Stock Market Volatility and Learning*

Klaus Adam Albert Marcet Juan Pablo Nicolini

January 25, 2008

Abstract

Introducing bounded rationality in a standard consumption-based asset pricing model with time separable preferences strongly improves empirical performance. Learning causes momentum and mean reversion of returns and thereby excess volatility, persistence of price-dividend ratios, long-horizon return predictability and a risk premium, as in the habit model of Campbell and Cochrane (1999), but for lower risk aversion. This is obtained, even though our learning scheme introduces just one free parameter and we only consider learning schemes that imply small deviations from full rationality. The findings are robust to the learning rule used and other model features. What is key is that agents forecast future stock prices using past information on prices.

JEL Class. No.: G12, D84

*Thanks go to Luca Dedola, George Evans, Katharina Greulich, Seppo Honkapohja, Ricardo Lagos, Bruce McGough, Bruce Preston, Jaume Ventura, Joachim Voth, and Raf Wouters for interesting comments and suggestions. We particularly thank Philippe Weil for a very interesting and stimulating discussion. Thanks go also to seminar participants at CERGE-EI in Prague, U. Autònoma de Barcelona, New York Fed, Fed. Reserve Bank of Minneapolis, NYU, MIT, Board of Governors, Carnegie-Mellon U., U. of Pennsylvania and Cambridge U, and to conference participants at CREI-CEPR Conference on 'Bubbles – Theory and Policy Implications' at U. Pompeu Fabra, ESSIM 2006 in Tarragona, Econometric Society European Meeting 2006, 4th IMOP Conference on Dynamic Macroeconomics in Hydra, SCE 2006 Conference, St. Louis Learning Workshop 2006, 2007 Dynare Conference in Paris, 1st Swiss National Bank Research Conference in Zurich, 2007 Bank of Finland - CEPR Conference in Helsinki, AEA Meetings 2008 in New Orleans. Davide Debortoli has supported us with outstanding research assistance and many suggestions. Marcet acknowledges support from CIRIT (Generalitat de Catalunya), DGES (Ministry of Education and Science, Spain), CREI, the Barcelona Economics program of XREA and the Wim Duisenberg fellowship from the European Central Bank. The views expressed herein are solely those of the authors and do not necessarily reflect the views of the European Central Bank. Author contacts: Klaus Adam (European Central Bank and CEPR) klaus.adam@ecb.int; Albert Marcet (Institut d'Anàlisi Econòmica CSIC) albert.marcet@iae.csic.es; Juan Pablo Nicolini (Universidad Torcuato di Tella) juanpa@utdt.edu.

"Investors, their confidence and expectations buoyed by past price increases, bid up speculative prices further, thereby enticing more investors to do the same, so that the cycle repeats again and again, ... "

Irrational Exuberance, Shiller (2005, p.56)

1 Introduction

The purpose of this paper is to show that a very simple asset pricing model is able to reproduce a variety of stylized facts if one allows for small departures from rational expectations. This result is somehow remarkable, since the literature in empirical finance had great difficulties in developing dynamic equilibrium rational expectations models accounting for all the asset pricing facts we consider.

We study the most standard asset pricing equation of dynamic equilibrium models determining stock prices, This equation arises, for example, in the representative agent endowment economy of Lucas (1978) or various other setups such as overlapping generations models. It is well known that the asset pricing implications of this equation under rational expectations (RE) are at odds with some basic asset pricing facts: in the data, the price dividend ratio is too volatile and persistent, stock returns are too volatile, long run excess stock returns are negatively related to the price dividend ratio, and the risk premium is too high. Our learning model introduces just one additional free parameter relative to the rational expectations counterpart. With this modification alone, our model quantitatively accounts for all these observations. Since the learning model reduces to the rational expectations model if the additional parameter is set to zero and since this parameter is close to zero throughout the paper, we consider the learning model to represent only a small departure from rationality. Nevertheless, the behavior of equilibrium prices differs considerably from that obtained under rational expectations, implying that the asset pricing implications of the standard model are not robust to small departures from rationality. As we document, this non-robustness is empirically encouraging, i.e., the model matches the data much better if this small departure from rationality is allowed for.

A large body of literature has documented that stock prices exhibit movements that are very hard to reproduce within the realm of rational expectations and the baseline model has been extended in a variety of directions to improve its empirical performance. A particular line of research introduced time-non-separable preferences for this purpose. After many papers and a couple of decades this line of research has succeeded: Campbell and Cochrane (1999) are able to reproduce all the facts, albeit at the cost of imposing a sophisticated habit specification for preferences and high effective degrees of risk aversion.

Our model retains simplicity and moderate curvature in utility, but instead deviates from rational expectations.

In contrast to the RE literature, the behavioral finance literature tried to understand the decision-making process of individual investors by means of surveys, experiments and micro evidence, exploring the intersection between economics and psychology. One of the main themes of this literature was to test the rationality hypothesis in asset markets, see Shiller (2005) for a non-technical summary. We borrow some of the economic intuition from this literature, but follow a different modeling approach: we aim for a model that is as close as possible to the standard model, with agents who formulate forecasts using statistical models that imply only small departures from rational expectations.

We first show that for a general class of learning schemes the model has the following properties: if *expectations* about stock price growth increase in a given period, the *actual* growth rate of prices has a tendency to increase beyond the fundamental growth rate, thereby reinforcing the initial belief of higher stock price growth. Learning thus imparts 'momentum' on stock prices and beliefs and this produces large and sustained deviations of the price dividend ratio from its mean, as can be observed in the data. Furthermore, we show that the model has mean reversion so that even if expectations are very high at some point they will eventually return to fundamentals. The model thus displays - for many learning schemes - something like the 'naturally occurring Ponzi schemes' described in Shiller's opening quote above.

In the baseline learning model, we assume agents form their expectations regarding future stock prices with the most standard learning scheme used in the literature: ordinary least squares (OLS).[1] We show that the equilibrium converges to rational expectations under OLS, but this process takes a very long time in this model. Then, the dynamics generated by learning along the transition cause prices to have the oscillations described in the previous paragraph for a very long period so that the model behaves in a very different way from RE.

As we mentioned, OLS is the most standard assumption to model the evolution of expectations functions in the learning literature and its limiting properties have been used extensively as a stability criterion to justify or discard RE equilibria. Yet, models of learning are still not commonly used to explain data or for policy analysis.[2] It still is the standard view in the economics research community that models of learning introduce too many degrees of freedom, so that it is easy to find a learning scheme that matches whatever observation one desires. One can deal with this important methodological issue in two ways: first, by using a learning scheme with as few free parameters as possible, and second, by imposing restrictions on the parameters of the learning scheme to only allow for small departures of rationality.[3] These considerations prompted

[1] We show that results are robust to using other standard learning rules.

[2] We will mention some exceptions along the paper.

[3] Marcet and Nicolini (2003) dealt with this issue by imposing bounds on the size of the mistakes agents can make in equilibrium. These bounds imposed discipline both on the type of learning rule and on the exact value of the parameters in the learning rule.

us to use an off-the-shelf learning scheme (OLS) that assumes agents revise their expectations in the direction of the last forecast error and that has only one free parameter. In addition, in the model at hand, OLS is the best estimator in the long run (as we prove), and to make the departure from rationality during the transition small, we assume in our quantitative analysis that initial beliefs are at the rational expectations equilibrium, and that agents initially have very strong - but less than complete - confidence in these initial beliefs. If agents had full confidence in these beliefs, the model would reduce to its RE counterpart.

Models of learning have been used before to explain some aspects of asset price behavior. Timmermann (1993, 1996), Brennan and Xia (2001) and Cogley and Sargent (2006) consider Bayesian learning to explain various aspects of stock prices. These authors consider agents who learn about the dividend process and assume agents to use the mean of the Bayesian posterior about some relevant unknown parameter to evaluate the discounted sum of dividends.[4] This approach is less able to explain asset price volatility: while agents' beliefs about the dividend process influence market prices, agents' beliefs remain unaffected by market outcomes because agents learn only about an exogenous driving process. One could say that expectations in these papers are 'anchored' by the dividend process. In the language of stochastic control, these models are not self-referential and this is why, in the language of Shiller, they can not give rise to 'naturally occurring Ponzi schemes'. In contrast, we largely abstract from learning about the dividend process and consider learning on the future stock price using past price observations, so that beliefs and prices are mutually determined. It is precisely the learning about future stock price growth and its self-referential nature that imparts momentum to expectations and is key in explaining stock price volatility.[5]

Other related papers by Bullard and Duffy (2001) and Brock and Hommes (1998) show that learning dynamics can converge to complicated attractors, if the RE equilibrium is unstable under learning dynamics.[6] Branch and Evans (2006) study a model where agents' algorithm to form expectations switches depending on which of the available forecast models is performing best. Marcet and Sargent (1992) also study convergence to RE in a model where agents use today's price to forecast the price tomorrow in a stationary environment with private information. By comparison, we look at learning about the stock price growth rate in an economy with dividend growth, we address more closely the data, and we do so in a model where the rational expectations equilibrium is stable under learning dynamics, so the departure from RE behavior occurs only along a transition related to the sample size of the observed data. Also related is Cárceles-Poveda and Giannitsarou (2007) who assume that agents know the

[4] In Cogley and Sargent (2006) the mean is driven by agents' initial pessimism.

[5] Timmerman (1996) also analyzes a case where learning is self-referential, but in his case agents use dividends to predict future price. He finds that this form of self-referential learning delivers lower volatility than settings with learning about the dividend process. It is thus crucial for our results that agents use information on past price behavior to predict future price.

[6] Stability under learning dynamics is defined in Marcet and Sargent (1989).

mean stock price and learn only about deviations from the mean; they find that the presence of learning does then not significantly alter the behavior of asset prices.[7]

The paper is organized as follows. Section 2 presents the stylized facts we focus on and the basic features of the underlying asset pricing model, showing that this model cannot explain the facts under the rational expectations hypothesis. In section 3 we take the simplest risk neutral model and assume instead that agents learn to forecast the growth rate of prices. We show that such a model can *qualitatively* deliver all the considered asset pricing facts for a very general class of learning schemes. We then show that least squares learning converges to rational expectations. We also explain how the deviations from rational expectations can be made arbitrarily small. In section 4 we present the baseline learning model with risk aversion and the baseline calibration procedure. Section 5 shows that the baseline model can *quantitatively* reproduce all the facts discussed in section 2. The robustness of our findings to various assumptions about the model, the learning rule, or the calibration procedure is illustrated in section 6.

Readers interested in obtaining a glimpse of the quantitative performance of the baseline model may - after reading section 2 - directly jump to table 4 in section 5.

2 Facts

This section describes stylized facts of U.S. stock price data and explains why it proved difficult to reproduce them using standard rational expectations models. The facts presented in this section have been extensively documented in the literature. We reproduce them here as a point of reference for our quantitative exercise in the latter part of the paper and using a single and updated data set.[8]

It is useful to start looking at the data through the lens of a simple dynamic stochastic endowment economy. There is an inelastically supplied stock that can be purchased or sold at any period in a competitive market. Each unit of the stock purchased in the previous period pays dividends D_t in period t. The dividend evolves according to

$$\frac{D_t}{D_{t-1}} = a\varepsilon_t \tag{1}$$

where $\log \varepsilon_t \sim ii\mathcal{N}(-\frac{s^2}{2}, s^2)$ and $a \geq 1$.[9] Obviously, this assumption guarantees that $E(\varepsilon_t) = 1$, $E\left(\frac{D_t}{D_{t-1}}\right) = a$ and $\sigma_{\frac{\Delta D}{D}} = s$.

[7] Cecchetti, Lam, and Mark (2000) determine the misspecification in beliefs about future consumption growth required to match the equity premium and other moments of asset prices.

[8] Details on the underlying data sources are provided in Appendix A.1.

[9] As documented in Mankiw, Romer and Shapiro (1985) and Campbell (2003), this is a reasonable first approximation to the empirical behavior of quarterly dividends in the U.S. It is also the standard assumption in the literature.

The price of the stock P_t depends on the particular model used. We focus, as most of the literature, on the pricing equation

$$P_t = \delta E_t \left[\frac{U'(D_{t+1})}{U'(D_t)} (P_{t+1} + D_{t+1}) \right]. \tag{2}$$

This formula for the equilibrium value of the asset price can be derived from many alternative models. For instance, it is the solution in the original formulation of Lucas (1978), in which homogeneous consumers live forever, have discount factor δ and U is the instantaneous utility function. It is also the equilibrium condition in a two-period overlapping generations model in which δ and U have the same interpretation.[10] Alternatively, it may be obtained from no-arbitrage relationships in partial equilibrium models. The specific underlying model is not relevant for the discussion in this section but will be important in section 3, when we introduce bounded rationality.

Equation (2) defines a mapping from the exogenous dividend process to the stochastic process of prices.[11] The nature of this mapping obviously depends on the way the intertemporal marginal rate of substitution moves with consumption. For example, in the standard case of power preferences, $U(C_t) = \frac{C_t^{1-\gamma}}{1-\gamma}$, equation (2) becomes

$$P_t = \delta E_t \left[\left(\frac{D_t}{D_{t+1}} \right)^{\gamma} (P_{t+1} + D_{t+1}) \right] \tag{3}$$

With rational expectations about next period price, the no-bubble equilibrium stock price satisfies[12]

$$P_t = \frac{\delta \beta^{RE}}{1 - \delta \beta^{RE}} D_t \tag{4}$$

where

$$\beta^{RE} = a^{1-\gamma} e^{-\gamma(1-\gamma)\frac{s^2}{2}} \tag{5}$$

$$E_t \left(\left(\frac{D_t}{D_{t+1}} \right)^{\gamma} P_{t+1}^{RE} \right) = \beta^{RE} P_t^{RE} \tag{6}$$

Equation (4) then implies that the price-dividend (PD) ratio in this model is constant over time and states. Figure 1 confronts this prediction with the actual evolution of quarterly PD ratio in the U.S.[13] To obtain annualized PD ratios one would have to divide the reported values by four. Compared to the simple

[10] A specific example with overlapping generations is provided in Appendix A.2.

[11] This is the case because the models mentioned above equate consumption C_t with dividends D_t. In the data consumption is much less volatile than dividends. This raises important issues that we discuss later in the paper. For the discussion in this section this distinction does not matter.

[12] To see that this is a RE equilibrium note that if prices satisfy (4), then (6) holds, and plugging this equation in (3) gives back (4).

[13] Throughout the paper we follow Campbell (2003) and account for seasonalities in dividend payments by averaging actual payments over the last 4 quarters.

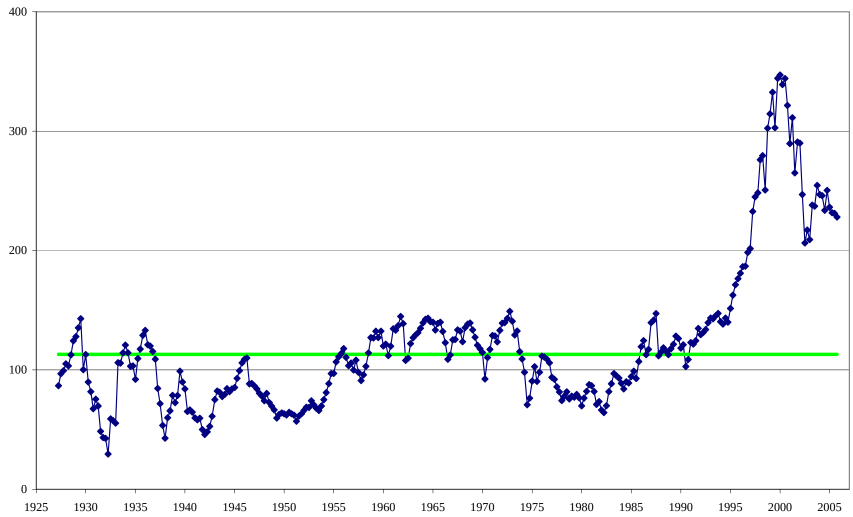

Figure 1: Quarterly U.S. price dividend ratio 1927:1-2005:4

model we just described, the historical PD ratio exhibits rather large fluctuations around its sample mean (the horizontal line in the graph). For example, the PD ratio takes on values below 30 in the year 1932 and values close to 350 in the year 2000. This large discrepancy between the prediction of the basic model and the data is also illustrated in table 1, which shows that the standard deviation of the PD ratio (σ_{PD}) is almost one half of its sample mean ($E(PD)$). This demonstrates the following asset pricing fact:

Fact 1: The PD ratio is very volatile.

It follows that matching the observed volatility of the PD ratio under rational expectations requires alternative preference specifications. Indeed, maintaining the assumptions of *i.i.d.* dividend growth and of a representative agent, the behavior of the marginal rate of substitution is the only degree of freedom left to the theorist. This explains the development of a large and interesting literature exploring time-non-separability in consumption or consumption habits. Introducing habits amounts to consider consumers whose instantaneous utility function is given by $U(\mathbb{C}_{\boldsymbol{t}})$, where $\mathbb{C}_{\boldsymbol{t}} = H(C_t, C_{t-1}, C_{t-2}, ..)$ is a function of current and past consumption. A simple habit model was studied by Abel (1990) who assumed

$$\mathbb{C}_{\boldsymbol{t}} = \frac{C_t}{C_{t-1}^{\kappa}}$$

with $\kappa \in (0,1)$.[14] In this case, the stock price under rational expectations is

$$\frac{P_t}{D_t} = A\left(a\varepsilon_t\right)^{\kappa(\gamma-1)} \tag{7}$$

for some constant A. This equation shows that this model can give rise to a volatile PD ratio. Yet, with ε_t being *i.i.d.* the PD ratio will display no autocorrelation, which is in stark contrast to the empirical evidence. As figure 1 illustrates, the PD ratio displays rather persistent deviations from its sample mean. Indeed, as table 1 shows, the quarterly autocorrelation of the PD ratio (denoted $\rho_{PD,-1}$) is very high. Therefore, this is the second fact we focus on:

Fact 2: The PD ratio is persistent.

The previous observations suggest that matching the volatility and persistence of the PD ratio under rational expectations would require a model that gives rise to a volatile and persistent marginal rate of substitution. This is the avenue pursued in Campbell and Cochrane (1999) who engineer preferences that can match the behavior of the PD ratio we observe in Figure 1. Their specification also helps in replicating the asset pricing facts mentioned later in this section, as well as other facts not mentioned here.[15] Their solution requires, however, imposing a very high degree of relative risk aversion and relies on a fairly sophisticated structure for the habit function $H(\cdot)$.[16]

In our model we maintain the assumption of standard time-separable consumption preferences with moderate degrees of risk aversion. Instead, we relax the rational expectations assumption by replacing the mathematical expectation in equation (2) by the most standard learning algorithm used in the literature. Persistence and volatility of the price dividend ratio will then be the result of adjustments in beliefs that are induced by the learning process.

Before getting into the details of our model, we want to mention three additional asset pricing facts about stock returns. These facts have received considerable attention in the literature and are qualitatively related to the behavior of the PD ratio, as we discuss below.

Fact 3: Stock returns are 'excessively' volatile.

Starting with the work of Shiller (1981) and LeRoy and Porter (1981) it has been recognized that stock prices are more volatile in the data than in standard models. Related to this is the observation that the volatility of stock returns (σ_{r^s}) in the data is almost four times the volatility of dividend growth

[14] Importantly, the main purpose of Abel's model was to generate an 'equity premium' - a fact we discuss below - not to reproduce the behavior of the price dividend ratio.

[15] They also match the pro-cyclical variation of stock prices and the counter-cyclical variation of stock market volatility. We have not explored conditional moments in our learning model, see also the discussion at the end of this section.

[16] They use a coefficient of relative risk aversion is 35 in steady state and higher still in states with 'low surplus consumption ratios'.

($\sigma_{\Delta D/D}$), see table 1.[17] The observed return volatility has been called 'excessive' mainly because the rational expectations model with time-separable preferences predicts approximately identical volatilities. To see this, let r_t^s denote the stock return

$$r_t^s = \frac{P_t + D_t - P_{t-1}}{P_{t-1}} = \left[\frac{\frac{P_t}{D_t} + 1}{\frac{P_{t-1}}{D_{t-1}}} \right] \frac{D_t}{D_{t-1}} - 1 \tag{8}$$

and since with time-separable preferences and *i.i.d.* dividend growth the PD ratio is constant and very large, the term in the square brackets above is approximately equal to one so σ_{r^s} is almost equal to $\sigma_{\Delta D/D}$.

From equation (8) follows that excessive return volatility is *qualitatively* related to Fact 1 discussed above, as return volatility depends partly on the volatility of the PD ratio.[18] Yet, *quantitatively* return volatility also depends on the volatility of dividend growth and - up to a linear approximation - on the first two moments of the cross-correlogram between the PD ratio and the rate of growth of dividends. Since the main contribution of the paper is to show the ability of the learning model to account for the *quantitative* properties of the data, we treat the volatility of returns as a separate asset pricing fact.

U.S. asset pricing facts, 1927:2-2005:4
(quarterly real values, growth rates & returns in percentage terms)

Fact 1	Volatility of PD ratio	$E(PD)$ σ_{PD}	113.20 52.98
Fact 2	Persistence of PD ratio	$\rho_{PD,-1}$	0.92
Fact 3	Excessive return volatility	σ_{r^s} $\sigma_{\frac{\Delta D}{D}}$	11.65 2.98
Fact 4	Excess return predictability	c_5^2 R_5^2	-0.0048 0.1986
Fact 5	Equity premium	$E\left[r^s\right]$ $E\left[r^b\right]$	2.41 0.18

Table 1: Stylized asset pricing facts

[17] This is not due to the fact that we took averages in order to account for seasonalities in quarterly dividends. Even at yearly frequency stock returns are about three times as volatile as dividend growth.

[18] Cochrane (2005) provides a detailed derivation of the qualitative relationship between facts 3 and 1 for *i.i.d.* dividend growth.

Fact 4: Excess stock returns are predictable over the long-run.

While stock returns are difficult to predict in general at short horizons, the PD ratio is negatively related to future excess stock returns in the long run. This is illustrated in Table 2, which shows the results of regressing future cumulated excess returns over different horizons on today's price dividend ratio.[19] The absolute value of the parameter estimate and the R^2 both increase with the horizon. As argued in Cochrane (2005, chapter 20), the presence of return predictability and the increase in the R^2 with the prediction horizon are *qualitatively* a joint consequence of Fact 2 and *i.i.d.* dividend growth. Nevertheless, we keep excess return predictability as an independent result, since we are again interested in the *quantitative* model implications. Yet, Cochrane also shows that the absolute value of the regression parameter increases approximately linearly with the prediction horizon, which is a quantitative result. For this reason, we summarize the return predictability evidence using the regression outcome for a single prediction horizon. We choose to include the 5 year horizon in table 1.[20]

Years	Coefficient on PD, c_s^2	R^2
1	-0.0008	0.0438
3	-0.0023	0.1196
5	-0.0048	0.1986
10	-0.0219	0.3285

Table 2: Excess stock return predictability

Fact 5: The equity premium puzzle.

Finally, and even though the emphasis of our paper is on moments of the PD ratio and stock returns, it is interesting to note that learning also improves the ability of the standard model to match the equity premium puzzle, i.e., the observation that stock returns - averaged over long time spans and measured in real terms - tend to be high relative to short-term real bond returns. The equity premium puzzle is illustrated in table 1, which shows the average quarterly real return on bonds $(E(r_t^b))$ being much lower than the corresponding return on stocks $(E(r_t^s))$.

Unlike Campbell and Cochrane (1999) we do not include in our list of facts any correlation between stock market data and real variables such as consumption or investment. In this sense, we follow more closely the literature in finance. In our model, it is the learning scheme that delivers the movement in stock prices, even in a model with risk neutrality in which the marginal rate of substitution is constant. This contrasts with the habit literature where the

[19] More precisely, the table reports results from OLS estimation of

$$X_{t,t+s} = c_s^1 + c_s^2 PD_t + u_t^s$$

for $s = 1, 3, 5, 10$, where $X_{t,t+s}$ is the observed real excess return of stocks over bonds between t and $t+s$. The second column of Table 2 reports estimates of c_s^2.

[20] We also used longer and shorter horizons. This did not affect our findings.

movement of stock prices is obtained by modeling the way the observed stochastic process for consumption generates movements in the marginal rate of substitution. The latter explains why the habit literature focuses on the relationship between particularly low values of consumption and low stock prices. Since this mechanism does not play a significant role in our model, we abstract from these asset pricing facts.

3 The risk neutral case

In this section we analyze the simplest asset pricing model assuming risk neutrality $\gamma = 0$. As in the rest of the paper, we will now maintain time-separable preferences $\mathbb{C}_t = C_t$. The goal of this section is to derive *qualitative* results and to show how the introduction of learning improves the performance compared to a setting with rational expectations. Sections 4 and 5 present a formal quantitative model evaluation, extending the analysis to risk-averse investors. With risk neutrality and rational expectations the model is far from explaining any of the asset pricing facts described in the previous section: the PD ratio is constant, stock returns are unforecastable (*i.i.d.*) at any horizon, approximately as volatile as dividend growth, i.e., do not display excess volatility. In addition, there is no equity premium.[21] For these reasons, the risk-neutral model is particularly suited to illustrate how the introduction of learning qualitatively improves model performance.

The consumer has beliefs about future variables (prices and dividends), summarized in expectations denoted $\widetilde{E}$, that we allow to be less than fully rational. Then, equation (3) becomes

$$P_t = \delta \; \widetilde{E}_t \left(P_{t+1} + D_{t+1} \right) \tag{9}$$

This asset pricing equation will be the focus of our analysis in this section.

The key feature that we exploit in this equation, is that today's stock price is determined by tomorrow's expected price. Alternatively, some papers in the learning literature[22] have studied stock prices when agents formulate expectations about the discounted sum of all future dividends and equilibrium prices satisfy

$$P_t = \widetilde{E}_t \sum_{j=1}^{\infty} \delta^j D_{t+j} \tag{10}$$

When agents learn about the discounted sum of dividends, the model has much lower chances to generate large asset price volatility. This is so because agents are learning about the parameters of the underlying dividend process, which evolves exogenously to the evolution of stock prices. While market prices then depend on agents' beliefs about the dividend process, the market outcomes fail

[21] We do not mention the serial correlation of PD ratio because the RE model implies a constant PD ratio so its serial correlation is undefined.

[22] For example, Timmermann (1993, 1996), Brennan and Xia (2001), Cogley and Sargent (2006).

to feed back into agents' expectations. This severely limits the ability of the model to generate interesting 'data-like' behavior. In contrast, the one-period formulation (9) requires agents to estimate the parameters driving the stock price process, i.e., to estimate the process of an endogenous variable. As a result, agents beliefs about the stock price process influence the actual behavior of stock prices and actual price behavior feeds back into agents beliefs. It is this self-referential nature of our model that makes it attractive in explaining the data.

It is well known that under RE the one-period ahead formulation of (9) is equivalent to the discounted sum expression for prices.[23] However, this fails to be the case under learning. There exist a number of theoretical reasons justifying the focus on the one-step-ahead formulation. A particular one is given in appendix A.2 which describes a two-period overlapping generations model where the optimality condition of agents is indeed given by (9). As shown in the appendix, individual agents do not care about the discounted sum of dividends (10) because they know that they will have to sell the stock in the second period of their life. This appendix makes the current paper self-contained in providing a justification for using (9).

In a companion paper (Adam, Marcet and Nicolini (2008)) we provide additional and more detailed theoretical reasons for preferring the one-step-ahead formulation (9). A summary follows.

First, note that a Bayesian agent would demand stocks until the first order condition

$$P_t = D_t \ E_t^{BAY} \left(\frac{\delta a}{1 - \delta a} \right) \tag{11}$$

holds, where E_t^{BAY} is the expectation based on the Bayesian posterior using information up to t. Since the denominator $1 - \delta a$ is close to zero, uncertainty about the true value of the mean dividend growth rate a is likely to make the above expectation very large, possibly even infinite. Specifically, the size of the valuation error depends critically on the support of agents' prior beliefs about a. Therefore, a small amount of prior uncertainty about the true value of a can induce the Bayesian agent to be willing to buy the stock at a very high or even infinite price, something that ex-post will appear as a very large 'error'. For this reason, a Bayesian agent who uses (11) to price the asset will not necessarily outperform the agents considered in this paper who use a simple learning scheme. This issue fails to show up in the literature studying discounted sums we mentioned above because it uses the pricing formula

$$P_t = \frac{\delta \ E_t^{BAY}(a)}{1 - \delta \ E_t^{BAY}(a)} \quad . \tag{12}$$

As is well recognized in this literature, this shortcut amounts to assuming that agents ignore their uncertainty about the true value of a. A fully Bayesian

[23] More precisely, equivalence is obtained if $\widetilde{E}_t[\cdot] = E_t[\cdot]$ and if the no-rational-bubble requirement $\lim_{j \to \infty} \delta^j \ E P_{t+j} = 0$ holds.

agent would use the entire posterior distribution of a to evaluate the expected discounted sum as in (11).

Second, we show that agents may care about forecasting future price as in (9) even if they live infinitely many periods. We demonstrate this using a heterogeneous agent model where markets are incomplete due to shortselling constraints. The market incompleteness drives a wedge between individual discount factor and market discount factor because agents occasionally strictly prefer to sell the asset at equilibrium prices. Assuming that agents are individually perfectly rational, in the sense that they know their utility and their constraints and that they maximize utility given prices, implies that a version of (9) holds. Moreover, these agents cannot simply iterate on their first order conditions (9) to obtain the infinite sum representation (10) of prices in terms of future dividends only, unless they have very detailed knowledge of how the stock market works. In particular, they need to know the market discount factor (which differs from their own discount factor) but also that there are no rational bubbles. Agents without such information will have to formulate expectations about future prices. We illustrate this point in appendix A.2 for the overlapping generations economy.

For all these reasons we believe that our one-period formulation (9) is an interesting avenue to explore.

3.1 Analytical results

In this section we show that the introduction of learning changes qualitatively the behavior of stock prices in the direction of improving the match with the stylized facts described above. At this point we consider a wide class of learning schemes that includes the standard rules used in the literature. This serves to prove that the effects we discuss occur in a general class of learning models. Later on we will restrict attention to learning schemes that forecast well within the model.

We first trivially rewrite the expectation of the agent by splitting the sum in the expectation:

$$P_t = \delta \widetilde{E}_t \left(P_{t+1}\right) + \delta \widetilde{E}_t \left(D_{t+1}\right) \tag{13}$$

We assume that agents know how to formulate the conditional expectation of the dividend $\widetilde{E}_t \left(D_{t+1}\right) = aD_t$. This assumption is for simplicity, it clarifies that it is learning about future prices that allows the model to better match the data. In Appendix A.5 we show that these theoretical pricing implications are very similar, and in section 6 we verify that the quantitative implications barely change when we introduce learning about dividends.

Agents are assumed to use a learning scheme in order to form a forecast $\widetilde{E}_t(P_{t+1})$ based on past information. Equation (4) shows that under rational expectations the growth of stock prices is constant, since $E_t \left[P_{t+1}^{RE}\right] = \beta^{RE} P_t^{RE}$. As we restrict our analysis to learning rules that behave close enough to rational expectations, we specify expectations under learning as

$$\widetilde{E}_t \left[P_{t+1}\right] = \beta_t \; P_t \tag{14}$$

where $\beta_t > 0$ denotes agents' time t estimate of gross stock price growth. For $\beta_t = a$, agents' beliefs coincide with rational expectations. Also, if agents' beliefs converge over time to the RE equilibrium ($\lim_{t\to\infty} \beta_t = a$), investors will realize in the long-run that they were correct in using the functional form (14). Yet, during the transition beliefs may deviate from rational ones.

It remains to specify how agents update their beliefs β_t. The focus of our analysis will be the following general learning mechanism

$$\Delta\beta_t = f_t\left(\frac{P_{t-1}}{P_{t-2}} - \beta_{t-1}\right) \tag{15}$$

for some exogenously chosen functions $f_t : R \to R$ with the properties

$$\begin{aligned} &f_t(0) = 0 \\ &f_t(\cdot)\text{increasing} \end{aligned}$$

We thus assume that beliefs adjust in the direction of the last prediction error, i.e., agents revise beliefs upwards (downwards) if their expectations last period underpredicted (overpredicted) the stock price growth that was actually realized in this period.[24] Arguably, a learning scheme that violates these conditions would appear quite unnatural.

As in many learning papers, we need to restrict the above learning scheme further so as to guarantee that beliefs remain bounded. In this section, we simply use a projection facility which assumes that agents ignore observations that would cause their expected price growth to be excessively high:

$$\Delta\beta_t = \begin{cases} f_t\left(\frac{P_{t-1}}{P_{t-2}} - \beta_{t-1}\right) & \text{if } \beta_{t-1} + f_t\left(\frac{P_{t-1}}{P_{t-2}} - \beta_{t-1}\right) < \beta^U \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

This projection facility has been used in many learning papers, including Timmermann (1993, 1996), Marcet and Sargent (1989), Evans and Honkapohja (2001) and Cogley and Sargent (2006). Throughout the paper we will refer to (16) as the "standard" projection facility. We give some interpretation of this after equation (17). In section 4 we use a different and less standard projection facility for numerical reasons. In the text we will often provide intuition referring to (15) since it is the equation that holds most periods, but in all formal results and in the proofs we use (16).

At this point, we show that the key features of the model emerge within this general specification. We will later specialize the learning scheme in section 3.2 to obtain quantitative and convergence results.

[24] Note that β_t is determined from observations up to period $t-1$ only. The assumption that the current price does not enter in the formulation of the expectations is common in the learning literature and is entertained for simplicity. Difficulties emerging with simultaneous information sets are discussed in Adam (2003).

3.1.1 Stock prices under learning

Given the perceptions β_t, the expectation function (14), and the assumption on perceived dividends, equation (13) implies that prices under learning satisfy

$$P_t = \frac{\delta a D_t}{1 - \delta \beta_t}. \tag{17}$$

Assuming that $\beta^U \leq \delta^{-1}$ guarantees that all prices are positive and finite. One interpretation is that if the last observed price growth implies beliefs that are too high, agents realize that this would prompt a crazy action (infinite stock demand) and they would ignore this observation, as dictated by the projection facility. Obviously, it is equivalent to require that the PD is less than the upper bound $U^{PD} \equiv \frac{\delta a}{1-\delta\beta^U}$, therefore an alternative interpretation is that if the PD is higher than this upper bound agents will start fearing a downturn and they will bring their expectations down. Yet another interpretation is that if PD is too high, some government agency will intervene to bring the price down.[25] In the simulations that we report below this facility is binding only rarely and the properties of the learning model are not sensitive to the precise value we assign to β^U, so in the remainder of this subsection 3.1 we give intuition on the dynamics of the model by focusing on (15).

Since β_t and ε_t are independent, the previous equation implies that

$$Var\left(\ln \frac{P_t}{P_{t-1}}\right) = Var\left(\ln \frac{1 - \delta\beta_{t-1}}{1 - \delta\beta_t}\right) + Var\left(\ln \frac{D_t}{D_{t-1}}\right), \tag{18}$$

and shows that price growth under learning is more volatile than dividend growth. Clearly, this occurs because the volatility of beliefs adds to the volatility generated by fundamentals. While this intuition is present in previous models of learning, e.g., Timmermann (1993), it will be particular to our case that under more specific learning schemes $Var\left(\ln \frac{1-\delta\beta_t}{1-\delta\beta_{t+1}}\right)$ is likely to be high and remain high for a long time.

Equation (17) shows that the PD ratio is monotonically related to beliefs β_t. One can thus understand the qualitative dynamics of the PD ratio by studying the belief dynamics. To derive these dynamics notice

$$\frac{P_t}{P_{t-1}} = T(\beta_t, \Delta\beta_t)\ \varepsilon_t \tag{19}$$

where

$$T(\beta, \Delta\beta) \equiv a + \frac{a\delta\ \Delta\beta}{1 - \delta\beta} \tag{20}$$

[25] To mention one such intervention that has been documented in detail, Voth (2003) explains how the German central bank intervened indirectly in 1927 to reduce lending to equity investors. This intervention was prompted by a "genuine concern about the 'exuberant' level of the stock market" on the part of the central bank and it caused stock prices to go down sharply. More recently, announcements by central bankers (the famous speech by Alan Greenspan on October 16th 1987) or interest rate increases may have played a similar role.

It follows directly from (19) that $E_t(P_{t+1}/P_t) = T(\beta_t, \Delta\beta_t)$. Therefore, $T(\beta_t, \Delta\beta_t)$ is the *actual* expected stock price growth given that the *perceived* price growth has been given by $\beta_t, \Delta\beta_t$. Substituting (19) in the law of motion for beliefs (16) we have that the dynamics of β_t $(t \geq 1)$ are fully described by a second-order stochastic difference equation

$$\Delta\beta_{t+1} = f_{t+1}\left(T(\beta_t, \Delta\beta_t)\varepsilon_t - \beta_t\right) \tag{21}$$

for given initial conditions (D_0, P_{-1}), and initial belief β_0. This equation can not be solved analytically due to non-linearities,[26] but it is still possible to gain qualitative insights into the belief dynamics of the model. We do this in the next section.

3.1.2 Deterministic dynamics

To discuss the dynamics of beliefs β_t under learning, we simplify matters by considering the deterministic case in which $\varepsilon_t \equiv 1$. Equation (21) then simplifies to

$$\Delta\beta_{t+1} = f_{t+1}\left(T(\beta_t, \Delta\beta_t) - \beta_t\right) \tag{22}$$

We thus restrict attention to the endogenous stock price dynamics generated by the learning mechanism rather than the dynamics induced by exogenous disturbances. Given the properties of f_t, equation (22) shows that beliefs are increasing whenever $T(\beta_t, \Delta\beta_t) > \beta_t$, i.e., whenever actual stock price growth exceeds expected stock price growth. Understanding the evolution of beliefs thus requires studying the T-mapping.

We start by noting that the actual stock price growth implied by T depends not only on the *level* of price growth expectations β_t but also on the *change* $\triangle\beta_t$. This imparts momentum on stock prices, leading to a feedback between expected and actual stock price growth. Formally we state the following result

Momentum: *If* $\beta_t \leq a$ *and* $\Delta\beta_t > 0$, *then*[27]

$$\Delta\beta_{t+1} > 0$$

It also holds if all inequalities are reversed.

Therefore, if agents arrived at the rational expectations belief $\beta_t = a$ from below ($\triangle\beta_t > 0$), the price growth generated by the learning model would keep growing and it would exceed the fundamental growth rate a. Just because agents' expectations have become more optimistic (in what a journalist would perhaps call a 'bullish' market), the price growth in the market has a tendency

[26] Notice that even in the case we consider in the next section where f_t is linear, the difference equation is non-linear due to the presence of T.

[27] This follows trivially from the fact that, given the definition of T, it follow that $\Delta\beta_t > 0$ implies $T(\beta_t, \Delta\beta_t) > a$ so that β has to increase. Of course, this holds if the projection facility does not apply

to be larger than the growth in fundamentals. Since agents will use this higher-than-fundamental stock price growth to update their beliefs in the next period, β_{t+1} will tend to overshoot a, which will further reinforce the upward tendency. Since beliefs are monotonically related to the PD ratio, see equation (17), there will be momentum in the stock price, resembling a 'naturally arising Ponzi process'. Conversely, if $\beta_t = a$ in a bearish market ($\Delta\beta_t < 0$), beliefs and price growth display downward momentum, i.e., a tendency to undershoot the RE value.

It can be shown, however, that stock prices and beliefs can not stay at levels unjustified by fundamentals forever and that after any deviation they will eventually move towards the fundamental value. Furthermore, it turns out that when the price starts going down it does so monotonically until it arrives at (or surpasses) the fundamental value. Formally, under some additional technical assumptions we have the following result

Mean reversion: [28] *If in some period t we have $\beta_t > a$, given $\eta > 0$ sufficiently small, there is a finite period $t'' > t$ such that $\beta_{t''} < a + \eta$.*

Furthermore, oscillations are monotonic in the sense that letting t' be the first period $t'' \geq t' \geq t$ such that $\Delta\beta_{t'} < 0$, then β is non-decreasing between t and t' and it is non-increasing between t' and t''.

Symmetrically, if $\beta_t < a$ eventually $\beta_{t''} > a - \eta$ and oscillations are monotonic.

Since η can be chosen arbitrarily small, the previous statement shows that even if beliefs are much higher than fundamentals they will eventually move back arbitrarily close to or below fundamentals. The monotone relationship between beliefs and the PD ratio implies mean reverting behavior of the PD ratio.

Momentum and mean reversion are also illustrated by the study of the phase diagram for the dynamics of the beliefs (β_t, β_{t-1}). Figure 2 illustrates the direction that beliefs move, according to equation (22).[29] The arrows in the figure thereby indicate the direction in which the difference equation is going to move for any possible state (β_t, β_{t-1}) and the solid lines indicate the boundaries of these areas.[30] Since we have a difference rather than a differential equation, we cannot plot the evolution of beliefs exactly. Nevertheless, the arrows suggest that the beliefs are likely to move in ellipses around the rational expectations equilibrium $(\beta_t, \beta_{t-1}) = (a, a)$. Consider, for example, point A in the diagram. Although at this point β_t is already below its fundamental value, the phase diagram indicates that beliefs will fall further. This is the result of the momentum induced by the fact that $\beta_t < \beta_{t-1}$ at point A. Beliefs can go, for example,

[28] See Appendix A.3 for the assumptions and the proof. Note also that, even though in the informal discussion of the text we ignore the projection facility, the proof in Appendix A.3 uses the projection facility.

[29] Appendix A.4 explains in detail the construction of the phase diagram.

[30] The vertical solid line close to δ^{-1} is meant to illustrate the restriction $\beta < \delta^{-1}$ imposed on the learning rule.

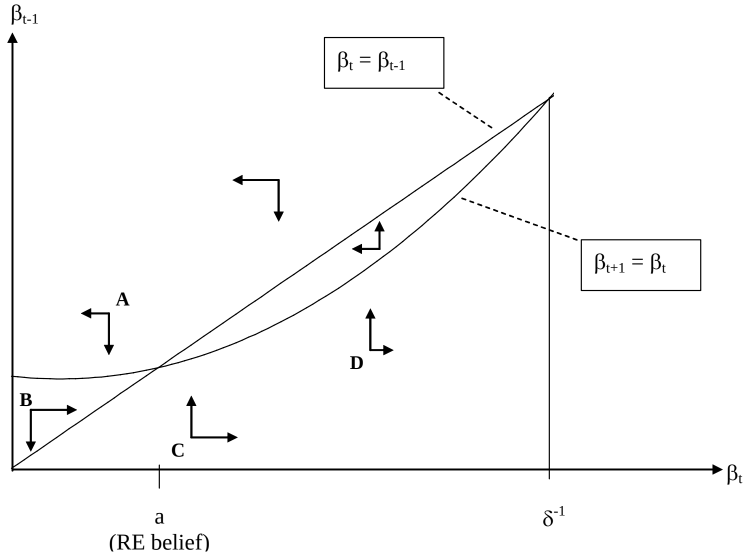

Figure 2: Phase diagram illustruating momentum and mean-reversion

to point B where they will start to revert direction and move on to points C then, to D etc.: beliefs thus display mean reversion. The elliptic movements imply that beliefs (and thus the PD ratio) are likely to oscillate in sustained and persistent swings around a.

Momentum together with the mean reversion allows the model to match the volatility and serial correlation of the PD ratio (facts 1 and 2). Also, according to our discussion around equation (18), momentum imparts variability to the ratio $\frac{1-\delta\beta_{t-1}}{1-\delta\beta_t}$ and is likely to deliver more volatile stock returns (fact 3). As discussed in Cochrane (2005), a serially correlated and mean reverting PD ratio should give rise to excess return predictability (fact 4).

These results held for a very general class of learning schemes. The next section specializes the learning scheme, allowing to prove asymptotic results and to study by simulation that introducing learning in the risk-neutral model causes a big improvement in the ability of the model to explain stock price volatility. It can also generate a sizable equity premium (fact 5).

3.2 The Risk Neutral Model with OLS

3.2.1 The learning rule

We specialize the learning rule by assuming the most common learning schemes used in the literature on learning about expectations. We assume the standard updating equation from the stochastic control literature

$$\beta_t = \beta_{t-1} + \frac{1}{\alpha_t}\left(\frac{P_{t-1}}{P_{t-2}} - \beta_{t-1}\right) \tag{23}$$

for all $t \geq 1$, for a given sequence of $\alpha_t \geq 1$, and a given initial belief β_0 which is given outside the model.[31] This assumes f_t to be linear. The values $1/\alpha_t$ are called the 'gain' sequence, they dictate how strongly beliefs are updated in the direction of the last prediction error. In this section, we assume the simplest possible specification:

$$\begin{aligned} &\alpha_t = \alpha_{t-1} + 1 \quad t \geq 2 \\ &\alpha_1 \geq 1 \quad \text{given.} \end{aligned} \tag{24}$$

With these assumptions the model evolves as follows. In the first period, β_0 determines the first price P_0 and, given the historical price P_{-1}, the first observed growth rate $\frac{P_0}{P_{-1}}$, which is used to update beliefs to β_1 using (23); the belief β_1 determines P_1 and the process evolves recursively in this manner. As in any self-referential model of learning, prices enter in the determination of beliefs and vice versa. As we argued in the previous section, it is this feature of the model that generates the dynamics we are interested in.

Using simple algebra, equation (23) implies

$$\beta_t = \frac{1}{t + \alpha_1 - 1}\left(\sum_{j=0}^{t-1} \frac{P_j}{P_{j-1}} + (\alpha_1 - 1)\ \beta_0\right).$$

Obviously, if $\alpha_1 = 1$ this is just the sample average of the stock price growth and, therefore, it amounts to OLS if only a constant is used in the regression equation. For the case where α_1 is an integer, this expression shows that β_t is equal to the average sample growth rate, if - in addition to the actually observed prices - we would have $(\alpha_1 - 1)$ observations of a growth rate equal to β_0. A high α_1 thus indicates that agents possess a high degree of 'confidence' in their initial belief β_0.

In a Bayesian interpretation, β_0 would be the prior mean of stock price growth, $(\alpha_1 - 1)$ the precision of the prior, and - assuming that the growth rate of prices is normally distributed and i.i.d. - the beliefs β_t would be equal to the posterior mean. One might thus be tempted to argue that β_t is effectively a Bayesian estimator. Obviously, this is only true for a 'Bayesian' placing probability one on $\frac{P_t}{P_{t-1}}$ being i.i.d. Since learning causes price growth to deviate from i.i.d. behavior along the transition, such priors fail to contain the 'grain of truth' (along the transition) that should be present in Bayesian analysis. Since the i.i.d. assumption will hold asymptotically (we will prove this later on), the learning scheme is a Bayesian estimator in the long run.[32]

[31] In the long-run the particular initial value β_0 is of little importance.

[32] In a proper Bayesian formulation agents would use a likelihood function with the property

For the case $\alpha_1 = 1$, β_t is just the sample average of stock price growth, i.e., agents have no confidence in their initial belief β_0. In this case β_0 matters only for the first period, but ceases to affect anything after the first piece of data has arrived. More generally, assuming a value for α_1 close to 1 would spuriously generate a large amount of price fluctuations, simply due to the fact that initial beliefs are heavily influenced by the first few observations and thus very volatile. Also, pure OLS assumes that agents have no faith whatsoever in their initial belief and possess no knowledge about the economy in the beginning.

In the spirit of restricting equilibrium expectations in our learning model to be close to rational, in the simulations of section 5 we set initial beliefs equal to the value of the growth rate of prices under RE

$$\beta_0 = a$$

and choose a very large α_1. Thus, we assume that beliefs start at the RE value, and that the initial degree of confidence in the RE belief is high, but not perfect. Clearly, in the limit as $1/\alpha_1 \to 0$ our learning model reduces to the RE model in all periods, so that the initial gain $1/\alpha_1$ can be interpreted as a measure of 'distance' of the learning model from to the rational expectations model. The maximum distance from RE is achieved for $1/\alpha_1 = 1$, i.e., pure OLS learning.

Finally, we need to introduce the projection facility in (23) as described before.

3.2.2 Asymptotic Rationality

In this section we study the limiting behavior of the model under learning. The literature on convergence of least squares learning shows that the T-mapping defined in equation (20) is central to whether or not the economy converges to RE.[33] It is now well established that in a large class of models convergence (divergence) of least squares learning to (from) RE equilibria is strongly related to stability (instability) of the associated o.d.e. $\dot{\beta} = T(\beta) - \beta$. Most of the literature considers models where the mapping from perceived to actual expectations does not depend on the change in perceptions, unlike in our case where T depends on $\Delta\beta_t$. Since for large t the gain $(\alpha_t)^{-1}$ is very small, we have that (23) implies $\Delta\beta_t \approx 0$. One could thus think of the relevant mapping for convergence in our paper as being $T(\cdot, 0) = a$ for all β. Asymptotically the T-map is thus flat and the differential equation $\dot{\beta} = T(\beta, 0) - \beta = a - \beta$ is stable. This suggests that convergence to the RE equilibrium value $\beta = a$ occurs and that it should be very fast, so that one might then conclude that the model under

that if agents use it to update their posterior, it turns out to be the true likelihood of the model in all periods. Since the 'correct' likelihood in each period depends on the way agents learn, it would have to solve a complicated fixed point. Finding such a truly Bayesian learning scheme is very difficult and the question remains how agents could have learned a likelihood that has such a special property. For these reasons Bray and Kreps (1987) concluded that models of self-referential Bayesian learning were unlikely to be a fruitful avenue of research.

[33] See Marcet and Sargent (1989) and Evans and Honkapohja (2001).

learning is likely to behave close to RE and there is not much to be gained from introducing learning into the standard asset pricing model.[34]

The following result shows that the above considerations are partly correct

Convergence of OLS *If $\beta^U > \beta^{RE}$, the learning scheme (23)-(24) of this section satisfies*

$$\beta_t \rightarrow a \text{ almost surely as } t \rightarrow \infty$$

for any initial conditions $\alpha_1 \geq 1, \beta_0 \in \left(0, \beta^U\right)$.

The proof is in Appendix A.7, it requires some mild technical assumptions on the support of ε. Note that we prove global convergence, while many results in the literature on self-referential learning are about local convergence.

The learning model thus satisfies 'Asymptotic Rationality' as defined in section III in Marcet and Nicolini (2003). It implies that agents using the learning mechanism will realize in the long run that they are using the best possible forecast, therefore, they would not have incentives to change their learning scheme in the long run.

Yet, the remainder of this paper shows that our model behaves very different from RE during the transition to the limit. This occurs even though agents are using an estimator that starts with strong confidence at the RE value, that converges to the RE value, and that will be the best estimator in the long run. The difference is so large that even this very simple version of the model together with the very simple learning scheme introduced in section 3.1 qualitatively matches the asset pricing facts much better than the model under RE.

3.2.3 Simulations

We now illustrate the previous discussion of the model under learning by reporting simulation results for certain parameter values. We compare outcomes with the RE solution to show in what dimensions the behavior of the model improves when learning is introduced.

We choose the parameter values for the dividend process (1) so as to match the observed mean and standard deviation of US dividends:

$$a = 1.0035, \quad s = 0.0298 \tag{25}$$

We bias results in favor of the RE version of the model by choosing the discount factor so that the RE model matches the average PD ratio we observe in the data.[35] This amounts to choosing

$$\delta = 0.9877.$$

[34] That convergence "should be fast" would follow from results in Marcet and Sargent (1995) and Evans and Honkapohja (2001), showing that the asymptotic speed of convergence depends on the size of $T'(\beta^{RE})$. Since in this model we have $T'(\beta^{RE}) = 0$ these results would seem to indicate that convergence should be quite fast.

[35] This differs from the latter part of the paper where we choose δ to match globally the moments of interest.

As we mentioned before, for the learning model we set $\beta_0 = a$. We also choose

$$1/\alpha_1 = 0.02$$

These starting values are chosen to insure that the agents' expectations will not depart too much from rationality: initial beliefs are equal to the RE value and the first quarterly observation of stock price growth obtains a weight of just 2%, with the remaining weight of 98% being placed on the RE 'prior'. With this value for α_1 (and without shocks) it takes more than twelve years for β_t to converge half way from the initial RE value towards the observed sample mean of stock price growth. Finally, we set the upper bound on β_t so that the quarterly PD ratio will never exceed $U^{PD} = 500$.

Table 3 shows the moments from the data together with the corresponding model statistics.[36] The column labeled US data reports the statistics discussed in section 2. It is clear from table 3 that the RE model fails to explain key asset pricing moments. Consistent with our earlier discussion the RE equilibrium is far away from explaining the volatility of asset prices and predictability of PD ratio.[37]

Statistic	US Data	RE model	Learning Model
$E(r^s)$	2.41	1.24	2.04
$E(r^b)$	0.18	1.24	1.24
$E(PD)$	113.20	113.20	86.04
σ_{r^s}	11.65	3.01	8.98
σ_{PD}	52.98	0.00	40.42
$\rho_{PD,-1}$	0.92	-	0.91
c_1^2	-0.0008	-	-0.0021
c_3^2	-0.0023	-	-0.0051
c_5^2	-0.0048	-	-0.0070
c_{10}^2	-0.0219	-	-0.0104
R_1^2	0.0438	0.00	0.1647
R_3^2	0.1196	0.00	0.2652
R_5^2	0.1986	0.00	0.2735
R_{10}^2	0.3285	0.00	0.3182

Table 3: Data and model under risk neutrality

Table 3 shows that introducing learning the model displays considerably higher volatility of stock returns, high volatility and high persistence of the

[36] Throghout the paper we compute model statistics as follows: for each model we use 5000 realizations of 295 periods each, i.e., the same length as the available data. The reported statistic is the average value of the statistics across simulations, which is a numerical approximation to the expected value of the statistic for this sample size.

[37] Since PD is constant under RE, the coefficients c_s^2 of the predictability regressions are undefined under RE. This is not the case for the R^2 values.

PD ratio, and the coefficients and R^2 of the excess predictability regressions all move strongly in the direction of the data. In particular, the model-implied regression coefficients become increasingly negative and the R^2 of the regressions increase with the forecast horizon. This is consistent with our earlier discussion about the price dynamics implied by learning. Clearly, the statistics of the learning model do not match the moments in the data quantitatively, but the purpose of the table is to show that allowing for small departures from rationality substantially improves the outcome. This is surprising, given that the model adds only one free parameter $(1/\alpha_1)$ relative to the RE model and that we made no effort to choose parameters to best match the moments in any way. Additional simulations we conducted show that the qualitative improvements provided by the model are very robust to changes in $1/\alpha_1$, as long as it is neither too small - in which case the model behaves as the RE model - nor too large - in which case the model delivers too much volatility.

Table 3 also shows that the learning model delivers a positive equity premium. The equity premium is not the focus of the paper, but here we go some way into explaining why the risk-neutral model delivers a risk premium. It proves useful to re-express the gross stock return between period 0 and period N as the product of three terms

$$\prod_{t=1}^{N} \frac{P_t + D_t}{P_{t-1}} = \underbrace{\prod_{t=1}^{N} \frac{D_t}{D_{t-1}}}_{=R_1} \cdot \underbrace{\left(\frac{PD_N + 1}{PD_0} \right)}_{=R_2} \cdot \underbrace{\prod_{t=1}^{N-1} \frac{PD_t + 1}{PD_t}}_{=R_3}.$$

The first term (R_1) is independent of the way expectations are formed, thus it cannot contribute to explaining the emergence of an equity premium in the learning model. The second term (R_2) can potentially generate an equity premium if the terminal price dividend ratio in the learning model (PD_N) is on average higher than under rational expectations.[38,39] Yet, our simulations show that the opposite is the case: under learning the terminal PD ratio in the sample is lower (on average) than under rational expectations; this term thus generates a negative premium under learning. The equity premium in Table 3 must thus be due to the behavior of the last component (R_3). This term gives rise to an equity premium via a *mean* effect and a *volatility* effect.

If PD is on average low, this drives R_3 up. Besides this mean effect, there exists also a volatility effect, which emerges from the convexity of $\frac{PD_t+1}{PD_t}$ as a function of the price dividend ratio. It implies that the equity premium is higher under learning for models with a volatile PD. Since the price dividend ratio has a higher variance than under rational expectations this contributes to generate a risk premium under learning.[40] The volatility effect suggests that

[38] The value of PD_0 is the same under learning and rational expectations since initial expectations in the learning model are set equal to the rational expectations value.

[39] The equity premium of Cogley and Sargent (2006) is due to this term.

[40] The data suggest that this convexity effect is only moderately relevant: for the US data 1927:2-2000:4, it is at most 0.16% in quarterly real terms, thus explains about 8% of the equity premium.

the inability to match the variability of the price dividend ratio and the equity premium are not independent facts and that models that generate insufficient variability of the price dividend ratio also tend to generate an insufficiently high equity premium.

4 Baseline model with risk aversion

The remainder of the paper shows that the learning model can also *quantitatively* account for the moments in the data, once we allow for moderate degrees of risk-aversion, and that this finding is robust. Here we present the baseline model with risk aversion, our simple baseline specification for the learning rule (OLS), and the baseline calibration procedure. The quantitative results are discussed in section 5, while section 6 illustrates the robustness of our quantitative findings to a variety of changes in the learning rule and the calibration procedure.

4.1 Learning under risk aversion

We now present the baseline learning model with risk aversion and show that the insights from the risk neutral model extend in a natural way to the case with risk aversion.

The investor's first-order conditions (3) together with the assumption that agents know the conditional expectations of dividends deliver the stock pricing equation under learning:[41,42]

$$P_t = \delta \widetilde{E}_t \left(\left(\frac{D_t}{D_{t+1}} \right)^{\gamma} P_{t+1} \right) + \delta E_t \left(\frac{D_t^{\gamma}}{D_{t+1}^{\gamma-1}} \right) \tag{26}$$

To specify the learning model, in close analogy to the risk-neutral case, we consider learning agents whose expectations in (26) are of the form

$$\widetilde{E}_t \left(\left(\frac{D_t}{D_{t+1}} \right)^{\gamma} P_{t+1} \right) = \beta_t P_t \tag{27}$$

where β_t denotes agents' best estimate of $E[(D_t/D_{t+1})^{\gamma} (P_{t+1}/P_t)]$, i.e., their expectations of *risk-adjusted* stock price growth. As implied by equation (6) if $\beta_t = \beta^{RE}$ agents have rational expectations and if $\beta_t \to \beta^{RE}$ the learning model will satisfy Asymptotic Rationality, where the expression for β^{RE} is given in equation (5).

As a baseline specification, we consider again the case where agents use OLS to formulate their expectations of future (risk-adjusted) stock price growth

$$\beta_t = \beta_{t-1} + \frac{1}{\alpha_t} \left[\left(\frac{D_{t-2}}{D_{t-1}} \right)^{\gamma} \frac{P_{t-1}}{P_{t-2}} - \beta_{t-1} \right] \tag{28}$$

[41] As in section 2, we impose the market clearing condition $C_t = D_t$ and will associate consumption with dividends in the data. This is not entirely innocuous as dividend growth in the data is considerably more volatile than consumption growth. Section 6 will consider the case $C_t \neq D_t$.

[42] Appendix A.2 shows how an overlapping generations model that implies equation (26).

where the gain sequence $1/\alpha_t$ continues to be described by (24).

Again, in the spirit of allowing for only small deviations from rationality, we restrict initial beliefs to be rational ($\beta_0 = \beta^{RE}$). Appendix A.7 shows that learning will cause beliefs to globally converge to RE, i.e., $\beta_t \to \beta^{RE}$ and $\left|P_t^{RE} - P_t\right| \to 0$ almost surely for any initial conditions $\beta_0 \in \left(0, \beta^U\right), \alpha_1 \geq 1$. The learning scheme thus satisfies Asymptotic Rationality.

For $\gamma = 0$ the setup above reduces to the risk-neutral model with learning studied in section 3. For $\gamma > 0$ the setup is analogous to that under risk neutrality, except that 1. the beliefs β now have an interpretation as risk-adjusted stock price growth rather than simple stock price growth; 2. The risk-adjustment factor $(D_{t-2}/D_{t-1})^\gamma$ now enters the belief updating formula (28). Since for sufficiently large γ the variance of realized risk-adjusted stock price growth under RE increases with γ, the latter implies that larger risk aversion is likely to generate more volatility in beliefs and, therefore, of actual prices under learning.[43] This will improve the ability of the learning model to match the moments in the data.

As in the risk-neutral case we need to impose a projection facility to insure that $\beta_t < \delta^{-1}$. To facilitate model calibration, described in the next section, we change the projection facility slightly. The projection facility described before (and standard in the learning literature) introduces a discontinuity in the simulated path. This makes it difficult to search over the parameter space as we will do in the next section. We assume instead that if β_t is too high the PD ratio (or, equivalently, adjusted beliefs) is given by an increasing, smooth function of β_t. The interpretation is either that individuals *downplay* observations that would entail too high a stock return (instead of completely *ignoring* these observations, as in the standard projection facility), or that a government agency intervenes if stock prices are too high and brings them down, but not exactly to the value of last period's PD ratio. This continuous projection facility is, in a way, more natural, and it insures differentiability of the solution with respect to parameter values, so that numerical searches over parameters that we perform in the next two sections will be better behaved. Details are described in appendix A.6.3. As before, the projection facility insures that the PD ratio will never exceed a value of 500.

Finally, we show that beliefs continue to display momentum and mean-reversion, similar to the case with risk-neutrality. Using equations (27), (26), and the fact that $E_t\left(D_t^\gamma D_{t+1}^{1-\gamma}\right) = \beta^{RE}\, D_t$ shows that stock prices under learn-

[43] The variance of risk adjusted stock price growth under rational expectations is

$$VAR\left(\left(\frac{D_{t-2}}{D_{t-1}}\right)^\gamma \frac{P_{t-1}^{RE}}{P_{t-2}^{RE}}\right) = a^{2(1-\gamma)} e^{(-\gamma)(1-\gamma)\frac{s^2}{2}} (e^{(1-\gamma)^2 s^2} - 1)$$

This variance reaches a minimum for $\gamma = 1$ and it increases with γ for $\gamma \geq 1$.

ing are given by

$$P_t = \frac{\delta\beta^{RE}}{1-\delta\beta_t} D_t \tag{29}$$

$$\frac{P_t}{P_{t-1}} = \left(1 + \frac{\delta\ \Delta\beta_t}{1-\delta\beta_t}\right)\ a\varepsilon_t \tag{30}$$

From equations (28) and (30) follows that the expected dynamics of beliefs in the risk averse model can be described by

$$E_{t-1}\Delta\beta_{t+1} = \frac{1}{\alpha_{t+1}}\left(T(\beta_t, \Delta\beta_t) - \beta_{t-1}\right) \tag{31}$$

where

$$T(\beta_t, \Delta\beta_t) \equiv \beta^{RE} + \frac{\beta^{RE}\delta\ \Delta\beta_t}{1-\delta\beta_t} \tag{32}$$

The updating equation (31) has the same structure as equation (22) and the T-map (32) is identical to (20), which has been studied before only with β^{RE} in place of a. The implications regarding momentum and mean reversion from section 3.1 thus directly apply to the expected belief dynamics in the model with risk-aversion.

We conclude that, qualitatively, the main features of the model under learning are likely to remain after risk aversion is introduced, but that the model now has an even larger chance to generate high volatility.

4.2 Baseline calibration procedure

This section describes and discusses our preferred calibration procedure. Recall that the parameter vector of our baseline learning model is $\theta \equiv (\delta, \gamma, 1/\alpha_1, a, s)$, where δ is the discount factor, γ the coefficient of relative risk aversion, α_1 the agents' initial confidence in the rational expectations value, and a and s the mean and standard deviation of dividend growth, respectively.

We choose the parameters (a, s) to match the mean and standard deviation of dividend growth in the data, as in equation (25). Since it is our interest to show that the model can match the volatility of stock prices for low levels of risk aversion we fix $\gamma = 5$.

This leaves us with *two* free parameters $(\delta, 1/\alpha_1)$ and *eight* remaining asset price statistics from table 1

$$\widehat{\mathcal{S}}' \equiv \left(\widehat{E}(r^s), \widehat{E}(PD), \widehat{\sigma}_{r^s}, \widehat{\sigma}_{PD}, \widehat{\rho}_{PD,-1}, \widehat{c}_2^5, \widehat{R}_5^2, \widehat{E}(r^b)\right)$$

where hatted variables indicate observed values as in the last column of Table 1. As discussed in detail in section 2, these statistics quantitatively capture the asset pricing observations we seek to explain. Our aim is to show that there are parameter values that make the model consistent with these moments.[44]

We could have proceeded further by fixing δ and/or $1/\alpha_1$ to match some additional moments exactly and use the remaining moments to test the model. This is the usual practice in papers using calibration. Yet, many of these moments have a rather large standard deviation in the data, see the column labeled "US Data std" in table 4 below. Related to this, the value of the moments can vary a lot depending on the precise sample period used. For example, the mean of the PD ratio would be estimated to be 99 using data up to 1996. Therefore, matching any of these moments exactly appears arbitrary, one obtains rather different parameter values depending on which moment is chosen for calibration and the sample period. For these reasons, we depart from the usual calibration practice and choose the values for $(\delta, 1/\alpha_1)$ so as to fit all eight moments in the vector $\widehat{\mathcal{S}}$ as well as possible. Of course, it is a challenging task for the model to match eight moments with just two parameters.

As in standard calibration exercises, we measure the goodness-of-fit using the t-ratios

$$\frac{\widehat{\mathcal{S}}_i - \widetilde{\mathcal{S}}_i(\theta^c)}{\widehat{\sigma}_{\mathcal{S}_i}} \tag{33}$$

where $\widehat{\mathcal{S}}_i$ denotes the i-th sample moment observed in the data, $\widetilde{\mathcal{S}}_i(\theta^c)$ the corresponding moment implied by the model at the calibrated parameter values θ^c, and $\widehat{\sigma}_{\mathcal{S}_i}$ the estimated standard deviation of the moment. As in standard calibration exercises, we conclude that the model's fit is satisfactory if the t-ratios are less than, say, two or three in absolute value. We choose the values for $(\delta, 1/\alpha_1)$ that minimize the sum of squared t-ratios, where the sum is over all eight moments. Therefore, moments with a larger standard deviation receive less weight and are matched less precisely. Notice that the calibration result is invariant to a potential rescaling of the moments. The details of the procedure are defined and explained in appendix A.6.

In the calibration literature it is standard to set the estimate of the standard deviation of the moments ($\widehat{\sigma}_{\mathcal{S}_i}$ in equation (33)) equal to the *model implied* standard deviation of the considered moment. This practice has a number of problems. First, it gives an incentive to the researcher to generate models with high standard deviations, i.e., unsharp predictions, as these appear to improve model fit because they artificially increase the denominator of the t-ratio; hence the criterion of fit varies across models and a model with high variability will artificially appear to be better. Second, to increase comparability with Campbell and Cochrane (1999) we choose a model with a constant risk-free rate, so that the model-implied standard deviation is $\widehat{\sigma}_{E(r^b)} = 0$. The above procedure would then require to match the average risk free rate exactly, not because the data suggests that this moment is known very precisely (it is not, see table 4), but

[44] Strictly speaking some of the elements of $\mathcal{S}$ are not 'moments', i.e., they are not a sample average of some function of the data. The R-square coefficient, for example, is a highly non-linear function of moments, rather than being a moment itself. This generates some slight technical problems discussed in appendix A.5. To be precise we should refer to $\widehat{\mathcal{S}}$ as 'statistics', but for simplicity we will proceed by refering to $\widehat{\mathcal{S}}$ as 'moments '.

only as a result of the modelling strategy.

To avoid all these problems we prefer to find an estimate of the standard deviation of each statistic $\widehat{\sigma}_{\mathcal{S}_i}$ that is based purely on data sources. This has the additional advantage that $\widehat{\sigma}_{\mathcal{S}_i}$ is constant across alternative models and thereby allows for model comparisons in a meaningful way. We show in appendix A.6 how to obtain consistent estimates of these standard deviations from the data. With these estimates we use these resulting t-ratios as our measure of 'fit' for each sample statistic and claim to have a good fit if this ratio is below two or three.

The procedure just described is in some ways close to estimation by the method of simulated moments (MSM), and using the t-ratios as measures of fit may appear like using test statistics. In appendix A.6 we describe how this interpretation would be correct under some additional assumptions and appealing to asymptotic theory, but we do not wish to interpret our procedure as a formal econometric method.[45]

Finally, since many economists feel uncomfortable with discount factors larger than 1, we restrict the search to $\delta \leq 1$. We relax this constraint in section 6.

In summary, we think of the method just described as a way to pick the parameters $(\delta, 1/\alpha_1)$ in a systematic way, such that the model has a good chance to meet the data, but where the model could also easily be rejected since there are many more moments than parameters, and where the fit criterion is analogous to the one used in the literature on MSM and calibration.

5 Quantitative results

We now evaluate the quantitative performance of the baseline learning model (using OLS and $\gamma = 5$) using the baseline calibration approach described in the previous section.

Our results are summarized in table 4 below. The second and third column of the table report, respectively, the asset pricing moments from the data that we seek to match and the standard deviation for each moment estimated with data sources alone using the procedure described in appendix A.6. Using the notation of the previous section the second column reports $\widehat{\mathcal{S}}_i$ and the third column $\widehat{\sigma}_{\mathcal{S}_i}$. The table shows that some of the moments are estimated rather imprecisely.

The calibrated parameters values of the learning model are reported at the bottom of the table. Notice that the gain parameter $1/\alpha_1$ is small, reflecting the tendency of the data to give large (but less than full) weight to the RE prior about stock price growth. As has been explained before, high values of $1/\alpha_1$

[45] This is because the distribution of the parameters and test statistics for these formal estimation methods relies on asymptotics, but asymptotically our baseline least squares learning model is indistinguishable from RE. Therefore, one would have to rely on short-sample distribution of statistics. Developing such distributions is well beyond the scope of this paper. This is less of a problem in the constant gain exercises below, since the model does not converge, so it is possible to use MSM for this learning algorithm if one relies on asymptotics.

would cause beliefs to be very volatile and give rise to too much volatility, this is why the calibration procedure chooses a low value for $1/\alpha_1$. The calibrated gain reported in the table implies that when updating beliefs in the initial period, the RE prior receives a weight of approximately 98.5% and the first quarterly observation a weight of about 1.5%. The discount factor is quite high.

Statistics	US Data	std	Model (OLS)	t-ratio
$E(r^s)$	2.41	0.45	2.41	0.01
$E(r^b)$	0.18	0.23	0.48	-1.30
$E(PD)$	113.20	15.15	95.93	1.14
σ_{r^s}	11.65	2.88	13.21	-0.54
σ_{PD}	52.98	16.53	62.19	-0.56
$\rho_{PD,-1}$	0.92	0.02	0.94	-1.20
c_5^2	-0.0048	0.002	-0.0067	0.92
R_5^2	0.1986	0.083	0.3012	-1.24
Parameters:	$\delta = .999, \quad 1/\alpha_1 = 0.015$			

Table 4: Moments and parameters. Baseline model and baseline calibration

The fourth column in table 4 reports the moments implied by the calibrated learning model and the fifth column the corresponding t-ratios. The learning model performs remarkably well. In particular, the model with risk aversion maintains the high variability and serial correlation of the PD ratio and the variability of stock returns, as in section 3. In addition, it now succeeds in matching the mean of the PD ratio and it also matches the equity premium quite well. As discussed in section 2 it is natural that the excess return regressions can be explained reasonably well once the serial correlation of the PD is matched.

Clearly, the point estimate of some model moments does not match exactly the observed moment in the data, but this tends to occur for moments that, in the short sample, have a large variance. This is shown in the last column of table 4 which reports the goodness-of-fit measures (t-ratios) for each considered moment. The t-ratios are all well below two and thus well within what is a 95% confidence interval, if this was a formal econometric test. Notice in particular that the calibration procedure chooses a value of δ that implies a risk-free interest rate that is more than twice as large as the point estimate in the data. Since the standard deviation of $\widehat{E}(r^b)$, reported in the third column of Table 4, is fairly large, one nevertheless obtains a low t-ratio.

In summary, the results of table 4 show that introducing learning substantially improves the fit of the model relative to the case with RE and is overall very successful in quantitatively accounting for the empirical evidence described in section 2. We find this result remarkable, given that we used the simplest version of the asset pricing model and combined it with the simplest available learning mechanism.

6 Robustness

This section shows that the quantitative performance of the model is robust to a number of extensions. We start by exploring alternative learning schemes, then we consider a different model, and finally we discuss alternative calibration procedures.

Learning about dividends. In the baseline model we assume agents know the conditional expectation of dividends. This was done to simplify the exposition and because learning about dividends has been considered in previous papers.[46] Since it may appear inconsistent to assume that agents know the dividend growth process but do not know how to forecast stock prices, we consider a model where agents learn simultaneously about dividend growth and stock price growth. In appendix A.5 we lay out the model and show that, while the analysis is more involved, the basic results do not change. Table 5 below shows the quantitative results with learning about dividends using the baseline calibration procedure described in section 5. It shows that introducing dividend learning does not lead to significant changes.

Constant gain learning. An undesirable feature of the OLS learning scheme is that volatility of stock prices decreases over time, which may seem counterfactual. Therefore, we depart from OLS and introduce a learning scheme with a constant gain, where the weight on the forecast error in the learning scheme is constant: $\alpha_t = \alpha_1$ for all t.[47] Beliefs then respond to forecast errors in the same way throughout the sample and stock price volatility does not decrease at the end of the sample period. As discussed extensively in the learning literature, such a learning scheme improves forecasting performance when there are changes in the environment. Agents who suspect the existence of switches in the average growth rate of prices or fundamentals, for example, may be reasonably expected to use a constant gain learning scheme. Table 5 reports the quantitative results for the constant gain model using the baseline calibration approach. For obvious reasons, stock prices are now more volatile, even if the initial gain is substantially lower than in the baseline case. Overall, the fit of the model is very good.[48] Figure 3 depicts three typical realizations of the PD ratio using the calibrated parameters. It illustrates that, similar to the actual data, stock prices in the model have a tendency to display sustained price increases that are followed by rather sharp price reductions.

[46] E.g., Timmermann (1993, 1996).

[47] The derivations for this model are as in section 4 and require only changing the evolution of α.

[48] We do not use constant gain as our main learning scheme because β_t does then not converge, i.e., we loose asymptotic rationality. Nevertheless, in a setup where there are trend switches in fundamentals, constant gain agents' forecasts will perform better than OLS. We leave this for future research.

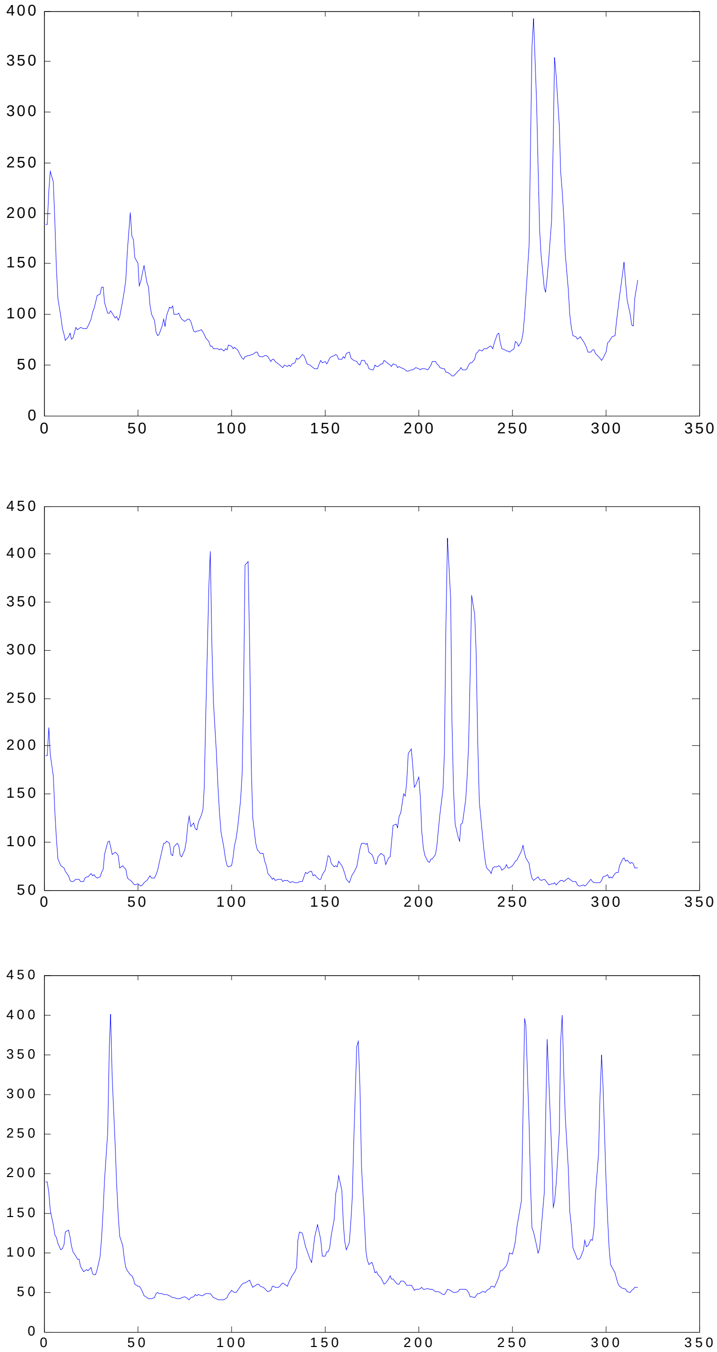

Figure 3: Simulated PD ratio from the estimated constant gain model (table 5)

Switching weights. We now introduce a learning model in which the gain switches over time, as in Marcet and Nicolini (2003).[49] The idea is to combine constant gain with OLS, using the former in periods where a large forecast error occurs and the latter when the forecast error is low. We report the quantitative results in Table 5, which are very similar to those with pure constant gain learning. The latter occurred because the model was frequently in 'constant gain mode'.

Statistic	US Data	Learning on Div.	t-ratio	Constant gain	t-ratio	Switching weights	t-ratio
$E(r^s)$	2.41	2.41	0.00	2.26	0.34	2.25	0.36
$E(r^b)$	0.18	0.48	-1.29	0.44	-1.11	0.44	-1.11
$E(PD)$	113.20	96.17	1.12	109.82	0.22	110.00	0.21
σ_{r^s}	11.65	13.23	-0.55	14.55	-1.00	14.51	-0.99
σ_{PD}	52.98	62.40	-0.57	74.60	-1.31	74.50	-1.30
$\rho_{PD,-1}$	0.92	0.94	-1.22	0.94	-0.81	0.94	-0.82
c_5^2	-0.0048	-0.0067	0.96	-0.0059	0.5344	-0.0059	0.5308
R_5^2	0.1986	0.2982	-1.20	0.2443	-0.5516	0.2454	-0.5650
Parameters:							
δ		0.999		1		1	
$1/\alpha_1$		0.015		0.00628		0.00626	

Table 5: Robustness, Part I

Consumption data. Throughout the paper we made the simplifying assumption of equating consumption with dividends ($C_t = D_t$). Then we calibrated this process to dividend data because when studying stock price volatility the data on dividends has to be brought out. However, it is well known that consumption growth is much less volatile than dividend growth, so that these two choices are likely to help in explaining volatility and risk premium. Therefore, we now allow for $C \neq D$ and calibrate the volatility of the consumption and dividend processes separately to the data. While the dividends process remains as before, we set

$$\frac{C_{t+1}}{C_t} = a\varepsilon_{t+1}^c \quad \text{for} \quad \ln \varepsilon_t^c \sim ii\mathcal{N}\left(-\frac{s_c^2}{2}; s_c^2\right)$$

The presence of two shocks modifies the equations for the RE version of the model in a well known way and we do not describe it in detail here.[50] We calibrate the consumption process following Campbell and Cochrane (1999) and we set $s_c = \frac{s}{7}$ and $\rho(\varepsilon_t^c, \varepsilon_t) = .2$.[51]

[49] The advantage of using switching gains, relative to using simple constant gain, is that under certain conditions, the learning model converges to RE, i.e., Asymptotic Rationality is preserved, while the learning scheme still reacts quickly if there is a change in the environment.

[50] Obviously, this would require replacing $U'(D_t)$ by $U'(C_t)$ in equation (2).

[51] We take these ratios and values from table 1 in Campbell and Cochrane (1999), which is based on a slightly shorter sample than the one used in this paper.

The quantitative results are reported in table 6 below, which shows that we do not fit the data as well as before and, in particular, for the calibrated parameters we do not match the risk-free rate. Equivalently, one could say that this simulation does not match the risk premium puzzle. We do not wish to make much of this quasi-rejection: first, it was not the objective of this paper to match perfectly all moments, second, the equity premium is not the main focus of this paper, one would have added a number of other features in the model if the objective would be to fit the risk premium. Furthermore, to a rational expectations fundamentalist, who would dismiss models of learning from the outset as being "non rigorous because they can always match any data", this shows that this is not always the case. One way to improve the fit of this model is considered below when we relax the constraint on δ.

Full weighting matrix from method of simulated moments We now investigate the robustness of our findings to changes in the calibration procedure. In an econometric MSM setup, one would have to find the parameters that minimize a quadratic form with a certain optimal weighting matrix, while our baseline calibration amounted to using a diagonal weighting matrix with $\widehat{\sigma}_{\mathcal{S}_i}^{-2}$ in the diagonal. Therefore, a more econometrically-oriented reader will think that the objective function to be minimized in the baseline calibration is not justified. For details see the discussion around equation (57) in appendix A.6. Table 6 shows the results when we use the optimal weighting matrix in MSM. We still obtain a good fit, the estimated confidence in initial beliefs is high. Some moments are matched less well, for example, the serial correlation of the PD ratio. This is natural since the weighting matrix does not bring down the t-ratios per se.

Statistics	US Data	$C \neq D$		OLS, Full matrix	
			t-ratio		t-ratio
$E(r^s)$	2.41	2.36	0.12	2.12	0.64
$E(r^b)$	0.18	1.76	-6.91	0.44	-1.11
$E(PD)$	113.20	63.56	3.28	102.43	0.71
σ_{r^s}	11.65	8.42	1.12	11.88	-0.08
σ_{PD}	52.98	30.14	1.38	61.07	-0.49
$\rho_{PD,-1}$	0.92	0.91	0.49	0.96	-1.94
c_5^2	-0.0048	-0.0073	1.2410	-0.0060	0.6207
R_5^2	0.1986	0.2641	-0.7911	0.3322	-1.6127
Parameters:					
δ		1		1	
$1/\alpha_1$		0.0178		0.0128	

Table 6: Robustness, Part II

Relaxing constraint on δ The attentive reader will have noticed that the constraint $\delta \leq 1$ of the baseline calibration is binding in many of the cases considered in tables 5 and 6. It turns out that this restriction is unnecessary because risk aversion causes agents to discount future dividends more heavily. As a result, the discounted sum of dividends can be finite even if $\delta > 1$.[52] Table 7 below shows how the model improves when δ is estimated without this constraint and the only constraints are that a rational expectations PD exists and it is below the price implied by β^L, where the differentiable projection facility starts to operate.[53] To save on space we only report the constant gain and $C \neq D$ models. Obviously, the fit of the model improves. In the case of constant gain, which already performed very well, there is not much room for improvement. The $C \neq D$ model now sustains a lower interest rate, with a t-ratio close to three, so that the equity premium is now much larger. We conclude that this model fits the data well.

Model-generated standard deviations As a final exercise we demonstrate the robustness of our finding to using t-ratios based on model-generated standard deviations. This is the preferred approach in most calibration exercises. Again, the fit of the model is quite good.[54]

	US Data	$C \neq D$		Constant gain		Model $\widehat{\sigma}_{\mathcal{S}_i}$	
Statistics			t-ratio		t-ratio		t-ratio
$E(r^s)$	2.41	2.01	0.89	2.26	0.34	2.28	0.52
$E(r^b)$	0.18	0.84	-2.89	0.31	-0.55	0.31	—
$E(PD)$	113.20	112.85	0.02	110.46	0.18	111.05	0.12
σ_{r^s}	11.65	10.43	0.42	14.77	-1.08	16.53	-1.40
σ_{PD}	52.98	61.16	-0.49	75.41	-1.36	77.04	-2.10
$\rho_{PD,-1}$	0.92	0.95	-1.43	0.94	-0.84	0.94	-0.77
c_5^2	-0.0048	-0.0089	2.0440	-0.0059	0.5622	-0.0061	1.5195
R_5^2	0.1986	0.2397	-0.4966	0.2412	-0.5151	0.2306	-0.6420
Parameters:							
δ		1.00906		1.000375		1.0013	
$1/\alpha_1$		0.0244		0.0063		0.0065	

Table 7: Robustness, Part III, δ unrestricted

[52] More precisely, in the context of the above simple model, it is obvious from (4) and (5) that all that is needed in order to have a finite price under RE is that $\delta < \left(\beta^{RE}\right)^{-1}$, and since risk aversion can bring β^{RE} below 1 this allows for $\delta > 1$. For a discussion, see Kocherlakota (1990).

[53] See appendix A.6.3 for details.

[54] For this case δ is chosen to match $\widehat{E}\left(r^b\right)$ as closely as possible subject to the constraint $PD^{RE} < PD^L$ where PD^L is the price dividend ratio implied by β^L, see appendix A.6.3 for details. It turns out that in this column the constraint is binding, so that one cannot match the real interest rate exactly.

The previous robustness exercises allowed for deviations from the baseline model and calibration. We found the results to be very robust and that the model continues to be able to explain the moments surprisingly well.

7 Conclusions and Outlook

A one parameter learning extension of a very simple asset pricing model strongly improves the ability of the model to quantitatively account for a number of asset pricing facts, even with moderate degrees of risk aversion. This outcome is remarkable, given the difficulties documented in the empirical asset pricing literature in accounting for these facts. The difficulties of rational expectations models suggest that learning processes may be more relevant for explaining empirical phenomena than previously thought.

While we relax the assumption of rational expectations, the learning scheme used here is a small deviation from full rationality. Introducing learning in the model amounts to adding only one new parameter in a basic model. It is surprising, therefore, that such a large improvement in the fit of the data can be achieved. Indeed, it seems that the most convincing case for models of learning can be made by explaining facts that appear 'puzzling' from the rational expectations viewpoint, as we attempt to do in this paper.

The simple setup presented in this paper could be extended in a number of interesting ways and also applied to study other substantive questions on stock prices. One avenue that we currently explore is to ask whether learning processes can account also for the otherwise puzzling behavior of exchange rates. Clearly, the ability of simple models of learning to explain puzzling empirical phenomena in more than one market would further increase confidence in that learning-induced small deviations from rationality are indeed economically relevant.

A Appendix

A.1 Data Sources

Our data is for the United States and has been downloaded from 'The Global Financial Database' (http://www.globalfinancialdata.com). The period covered is 1925:4-2005:4. For the subperiod 1925:4-1998:4 our data set corresponds very closely to Campbell's (2003) data (http://kuznets.fas.harvard.edu/~campbell/data.html).

In the calibration part of the paper we use moments that are based on the same number of observations. Since we seek to match the return predictability evidence at the five year horizon (c_5^2 and R_5^2) we can only use data points up to 2000:4. For consistency the effective sample end for all other moments reported in table 1 has been shortened by five years to 2000:4. In addition, due to the seasonal adjustment procedure for dividends described below and the way we compute the standard errors for the moments described in appendix A.6, the effective starting date was 1927:2.

To obtain real values, nominal variables have been deflated using the 'USA BLS Consumer Price Index' (Global Fin code 'CPUSAM'). The monthly price series has been transformed into a quarterly series by taking the index value of the last month of the considered quarter.

The nominal stock price series is the 'SP 500 Composite Price Index (w/GFD extension)' (Global Fin code '_SPXD'). The weekly (up to the end of 1927) and daily series has been transformed into quarterly data by taking the index value of the last week/day of the considered quarter. Moreover, the series has been normalized to 100 in 1925:4.

As nominal interest rate we use the '90 Days T-Bills Secondary Market' (Global Fin code 'ITUSA3SD'). The monthly (up to the end of 1933), weekly (1934-end of 1953), and daily series has been transformed into a quarterly series using the interest rate corresponding to the last month/week/day of the considered quarter and is expressed in quarterly rates, i.e., not annualized.

Nominal dividends have been computed as follows

$$D_t = \left(\frac{I^D(t)/I^D(t-1)}{I^{ND}(t)/I^{ND}(t-1)} - 1 \right) I^{ND}(t)$$

where I^{ND} denotes the 'SP 500 Composite Price Index (w/GFD extension)' described above and I^D is the 'SP 500 Total Return Index (w/GFD extension)' (Global Fin code '_SPXTRD '). We first computed monthly dividends and then quarterly dividends by adding up the monthly series. Following Campbell (2003), dividends have been deseasonalized by taking averages of the actual dividend payments over the current and preceding three quarters.

A.2 OLG foundations

This appendix derives the asset pricing equations (9) and (26) analyzed in the main text from an overlapping generations (OLG) model. Specifically, the appendix shows that these equations are first order conditions implied by utility maximizing agents that hold a consistent (but not necessarily rational) set of beliefs. It also shows that - unless investors possess a considerable amount of additional information - agents' beliefs about prices fail to be determined by their beliefs about the dividend process.

Consider an OLG model with constant population and cohorts living for two periods. Each cohort consists of a unit mass of identical representative agents. Young agents own one unit of productive time and inelastically supply it to work. Old agents do not work but consume their savings. Agents' preferences are given by

$$\frac{X_t^{1-\gamma}}{1-\gamma} + \delta \frac{Y_{t+1}^{1-\gamma}}{1-\gamma} \tag{34}$$

where X denotes consumption when young and Y consumption when old. There is a unit mass of firms (Lucas trees) producing the consumption good. Each agent of the initial old generation owns one of those firms. Every period, old agents operate the firm by hiring labor, producing, paying wages and consuming the profits/dividends. In addition, they sell the firm to the new generation. All markets are competitive. Firms' production function is

$$A_t N_t^{\alpha}$$

with $\alpha \in (0,1)$ and N_t denoting labor input. There is no entry of firms (trees) and due to decreasing returns there will be profits/dividends in equilibrium. Total factor productivity follows

$$A_{t+1} = a A_t \varepsilon_{t+1}$$

Investors face the following budget constraint

$$X_t + P_t^b B_t + S_t P_t = W_t + F_t \tag{35}$$

$$Y_{t+1} = B_t + S_t(P_{t+1} + D_{t+1}) + G_{t+1} \tag{36}$$

where S_t denotes the number of stocks/firms owned by the consumer, B_t real bond holdings, P_t the real (ex-dividend) price of the stock, P_t^b the real price of bonds, W_t the real wage/labor income, and F_t and G_{t+1} additional exogenous endowments of the consumption good. For simplicity, we assume

$$F_t = \phi A_t$$
$$G_{t+1} = \phi A_{t+1}$$

for some $\phi \geq 0$.[55] Households maximize the expected value of (34) subject to (35) and (36), given their beliefs about the future values of $(P_{t+1}, D_{t+1}, G_{t+1})$. Since we do not require these beliefs to be rational, we denote the expectations operator induced by these beliefs by $\widetilde{E}_t\left[\cdot\right]$.

We now derive the optimality and market clearing conditions and show that these imply the asset pricing equations analyzed in the main text. Profit maximization by firms implies

$$\alpha A_t N_t^{\alpha-1} = W_t$$

and the household's optimality conditions are given by

$$P_t^b x_t^{-\sigma} = \delta \widetilde{E}_t\left(y_{t+1}^{-\sigma}\right) \tag{37}$$

$$P_t x_t^{-\sigma} = \delta \widetilde{E}_t\left(y_{t+1}^{-\sigma}\left(P_{t+1} + D_{t+1}\right)\right) \tag{38}$$

Clearing of bond and stock markets requires

$$B_t = 0$$
$$S_t = 1$$

[55] These endowments could be generated by additional inelastic labor supply and the presence of firms that operate the constant returns to scale technology $\alpha A_t N_t$. Since these firms generate zero profits, one can abstract from ownership.

while clearing of labor and goods markets implies

$$N_t = 1$$
$$X_t + Y_t = A_t + F_t + G_t$$

Labor market clearing and firm optimality deliver

$$\alpha A_t = W_t \tag{39}$$

showing that profits/dividends are

$$D_t = (1-\alpha)A_t \tag{40}$$

Dividends thus follow the process assumed in the main text, provided the initial productivity level is appropriately scaled to match the initial level of dividends.

Assuming ϕ sufficiently large, i.e., that dividends make up only a small share of old age consumption, and using the budget constraints, the market clearing conditions, and equations (39) and (40), allows to make the following approximation

$$\frac{Y_{t+1}}{X_t} = \frac{A_{t+1}}{A_t} = \frac{D_{t+1}}{D_t}$$

The asset pricing equation (38) then simplifies to

$$P_t = \delta \widetilde{E}_t \left(\left(\frac{D_{t+1}}{D_t} \right)^{-\sigma} (P_{t+1} + D_{t+1}) \right) \tag{41}$$

which is identical to equation (26) in the main text and equal to (9) for $\sigma = 0$. Except for equation (37) which prices short-term bonds, equation (41) is the only restriction implied by household optimality. Therefore, individual rationality does not imply that beliefs about the dividend process impose restrictions on agents' beliefs about the price process, and in particular agents' beliefs about price need not be consistent with their beliefs about the discounted sum of dividends. To obtain the latter, investors have to know a lot more. Specifically, they need to know

- that assets in future periods are priced by an equation corresponding to equation (41),
- that they can safely apply the law of iterated expectations to formulate expectations about the expectations of future generations, i.e., $\widetilde{E}_t \left[\widetilde{E}_{t+j} [\cdot] \right] = \widetilde{E}_t [\cdot]$ for all $j > 0$, and
- that the limiting condition $\lim_{j\to\infty} \widetilde{E}_t \left[\delta^j \left(\frac{D_{t+j}}{D_{t+j-1}} \right)^{-\sigma} P_{t+j} \right] = 0$ is valid.

Knowing the previous allows agents to iterate on (41) so as to derive

$$P_t = \widetilde{E}_t \left(\sum_{j=1}^{\infty} \delta^j \left(\frac{D_{t+j}}{D_{t+j-1}} \right)^{-\sigma} D_{t+j} \right)$$

which shows that beliefs about the price process are then determined by agents' beliefs about the dividend process.

A.3 Proof of mean reversion

To prove mean reversion for the general learning scheme of (16) we need the following additional technical assumptions on the updating function f_t:

Assumption 1 *There is a* $\overline{\eta} > 0$ *such that* f_t *is differentiable in the interval* $(-\overline{\eta}, \overline{\eta})$ *and, letting*

$$\mathcal{D}_t \equiv \inf_{\overline{\Delta} \in (-\overline{\eta}, \overline{\eta})} \frac{\partial f_t(\overline{\Delta})}{\partial \Delta} \quad ,$$

we have

$$\sum_{t=0}^{\infty} \mathcal{D}_t = \infty$$

This is satisfied by all the learning mechanisms considered in this paper and by most algorithms used in the stochastic control literature. For example, it is guaranteed in the OLS case $\mathcal{D}_t = 1/(t + \alpha_1)$ and constant gain where $\mathcal{D}_t = 1/\alpha_1$ for all t. If the assumption fails and $\sum \mathcal{D}_t < \infty$, then beliefs get 'stuck' away from the fundamental value simply because updating of beliefs ceases to incorporate new information for t large enough. In this case convergence can not occur, and agents make systematic mistakes forever. One could thus claim that algorithms with $\sum \mathcal{D}_t < \infty$ are not satisfactory learning schemes. We also need:

Assumption 2 *a)* $D_t \geq 0$ *for all* t*; b)* $\beta_0 \in \left(0, \beta^U\right)$ *; c) the learning rule satisfies* $f(-z) \geq -z$ *for all* $z \in \left[0, \beta^U\right]$.

Assumptions 2 b) and c) insure that $\beta_t \geq 0$ for all t, i.e., that agents predict future stock prices to be non-negative. This together with 2 a) insures that actual stock prices are indeed always non-negative. Note that OLS and constant gain learning satisfy 2 c) for $\alpha_1 \geq 1$, as is assumed throughout the paper.

Clearly, we have $0 \leq \beta_t \leq \beta^U$ for all t.

We start proving mean reversion for the case $\beta_t > a$. Fix $\eta > 0$ small enough that $\eta < \min(\overline{\eta}, (\beta_t - a)/2)$ where $\overline{\eta}$ is as in assumption 1.

We first prove that there exists a finite $t' \geq t$ such that

$$\Delta\beta_{\widetilde{t}} \geq 0 \text{ for all } \widetilde{t} \text{ such that } t < \widetilde{t} < t', \text{ and} \quad (42)$$

$$\Delta\beta_{t'} < 0 \quad (43)$$

To prove this, choose $\epsilon = \eta\left(1 - \delta\beta^U\right)$. It cannot be that $\Delta\beta_{\widetilde{t}} \geq \boldsymbol{\epsilon}$ for all $\widetilde{t} > t$, since $\epsilon > 0$ and this would contradict the bound $\beta_{\widetilde{t}} \leq \beta^U$. Therefore $\Delta\beta_{\overline{t}} < \boldsymbol{\epsilon}$ for some finite $\overline{t} \geq t$. Take $\overline{t} \geq t$ to be the *first* period where $\Delta\beta_{\overline{t}} < \boldsymbol{\epsilon}$.

There are two possible cases: either *i)* $\Delta\beta_{\overline{t}} < 0$ or *ii)* $\Delta\beta_{\overline{t}} \geq 0$.

In case *i)* we have (42) and (43) hold if we take $t' = \bar{t}$.

In case *ii)* β_t can not decrease between t and $\bar{t}$ so that

$$\beta_{\bar{t}} \geq \beta_t > a + \eta$$

Furthermore, we have

$$\begin{aligned} T(\beta_{\bar{t}}, \Delta\beta_{\bar{t}}) &= a + \frac{\Delta\beta_{\bar{t}}}{1 - \delta\beta_{\bar{t}}} < a + \frac{\boldsymbol{\epsilon}}{1 - \delta\beta_{\bar{t}}} \\ &< a + \frac{\boldsymbol{\epsilon}}{1 - \delta\beta^U} = a + \eta \end{aligned}$$

where the first equality follows from the definition of T, the first inequality uses $\Delta\beta_{\bar{t}} < \boldsymbol{\epsilon}$ and the second inequality that $\beta_t < \beta^U$ and the last equality follows from the choice for $\boldsymbol{\epsilon}$. The previous two relations imply

$$\beta_{\bar{t}} > T(\beta_{\bar{t}}, \Delta\beta_{\bar{t}})$$

This together with (22) and the properties for f_t gives

$$\beta_{\bar{t}} + f_{\bar{t}+1}\left(T(\beta_{\bar{t}}, \Delta\beta_{\bar{t}}) - \beta_{\bar{t}}\right) < \beta_t < \beta^U$$

so that the projection facility does not apply at $\bar{t} + 1$. Therefore

$$\Delta\beta_{\bar{t}+1} = f_{\bar{t}+1}\left(T(\beta_{\bar{t}}, \Delta\beta_{\bar{t}}) - \beta_{\bar{t}}\right) < 0$$

Therefore in case *ii)* we have that (42) and (43) hold for $t' = \bar{t} + 1$.

This shows that (42) and (43) hold for a finite t'. Now we need to show that from then on beliefs decrease and, eventually, they go below $a + \eta$.

First, notice that given any $j \geq 0$, if

$$\begin{aligned} \Delta\beta_{t'+j} &< 0 \quad \text{and} && (44) \\ \beta_{t'+j} &> a + \eta && (45) \end{aligned}$$

then

$$\begin{aligned} \Delta\beta_{t'+j+1} &= f_{t'+j+1}\left(a + \frac{\Delta\beta_{t'+j}}{1 - \delta\beta_{t'+j}} - \beta_{t'+j}\right) < f_{t'+j+1}\left(a - \beta_{t'+j}\right) && (46) \\ &< f_{t'+j+1}(-\eta) \leq -\eta\mathcal{D}_{t'+j+1} \leq 0 && (47) \end{aligned}$$

where the first inequality follows from (44), the second inequality from (45) and the third from the mean value theorem and $\mathcal{D}_{t'+j+1} \geq 0$. Assume, towards a contradiction, that (45) holds for all $j \geq 0$. Since (44) holds for $j = 0$, it follows by induction that $\Delta\beta_{t'+j} \leq 0$ for all $j \geq 0$ and, therefore, that (47) would hold for all $j \geq 0$. This would imply

$$\beta_{t'+j} = \sum_{i=1}^{j} \Delta\beta_{t'+i} + \beta_{t'} \leq -\eta \sum_{i=1}^{j} \mathcal{D}_{t'+i} + \beta_{t'}$$

for all $j > 0$. Assumption 1 above would then imply $\beta_t \to -\infty$ showing that (45) can not hold for all j. Therefore there is a finite j such that $\beta_{t'+j}$ will go below $a + \eta$ and β is decreasing from t' until it goes below $a + \eta$.

For the case $\beta_t < a - \eta$ we need to make the additional assumption that $\beta^U > a$. Then, choosing $\boldsymbol{\epsilon} = \eta$ we can use a symmetric argument to make the proof.

A.4 Details on the phase diagram

The second order difference equation (22) describing the deterministic evolution of beliefs allows to construct non-linear first-order learning dynamics in the (β_t, β_{t-1}) plane. For clarity, we define $x_t' \equiv (x_{1,t}, x_{2,t}) \equiv (\beta_t, \beta_{t-1})$, whose dynamics are given by

$$x_{t+1} = \begin{pmatrix} x_{1,t} + f_{t+1}\left(a + \frac{a\delta(x_{1,t}-x_{2,t})}{1-\delta x_{1,t}} - x_{1,t}\right) \\ x_{1,t} \end{pmatrix}$$

The points in the phase diagram where there is no change are the following: we have $\Delta x_2 = 0$ at points $x_1 = x_2$, so that the 45^o line gives the point of no change in x_2, and $\Delta x_2 > 0$ above this line. We have $\Delta x_1 = 0$ for $x_2 = \frac{1}{\delta} - \frac{x_1(1-\delta x_1)}{a\delta}$, this is the curve labelled "$\beta_{t+1} = \beta_t$" in Figure 2 and $\Delta x_1 > 0$ below this curve. So the zeroes for Δx_1 and Δx_2 intersect are at $x_1 = x_2 = a$ which is the REE and, interestingly, at $x_1 = x_2 = \delta^{-1}$ which is the limit of rational bubble equilibria. These results give rise to the phase diagram shown in figure 2.

A.5 Model with learning about dividends

This section considers agents who learn to forecast future dividends in addition to forecast future price. We make the arguments directly for the general model with risk aversion from section 4. Equation (26) then becomes

$$P_t = \delta \widetilde{E}_t \left(\left(\frac{C_t}{C_{t+1}} \right)^\gamma P_{t+1} \right) + \delta \widetilde{E}_t \left(\frac{D_t^\gamma}{D_{t+1}^{\gamma-1}} \right)$$

Under RE one has

$$\begin{aligned} E_t \left(\frac{D_t^\gamma}{D_{t+1}^{\gamma-1}} \right) &= E_t \left(\frac{D_{t+1}^{1-\gamma}}{D_t^{-\gamma}} \right) = E_t \left(\left(\frac{D_{t+1}}{D_t} \right)^{1-\gamma} \right) D_t \\ &= E_t \left((a\varepsilon)^{1-\gamma} \right) D_t \\ &= \beta^{RE} D_t \end{aligned}$$

This justifies that learning agents will forecast future dividends according to

$$\widetilde{E}_t \left(\frac{D_{t+1}^{1-\gamma}}{D_t^{-\gamma}} \right) = \varphi_t D_t$$

where φ_t is agents's best estimate of $\widetilde{E}_t\left(\left(\frac{D_{t+1}}{D_t}\right)^{1-\gamma}\right)$, which can be interpreted as risk-adjusted dividend-growth. In close analogy to the learning setup for future price we assume that agents' estimate evolves according to

$$\varphi_t = \varphi_{t-1} + \frac{1}{\alpha_t}\left(\left(\frac{D_{t-1}}{D_{t-2}}\right)^{1-\gamma} - \varphi_{t-1}\right) \tag{48}$$

which can be given a proper Bayesian interpretation. In the spirit of allowing for only small deviations from rationality, we assume that the initial belief is correct

$$\varphi_0 = \beta^{RE}.$$

Moreover, the gain sequence α_t is the same as the one used for updating the estimate for β_t. Learning about β_t remains to be described by equation (28). With these assumptions realized price and price growth are given by

$$\begin{aligned} P_t &= \frac{\delta\varphi_t}{1-\delta\beta_t} D_t \\ \frac{P_t}{P_{t-1}} &= \frac{\varphi_t}{\varphi_{t-1}}\left(1 + \frac{\delta\triangle\beta_t}{1-\delta\beta_t}\right) a\varepsilon_t \end{aligned}$$

The map T from perceived to actual expectations of the risk-adjusted price growth $\frac{P_{t+1}}{P_t}\left(\frac{D_t}{D_{t+1}}\right)^{\gamma}$ in this more general model is given by

$$T(\beta_{t+1}, \Delta\beta_{t+1}) \equiv \frac{\varphi_{t+1}}{\varphi_t}\left(\beta^{RE} + \frac{\beta^{RE}\delta\ \Delta\beta_{t+1}}{1-\delta\beta_{t+1}}\right) \tag{49}$$

which differs from (32) only by the factor $\frac{\varphi_{t+1}}{\varphi_t}$. From (48) it is clear that $\frac{\varphi_{t+1}}{\varphi_t}$ evolves exogenously and that $\lim_{t\to\infty} \frac{\varphi_{t+1}}{\varphi_t} = 1$ since $\lim_{t\to\infty}\varphi_t = \beta^{RE}$ and $\alpha_t \to \infty$. Thus, for medium to high values of α_t and initial beliefs not too far from the RE value, the T-maps with and without learning about dividends are very similar.

For the deterministic setting with risk-neutrality considered in section 3, one has $\varphi_t = \varphi_0 = a$ and $\beta^{RE} = a$ so that (49) becomes identical to (20).

A.6 Calibration procedure

This section describes the details of our calibration approach and explains how we estimate the standard deviation of the sample statistics reported in table 4.

Let N be the sample size, $(\mathbf{y}_1, ...\mathbf{y}_N)$ the observed data sample, with $\mathbf{y}_t$ containing m variables. In the text we referred to all data statistics we used as "moments" even though this is not exactly correct (see footnote 44). In this section we have to distinguish between proper moments and functions of these moments, to which we just refer as statistics

We consider the sample statistics $\mathcal{S}(M_N)$ which are a function of the sample moment M_N. Here $\mathcal{S} : R^q \to R^s$ is a *statistic function* that maps moments into the considered statistics. For a given function $h : R^m \to R^q$, the sample moments on which our statistics are based are defined as $M_N \equiv \frac{1}{N}\sum_{t=1}^{N} h(\mathbf{y}_t)$. The explicit expressions for $h(\cdot)$ and $\mathcal{S}(\cdot)$ for our particular application are stated in A.6.1 below.

In the main text we have denoted the observed sample statistics as $\widehat{\mathcal{S}} \equiv \mathcal{S}(M_N)$.

We now explain how we compute the corresponding model statistics for a given model parameterization $\theta \in R^n$. Let ω^s denote a realization of shocks and $(y_1(\theta, \omega^s), ...y_N(\theta, \omega^s))$ the random variables corresponding to a history of length N generated by the model for shock realization ω^s. Furthermore, let

$$M_N(\theta, \omega^s) \equiv \frac{1}{N}\sum_{t=1}^{N} h(y_t(\theta, \omega^s))$$

denote the model moment for realization ω^s and

$$\mathcal{S}(M_N(\theta, \omega^s))$$

the corresponding model statistics for this realization. The model statistics we wish to report are the expected value of the statistic across possible shock realizations and for a sample size N:

$$\widetilde{\mathcal{S}}(\theta) \equiv E\left[\mathcal{S}(M_N(\theta, \omega^s))\right]$$

One can obtain a numerical approximation to the theoretical model statistic $\widetilde{\mathcal{S}}(\theta)$ by averaging (for a given a parameter vector θ) across a large number of simulations of length N the statistics $\mathcal{S}(M_N(\theta, \omega^s))$ implied by each simulation. We report this average in the tables of the main text.

Now that we have explained how to compute statistics in the data and the model, we explain how we calibrate the parameters so as to match the model statistics to the statistics of the data. Let $\widehat{\mathcal{S}}_i = \mathcal{S}_i(M_N)$ denote the i-th statistic from the data and let $\widehat{\sigma}_{\mathcal{S}_i}$ be an estimate for the standard deviation of the i-th statistic. How we obtain $\widehat{\sigma}_{\mathcal{S}_i}$ will be explained in detail below. The baseline parameter choice $\widehat{\theta}_N$ is then found as follows

$$\widehat{\theta}_N \equiv \arg\min_{\theta} \sum_{i=1}^{s} \left(\frac{\widehat{\mathcal{S}}_i - \widetilde{\mathcal{S}}_i(\theta)}{\widehat{\sigma}_{\mathcal{S}_i}}\right)^2 \tag{50}$$

subject to the restrictions on the parameters described in the text, that is, a, s are given by (25), $\delta \leq 1$, and $\gamma = 5$. Our procedure thus tries to match the model statistics as closely as possible to the data statistics, but gives less weight to statistics with a larger standard deviation. Notice that the calibration result is invariant to a rescaling of the variables of interest. Of course, the number of parameters should be less than the number of statistics s. In order to avoid a certain singularity it will be required, in addition, that $s \leq q$.

In order to solve the minimization problem (50) with standard numerical procedures we slightly modify the projection facility described in (16) to insure that the objective function in (50) is continuously differentiable. Appendix A.6.3 describes this in detail.

We now explain how we obtain the estimate for the standard deviation of the i-th statistic $\widehat{\sigma}_{\mathcal{S}_i}$. We start by discussing desirable asymptotic properties of such an estimate and then explain how it has been constructed.

We would like to have an estimator $\widehat{\sigma}_{\mathcal{S}_i}$ that converges to the standard deviation of the statistic sufficiently fast, so that asymptotically the t-ratio has a standard normal distribution. More precisely, assuming stationarity, let $M_0 = E\left[h(y_t(\theta_0, \omega^s))\right]$ denote the theoretical moment at the true parameter value, we require

$$\sqrt{N}\frac{\widehat{\mathcal{S}}_i - \mathcal{S}_i(M_0)}{\widehat{\sigma}_{\mathcal{S}_i}} \to \mathcal{N}(0,1) \quad \text{in distribution} \tag{51}$$

as $N \to \infty$. Once we have such an estimator $\widehat{\sigma}_{\mathcal{S}_i}$, it is justified to interpret t-ratios as goodness of fit measures that should be below two or three in absolute value.

For this purpose we find an estimate for the full covariance matrix $\widehat{\Sigma}_{\mathcal{S},N}$ of model statistics from a sample of N observations such that

$$\widehat{\Sigma}_{\mathcal{S},N} \to \Sigma_{\mathcal{S}} \quad \text{almost surely, for} \tag{52}$$

$$\sqrt{N}\left[\mathcal{S}(M_N) - \mathcal{S}(M_0)\right] \to \mathcal{N}(0, \Sigma_{\mathcal{S}}) \tag{53}$$

as $N \to \infty$. Then, taking $\widehat{\sigma}^2_{\mathcal{S}_i}$ in (50) to be the diagonal entries of $\widehat{\Sigma}_{\mathcal{S},N}$ insures that (51) is satisfied, as required. Now we need to build $\widehat{\Sigma}_{\mathcal{S},N}$ that satisfies (52). For this purpose we first find an expression for $\Sigma_{\mathcal{S}}$ under some additional assumptions.

Assume y to be stationary and ergodic, $\mathcal{S}$ to be continuously differentiable at M_0, and that the matrix

$$S_w \equiv \sum_{j=-\infty}^{\infty} E\left[\left(h(y_t) - M_0\right)\left(h(y_{t-j}) - M_0\right)'\right] \tag{54}$$

is finite. We then have $\Sigma_{\mathcal{S}}$ in (52) given by

$$\Sigma_{\mathcal{S}} = \frac{\partial \mathcal{S}(M_0)}{\partial M'} S_w \frac{\partial \mathcal{S}'(M_0)}{\partial M} \tag{55}$$

This follows from standard arguments: by the mean value theorem

$$\sqrt{N}\left[\mathcal{S}(M_0) - \mathcal{S}(M_N)\right] = \frac{\partial \mathcal{S}(\overline{M}_N)}{\partial M'}\sqrt{N}\left[M_0 - M_N\right] \tag{56}$$

where $\overline{M}_N$ is a certain convex combination of M_N and M_0.[56] Under stationarity and ergodicity of y, we have $M_N \to M_0$ a.s. by the ergodic theorem. Since $\overline{M}_N$

[56] As is well known, a different $\overline{M}_N$ is needed for each row of $\mathcal{S}$ but this issue is inconsequential for the proof and we ignore it here.

is between M_N and M_0, this implies $\overline{M}_N \to M_0$ a.s. and, since $\frac{\partial \mathcal{S}(\cdot)}{\partial M'}$ has been assumed continuous at M_0 we have that

$$\frac{\partial \mathcal{S}(\overline{M}_N)}{\partial M'} \to \frac{\partial \mathcal{S}(M_0)}{\partial M'} \text{ a.s.}$$

From the central limit theorem

$$\sqrt{N}\,(M_0 - M_N) \to \mathcal{N}(0, S_w) \text{ in distribution}$$

Plugging the previous two relationships into (56) shows that if $\Sigma_{\mathcal{S}}$ is given by (55) then (53) holds and, therefore the t-ratio have a standard normal limiting distribution.

Given the expression that we have found for Σ we have that given an estimate $S_{w,N}$ that converges a.s. to S_w we can use as our estimator

$$\widehat{\Sigma}_{\mathcal{S},N} \equiv \frac{\partial \mathcal{S}(M_N)}{\partial M'} S_{w,N} \frac{\partial \mathcal{S}'(M_N)}{\partial M}$$

An explicit expression for $\partial \mathcal{S}(M_N)/\partial M'$ is given in appendix A.6.2. It now only remains to find the estimates $S_{w,N}$ from the data. We follow standard practice and employ the Newey West estimator, which truncates the infinite sum in (54) and weighs the autocovariances in a particular way. This is standard and we do not describe the details here. It is the diagonal terms of $\widehat{\Sigma}_{\mathcal{S},N}$ that we use for the denominator in the t-ratio.

Our baseline procedure for choosing parameter values described above can be thought of as a hybrid between the method of simulated moments (MSM) and calibration. It differs from fully-fledged MSM (described below) because we do not perform any formal estimation, we do not attempt to use an optimal weighting matrix, and because we do not think of this as an exercise in accepting or rejecting the model. Instead, our procedure is simply a way of systematically choosing parameter values that allows us to display the behavior of the model and to interpret the t-ratios as giving a measure of goodness of fit.

We also differ from calibration because we do not pin down each parameter with a given moment and use the remaining moments to test the model. Instead, we let the algorithm find the parameters that best fit the statistics considered. Moreover, in our procedure the standard deviation of the moment $\widehat{\sigma}_{\mathcal{S}_i}$ is computed from the data, we already commented on the advantages of this option in the text, when we discussed the baseline calibration.

In addition to the baseline calibration procedure above, we engage in a robustness exercise, reported under the heading 'full matrix' in table 6, which is a version of MSM. In particular, we choose parameters to solve

$$\widehat{\theta}_N \equiv \arg\min_{\theta} \left[\mathcal{S}(M_N(\theta)) - \mathcal{S}(M_N)\right]' \; \widehat{\Sigma}^{-1}_{\mathcal{S},N} \; \left[\mathcal{S}(M_N(\theta)) - \mathcal{S}(M_N)\right] \tag{57}$$

subject to the constraints in the text. The only difference from standard practice in MSM is that instead of using model moments computed from long simulations of the model we use averages of short run simulations $\mathcal{S}(M_N(\theta))$. This way of fitting the model is less intuitive but generally has the advantage that $\widehat{\Sigma}_{\mathcal{S},N}^{-1}$ is an optimal weighting matrix so the estimate should be closer to the true model parameter if the model was the true one and if asymptotic distribution is to be trusted. One problem we encountered is that $\widehat{\Sigma}_{\mathcal{S},N}$ is nearly singular and it is well known that in this case the weighting matrix in short samples does not produce good results. While the literature suggests ways to address this problem, this is clearly beyond the scope of this paper.

A.6.1 The statistic and moment functions

This section gives explicit expressions for the statistic function $\mathcal{S}(\cdot)$ and the moment functions $h(\cdot)$ introduced in appendix A.6.

The underlying sample moments are

$$M_N \equiv \begin{bmatrix} M_{1,N} \\ . \\ . \\ . \\ M_{9,N} \end{bmatrix} \equiv \frac{1}{N}\sum_{t=1}^{N} h(\mathbf{y}_t)$$

where $h(\cdot)$ and $\mathbf{y}_t$ are defined as

$$h(\mathbf{y}_t) \equiv \begin{bmatrix} r_t^s \\ PD_t \\ (r_t^s)^2 \\ (PD_t)^2 \\ PD_t\ PD_{t-1} \\ r_{t-20}^{s,20} \\ \left(r_{t-20}^{s,20}\right)^2 \\ r_{t-20}^{s,20} PD_{t-20} \\ r_t^b \end{bmatrix}, \qquad \mathbf{y}_t \equiv \begin{bmatrix} PD_t \\ D_t/D_{t-1} \\ PD_{t-1} \\ D_{t-1}/D_{t-2} \\ \vdots \\ PD_{t-19} \\ D_{t-19}/D_{t-20} \\ PD_{t-20} \\ r_t^b \end{bmatrix}$$

where $r_t^{s,20}$ denotes the stock return over 20 quarters, which can be computed using from y_t using $(PD_t, D_t/D_{t-1}, ..., PD_{t-19}, D_{t-19}/D_{t-20})$.

The eight statistics we consider can be expressed as function of the moments:

$$\mathcal{S}(M) \equiv \begin{bmatrix} E(r_t^s) \\ E(PD_t) \\ \sigma_{r_t^s} \\ \sigma_{PD_t} \\ \rho_{PD_t,-1} \\ c_2^5 \\ R_5^2 \\ E(r_t^b) \end{bmatrix} = \begin{bmatrix} M_1 \\ M_2 \\ \sqrt{M_3 - (M_1)^2} \\ \sqrt{M_4 - (M_2)^2} \\ \frac{M_5-(M_2)^2}{M_4-(M_2)^2} \\ c_2^5(M) \\ R_5^2(M) \\ M_9 \end{bmatrix}$$

where the functions $c_2^5(M)$ and $R_5^2(M)$ defining the OLS and R^2 coefficients of the excess returns regressions, respectively, are

$$c^5(M) \equiv \begin{bmatrix} 1 & M_2 \\ M_2 & M_4 \end{bmatrix}^{-1} \begin{bmatrix} M_6 \\ M_8 \end{bmatrix}$$
$$R_5^2(M) \equiv 1 - \frac{M_7 - [M_6, M_8]\, c^5(M)}{M_7 - (M_6)^2}$$

A.6.2 Derivatives of the statistic function

This appendix gives explicit expressions for $\partial \mathcal{S}/\partial M'$ using the statistic function stated in appendix A.6.1. We need this in order to find $\widehat{\Sigma}_{\mathcal{S},N}$. Straightforward but tedious algebra shows

$$\frac{\partial \mathcal{S}_i}{\partial M_j} = 1 \qquad \text{for } (i,j) = (1,1), (2,2), (8,9)$$
$$\frac{\partial \mathcal{S}_i}{\partial M_i} = \frac{1}{2\mathcal{S}_i(M)} \quad \text{for} \quad i = 3, 4$$
$$\frac{\partial \mathcal{S}_i}{\partial M_j} = \frac{-M_j}{\mathcal{S}_i(M)} \quad \text{for } (i,j) = (3,1), (4,2)$$
$$\frac{\partial \mathcal{S}_5}{\partial M_2} = \frac{2M_2(M_5 - M_4)}{(M_4 - M_2^2)^2}, \qquad \frac{\partial \mathcal{S}_5}{\partial M_5} = \frac{1}{M_4 - M_2^2}, \qquad \frac{\partial \mathcal{S}_5}{\partial M_4} = -\frac{M_5 - M_2^2}{(M_4 - M_2^2)^2}$$
$$\frac{\partial \mathcal{S}_6}{\partial M_j} = \frac{\partial c_2^5(M)}{\partial M_j} \quad \text{for } i = 2, 4, 6, 8$$
$$\frac{\partial \mathcal{S}_7}{\partial M_j} = \frac{[M_6, M_8]\, \frac{\partial c^5(M)}{\partial M_j}}{M_7 - M_6^2} \quad \text{for } j = 2, 4$$
$$\frac{\partial \mathcal{S}_7}{\partial M_6} = \frac{\left[c_1^5(M) + [M_6, M_8]\, \frac{\partial c^5(M)}{\partial M_6}\right]\left(M_7 - M_6^2\right) - 2M_6\, [M_6, M_8]\, c^5(M)}{\left(M_7 - M_6^2\right)^2}$$
$$\frac{\partial \mathcal{S}_7}{\partial M_7} = \frac{M_6^2 - [M_6, M_8]\, c^5(M)}{\left(M_7 - M_6^2\right)^2}$$
$$\frac{\partial \mathcal{S}_7}{\partial M_8} = \frac{c_2^5(M) + [M_6, M_8]\, \frac{\partial c^5(M)}{\partial M_8}}{M_7 - M_6^2}$$

Using the formula for the inverse of a 2x2 matrix

$$c^5(M) = \frac{1}{M_4 - M_2^2} \begin{bmatrix} M_4 M_6 - M_2 M_8 \\ M_8 - M_2 M_6 \end{bmatrix}$$

we have

$$\frac{\partial c^5(M)}{\partial M_2} = \frac{1}{M_4 - M_2^2}\left(2M_2 c^5(M) - \begin{bmatrix} M_8 \\ M_6 \end{bmatrix}\right)$$
$$\frac{\partial c^5(M)}{\partial M_4} = \frac{1}{M_4 - M_2^2}\left(-c^5(M) + \begin{bmatrix} M_6 \\ 0 \end{bmatrix}\right)$$
$$\frac{\partial c^5(M)}{\partial M_6} \equiv \frac{1}{M_4 - M_2^2}\begin{bmatrix} M_4 \\ -M_2 \end{bmatrix}$$
$$\frac{\partial c^5(M)}{\partial M_8} \equiv \frac{1}{M_4 - M_2^2}\begin{bmatrix} -M_2 \\ 1 \end{bmatrix}$$

All remaining terms $\partial \mathcal{S}_i / \partial M_j$ not listed above are equal to zero.

A.6.3 Differentiable projection facility

As discussed in the main text, we need to introduce a feature that prevents perceived stock price growth from being higher than δ^{-1}, so as to insure a finite stock price. In addition, it is convenient for our calibration exercises if the learning scheme is a continuous and differentiable function, see the discussion in appendix A.6. The standard projection facility as in (16) causes a simulated series for a given realization $P_t(\theta.\omega^s)$ to be discontinuous in θ, because the price will jump at a parameter value where the facility is exactly binding.

We thus introduce a projection facility that 'phases in' more gradually. We define

$$\beta_t^* = \beta_{t-1} + \frac{1}{\alpha_t}\left[\left(\frac{D_{t-1}}{D_{t-2}}\right)^{-\gamma}\frac{P_{t-1}}{P_{t-2}} - \beta_{t-1}\right] \tag{58}$$

and modify the updating scheme (28) to

$$\beta_t = \begin{cases} \beta_t^* & \text{if } \beta_t^* \leq \beta^L \\ \beta^L + w(\beta_t^* - \beta^L)(\beta^U - \beta^L) & \text{otherwise} \end{cases} \tag{59}$$

where β^U is the upper bound on beliefs, chosen to insure that the implied PD ratio is always less than a certain upper bound $U^{PD} \equiv \frac{\delta a}{1-\delta\beta^U}$, where $\beta^L < \beta^U$ is some arbitrary level of beliefs above which the projection facility starts to operate, and $w(\cdot) : R^+ \to [0,1]$ is a weighting function. Since $w(\beta_t^*)$ is between zero and one this formula insures that the beliefs are below β^U. We further require that w is increasing, $w(0) = 0$ and $w(\infty) = 1$, and we want to insure that the resulting beliefs are continuously differentiable w.r.t. β_t^* at the point β^L.

In particular, we define

$$w(x) = 1 - \frac{\beta^U - \beta^L}{x + \beta^U - \beta^L}.$$

With this weighting function

$$\lim_{\beta_t^* \nearrow \beta^L} \beta_t = \lim_{\beta_t^* \searrow \beta^L} \beta_t = \beta^L$$

$$\lim_{\beta_t^* \searrow \beta^L} \frac{\partial \beta_t}{\partial \beta_t^*} = 1$$

$$\lim_{\beta_t^* \to \infty} \beta_t = \beta^U$$

In our numerical applications we choose β^U so that the implied PD ratio never exceeds $U^{PD} = 500$ and $\beta^L = \delta^{-1} - 2(\delta^{-1} - \beta^U)$, which implies that the dampening effect of the projection facility starts to come into effect for values of the PD ratio above 250.

Figure 4 shows how the standard projection facility operates versus the continuous projection facility proposed in this appendix. It displays the discontinuity introduced by the standard projection facility: once the regular update β_t^* exceeds its upper bound β^U (the upper horizontal line in the graph), β_t will not be updated but remain at its previous level (assumed to be 1.007). The continuous projection facility smoothly dampens the update once β_t^* starts exceeding β^U (the lower horizontal line) with the dampening factor increasing so that β_t will never exceed β^U. The graph also shows that for most β_t^* the projection facility is irrelevant. For this graph $\beta^{RE} = 1.0035$.

A.7 Convergence of least squares to RE

We show global convergence when agents use least squares learning and they have risk aversion as in section 4. The proof shows global convergence, that is, it obtains a stronger result than is usually found in many applications using the associated o.d.e. approach. The proof below is for the standard projection facility.

We assume $D_t \geq 0$ with probability one, which requires that $\varepsilon_t \geq 0$. We also need $\varepsilon_t^{1-\gamma}$ to be bounded, formally we assume existence of some positive $U^\varepsilon < \infty$ such that

$$\text{Prob}(\varepsilon_t^{1-\gamma} < U^\varepsilon) = 1$$

This excludes log-normality (except for the case of log utility) but it still allows for a rather general distribution for ε_t. Obviously, if $\gamma < 1$ this is satisfied if ε_t is bounded above a.s. by a finite constant, and if $\gamma > 1$ this is satisfied if ε_t is bounded away from zero. The highest PD ratio that can be achieved with the projection facility is denoted as $U^{PD} \equiv \frac{\delta\beta^U}{1-\delta\beta^U} < \infty$.

We first show that the projection facility will almost surely cease to be binding after some finite time. In a second step, we prove that β_t converges to β^{RE} from that time onwards.

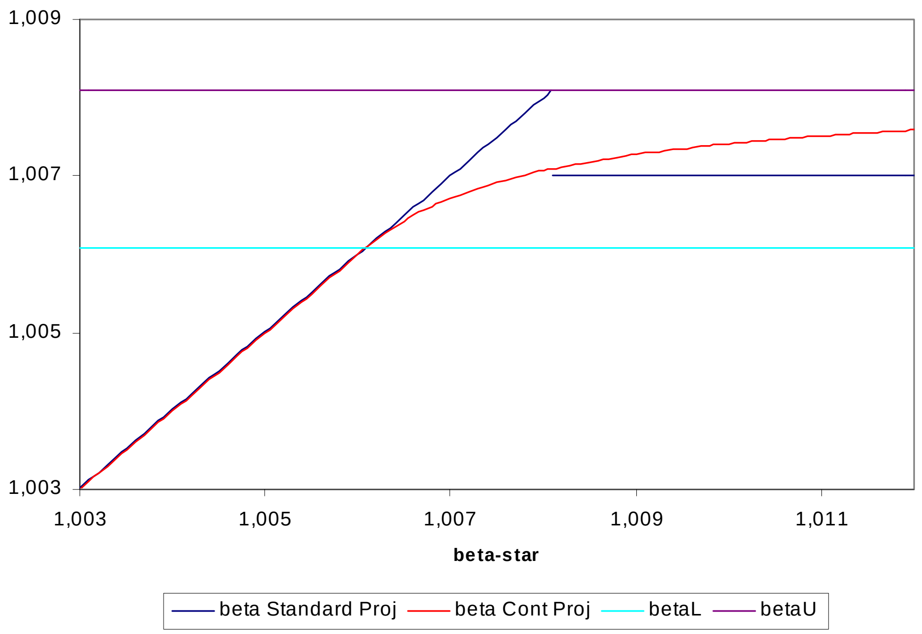

Figure 4: Projection facilities

The standard projection facility implies

$$\Delta\beta_t = \begin{cases} \alpha_t^{-1}\left((a\varepsilon_{t-1})^{-\gamma}\frac{P_{t-1}}{P_{t-2}} - \beta_{t-1}\right) & \text{if } \beta_{t-1} + \alpha_t^{-1}\left((a\varepsilon_{t-1})^{-\gamma}\frac{P_{t-1}}{P_{t-2}} - \beta_{t-1}\right) < \beta^U \\ 0 & \text{otherwise} \end{cases} \tag{60}$$

If the lower equality applies one has $\alpha_t^{-1}(a\varepsilon_{t-1})^{-\gamma}\frac{P_{t-1}}{P_{t-2}} \geq \beta_{t-1} \geq 0$ and this shows the following inequality

$$\beta_t \leq \beta_{t-1} + \alpha_t^{-1}\left((a\varepsilon_{t-1})^{-\gamma}\frac{P_{t-1}}{P_{t-2}} - \beta_{t-1}\right) \tag{61}$$

holds for all t a.s., whether or not the projection facility is binding at t. We also have that

$$\left|\beta_t - \beta_{t-1}\right| \leq \alpha_t^{-1}\left|(a\varepsilon_{t-1})^{-\gamma}\frac{P_{t-1}}{P_{t-2}} - \beta_{t-1}\right| \tag{62}$$

holds for all t a.s., because if $\beta_t < \beta^U$ this holds with equality and if $\beta_{t-1} + \alpha_t^{-1}\left((a\varepsilon_{t-1})^{-\gamma}\frac{P_{t-1}}{P_{t-2}} - \beta_{t-1}\right) \geq \beta^U$ then $\left|\beta_t - \beta_{t-1}\right| = 0$.

Substituting recursively backwards in (61) for past β's delivers the first line in

$$\begin{aligned} \beta_t &\leq \frac{1}{t-1+\alpha_1}\left((\alpha_1-1)\ \beta_0 + \sum_{j=0}^{t-1}(a\varepsilon_j)^{-\gamma}\frac{P_j}{P_{j-1}}\right) \\ &= \underbrace{\frac{t}{t-1+\alpha_1}\left(\frac{(\alpha_1-1)\ \beta_0}{t} + \frac{1}{t}\sum_{j=0}^{t-1}(a\ \varepsilon_j)^{1-\gamma}\right)}_{=T_1} + \underbrace{\frac{1}{t-1+\alpha_1}\left(\sum_{j=0}^{t-1}\frac{\delta\ \Delta\beta_j}{1-\delta\beta_j}(a\varepsilon_j)^{1-\gamma}\right)}_{=T_2} \end{aligned} \tag{63}$$

a.s., where the second line follows from (30). Clearly, $T_1 \to 1(0+E((a\ \varepsilon_j)^{1-\gamma})) = \beta^{RE}$ as $t \to \infty$ a.s. Also, if we can establish $|T_2| \to 0$ a.s. this will show that β_t will eventually be bounded away from its upper bound or, more formally, that $\limsup_{t\to\infty}\beta_t \leq \beta^{RE} < \beta^U$. This is achieved by noting that

$$\begin{aligned} |T_2| &\leq \frac{1}{t-1+\alpha_1}\sum_{j=0}^{t-1}\frac{\delta\,(a\ \varepsilon_j)^{1-\gamma}}{1-\delta\beta_j}\left|\Delta\beta_j\right| \\ &\leq \frac{U^\varepsilon}{t-1+\alpha_1}\sum_{j=0}^{t-1}\frac{a^{1-\gamma}\delta\ \left|\Delta\beta_j\right|}{1-\delta\beta_j} \\ &\leq \frac{U^\varepsilon}{t-1+\alpha_1}\frac{\delta a^{1-\gamma}}{1-\delta\beta^U}\sum_{j=0}^{t-1}\left|\Delta\beta_j\right| \end{aligned} \tag{64}$$

a.s., where the first inequality results from the triangle inequality and the fact that both ε_j and $\frac{1}{1-\delta\beta_j}$ are positive, the second inequality follows from the a.s.

bound on ε_j, and the third inequality from $\beta_t \leq \beta^U$. Next, observe that a.s. for all t

$$(a\varepsilon_t)^{-\gamma} \frac{P_t}{P_{t-1}} = \frac{1-\delta\beta_{t-1}}{1-\delta\beta_t} (a\varepsilon_t)^{1-\gamma} < \frac{(a\varepsilon_t)^{1-\gamma}}{1-\delta\beta_t} < \frac{a^{1-\gamma}U^{\varepsilon}}{1-\delta\beta^U} \tag{65}$$

where the equality follows from (29), the first inequality from $\beta_{t-1} > 0$, and the second inequality from the bounds on ε and β_t. Applying the triangle inequality in the right side of equation (62), using (65) and $\beta_{t-1} < \beta^U$ gives the inequality in

$$\begin{aligned}\frac{1}{t-1+\alpha_1}\sum_{j=0}^{t-1}\left|\Delta\beta_j\right| &\leq \frac{1}{t-1+\alpha_1}\sum_{j=0}^{t-1}\alpha_j^{-1}\left(\frac{a^{1-\gamma}U^{\varepsilon}}{1-\delta\beta^U}+\beta^U\right)\\ &= \left(\frac{a^{1-\gamma}U^{\varepsilon}}{1-\delta\beta^U}+\beta^U\right)\frac{1}{t-1+\alpha_1}\sum_{j=0}^{t-1}\frac{1}{j-1+\alpha_1}\end{aligned} \tag{66}$$

the equality follows from simple algebra. Now, for any $\zeta > 0$

$$t^{-1}\sum_{i=0}^{t} i^{-1} = t^{\zeta-1}\sum_{i=0}^{t} t^{-\zeta} i^{-1} \leq t^{-1+\zeta}\sum_{i=0}^{t} i^{-(1+\zeta)} \to 0 \text{ as } t\to\infty$$

where the convergence follows from the well known fact that the over-harmonic series $\sum_{i=0}^{t} i^{-(1+\zeta)}$ is convergent. This and the fact that the large parenthesis in (66) is finite implies

$$\frac{1}{t-1+\alpha_1}\sum_{j=0}^{t-1}\left|\Delta\beta_j\right| \to 0 \text{ for all } t \text{ a.s.}$$

Then (64) implies that $|T_2| \to 0$ a.s. as $t \to \infty$. Taking the lim sup on both sides of (63), it follows from $T_1 \to \beta^{RE}$ and $|T_2| \to 0$ that

$$\limsup_{t\to\infty} \beta_t \leq \beta^{RE} < \beta^U$$

a.s. The projection facility is thus binding finitely many periods with probability one.

We now proceed with the second step of the proof. Consider for a given realization a finite period $\bar{t}$ where the projection facility is not binding for all $t > \bar{t}$. Then the upper equality in (60) holds for all $t > \bar{t}$ and simple algebra gives

$$\begin{aligned}\beta_t &= \frac{1}{t-\bar{t}+\alpha_{\bar{t}}}\left(\sum_{j=\bar{t}}^{t-1}(a\varepsilon_j)^{-\gamma}\frac{P_j}{P_{j-1}}+\alpha_{\bar{t}}\,\beta_{\bar{t}}\right)\\ &= \frac{t-\bar{t}}{t-\bar{t}+\alpha_{\bar{t}}}\left(\frac{1}{t-\bar{t}}\sum_{j=\bar{t}}^{t-1}(a\varepsilon_j)^{1-\gamma}+\frac{1}{t-\bar{t}}\sum_{j=\bar{t}}^{t-1}\frac{\delta\,\Delta\beta_j}{1-\delta\beta_j}(a\varepsilon_j)^{1-\gamma}+\frac{\alpha_{\bar{t}}}{t-\bar{t}}\,\beta_{\bar{t}}\right)\end{aligned} \tag{67}$$

for all $t > \bar{t}$. Equations (61) and (62) now hold with equality for all $t > \bar{t}$. Similar operations as before then deliver

$$\frac{1}{t-\bar{t}} \sum_{j=\bar{t}}^{t-1} \frac{\delta \, \Delta\beta_j}{1-\delta\beta_j} \left(a\varepsilon_j\right)^{1-\gamma} \to 0$$

a.s. for $t \to \infty$. Finally, taking the limit on both sides of (67) establishes

$$\beta_t \to a^{1-\gamma} E(\varepsilon_t^{1-\gamma}) = \beta^{RE}$$

a.s. as $t \to \infty$.■

References

Abel, A. B. (1990): "Asset Prices under Habit Formation and Catching Up with the Joneses," *American Economic Review*, 80, 38–42.

Adam, K. (2003): "Learning and Equilibrium Selection in a Monetary Overlapping Generations Model with Sticky Prices," *Review of Economic Studies*, 70, 887–908.

Adam, K., A. Marcet, and J. P. Nicolini (2008): "Individual Rationality and Asset Prices," *mimeo.*

Branch, W., and G. W. Evans (2006): "Asset Return Dynamics and Learning," University of Oregon mimeo.

Bray, M. M., and D. M. Kreps (1987): "Rational Learning and Rational Expectations," in *Arrow and the Ascent of Modern Economic Theory*, ed. by G. Feiwel, pp. 597–625. New York University Press, New York.

Brennan, M. J., and Y. Xia (2001): "Stock Price Volatility and Equity Premium," *Journal of Monetary Economics*, 47, 249–283.

Brock, W. A., and C. H. Hommes (1998): "Heterogeneous Beliefs and Routes to Chaos in a Simple Asset Pricing Model," *Journal of Economic Dynamics and Control*, 22, 1235–1274.

Bullard, J., and J. Duffy (2001): "Learning and Excess Volatility," *Macroeconomic Dynamics*, 5, 272–302.

Campbell, J. Y. (2003): "Consumption-Based Asset Pricing," in *Handbook of Economics and Finance*, ed. by G. M. Constantinides, M. Harris, and R. Stulz, pp. 803–887. Elsevier, Amsterdam.

Campbell, J. Y., and J. H. Cochrane (1999): "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, 205–251.

Carceles-Poveda, E., and C. Giannitsarou (2007): "Asset Pricing with Adaptive Learning," *Review of Economic Dynamics (forthcoming).*

Cecchetti, S., P.-S. Lam, and N. C. Mark (2000): "Asset Pricing with Distorted Beliefs: Are Equity Returns Too Good to Be True?," *American Economic Review*, 90, 787–805.

Cochrane, J. H. (2005): *Asset Pricing (Revised Edition).* Princeton University Press, Princeton, USA.

Cogley, T., and T. J. Sargent (2006): "The Market Price of Risk and the Equity Premium: A Legacy of the Great Depression?," *New York University mimeo.*

Evans, G. W., and S. Honkapohja (2001): *Learning and Expectations in Macroeconomics.* Princeton University Press, Princeton.

Kocherlakota, N. (1990): "On the 'Discount' Factor in Growth Economies," *Journal of Monetary Economics*, 25, 43–47.

LeRoy, S. F., and R. Porter (1981): "The Present-Value Relation: Test Based on Implied Variance Bounds," *Econometrica*, 49, 555–574.

Lucas, R. E. (1978): "Asset Prices in an Exchange Economy," *Econometrica*, 46, 1426–1445.

Mankiw, G., D. Romer, and M. D. Shapiro (1985): "An Unbiased Reexamination of Stock Market Volatility," *Journal of Finance*, 40(3), 677–687.

Marcet, A., and J. P. Nicolini (2003): "Recurrent Hyperinflations and Learning," *American Economic Review*, 93, 1476–1498.

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

——— (1992): "The Convergence of Vector Autoregressions to Rational Expectations Equilibria," in *Macroeconomics: A Survey of Research Strategies*, ed. by A. Vercelli, and N. Dimitri, pp. 139–164. Oxford University Press, Oxford.

——— (1995): "Speed of Convergence of Recursive Least Square Learning with ARMA Perceptions," in *Learning and Rationality in Economics*, ed. by M. Salmon, and A. Kirman, pp. 179–215. Blackwell, Oxford, UK.

Shiller, R. J. (1981): "Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?," *American Economic Review*, 71, 421–436.

——— (2005): *Irrational Exuberance: Second Edition.* Princeton University Press, Princeton, USA.

Timmermann, A. (1993): "How Learning in Financial Markets Generates Excess Volatility and Predictability in Stock Prices," *Quarterly Journal of Economics*, 108, 1135–1145.

——— (1996): "Excess Volatility and Predictability of Stock Prices in Autoregressive Dividend Models with Learning," *Review of Economic Studies*, 63, 523–557.

Voth, H. J. (2003): "With a Bang, Not a Whimper: Pricking Germany's 'Stockmarket Bubble' in 1927 and the Slide Into Depression," *Journal of Economic History*, 63, 65–99.