac{\sigma_u^2}{\sigma_\varepsilon^2}}.
\end{aligned}
$$

D   Proof of uniqueness of β^* (Proposition 2)

Using the first-order autocorrelation $F(\beta)$ in (3.11), it can be calculated that

$$F''(\beta) = \frac{2b_1(1-\rho^2)}{(\rho b_1 \beta^2 + 1)^2} - \frac{8\rho b_1^2 \beta^2(1-\rho^2)}{(\rho b_1 \beta^2 + 1)^3} = \frac{2b_1(1-\rho^2)(1 - 3\rho b_1 \beta^2)}{(\rho b_1 \beta^2 + 1)^3}.$$

In the case $b_1 > 0$, if $\rho \le \frac{1}{3b_1}$, then $1 - 3\rho b_1 \beta^2 \ge 1 - \beta^2 > 0$. Thus $G''(\beta) = F''(\beta) > 0$. Note that $G(0) > 0$, $G'(0) = -1 < 0$ and $G(1) < 0$, $G'(1) = \frac{2b_1(1-\rho^2)}{(b_1\rho+1)^2} - 1$. Hence if $G'(1) \le 0$, then $G'(\beta^*) < 0$. If $G'(1) > 0$, then there exists a minimal point β_1 such that $G'(\beta_1) = 0$. Moreover, since $G(1) < 0$, then $G(\beta_1) < 0$ (otherwise, $G(1) \ge G(\beta_1) \ge 0$, which is contradictory to $G(1) < 0$). Hence $\beta^*(\in (0, \beta_1))$ is unique and $G'(\beta^*) < 0$, hence $0 < F'(\beta^*) < 1$.

If $\rho > \frac{1}{3b_1}$, then $G''(\beta)\big|_{\beta=\sqrt{1/(3b_1\rho)}} = F''(\beta)\big|_{\beta=\sqrt{1/(3b_1\rho)}} = 0$ and $G'(\beta)\big|_{\beta=\sqrt{1/(3b_1\rho)}}$ is maximal. Thus in the case that $\rho > \frac{1}{3b_1}$,

$$
\begin{aligned}
G'(\beta) &= F'(\beta) - 1 \\
&= \frac{2b_1\beta(1-\rho^2)}{(\rho b_1 \beta^2 + 1)^2} - 1 \\
&\le \frac{2b_1 \frac{1}{\sqrt{3b_1\rho}}(1-\rho^2)}{(b_1\rho\frac{1}{3b_1\rho} + 1)^2} - 1 \\
&= \frac{3\sqrt{3}\sqrt{b_1}(1-\rho^2)}{8\sqrt{\rho}} - 1 \\
&< \frac{3\sqrt{3}\sqrt{b_1}(1-\frac{1}{9b_1^2})}{8\frac{1}{\sqrt{3b_1}}} - 1 \\
&= \frac{-(1-b_1)(1+9b_1)}{8b_1} < 0.
\end{aligned}
$$

Furthermore, it is easy to see that $F(\beta)$ only depends on β^2 and $F(\beta) > 0$. Hence $G(\beta) = F(\beta) - \beta > 0$ for $\beta \in [-1, 0]$. So for $b_1 > 0$ there is a unique β^* satisfying $0 < F'(\beta^*) < 1$.

In the case $b_1 \le 0$, since $F'(\beta) = \frac{2b_1\beta(1-\rho^2)}{(\rho b_1 \beta^2 + 1)^2} \le 0$ for $\beta \in [0, 1]$, then $G'(\beta) = F'(\beta) - 1 < 0$. Thus $G(\beta)$ is monotonically decreasing and hence β^* is unique within the interval $(0, 1)$ satisfying $F'(\beta^*) < 1$. Moreover, for $\beta \in [-1, 0)$, $G''(\beta) = F''(\beta) \le 0$. It is easy to see further $G(-1) = F(-1) + 1 > 0$, $G(0) > 0$ and $G'(0) = -1$. For $\beta \in [-1, 0)$, $G(\beta)$ is decreasing or first increasing and then decreasing. In any case there is no solution for $G(\beta) = 0$ within the interval $[-1.0]$. So for $b_1 \le 0$ β^* is unique satisfying $\beta^* \in (0, 1)$ and $F'(\beta^*) < 1$.

Therefore β^* is unique, which is within the interval $(0, 1)$ and satisfies $F'(\beta^*) < 1$.

E   Proof of Proposition 3 (at most 3 SCEE)

A straightforward computation, using (3.9), yields $G(\beta) = F(\beta) - \beta =$

$$= \frac{b_2^2(b_1\beta^2 + \rho) + b_1\beta^2(1-\rho^2)(1-b_1\beta^2\rho)R_v}{b_2^2(b_1\beta^2\rho + 1) + (1-\rho^2)(1-b_1\beta^2\rho)R_v} - \beta$$

$$= \frac{-b_1^2\rho(1-\rho^2)R_v\beta^4 - b_1\rho[b_2^2-(1-\rho^2)R_v]\beta^3 + b_1[b_2^2+(1-\rho^2)R_v]\beta^2 - [b_2^2+(1-\rho^2)R_v]\beta + b_2^2\rho}{b_1\rho[b_2^2-(1-\rho^2)R_v]\beta^2 + [b_2^2+(1-\rho^2)R_v]},$$

where $R_v = \frac{\sigma_u^2}{\sigma_\varepsilon^2}$.

If $b_1\rho[b_2^2 - (1-\rho^2)R_v] \geq 0$, then $b_1\rho[b_2^2 - (1-\rho^2)R_v]\beta^2 + [b_2^2 + (1-\rho^2)R_v] > 0$ for any β. Thus $G(\beta) = 0$ is equivalent to the 4-th order polynomial equation $\overline{G}(\beta) = -b_1^2\rho(1-\rho^2)R_v\beta^4 - b_1\rho[b_2^2-(1-\rho^2)R_v]\beta^3 + b_1[b_2^2+(1-\rho^2)R_v]\beta^2 - [b_2^2+(1-\rho^2)R_v]\beta + b_2^2\rho = 0$. Hence, there are at most four real solutions. Since $\overline{G}(-1) = b_2^2(1+b_1)(1+\rho) + (1+b_1)(1-b_1\rho)(1-\rho^2)R_v > 0$ and $\overline{G}(\beta) \to -\infty$ as $\beta \to -\infty$ due to negative coefficient of β^4, there exists one solution within the interval $(-\infty, -1)$. So there are at most three solutions for $\overline{G}(\beta) = 0$, i.e. $G(\beta) = 0$, within the interval $[-1, 1]$.

If $b_1\rho[b_2^2 - (1-\rho^2)R_v] < 0$, then there are two singularities of $G(\beta)$, i.e., two solutions for $b_1\rho[b_2^2 - (1-\rho^2)R_v]\beta^2 + [b_2^2 + (1-\rho^2)R_v] = 0$, given by $\beta_{1,2} = \pm\sqrt{\frac{b_2^2+(1-\rho^2)R_v}{b_1\rho[(1-\rho^2)R_v-b_2^2]}}$. It is easy to see that $|\beta_{1,2}| > 1$. For $\beta \in (\beta_2, \beta_1)$, if $\beta \to \beta_2$, then $\overline{G}(\beta) \to b_2^2\rho(1 - b_1^2\beta_2^4)$. Note that $b_2^2\rho(1-b_1^2\beta_2^4) < 0$. Thus $G(\beta) \to -\infty$ as $\beta \to \beta_2$ for $\beta \in (\beta_2, \beta_1)$. As discussed above, $G(-1) > 0$. Hence there exists one solution for $G(\beta) = 0$ within $(\beta_2, -1)$. Furthermore, in the interval $(\beta_2, \beta_1) \supset [-1, 1]$, $G(\beta) = 0$ is equivalent to $\overline{G}(\beta) = 0$. So there are at most three solutions for $G(\beta) = 0$ within the interval $[-1, 1]$.

Therefore there are at most three zeros for $G(\beta) = F(\beta) - \beta$ within the interval $[-1, 1]$. That is, at most three first-order stochastic consistent expectations equilibrium (SCEE) (α^*, β^*) coexist.

F   Proof of Proposition 4 (stability SAC-learning)

Set $\gamma_t = (1+t)^{-1}$. For the state dynamics equations in (3.12) and (2.8)[26], since all functions are smooth, the SAC-learning rule satisfies the conditions (A.1-A.3) of Section 6.2.1 in Evans and Honkapohja (2001, p.124).

In order to check the conditions (B.1-B.2) of Section 6.2.1 in Evans and Honkapohja

---

[26]For convenience of theoretical analysis, one can set $S_{t-1} = R_t$.

(2001, p.125), we rewrite the system in matrix form by

$$X_t = A(\theta_{t-1})X_{t-1} + B(\theta_{t-1})W_t,$$

where $\theta'_t = (\alpha_t, \beta_t, R_t)$, $X'_t = (1, x_t, x_{t-1}, y_t)$ and $W'_t = (1, u_t, \varepsilon_t)$,

$$A(\theta) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ b_0 + b_1\alpha(1 - \beta^2) + b_2 a & b_1\beta^2 & 0 & b_2\rho \\ 0 & 1 & 0 & 0 \\ a & 0 & 0 & \rho \end{pmatrix},$$

$$B(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & b_2 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

As shown in Evans and Honkapohja (2001, p.186), $A(\theta)$ and $B(\theta)$ clearly satisfy the Lipschitz conditions and B is bounded. Since u_t and ε_t are assumed to have bounded moments, condition (B.1) is satisfied. Furthermore, the eigenvalues of matrix $A(\theta)$ are 0 (double), ρ and $b_1\beta^2$. According to the assumption $|\beta| \leq 1$, $|b_1| < 1$ and $0 < \rho < 1$, all eigenvalues of $A(\theta)$ are less than 1 in absolute value. Then it follows that there is a compact neighborhood including the SCEE solution (α^*, β^*) on which the condition that $|A(\theta)|$ is bounded strictly below 1 is satisfied.

Thus the technical conditions for Section 6.2.1 of Chapter 6 in Evans and Honkapohja (2001) are satisfied. Moreover, since x_t is stationary under the condition $|\beta| \leq 1$, $|b_1| < 1$ and $0 < \rho < 1$, then the limits

$$\sigma^2 := \lim_{t \to \infty} E(x_t - \alpha)^2, \qquad \sigma^2_{xx_{-1}} := \lim_{t \to \infty} E(x_t - \alpha)(x_{t-1} - \alpha)$$

exist and are finite. Hence according to Section 6.2.1 of Chapter 6 in Evans and Honkapohja (2001, p.126), the associated ODE is

$$\begin{cases} \dfrac{d\alpha}{d\tau} = \bar{x}(\alpha, \beta) - \alpha, \\[2mm] \dfrac{d\beta}{d\tau} = R^{-1}[\sigma^2_{xx_{-1}} - \beta\sigma^2], \\[2mm] \dfrac{dR}{d\tau} = \sigma^2 - R. \end{cases}$$

38

That is,

$$\begin{cases} \dfrac{d\alpha}{d\tau} = \dfrac{b_0 + b_1\alpha(1-\beta^2) + b_2\bar{y}}{1 - b_1\beta^2} - \alpha = \dfrac{b_0 + \alpha(b_1 - 1) + b_2\bar{y}}{1 - b_1\beta^2}, \\[3mm] \dfrac{d\beta}{d\tau} = F(\beta) - \beta = \dfrac{b_2^2(b_1\beta^2 + \rho) + b_1\beta^2(1-\rho^2)(1-b_1\beta^2\rho)\frac{\sigma_u^2}{\sigma_\varepsilon^2}}{b_2^2(b_1\beta^2\rho + 1) + (1-\rho^2)(1-b_1\beta^2\rho)\frac{\sigma_u^2}{\sigma_\varepsilon^2}} - \beta. \end{cases} \tag{F.1}$$

Furthermore,

$$JG(\alpha^*, \beta^*) = \begin{pmatrix} \dfrac{-(1-b_1)}{1-b_1(\beta^*)^2} & 0 \\[3mm] 0 & F'(\beta^*) - 1 \end{pmatrix}.$$

Hence a SCEE corresponds to a fixed point of the ODE (F.1). Furthermore, the SAC-learning (α_t, β_t) converges to the stable SCEE (α^*, β^*) as time t tends to ∞. In the special case $\sigma_u = 0$ or $b_2 \to \infty$, based on Proposition 2 and Appendix D, the SCEE (α^*, β^*) is unique and stable with $F'(\beta^*) - 1 < 0$. Thus the SAC-learning (α_t, β_t) converges to the unique (locally) stable SCEE (α^*, β^*) as time t tends to ∞.

G   Dependence of F on parameters

In this appendix we show that the partial derivatives of the first-order autocorrelation coefficient $F(\beta)$ in (4.16) of the implied ALM (4.15) satisfy (4.18).

Based on (4.25), $F(\beta) = \delta\beta^2 + \dfrac{\rho(1-\delta^2\beta^4)}{(\delta\beta^2\rho+1)+(1-\rho^2)(1-\delta\beta^2\rho)\frac{1}{\gamma^2}\frac{\sigma_u^2}{\sigma_\varepsilon^2}} > 0$. As shown in the first paragraph in the Subsection 4.2.4,

$$\frac{\partial F}{\partial \gamma} = \frac{2\rho(1-\delta^2\beta^4)(1-\rho^2)(1-\delta\beta^2\rho)\frac{\sigma_u^2}{\sigma_\varepsilon^2}}{\gamma^3\left[(\delta\beta^2\rho+1)+(1-\rho^2)(1-\delta\beta^2\rho)\frac{1}{\gamma^2}\frac{\sigma_u^2}{\sigma_\varepsilon^2}\right]^2} > 0.$$

Denote $\frac{\sigma_u^2}{\sigma_\varepsilon^2}$ by ξ.

$$\frac{\partial F}{\partial \xi} = \frac{-\rho(1-\delta^2\beta^4)(1-\rho^2)(1-\delta\beta^2\rho)\frac{1}{\gamma^2}}{\left[(\delta\beta^2\rho+1)+(1-\rho^2)(1-\delta\beta^2\rho)\frac{1}{\gamma^2}\frac{\sigma_u^2}{\sigma_\varepsilon^2}\right]^2} < 0.$$

Now consider the parameter ρ. It can be calculated that

$$\frac{\partial F}{\partial \rho} = \frac{(1-\delta^2\beta^4)\left[1+(1+\rho^2-2\delta\beta^2\rho^3)\frac{1}{\gamma^2}\frac{\sigma_u^2}{\sigma_\varepsilon^2}\right]}{\left[(\delta\beta^2\rho+1)+(1-\rho^2)(1-\delta\beta^2\rho)\frac{1}{\gamma^2}\frac{\sigma_u^2}{\sigma_\varepsilon^2}\right]^2}.$$

In the following we will show that

$$1 + \rho^2 - 2\delta\beta^2\rho^3 > 0$$

for any given $\delta \in (0,1), \rho \in [0,1)$ and $\beta \in [-1,1]$. Thus $\frac{\partial F}{\partial \rho} > 0$.

Let $h(\rho)$ denote the 3rd order polynomial $-2\delta\beta^2\rho^3 + \rho^2 + 1$. It is easy to see that $h(0) = 1 > 0, h(-\infty) \rightarrow +\infty, h(+\infty) \rightarrow -\infty$. Moreover $h'(\rho) = 2\rho(1 - 3\delta\beta^2\rho)$. That is, there are two values $\rho = 0, \frac{1}{3\delta\beta^2}$ such that $h'(0) = 0$ and $h'(\frac{1}{3\delta\beta^2}) = 0$. Moreover, $h''(0) = 2 > 0$ and $h''(\frac{1}{3\delta\beta^2}) = -2 < 0$. Hence within the interval $[0,1]$, $h(\rho)$ is monotonically increasing or first increasing and then decreasing. In any case since $h(0) = 1 > 0$ and $h(1) = 2(1 - \delta\beta^2) > 0$, then $h(\rho) > 0$ for any $\rho \in [0,1]$. Hence $\frac{\partial F}{\partial \rho} > 0$.

Finally, for δ, it can be calculated that

$$\frac{\partial F}{\partial \delta} = \frac{\beta^2 \left[\eta(\eta - 1)\rho^2\delta^2\beta^4 - 2\eta(1+\eta)\rho\delta\beta^2 + (1+\eta)^2 + \rho^2(\eta - 1) \right]}{\left[(\delta\beta^2\rho + 1) + \eta(1 - \delta\beta^2\rho) \right]^2},$$

where $\eta = (1-\rho^2)\frac{1}{\gamma^2}\frac{\sigma_u^2}{\sigma_\varepsilon^2} \geq 0$. If $\eta = 0$, it is easy to get $\frac{\partial F}{\partial \delta} = \frac{\beta^2[1-\rho^2]}{(\delta\beta^2\rho+1)^2} > 0$ for $\beta \in (0,1)$ and $\rho \in [0,1)$. In the following we assume $\eta > 0$. Let $g(\delta)$ denote the 2nd order polynomial $\eta(\eta - 1)\rho^2\delta^2\beta^4 - 2\eta(1+\eta)\rho\delta\beta^2 + (1+\eta)^2 + \rho^2(\eta - 1)$. We will show $g(\delta) > 0$ for any $\delta \in [0,1)$. If $\eta = 1$, then $g(\delta) = 4(1 - \delta\beta^2\rho) > 0$ for $\rho \in [0,1), \beta \in (0,1)$ and $\delta \in [0,1)$. If $\eta > 1$, then the symmetric axis of the 2nd order polynomial $g(\delta)$ is $\delta = \frac{\eta+1}{(\eta-1)\rho\beta^2} > 1$ and the coefficient $\eta(\eta - 1)\rho^2\beta^4 > 0$. If $\eta < 1$, then the symmetric axis of the 2nd order polynomial $g(\delta)$ is $\delta = \frac{\eta+1}{(\eta-1)\rho\beta^2} < 0$ and the coefficient $\eta(\eta-1)\rho^2\beta^4 < 0$. Hence no matter if $\eta > 1$ or $\eta < 1$, $g(\delta)$ decreases within the interval $[0,1)$. That is, if $g(1) \geq 0$, then $g(\delta) > 0$ for any $\delta \in [0,1)$.

Note that $g(1) = \eta(\eta - 1)\rho^2\beta^4 - 2\eta(1+\eta)\rho\beta^2 + (1+\eta)^2 + \rho^2(\eta - 1) := \tilde{g}(\beta^2)$. This is a 2nd order polynomial with respect to β^2. Similarly since the symmetric axis of the 2nd order polynomial $\tilde{g}(\beta^2)$ is $\beta^2 = \frac{\eta+1}{(\eta-1)\rho} > 1$ for $\eta > 1$ and $\beta^2 = \frac{\eta+1}{(\eta-1)\rho} < 0$ for $\eta < 1$, then $\tilde{g}(\beta^2)$ decreases within the interval $[0,1]$ no matter if $\eta > 1$ or $\eta < 1$. Thus we just need to prove $\tilde{g}(1) \geq 0$. Note that $\tilde{g}(1) = \eta(\eta - 1)\rho^2 - 2\eta(1 + \eta)\rho + (1 + \eta)^2 + \rho^2(\eta - 1) = (\eta^2 - 1)\rho^2 - 2\eta(1 + \eta)\rho + (1 + \eta)^2 := \hat{g}(\rho)$. Similarly since the symmetric axis of the 2nd order polynomial $\hat{g}(\rho)$ is $\rho = \frac{\eta+1}{\eta-1} > 1$ for $\eta > 1$ and $\rho = \frac{\eta+1}{\eta-1} < 0$ for $\eta < 1$. Hence $\hat{g}(\rho)$ decreases within the interval $[0,1)$ no matter if $\eta > 1$ or $\eta < 1$. That is, if $\tilde{g}(1) \geq 0$, then $\tilde{g}(\rho) > 0$ for any $\rho \in [0,1)$. In fact, $\tilde{g}(1) = (\eta^2 - 1) - 2\eta(1 + \eta) + (1 + \eta)^2 = 0$. Thus based on the above analysis, for any $\rho \in [0,1)$, we have $\tilde{g}(1) = \hat{g}(\rho) > 0$, and hence $g(1) = \tilde{g}(\beta^2) > \tilde{g}(1) > 0$ for any $\beta^2 \in (0,1)$. That is, for any $\delta \in [0,1)$, $g(\delta) > 0$. Therefore $\frac{\partial F}{\partial \delta} > 0$.

References

[1] Adam, K., 2007. Experimental evidence on the persistence of output and inflation. Economic Journal 117, 603-636.

[2] Adam, K., Marcet, A., 2011. Internal rationality, imperfect market knowledge and asset prices. Journal of Economic Theory 146, 1224-1252

[3] Adam, K., Marcet, A., 2012. Booms and busts in asset prices. Working Paper, University of Mannheim.

[4] Assenza, T., Heemeijer, P., Hommes, C. and Massaro, D., 2012. Individual expectations and aggregate macro behavior. CeNDEF Working Paper, University of Amsterdam.

[5] Berardi, M., 2007. Heterogeneity and misspecification in learning. Journal of Economic Dynamics & Control 31, 3203-3227.

[6] Bernanke, B. and Gertler, M., 1999. Monetary policy and asset market volatility. Federal Reserve Bank of Kansas City Economic Review, Fourth Quarter 1999, 17-51.

[7] Bernanke, B. and Gertler, M., 2001. Should central banks respond to movements in asset prices? American Economic Review 91, No. 2, *Papers and Proceedings*, 253-257.

[8] Beshears, J., Choi, J.J., Fuster, A., Laibson, D. and Madrian, B.C. (2013), What Goes Up Must Come Down? Experimental Evidence on Intuitive Forecasting, *American Economic Review* 103, Paper and Proceedings, May 2013, pp. 570-574.

[9] Branch, W.A., 2004. The theory of rationally heterogeneous expectations: evidence from survey data on inflation expectations. Economic Journal 114, 592-621.

[10] Branch, W.A., 2006. Restricted perceptions equilibria and learning in macroeconomics, in: Colander, D. (Ed.), Post Walrasian Macroeconomics: Beyond the Dynamic Stochastic General Equilibrium Model. Cambridge University Press, New York, pp. 135-160.

[11] Branch, W.A., Evans, G.W., 2006. Intrinsic heterogeneity in expectation formation. Journal of Economic Theory, 127, 264-295.

[12] Branch, W.A., Evans, G.W., 2007. Model uncertainty and endogenous volatility. Review of Economic Dynamics 10, 207-237.

[13] Branch, W.A., Evans, G.W., 2010. Asset return dynamics and learning. Review of Financial Studies 23(4), 1651-1680.

[14] Branch, W.A., Evans, G.W., 2011. Learning about risk and return: a simple model of bubbles and crashes. American Economic Journal: Macroeconomics 3(3), 159-191.

[15] Branch, W.A., Evans, G. and McGough, B. 2012. Finite horizon learning. In: Sargent, T.J. and Vilmunen, J. (Eds.), *Macroeconomics at the Service of Public Policy*, Chapter 8, Oxford University Press.

[16] Branch, W.A., McGough, B., 2005. Consistent expectations and misspecification in stochastic non-linear economies. Journal of Economic Dynamics & Control 29, 659-676.

[17] Branch, W.A., McGough, B., 2009. Monetary Policy in a new Keynesian Model with Heterogeneous Expectations. Journal of Economic Dynamics and Control, 33, 1036–1051.

[18] Brock, W., Hommes, C., 1998. Heterogeneous beliefs and routes to chaos in a simple asset pricing model. Journal of Economic Dynamics & Control 22, 1235-1274.

[19] Bullard, J., 2006. The learnability criterion and monetary policy. Federal Reserve Bank of St. Louis Review 88, 203-217.

[20] Bullard, J., Evans, G.W., Honkapohja, S., 2008. Monetary policy, judgment and near-rational exuberance. American Economic Review 98, 1163-1177.

[21] Bullard, J., Evans, G.W., Honkapohja, S., 2010. A model of near-rational exuberance. Macroeconomic Dynamics 14, 166-188.

[22] Bullard, J., Duffy, J., 2001. Learning and excess volatility. Macroeconomic Dynamics 5, 272-302.

[23] Bullard, J. and Mitra, K. 2002. Learning about monetary policy rules. Journal of Monetary Economics 49, 1105-1129.

[24] Calvo, G., (1983). Staggered Contracts in a Utility-Maximizing Framework, Journal of Monetary Economics, 12, 383–398.

[25] Campbell, J.Y., Lo, A.W., and MacKinlay, A.G., 1997. The Econometrics of Financial Markets. Princeton University Press, Princeton.

[26] Enders, W., 2010. Applied Econometric Time Series (3rd ed.). John Wiley & Sons, Inc., USA.

[27] Evans, G.W., Honkapohja, S., 2001. Learning and Expectations in Macroeconomics. Princeton University Press, Princeton.

[28] Evans, G.W., Honkapohja, S., 2003. Expectations and the stability problem for optimal monetary policeis. Review of Economic Studies 70, 807-824.

[29] Evans, G.W., Honkapohja, S., 2006. Monetary policy, expectations and commitment. Scandinavian Journal of Economics 108, 15-38.

[30] Evans, G.W., Honkapohja, S., 2011. Learning as a rational foundation for macroeconomics and finance, forthcoming in Rethinking Expectations: The Way Forward for Macroeconomics, Roman Frydman and Edmund S. Phelps (eds.), Princeton University Press, 2013.

[31] Evans, G. and McGough, B. 2013. Learning to optimize. Working Paper, University of Oregon.

[32] Evans, G. and Mitra, K. 2013. E-stability in the stochastic Ramsey model. Economics Letters 118, 407-410.

[33] Fuhrer, J.C., 2006. Intrinsic and inherited inflation persistence. International Journal of Central Banking 2, 49-86.

[34] Fuhrer, J.C., 2009. Inflation persistence. Working paper.

[35] Fuster, A., Laibson, D. and Mendel, B., 2010. Natural expectations and macroeconomic fluctuations. Journal of Economic Perspectives 24, 67-84.

[36] Fuster, A., Hebert, B., Laibson, D., 2011. Natural expectations, macroeconomic Dynamics, and asset pricing. Forthcoming NBER Macroeconomics Annual 26.

[37] Fuster, A., Hebert, M., Laibson, D., 2012. Investment Dynamics with Natural Expectations. International Journal of Central Banking (Special Issue in Honor of Benjamin Friedman) 8(1), 243-265.

[38] Gali, J., Gertler, M., López-Salido, J.D., 2001. European inflation dynamics. European Economic Review 45, 1237-1270.

[39] Grandmont, J.M., 1998. Expectation formation and stability in large socio-economic systems. Econometrica 66, 741-781.

[40] Hamilton, J.D., 1994. Time Series Analysis. Princeton Univeristy Press.

[41] Heemeijer, P., Hommes, C.H., Sonnemans, J., Tuinstra, J., 2009. Price stability and volatility in markets with positive and negative expectations feedback. Journal of Economic Dynamics & Control 33, 1052-1072.

[42] Hommes, C.H., 2011. The heterogeneous expectations hypothesis: some evidence for the lab. Journal of Economic Dynamics & Control 35, 1-24.

[43] Hommes, C.H., Sonnemans, J., Tuinstra, J., van de Velden, H., 2005. Coordination of expectations in asset pricing experiments. Review of Financial Studies 18, 955-980.

[44] Hommes, C.H., Sorger, G., 1998. Consistent expectations equilibria. Macroeconomic Dynamics 2, 287-321.

[45] Hommes, C.H., Rosser, J.B., 2001. Consistent expectations equilibria and complex dynamics in renewable resource markets. Macroeconomic Dynamics 5, 180-203.

[46] Hommes, C.H., Sorger, G., Wagener, F., 2013. Consistency of linear forecasts in a nonlinear stochastic economy. In: Bischi, G.I., Chiarella, C. and Sushko, I. (Eds.), *Global Analysis of Dynamic Models in Economics and Finance*, Springer-Verlag Berlin Heidelberg, pp. 229-287.

[47] Honkapohja, S., Mitra, K. and Evans, G. 2013. Notes on agents' behavioral rules under adaptive learning and recent studies of monetary policy. In: Sargent, T.J. and Vilmunen, J. (Eds.), *Macroeconomics at the Service of Public Policy*, Chapter 4, Oxford University Press.

[48] Lansing, K.J., 2009. Time-varing U.S. inflation dynamics and the new Keynesian Phillips curve. Review of Economic Dynamics 12, 304-326.

[49] Lansing, K.J., 2010. Rational and near-rational bubbles without drift. Economic Journal 120, 1149-1174.

[50] Lucas, R., 1972. Expectations and the neutrality of moeny. Journal of Economic Theory 4, 103-124.

[51] Milani, F., 2007. Expectations, learning and macroeconomic persistence. Journal of Monetary Economics 54, 2065-2082.

[52] Muth, J.E., 1961. Rational expectaions and the theory of price movements. Econometrica 29, 315-335.

[53] Nelson, C., 1972. The prediction performance of the FRB-MIT-PENN model of the US economy. American Economic Review 62, 902-917.

[54] Pfajfar, D., and B. Zakelj, 2011. Inflation Expectations and Monetary Policy Design: Evidence from the Laboratory. CentER Discussion Paper, Tilburg University.

[55] Preston, B., 2005. Learning About Monetary Policy Rules When Long-Horizon Expectations Matter. International Journal of Central Banking 1(2), 81-126.

[56] Sargent, T.J., 1991. Equilibrium with signal extraction from endogenous variables. Journal of Economic Dynamics & Control 15, 245-273.

[57] Sargent, T.J., 1999. The Conquest of American Inflation. Princeton University Press, Princeton, NJ.

[58] Sögner, L., Mitlöhner, H., 2002. Consistent expectations equilibria and learning in a stock market. Journal of Economic Dynamics & Control 26, 171-185.

[59] Stock, J.H. and Watson, M.W., 2007. Why has inflation become harder to forecast? Journal of Money, Credit and Banking 39, 3-34.

[60] Tallman, E. W., 2003. Monetary policy and learning: some implications for policy and research. Federal Reserve Bank of Atlanta, ECONOMIC REVIEW, Third Quater 2003.

[61] Timmermann, A., 1993. How learning in financial markets generates excess volatility and predictability in stock prices. Quarterly Journal of Economics 108, 1135-1145.

[62] Timmermann, A., 1996. Excess volatility and predictability in stock prices in autoregressive dividend models with learning. Review of Economic Studies 63, 523-557.

[63] Tuinstra, J., 2003. Beliefs equilibria in an overlapping generations model. Journal of Economic Behavior & Organization 50, 145-164.

[64] Woodford, M., 1990. Learning to believe in sunspots. Econometrica 58, 277 307.

[65] Woodford, M., 2003. Interest and Prices: Foundations of a Theory of Monetary Policy. Princeton University Press, Princeton, NJ.

[66] Woodford, M., 2013. Macroeconomic Analysis without the Rational Expectations Hypothesis. Forthcoming in Annual Review of Economics.