Comments on "Natural Expectations, Macroeconomic Dynamics and Asset Pricing"

George W. Evans
University of Oregon and University of St. Andrews

July 22, 2011

1 Introduction

Expectations clearly play a central role in modern macroeconomics. Households and firms are assumed to be dynamic optimizers, making decisions about work, consumption, savings, production and investment, based in part on current economic conditions, but also to a great extent on the future state of the economy. Thus, in particular, household saving and portfolio decisions depend on expected future interest rates, inflation and taxes and on the likely future trajectory of equity dividends and prices. Because of the key role of expectations in economics and finance, theories of expectations have been central to modern economic theory. Since the rational expectations (RE) revolution of the 1970s, associated with John Muth, Robert Lucas and Thomas Sargent, the benchmark theory has been that expectations are formed rationally, in the sense that they are consistent with the true model and yield forecast errors that are orthogonal to agents' information sets.

In their paper in this volume, Andreas Fuster, Benjamin Hebert and David Laibson (FHL) present an asset-pricing model in which RE is replaced by "Natural Expectations" (NE). Under NE agents misspecify the time-series model in a "natural" way: they chose a parsimonious model of dividends that omits longer lags. This captures short-run dynamics but misses long-run mean reversion. An earlier paper, Fuster et al. (2010) made similar arguments about other macroeconomic time-series.[1] Taken together, the

[1] Actually in Fuster et al. (2010) the term natutal expectations is used to denote an average between RE and what is called NE in the current paper.

two papers suggest the even bolder possibility of NE as a general stylized description of expectation formation.

In the current paper FHL insert a NE dividend forecasting equation into a Lucas-type consumption-based asset-pricing model with CARA preferences and habit persistence. Using this set-up, FHL can replicate a number of stylized facts and puzzles about asset price data and consumption. These include the findings of excess volatility of stock prices, that excess returns are negatively predicted by lagged excess returns, price to earnings ratios and consumption growth, and the existence of a large equity premium.

2 Outline of their Argument

FHL consider a variation of the Lucas-type "tree" model of asset prices. In keeping with Lucas (1978) there is an endowment economy with a single risky asset, trees, which provide an exogenous stochastic dividend of the perishable, homogeneous consumption good. The principal variation is that FHL consider an open economy version in which agents can borrow or lend internationally at a fixed interest rate R. In addition, exponential (CARA) preferences with habits are used. Finally, Δd_t, the first difference in dividends, is assumed to follow to follow a stationary $AR(p)$ process for some $p > 0$.

The exponential preferences give a type of mean-variance set-up, and FHL show that the asset price p_t satisfies

$$p_t = \sum_{s=1}^{\infty} R^{-s} E_t d_{t+1} - \frac{R\alpha \times Var_t \Delta c_{t+1}}{(1 - R^{-1}\gamma)(R-1)^2}.$$

FHL also show how to obtain closed-form expressions for p_t and for consumption, c_t, given beliefs about the $AR(p)$ process for Δd_t. Here R is the inverse of the discount factor, α is the CARA measure of risk aversion, and γ is the habit-persistence parameter.

The key assumptions of the FHL model concern the stochastic process *actually* followed by dividends and the *perceived* process followed by dividends. The true dividend process is assumed to be a high-order stationary $AR(p)$ for the first difference in dividends d_t, i.e. $\Delta d_t = \sum_{i=1}^{p} \rho_i \Delta d_{t-i} + \varepsilon_t$, or

$$\Delta d_t = (1 - \Phi(L))^{-1} \varepsilon_t,$$

where ε_t is white noise and $\Phi(L) = \sum_{i=1}^{p} \rho_i L^i$. Furthermore $\Phi(L)$ is assumed to be such that there is a hump-shaped impulse response function for dividend levels $\partial d_{t+j}/\partial \varepsilon_t$. Put differently, d_t is assumed to have a unit root with dynamics that lead to long-run mean reversion.

Evidence for this is given in their Figure 2, which presents $\partial d_{t+j}/\partial \varepsilon_t$ based on empirical estimates of $\Phi(L)$ for $AR(p)$ processes with alternative values of p. (In the empirical work, FHL use earnings rather than dividends). The long-run level of persistence is given by

$$\lim_{j \to \infty} \partial d_{t+j}/\partial \varepsilon_t = (1 - \Phi(1))^{-1},$$

and for $p \geq 15$ we have $0 < (1 - \Phi(1))^{-1} < 1$. Thus there is *mean reversion* in the sense that an innovation ε_t has a reduced permanent impact. FHL assume that large values of p, e.g. $p = 40$, correspond to the truth.

In contrast, for $p \leq 10$ estimates of long-run persistence are $(1-\Phi(1))^{-1} > 1$. That is, based on low order AR(n) estimates, one would come to the conclusion that one should *extrapolate* innovations in dividends – that the long-run effects are larger than the impact effect. FHL assume that low-order AR(n) estimates correspond to the *perceived* dividend process, i.e. to the view held by economic agents.

The essence of FHL's approach is thus that the beliefs of agents differ from the truth and do so in a particular way. Agents use simpler low-order time-series models that lead them to accentuate the importance of short-run trends and to neglect longer-run corrections in which dividends revert toward an underlying trend. This central feature leads to the empirical implications noted above.

What is the rationale for the discrepancy they assume between truth and perception? FHL give two types of argument – statistical/econometric and psychological. The statistical argument is that econometricians have often argued that in forecasting there is an advantage in using parsimonious models in preference to more complex models. Furthermore standard statistical procedures for model selection based on AIC and, especially, BIC often select low-order models. The psychological argument is that agents for a variety of reasons prefer to use simple, parsimonious models in preference to complex models, when trying to understand the world and make decisions.

FHL are usually careful not to be too dogmatic on this point. In essence they say: there is some evidence, as seen in their Table 1 and Figure 2, that a higher-order AR(n) process of Δd_t might well be correct, while agents might plausibly believe in a low-order process. Their paper then explores the full implications of this assumption.

3 Links to the Macro Learning Literature

If the truth is that p is large, but agents believe that p is small, then clearly agents do not have RE. There is a now extensive macro literature, which started around 1980, in which RE is replaced, e.g., by adaptive or econometric learning. See, for example, Sargent (1993) and Evans and Honkapohja (2001, 2009).

A major argument for the adaptive learning approach is what might be called the *cognitive consistency principle*.[2] According to this principle, agents should be assumed to have the same level of rationality as the economic modeler or policymaker (in contrast to both old-style adaptive expectations and to RE). On the adaptive learning approach agents are assumed to make forecasts in the same way that econometricians do – formulating models, estimating their parameters and updating estimated coefficients over time as new data become available. When parameters are updated using a form of least-squares, this is known as least-squares (LS) learning.

The early macro learning literature focused on whether or not LS learning would converge over time to RE in *self-referential* models, in which the variables being forecasted are affected by the forecasts. Conditions were worked out that determined whether or not REE (RE equilibria) were indeed stable under LS learning. Stability under learning could then be used as a selection criteria in models with multiple REE, since in some cases only a subset of REE were stable under learning.

More recently, another major strand has been to show how learning can generate transitory or persistent "learning dynamics," i.e. dynamics different from RE. Much of the recent macro learning literature has emphasized learning dynamics induced by one or more of the following factors: (i) misspecified forecasting models (misspecified "perceived laws of motion" or PLMs), (ii) discounted LS learning (downweighting past data due to concern about unknown structural change), and (iii) dynamic predictor selection (selecting between alternative PLMs based on past performance) or Bayesian model averaging. Applications of the approach that emphasize learning dynamics

[2]See Evans and Honkapohja (2009, 2011).

include: the rise and fall of inflation in the US, hyperinflation, business cycles, output and inflation inertia, optimal monetary and fiscal policy, and asset price anomalies.

One useful concept from the recent macro learning literature has been that of a *Restricted Perceptions Equilibrium* (RPE), in which agents make the best forecast they can, given their misspecified PLM.[3] A special case of interest has been models that are underparameterized, either in terms of variables or lag lengths. The argument here has been precisely that econometricians recognize the value of parsimoniously specified models, and thus the cognitive consistency principle dictates that we should examine the implications of underparameterization. One can, for example, work out stability conditions for an RPE when agents use LS learning to update coefficients of a particular underparameterized model.

Thus the FHL approach fits well with the recent macro learning literature. The principal contribution of the FHL paper, in this context, is that it posits a particular, plausible type of misspecification by agents, which can arguably explain several puzzling features of asset prices, and which may also be of more general applicability.

4 Discussion

I certainly find plausible FHL's key assumption that economic agents underparameterize their forecasting models. This assumption is consistent with the cognitive consistency principle, given that many applied econometricians place value on parsimony and recognize the likelihood of misspecification. This hypothesis also fits well the observation that many economists believe there is long-run mean reversion that is nonetheless difficult to detect.[4] That is, the misspecification that FHL assume is particularly plausible.

Other aspects of the FHL model are also attractive: the closed-form solutions under CARA preferences, for the class of perceived dividend processes examined, is likely to be more generally useful, and the simultaneous fit of a range of stylized facts is impressive.

It is therefore hard not to like this paper: the model is both simple and

[3] Closely related concepts are those of "self-confirming equilibria" and "consistent expectations equilibria."

[4] Estimating long-run persistence $(1-\Phi(1))^{-1}$ is equivalent to estimating the spectrum of d_t at zero, and is understood to be subject to great uncertainty.

powerful. However, of course the model has some weaknesses, several of which are noticeable from the macro learning viewpoint. This in turn suggests a number of natural extensions.

4.1 Criticisms

In my critical discussion I will focus on three main issues. The first concerns the information set available to agents when making forecasts. By assumption Δd_t is an exogenous univariate process, which leads FHL to examine alternative univariate forecasting models. However, within macroeconomics the norm is to consider multivariate forecasting models, and this issue is pertinent to the question of long-run persistence and to the plausibility of the form of underparameterization assumed. For example, in the early discussion of long-run GDP persistence, Campbell and Mankiw (1987) focussed on univariate techniques, and found persistence levels greater than one. However, both the unemployment rate and the consumption - output ratio Granger cause output growth, and lower levels of persistence, with mean reversion, are found in multivariate models, e.g. see Evans (1989) and Evans and Reichlin (1994). In the current context Timmermann (1994), for example, has argued that stock prices Granger cause dividends. Thus a simple bivariate forecasting model might lead to different persistence results. The issue is whether simple, i.e. low-order vector autoregressions might show long-run mean reversion more clearly, in which case this feature of the data would be less plausibly missed by economic agents. Of course, many agents might still in practice use "natural" low-order univariate models, but a heterogeneous expectations model might then be more realistic.

My second concern is the fixed parameter assumption of FHL. Suppose first that we agree that agents plausibly underparameterize Δd_t as an AR(1). From the learning viewpoint this leads to the corresponding RPE as the appropriate equilibrium to which the system would, if stable, converge. However, the cognitive consistency principle suggests that agents would not know the parameters of this process *a priori*, but, like real-world econometricians, would estimate the parameters and update their estimates over time. Furthermore, if agents are concerned about potential structural change, they might discount older data, leading to persistent learning dynamics around the RPE. This particular issue could easily be addressed by simulations in which fixed parameter natural expectations were replaced by discounted LS learning with the same AR(1) PLM.

Related to both of the previous two points, if long-run estimates of persistence are crucial for good decision-making in their portfolio choices, one might expect agents to focus on this issue in their choice of forecasting models. They might estimate mean reversion directly and allow for uncertainty concerning its value in their decisions. Alternatively, they might adopt decision-making rules that are robust to errors in this dimension, along the lines of Hansen and Sargent (2007).

The third issue, which is probably most central from the learning perspective, is that the FHL model is not self-referential. Agents simply forecast dividends, which is treated as an exogenous process, and do not have a forecasting model for stock prices. Some learning models emphasize short-horizon decision-making in which the demand for stocks depends on short-horizon expected returns, and possibly also on the estimated conditional variance of returns. See, e.g., Brock and Hommes (1998), Lansing (2010), Adam, Marcet and Nicolini (2010) and Branch and Evans (2011). Indeed one way to formulate the most basic risk-neutral model of stock-prices is to assume that prices are determined by the sum of expected dividend and expected stock price in the coming period. These models are self-referential in the sense that asset prices today depend on the expected price tomorrow, so that the evolution of the variable being forecasted depends on the expectations themselves.

Self-referential models, because of this feedback, give a much greater role to expectations, and this makes more likely asset-price bubbles: self-fulfilling or nearly self-fulfilling asset price movements with complex dynamics in which prices can become detached from fundamentals for extended periods. My own view is that this dynamic plays a central role in asset prices.

4.2 Example: a simple model of bubbles

An example of the scope for dramatic learning dynamics in self-referential asset price models is given in my work with William Branch presented in Branch and Evans (2011). We use a simple mean-variance linear asset pricing model. The set-up can come from an overlapping generations model in which agents have two period planning horizons, CARA preferences and a choice between a risky stock and a risk-free asset. The central equation is

$$p_t = \beta E_t^* \left(p_{t+1} + d_{t+1}\right) - \beta a \sigma_t^2 z_{st},$$

where z_{st} is the iid random supply of the risky asset, E_t^* denotes the subjective expectations of agents and σ_t^2 is their estimate of the conditional variance of

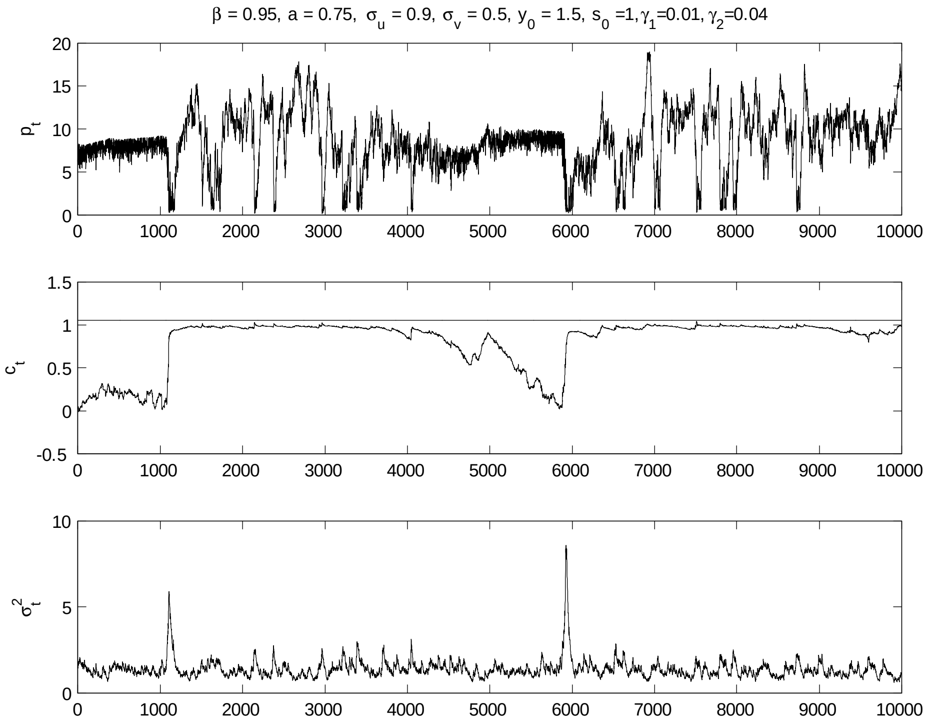

Figure 1: Asset price dynamics in the Branch and Evans (2011) model.

returns. We assume the dividend process is known and that agents therefore need estimates of the price process to make their decisions.

With *iid* dividend and supply shocks, the REE for p_t is a constant + white noise. Under learning, agents forecast p_t as an AR(1) using discounted LS and they estimate σ_t^2 using a simple recursive algorithm. Because agents discount past data, prices under learning will occasionally break free from their fundamentals and exhibit bubbles and crashes. This results from the self-referential feature of the model.

An illustrative simulation is shown in the accompanying Figure (see Branch and Evans (2011) for analysis and for other simulations). The figure shows the realized price p_t under learning and also the time series of estimates of two key learning parameters, the AR(1) coefficient c_t and the estimate of the conditional variance σ_t^2. The figure shows the price process initially very close to the REE, which is a constant plus white noise in our setup. However, under learning, asset prices occasionally break free into a bubble regime in which stock prices are believed to follow a pure random walk ($c = 1$). In this regime p_t is particularly sensitive to changes in the estimate of risk σ_t^2.

In summary, self-referential learning models have great scope for generating some of the more extreme partially self-fulfilling movements of stock prices often described as bubbles and crashes. Intuitively the reason for this is that the stock price p_t depends on expected price $E_t^* p_{t+1}$with a coefficient $\beta < 1$ that is close to one.

4.3 Other types of learning dynamics

In various settings learning dynamics have also been shown in self-referential models to lead to: (i) inertia of inflation and output, as in Orphanides and Williams (2007) and Milani (2007), (ii) overshooting and non-monotone IRFs, e.g. Eusepi and Preston (2011) and Evans, Honkapohja and Mitra (2009) and (iii) regime-switching and parameter drift, e.g. Sargent (1999), Cogley and Sargent (2005) and Branch and Evans (2007). For numerous examples and references, see Evans and Honkapohja (2011).

With this in mind, consider again the question of whether it is plausible that agents believe Δd_t is a specific constant coefficient $AR(p)$ process with known parameters. Parameter drift and regime switching appear to be standard features of the data, as emphasized by Sims and Zha (2006), Cogley and Sargent (2005) and Sargent et al. (2006). The cognitive consistency principle suggests that agents should therefore allow for the possibility of structural change in their parameter estimation and through model selection, model averaging or robust decision making.

In the FHL set-up, under RE the risk-premium is very small. In the late 1990s some people argued ("Dow 30,000") that the rise of the stock market was due to a recognition that the risk premium was too high. An implication of FHL is that this view is fundamentally correct. Is it not, however, more plausible to believe that the risk-premium reflects the uncertainty that economists and agents share?

5 Conclusions

Although I have indicated a number of reservations, overall I find the FHL story very attractive. The set-up is conceptually simple, and it is based on the plausible premise that agents underparameterize their forecasting model. This is in line with standard econometric advice to estimate parsimonious models, as well as evidence from psychology that people are inclined to make decisions based on simple heuristics. The FHL model is disciplined and delivers a number of important empirical implications that appear to be in line with the data.

I would prefer to extend the model to include additional insights from the adaptive learning literature, but even as it stands the FHL model provides an impressive but simple benchmark model of asset-price behavior, which is sure to receive considerable attention.

References

[1] Adam, Klaus, Albert Marcet, and Juan Pablo Nicolini (2010), "Learning and Stock Market Volatility," Working Paper.

[2] Branch, William A. and George W. Evans (2007), "Model Uncertainty and Endogenous Volatility," *Review of Economic Dynamics*, 10, 207-237.

[3] Branch, William A. and George W. Evans (2011), "Learning about Risk and Returns: a simple model of bubbles and crashes," *American Economic Journal: Macroeconomics*, 3, 159-191.

[4] Brock, William A. and Cars H. Hommes (1998), "Heterogenous Beliefs and Routes to Chaos in a Simple Asset Pricing Model," *Journal of Economic Dynamics and Control*, 22, 1235-1274.

[5] Campbell, John Y. and N. Gregory Mankiw (1987), *Quarterly Journal of Economics*, 102, 857-880.

[6] Cogley, Timothy and Thomas J. Sargent (2005), "The Conquest of US Inflation: Learning and Robustness to Model Uncertainty," *Review of Economic Dynamics,* 8, 528–63.

[7] Eusepi, Stefano and Bruce Preston (2011), "Expectations, Learning and Business Cycle Fluctuations," *American Economic Review*, forthcoming.

[8] Evans, George W. (1989), "Output and Unemployment Dynamics in the United States: 1950-1985," *Journal of Applied Econometrics*, 4, 213-237.

[9] Evans, George W., and Seppo Honkapohja (2001), *Learning and Expectations in Macroeconomics*, Princeton University Press, Princeton, NJ.

[10] Evans, George W., and Seppo Honkapohja (2009), "Learning and Macroeconomics" *Annual Review of Economics*, 1, 421-451.

[11] Evans, George W., and Seppo Honkapohja (2011), "Learning as a Rational Foundation for Macroeconomics and Finance," working paper.

[12] Evans, George W., and Honkapohja, Seppo and Kaushik Mitra (2009), "Anticipated Fiscal Policy and Learning," *Journal of Monetary Economics*, 56, 930-953.

[13] Evans, George W. and Lucrezia Reichlin (1994), "Information, Forecasts and Measurement of the Business Cycle," *Journal of Monetary Economics*, 33, 233-254.

[14] Fuster, A., D. Laibson, and B. Mendel (2010), "Natural Expectations and Macroeconomic Fluctuations," *Journal of Economic Perspectives*, 24, 67–84.

[15] Hansen, Lars P. and Thomas J. Sargent (2007), *Robustness*, Princeton University Press, Princeton, NJ.

[16] Lansing, Kevin (2010), "Rational and Near-rational Bubbles without Drift," *Economic Journal*, 120, 1149–1174.

[17] Lucas, Robert E. (1978), "Asset Prices in an Exchange Economy," *Econometrica*, 46, 1429-1445.

[18] Milani, Fabio (2007), "Expectations, Learning and Macroeconomic Persistence," *Journal of Monetary Economics*, 54, 2065-2082..

[19] Orphanides, Athanasios and John C. Williams, "Robust Monetary Policy with imperfect Knowledge," *Journal of Monetary Economics*, 53, 1406-1435.

[20] Sargent, Thomas J. (1993), *Bounded rationality in Macroeconomics*. Oxford: Oxford University Press.

[21] Sargent, Thomas J. (1999), *The Conquest of American Inflation*. Princeton, NJ: Princeton University Press.

[22] Sargent, Thomas J., Williams, Noah and Tao Zha (2006), "Shocks and Government Beliefs: The Rise and Fall of American Inflation," *American Economic Review* (2006), 1193–1224.

[23] Sims, Christopher and Tao Zha (2006), "Were There Regime Switches in US Monetary Policy?" *American Economic Review*, 96, 54–81.

[24] Timmermann, Allan (1994), "Can Agents Learn to Form Rational Expectations? Some Results on Convergence and Stability of Learning in the UK Stock Market," *Economic Journal,* 104, 777-797.