

Stock Market Volatility and Learning*

Klaus Adam Albert Marcet Juan Pablo Nicolini

September 2006

Abstract

Introducing learning into a standard consumption based asset pricing model with constant discount factor considerably improves its empirical performance. Learning causes momentum and mean reversion of returns and thereby excess volatility, long-horizon return predictability, and low frequency deviations from rational expectations (RE) prices. Learning also generates the possibility of price bubbles and - for overvalued prices - stock market ‘crashes’, i.e., sudden and strong price decreases with prices having a tendency to fall below their RE value. No symmetric stock market increases occur when prices are undervalued. Besides these qualitative features, learning considerably improves the ability to quantitatively match a range of standard asset pricing moments. Estimating the learning model using the method of simulated moments and U.S. asset price data (1926:1-1998:4), we show that it passes the test for the overidentifying restrictions at conventional significance levels. This is the case although learning introduces just one additional parameter into a standard (Lucas) asset pricing model, which fails to pass the overidentifying test at significance levels above machine precision.

JEL Class. No.: G12

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 very small departures from rationality. The result is somehow remarkable, since the literature in empirical finance has had a very hard time in developing dynamic equilibrium rational expectations models that can account for some of those facts. For

*Thanks go to Luca Dedola and Jaume Ventura for interesting comments and suggestions. Marcet acknowledges support from CIRIT (Generalitat de Catalunya), DGES (Ministry of Education and Science), 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, Universitat Pompeu Fabra) albert.marcet@upf.edu; Juan Pablo Nicolini (Universidad Torcuato di Tella) juanpa@utdt.edu.

example, Campbell and Cochrane (1999) show that a habit-persistence model is able to match US data only after imposing a multiple-parameter complex specification for the formation of habit in preferences.¹

It has long been recognized that stock prices exhibit movements that cannot be reproduced within the realm of rational expectation models: the risk premium is too high, stock prices are too volatile, the price/dividend ratio is too persistent and volatile, stock returns are unpredictable in the short run but negatively related to the price/dividend ratio in the long run, and there are stock market crashes. A very large body of literature has been devoted to documenting these empirical observations and to finding extensions of the standard model that will improve its empirical performance. A quick (and, therefore, unfair) summary is that it is not possible to find reasonable extensions of the basic model that will get close to explaining all these facts², unless a large number of parameter is added to the model, as in Campbell and Cochrane (1999). Instead, we follow a different approach: we replace the full rationality assumption by the most standard scheme used in the learning literature³: least squares learning (OLS). We show that with this modification, the model can replicate the data surprisingly well.

In this model, least squares learning has the property that in the long run the equilibrium converges to rational expectations, but this process takes a very long time, and the dynamics generated by learning along the transition cause prices to be very different from the rational expectations (RE) prices. The reason is that if expectations about stock price growth have increased, the actual growth rate of prices has a tendency to increase beyond the growth of fundamentals, thereby reinforcing the belief in a higher stock price growth. Learning thus imparts ‘momentum’ on stock prices and beliefs and produces large and sustained deviations of the price/dividend ratio, as they are observed in the data. Our model also produces - not rational - ‘bubbles’, meaning large increases in stock prices that do not seem justified by increases in fundamentals.⁴ Stock prices can be very high precisely because agents believe in higher stock price growth and the market behavior reinforces this belief. The high volatility of stock price growth and the predictability of stock returns in the long run follow from this behavior. We also find that once price embarks on a ‘bubble’ path, small changes in fundamentals (dividends) can trigger a market ‘crash’ that will end the bubble, meaning a sudden-large drop in stock prices.⁵

As we mentioned, OLS is the most standard assumption to model expectations in the learning literature. Although the limiting properties of least squares

¹Habit-persistence models with more natural specifications were unable to reproduce the data, see our discussion of Abel (1990) in section 4 .

²Campbell (2003) is a recent summary of this literature.

³See Bray (1982), Marcet and Sargent (1989), or Evans and Honkapoja (2001) for a survey.

⁴This is, of course, different from the rational bubbles described, for example, in Santos and Woodford (1997).

⁵Such price decreases can also be triggered by the learning dynamics themselves, i.e., without any change in fundamentals.

learning have been used extensively as a stability criterion to justify or discard RE equilibria, they are not commonly used to explain data or for policy analysis.⁶ It still is the standard view in the economics research literature 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 crucial methodological issue in two ways: first, by using a learning scheme with as few free parameters as possible, second, by imposing restrictions on the parameters of the learning scheme to only allow for small departures of rationality. In order to illustrate the effect of learning on the implications of the model in the simplest possible way, we adopted the first alternative: to use an off the shelf scheme (i.e., OLS) that has only one parameter.⁷ Still, in the model at hand, OLS performs reasonably well, it is the best estimator in the long run, and in order to minimize departures from rationality, we assume that initial beliefs are at the rational expectations equilibrium, and that agents have a strong confidence in these beliefs.

Models of learning have been used before to explain some aspects of asset pricing. Timmermann (1993, 1996), Brennan and Xia (2001), Cogley and Sargent (2006), show that Bayesian learning can help explain various aspects of stock prices. They assume that agents learn about the dividend process and they use the Bayesian posterior on the dividend process to estimate the discounted sum of dividends that would determine the stock price under RE. Therefore the belief of agents influences the market outcome, but agents' beliefs are not affected by market outcomes. In the language of stochastic control these models are not self-referential. By comparison, we abstract from learning about the dividend process and consider learning about the stock price process instead, so that beliefs affect prices and vice versa; it is precisely the learning about stock price growth and its self-referential nature that imparts the momentum to expectations and, therefore, is key in explaining the data. Other papers have pointed out that models of learning about stock prices can give rise to complicated stock price behavior, among others, Bullard and Duffy (2001) and Brock and Hommes (1998) show that learning dynamics can converge to complicated attractors, whenever the RE equilibrium is unstable under learning dynamics.⁸ By comparison, we address more closely the data in a model where the rational expectations equilibrium is stable under learning dynamics, and the strong departure from RE behavior occurs along the transition. Also related is Cárceles-Poveda and Giannitsarou (2006); they assume, in effect, that agents know the mean stock price and study deviations from the mean, their finding is that the presence of learning does not alter significantly the behavior of asset prices when agents learn about the effect of deviations from the mean. In the present paper we concentrate on agents that learn about the mean growth rate

⁶We will mention some exceptions along the paper.

⁷Marcet and Nicolini (2003) used a less standard scheme that combines OLS with tracking, but imposed "rational expectations-like" bounds on the size of the mistakes agents can make in equilibrium.

⁸Stability under learning dynamics is defined in Marcet and Sargent (1989).

of the stock price.⁹

In addition to studying the qualitative features introduced by learning, we also evaluate the ability of our model to quantitatively account for the behavior of U.S. stock markets. In particular, we formally estimate and test the model with learning using the method of simulated moments (MSM). We show that the model quantitatively matches the volatility of stock prices and returns, the volatility and persistence of the price dividend ratio, the evidence on stock return predictability over long horizons, the risk premium and, in a sense, it displays crashes. The match is surprisingly good, even though the model is the simplest possible equilibrium model with the most basic OLS learning which introduces one single additional parameter.

For the purposes of comparison, we also show the results of estimating a RE model with time-varying discount factors generated by habit persistence as in Abel (1990) which has the same number of parameters as the learning model. This RE model grossly fails to capture most of the evidence mentioned. We have to modify the standard MSM procedure that focuses on long run moments since, in our case, the learning model behaves just like RE in the long run. We adapt the standard MSM method in order to take into account short sample behavior of the model.

The paper is organized as follows. Section 2 documents various asset pricing facts that have been described in the literature and that this paper is concerned with. Section 3 presents a simple learning-based asset pricing model and derives analytical results about the behavior of stock prices under learning. Section 4 extends the simple model to the case with risk aversion and habit persistence and presents our estimation procedure. In section 5 we report the estimation outcomes for the extended learning model and - for comparison - for an RE model with habit persistence. Section 6 concludes. Technical material is contained in an appendix.

2 Facts

We are concerned with basic asset pricing facts that have been well documented in the literature. For completeness we reproduce these facts here using a single data set for the U.S. covering the period 1926:1-1998:4.¹⁰ Table 1 provides a first set of facts that we briefly discuss.¹¹

⁹Cecchetti, 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.

¹⁰The data is provided by Campbell (2003) and based on NYSE/AMEX value-weighted portfolio returns taken from CRSP stock file indices. It can be downloaded at <http://kuznets.fas.harvard.edu/~campbell/data.html>. Following standard practice we use lagged dividends to compute the price dividend ratio, causing the effective sample to start in 1927:1.

¹¹The table reports quarterly real values with returns and growth rates being expressed in percentage points. Real values are computed using the CPI deflator provided by Campbell

1. **Equity premium** Stock returns - averaged over long time spans and measured in real terms - tend to be high relative to short-term real bond returns.¹² The latter tend to be positive but fairly close to zero on average.¹³
2. **Stock Price Volatility.** Stock prices are much more volatile than dividends.¹⁴ This fact is recently summarized by the related observation that stock returns are much more volatile than dividend growth.¹⁵
3. **Price Dividend Ratio.** The price dividend ratio (PD) is high on average, very volatile and displays very persistent fluctuations. Figure 1 depicts the U.S. price dividend ratio. It illustrates the presence of large low frequency deviations of the PD ratio from its sample mean (bold horizontal line in the graph).

U.S. data, 1927:1-1998:4
(quarterly real values)

First moments	Symbol	Value
Av. stock return	$E(r^s)$	2.36
Av. bond return	$E(r^B)$	0.16
Av. PD ratio	$E(PD)$	105.4
Av. dividend growth	$E\left(\frac{\Delta D}{D}\right)$	0.346
Second moments		
StdDev stock return	σ_{r^s}	11.5
StdDev bond return	σ_{r^B}	1.35
StdDev PD ratio	σ_{PD}	35.4
StdDev dividend growth	$\sigma_{\frac{\Delta D}{D}}$	3.63
Autocorrel. PD ratio	$\rho(PD_t, PD_{t-1})$	0.95

Table 1: Asset pricing moments

4. **Stock Return predictability.** While stock returns are generally difficult to predict, the PD ratio is negatively related to future excess stock returns in the long run.¹⁶ Table 2 shows the results of regressing future cumulated

(2003). All variables are in levels. Using the log values instead gives rise to a very similar picture.

¹²Mehra and Prescott (1985).

¹³Weil (1989).

¹⁴Shiller (1981) and LeRoy and Porter (1981).

¹⁵In table 1 quarterly stock returns are about three times as volatile as quarterly dividend growth, where quarterly dividend growth is averaged over the last 4 quarters so as to eliminate seasonalities, as in Campbell (2003). In any case, stock returns are also about three times as volatile as dividend growth at yearly frequency.

¹⁶Poterba and Summers (1988), Campbell and Shiller (1988), and Fama and French (1988).

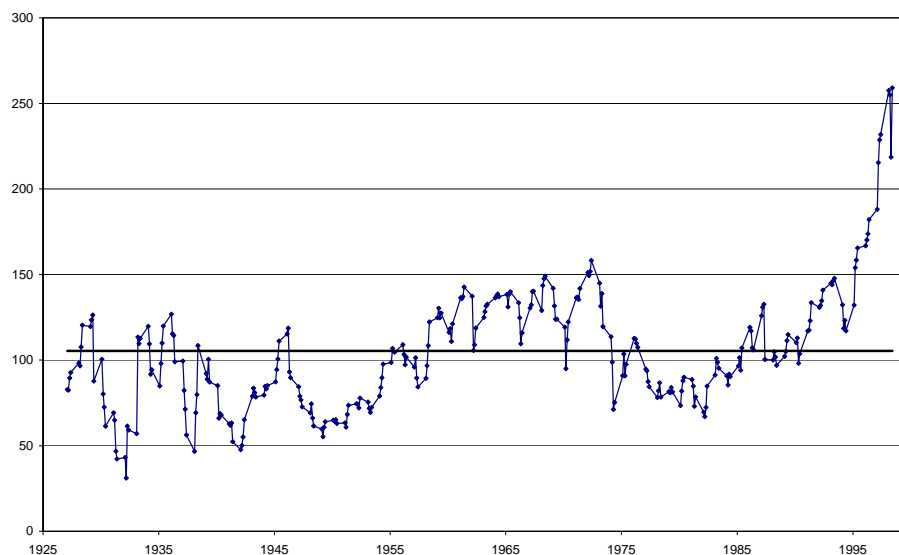


Figure 1: Quarterly U.S. price dividend ratio 1927:1-1998:4

excess returns over different horizons on today's price dividend ratio.¹⁷ As has been reported before, the R^2 increases for longer horizons, and the regression coefficients become increasingly negative.¹⁸ This suggests the presence of low frequency components in excess stock returns, i.e., the presence of long and sustained increases and downturns of stock prices that are related to the PD. At the same time, the price dividend ratio has no clear ability to forecast future dividends, future earnings, or future real

¹⁷The table reports results from OLS estimation of

$$X_{t,t+s} = c_0^s + c_1^s PD_t + u_t^s$$

for $s = 4, 20, 40, 60$ quarters 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_1^s . As in Campbell (2003) the price dividend ratio is the price divided by average dividend payments in the last 4 quarters.

¹⁸Whether the coefficients are significantly different from zero is a non-trivial question because the price dividend ratio is highly autocorrelated, see the discussion in Campbell and Yogo (2005).

interest rates.¹⁹

Years	Coefficient on PD	R^2
1	-0.0017	0.05
5	-0.0118	0.34
10	-0.0267	0.46
15	-0.0580	0.53

Table 2: Excess stock return predictability (1927:1-1998:4)

5. Stock market crashes. Stock markets occasionally experience ‘crashes’, i.e., strong and sudden price decreases, which seem to occur after a period of strong asset price increases. Table 3 lists the crashes identified by Mishkin and White (2002) for the S&P 500 over the period 1947:2-1998:4. A stock market crash is defined as a nominal price decrease by more than 20% occurring in a short period of time (generally less than 3 months). There are four episodes with such strong reductions in prices, with the stock market crash in October 1987 probably being the most uncontroversial one. The stock market crashes listed in table 3 are clearly identifiable as sharp decreases of the price dividend ratio in figure 1, which suggests that crashes are not the result of changes in fundamentals (dividends) only.

Start	End	Total Change
Dec 1961	June 1962	-22.5%
Nov 1968	June 1970	-30.9%
Jan 1973	Dec 1974	-45.7%
Aug 1987	Dec 1987	-26.8%

Table 3: Stock Market Crashes in the S&P 500 (1947:1-1998:4)

A very large body of literature generalizes the basic asset pricing model under RE to explain some of these facts. A rough summary of the literature is that some papers have been able to explain some of these facts, providing a better understanding of what drives some of the above fluctuations. With the possible exception of the highly parameterized model of Campbell and Cochrane (1999) mentioned before, none of these papers have come close to explaining all of the observations above.²⁰

¹⁹Campbell (2003).

²⁰See Campbell (2003) for a summary.

3 A Simple Model of Stock Prices

In this section we consider the simplest risk-neutral asset pricing model. As is well known, this model fails to explain basic observations under RE. Precisely for this reason it is useful for investigating how asset pricing behavior is changed once learning is introduced. The emphasis in this section is on qualitative results that can be obtained from analytical reasoning. Section 4 extends the analysis to the case with risk-averse investors and evaluates the quantitative performance of the model under learning and RE.

Consider a stock that yields exogenous dividend D_t each period. For simplicity we assume (log) dividends to follow a unit root process

$$\frac{D_t}{D_{t-1}} = a\varepsilon_t \quad (1)$$

where $\varepsilon_t > 0$ is an iid shock with $E(\varepsilon_t) = 1$. In some cases we make the additional assumption $\log \varepsilon_t \sim N(-\frac{s^2}{2}, s^2)$. The expected growth rate of dividends is given by $a \geq 1$. As documented in Mankiw, Romer and Shapiro (1985) or Campbell (2003), process (13) provides a reasonable approximation to the empirical behavior of quarterly dividends in the U.S.

The consumer has beliefs about future variables, these beliefs are summarized in expectations denoted \tilde{E} which we allow to be less than fully rational. Prices satisfy

$$P_t = \delta \tilde{E}_t (P_{t+1} + D_{t+1}) \quad (2)$$

where P_t is stock price and δ some discount factor.

Equation (2) will be the focus of our analysis in this section. It will be derived from an equilibrium model with infinitely lived agents that we describe more formally in section 4. Although the infinite horizon model has been the focus of the literature, equation (2) can also be derived from many other models, e.g., from a simple no-arbitrage condition with risk-neutral investors if δ denotes the inverse of the short-term gross interest rate, or from an overlapping generations model with risk-neutral agents, etc. The key to equation (2) is that investors formulate expectations about the future payoff $P_{t+1} + D_{t+1}$ and for investors' choice to be in equilibrium today's price has to equal next period's discounted expected payoff.

Some papers in the learning literature have studied stock prices when agents formulate expectations about the discounted sum of all future dividends.²¹ These papers set

$$P_t = \tilde{E}_t \sum_{j=1}^{\infty} \delta^j D_{t+j} \quad (3)$$

²¹Timmermann (1993, 1996), Brennan and Xia (2001), Cogley and Sargent (2006).

and evaluate the expectation based on the Bayesian posterior distribution of the parameters in the dividend process. It is well known that under RE and some limiting condition on price growth the one-period ahead formulation of (2) is equivalent to the discounted sum expression for prices.²² However, under learning this is not the case.

If agents learn about price according to (3), the posterior is about parameters of an exogenous variable, namely the dividend process. As a result, market prices will not influence expectations and learning will not be self-referential. While this allows for straightforward formulation of Bayesian posteriors, the lack of feedback from market prices to expectations limits the ability of the model to generate interesting ‘data-like’ behavior. Using the formulation (2) requires agents to have a model of next period’s price directly and forces them to estimate the parameters of their model using stock price data. Our point will be that it is precisely when agents formulate expectations on future prices using past prices to satisfy (2) that there is a large effect of learning and that many moments of the data are matched better. It is in fact this self-referential nature of our model that makes it attractive in explaining the data.²³

Focusing on (2) instead of (3) can be justified by a number of arguments based on principles. Informally, one can say that most participants in the stock market care much more about the selling price of the stock than about the discounted dividend stream, a feature that may be caused by short investment horizons.²⁴ More formally, it is the case that evaluating (3) in a fully rational Bayesian sense is computationally extremely costly. Indeed, the literature on Bayesian learning has used various short-cuts for evaluating the discounted sum.²⁵ The pricing implications of these short-cuts are unclear at best and can

²²For $\tilde{E}_t[\cdot] = E_t[\cdot]$ this limiting condition is the no-rational-bubble requirement $\lim_{j \rightarrow \infty} \delta^j E_t P_{t+j} = 0$.

²³Timmerman (1996) considers self-referential learning assuming that agents use dividends to predict both future price and future dividend. While this generates a self-referential learning model, it also generates close to unit eigenvalues in the mapping from perceived to actual parameters. This causes learning dynamics to become extremely slow and not contribute significantly to return dynamics.

²⁴It is possible to formally justify the interest in predicting future price in the framework of an overlapping generations model. We do not pursue this further in this paper.

²⁵For example, Timmermann (1996) assumes that agents form a Bayesian posterior $E_t^{Bay}[\rho]$ for the serial correlation of dividends ρ and treat it as a point estimate such that (3) can be evaluated as $P_t = \sum_{j=1}^{\infty} \delta^j \left[E_t^{Bay}(\rho) \right]^j D_t$. While this is a valuable simplification, it is not a fully rational model because under rational expectations $E_t^{Bay}(\rho^j) \neq \left[E_t^{Bay}(\rho) \right]^j$. Related to this is the observation that simply iterating optimal one-step forecasts does not produce optimal multi-step forecasts. Adam (2005) provides experimental evidence showing that agents cease to iterate on one-step forecasts once they become gradually aware that they use a possibly misspecified forecasting model.

be extreme under some circumstances.²⁶ Also, the discounted sum formula implicitly assumes that agents know perfectly the process for the *market* interest rate, therefore it either assumes a lot of knowledge about interest rates on the part of the agents or it ignores issues of learning about the interest rate.²⁷ For all these reasons we conclude that our one-period formulation in terms of prices is an interesting avenue to explore.

3.1 RE equilibrium

If agents hold rational expectations (RE) about future prices and dividends ($\tilde{E}_t[\cdot] = E_t[\cdot]$), equations (2) and (1) imply

$$P_t^{RE} = \frac{\delta a}{1 - \delta a} D_t. \quad (4)$$

This RE equilibrium misses all asset pricing facts mentioned before in section 2. In particular, the model with risk neutrality generates a zero equity premium, violating fact 1.²⁸ In addition, since $\frac{P_t^{RE}}{P_{t-1}^{RE}} = \frac{D_t}{D_{t-1}}$, average price growth is exactly equal to average dividend growth, and approximately equal to mean stock returns.²⁹ The volatility of stock returns is thus roughly equal to the volatility of dividend growth, which contrasts with fact 2. The model predicts a constant price dividend ratio, therefore fails to explain fact 3. Since $\frac{P_t + D_t}{P_{t-1}} = \frac{\varepsilon_t}{\delta}$ stock returns are i.i.d., implying no predictability of returns at any horizon, unlike suggested by fact 4. Finally, stock prices are proportional to dividends, so there cannot be ‘crashes’ without sudden corresponding reductions in dividends, violating fact 5.

Obviously, it is possible to do better than the simple risk neutral model maintaining RE. Yet, precisely because the risk neutral case fails so strongly, it constitutes the most useful setting for demonstrating the potential of a very simple self-referential learning model to match the data. Later sections will offer a more detailed quantitative comparison of learning models with other more general RE models that have the chance of meeting some of the facts mentioned in section 2.

²⁶For example given in the previous footnote, $\lim_{T \rightarrow \infty} \sum_{j=1}^T \delta^j \left[E_t^{Bay}(\rho) \right]^j D_t$ may converge, while the properly evaluated sum $\lim_{T \rightarrow \infty} \sum_{j=1}^T \delta^j E_t^{Bay}(\rho^j) D_t$ may diverge to infinity. See Weitzman (2005) for a related point.

²⁷This point can be formalized in a model of heterogeneous agents where the market interest rate is not equal to the discount factor of a single agent. In that case, the agent’s knowledge about his/her own discount factor does not imply knowledge of the market interest rate.

²⁸Mehra and Prescott (1985) show that introducing reasonable degrees of risk aversion do not solve this problem.

²⁹This follows from $\frac{P_t + D_t}{P_{t-1}} = \frac{1 + PD_t}{PD_{t-1}} \frac{D_t}{D_{t-1}} \approx \frac{D_t}{D_{t-1}}$ where $PD_j = \frac{\delta a}{1 - \delta a}$ is the quarterly price dividend ratio, which tends to be large.

3.2 Learning Mechanism

In this section we introduce self-referential learning into the asset pricing model with risk neutrality. We want to study learning schemes that forecast reasonably well within the model. For this reason, we introduce a number of features in the formulation of the expectations in (2) insuring that learning agents do not make large forecasting errors within the model.

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

$$P_t = \delta \tilde{E}_t(P_{t+1}) + \delta \tilde{E}_t(D_{t+1}) \quad (5)$$

We assume that agents know how to formulate the conditional expectation of the dividend $\tilde{E}_t(D_{t+1}) = aD_t$, which amounts to assuming that agents have rational expectations about the dividend process. This may appear inconsistent with our assumption regarding expectations formation about prices, but the results we obtain are very similar when agents are also learning to forecast dividends.³⁰ We maintain this assumption in the paper for simplicity and because it allows us to highlight the effect of the self-referential component of the model.

As mentioned before, agents are assumed to use a learning scheme to form $\tilde{E}_t(P_{t+1})$ using past information. Equation (4) shows that under rational expectations $E_t\left[\frac{P_{t+1}}{P_t}\right] = a$. This justifies specifying the expectations under learning as

$$\tilde{E}_t[P_{t+1}] = \beta_t P_t \quad (6)$$

where β_t is some estimator of stock price growth based on past observations. It is clear that if the model converges to the RE equilibrium, agents will realize that this is a good way to forecast future prices in the long run. In this way, this learning scheme has a chance of satisfying Asymptotic Rationality, as defined in Marcet and Nicolini (2003). As long as the model converges to RE - we prove this to be the case later on - agents' forecasts are optimal in the limit.

We now have to specify how past information is taken into account when updating the estimator β_t . We start by presenting the updating mechanics and thereafter offer an interpretation. The learning mechanism is assumed to satisfy the standard equation in stochastic control

$$\beta_t = \beta_{t-1} + \frac{1}{\alpha_t} \left(\frac{P_{t-1}}{P_{t-2}} - \beta_{t-1} \right) \quad (7)$$

for all $t \geq 1$, for a given sequence of α_t , and a given initial belief β_0 which is given outside the model.³¹ The sequence $(\alpha_t)^{-1}$ is called the 'gain' sequence

³⁰Appendix ?? shows that the conclusions of the paper are robust to assuming that agents also learn about how to forecast dividends. Imposing RE about dividends implicitly assumes that learning about dividends has converged already. Since dividend growth follows an exogenous process, learning the parameters governing the dividend process is fairly easy for agents.

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

and dictates how the last prediction error is incorporated into beliefs.³² The assumed gain sequence is

$$\begin{aligned}\alpha_t &= \alpha_{t-1} + 1 \quad t \geq 2 \\ \alpha_1 &\geq 1 \quad \text{given.}\end{aligned}\tag{8}$$

With these assumptions the model evolves as follows. In the first period β_0 determines the first price P_0 ; using the previous price level one finds the first observed growth rate $\frac{P_0}{P_{-1}}$, which is used to update beliefs to β_1 using (7); the belief β_1 determines P_1 and the process continues to evolve recursively in this manner. As in any self-referential model of learning, prices enter in the determination of beliefs and vice versa.

Using simple algebra equation (7) implies

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

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 . The initial gain α_1 is thus a measure of the degree of ‘confidence’ agents place on 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 arguing 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, such priors fail to contain the ‘grain of truth’ typically assumed to be present in Bayesian analysis. While the i.i.d. assumption will hold asymptotically (we will prove this later on), it is violated under the transition dynamics. In a proper Bayesian formulation, therefore, agents would use a likelihood function with the property that if agents use it to update their posterior, it turns out to be the true likelihood of the model in all periods. Most likely, β_t would have to depend on the past in a complicated non-linear way and only in the limit would the Bayesian use a simple average as has been assumed above. Since the ‘correct’ likelihood in each period 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

³²Note 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 it is entertained for simplicity.

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.

For the case $\alpha_1 = 1$ the belief β_t is given by the sample average of stock price growth, i.e., the OLS estimate of the mean growth rate. The initial belief β_0 then matter only for the first period, but ceases to affect beliefs after the first piece of data has arrived. More generally, assuming a low value for α_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. Therefore, in the spirit of using initial beliefs that have a chance of being near-rational we set initial beliefs equal to the RE belief

$$\beta_0 = a$$

and choose a high initial weight α_1 for these beliefs. As a result, initial beliefs will be ‘close’ to the beliefs that support the RE equilibrium.

We can summarize as follows. We assume agents to formulate their beliefs by an average of OLS and their initial (correct under RE) belief, with the relative weight given by the number of observations and $(\alpha_1 - 1)$, respectively.

3.3 Stock prices under learning

Given the perceptions β_t , the expectation function (6), and the assumption on perceived dividends, equation (5) implies that prices under learning satisfy³³

$$P_t = \frac{\delta a D_t}{1 - \delta \beta_t}. \quad (9)$$

Since β_t is independent of ε_t the previous equation implies that

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

which shows that prices growth under learning is more volatile than dividend growth. This intuition is present in previous models of learning, e.g., Timmermann (1993). Particular to our case will be the fact that $\text{Var} \left(\ln \frac{1 - \delta \beta_t}{1 - \delta \beta_{t+1}} \right)$ is very high and will remain high for a long time, so that the volatility of prices will be increased by a large amount for long periods of time.

³³For this equation to be valid we need $\beta_t \in (0, \delta^{-1})$, otherwise there exist no market clearing price. Since prices are positive, β_t is always positive, but the model has to somehow be modified to avoid β_t from becoming larger than δ^{-1} . We will discuss this issue in more detail later on. For the moment, we assume that beliefs that satisfy this inequality.

Simple algebra gives

$$\frac{P_t}{P_{t-1}} = T(\beta_t, \Delta\beta_t) \varepsilon_t \quad (10)$$

where

$$T(\beta, \Delta\beta) \equiv a + \frac{a\delta \Delta\beta}{1 - \delta\beta} \quad (11)$$

Substituting (10) in the law of motion for beliefs (7) delivers an equation describing the whole evolution of β_t as a function of the shocks ε_t and the initial belief β_0 . Prices can then be determined from equation (9). The dynamics of β_t are thus governed by a second order stochastic non-linear difference equation. This equation can not be solved analytically, but it is possible to give considerable insights in the behavior of the model using analytic reasoning.

3.3.1 Asymptotic Rationality

We start by studying the limiting behavior of the model, drawing on results from the literature on least squares learning. This literature shows that the T -mapping defined in equation (11) is central to stability of RE equilibria under learning.³⁴ 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 (7) 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 $\dot{\beta} = T(\beta) - \beta = a - \beta$ stable. This seems to indicate that beliefs should converge to the RE equilibrium value $\beta = a$ relatively quickly. One might then conclude that there is not much to be gained from introducing learning into the standard asset pricing model.

Appendix D shows in detail that the above approximations are correct and that learning globally converge to the RE equilibrium in this model, i.e., $\beta_t \rightarrow a$. 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, would not have incentives to change their learning scheme.

In the remainder of the paper we show that the model here behaves very different from RE during the transition to the limit. This occurs although agents are using an estimator that starts at the RE value, that will be the best estimator in the long run, and that converges to the RE value. The difference is so large that even the very simple version of the model together with the very simple learning scheme introduced in section 3.2 explains the data much better than

³⁴See Marcet and Sargent (1989) and Evans and Honkapohja (2001)

the model under RE. This brings about the general point that concentrating on the limiting properties of least squares learning may undervalue the potential for models of learning to explain the behavior of the economy.³⁵

3.3.2 Mean dynamics

We now describe the transition behavior of the model under learning by studying its mean dynamics conditional on past information. Since β_{t+1} is a function of the shock ε up to period t , we study $E_{t-1}\beta_{t+1}$ to examine the expectation of β_{t+1} before it is actually known. In particular, we will be interested in finding $E_{t-1}\Delta\beta_{t+1}$. Using (10) we have $E_{t-1}\left(\frac{P_t}{P_{t-1}}\right) = T(\beta_t, \Delta\beta_t)$, where T is the *actual* expected stock price growth as a function of current and past beliefs. Using this observation and conditioning on both sides of (7) we obtain

$$E_{t-1}\Delta\beta_{t+1} = \frac{1}{\alpha_{t+1}} [T(\beta_t, \Delta\beta_t) - \beta_t] \quad (12)$$

where E_{t-1} denotes actual conditional expectations given that prices are determined within the model of learning. Equation (12) shows that β_{t+1} is expected to adjust towards $T(\beta_t, \Delta\beta_t)$. For example, if history generated beliefs such that $T(\beta_t, \Delta\beta_t) > \beta_t$ then we expect the perceptions β_t to increase. The gain α thereby determines the size of the updating step only. Understanding how beliefs are expected to evolve under learning thus requires studying the T -mapping. Below we derive a number of results about the map T , which are followed by an interpretation of their implications.

We start by noting that actual expected stock price growth depends not only on the *level* of price growth expectations β_t but also on the *change* $\Delta\beta_t$:

Result 1: For all $\beta \in (0, \delta^{-1})$

$$\begin{aligned} T(\beta, \Delta\beta) &> a && \text{if } \Delta\beta > 0 \\ T(\beta, \Delta\beta) &< a && \text{if } \Delta\beta < 0 \end{aligned}$$

Therefore, if agents arrived at the rational expectations belief $\beta_t = a$ from below ($\Delta\beta_t > 0$), the price growth generated by the learning model exceeds the fundamental growth rate a in expectations. We can state this formally as

$$E_{t-1}(\Delta\beta_{t+1} \mid \beta_t = a, \Delta\beta_t > 0) > 0$$

Just because agents' expectations have become more optimistic (in what a journalist would perhaps call a 'bullish' market), the price growth in the market

³⁵Some papers, including Marcet and Sargent (1995) and Ferrero (2004), have emphasized that least squares learning converges slowly to RE if $\partial T(\beta)/\partial\beta$ is close to one, but converges much faster if $\partial T(\beta)/\partial\beta < 1/2$. In the current model we have $\partial T(\beta, 0)/\partial\beta = 0$ indicating fast convergence. Our findings show that values of $\partial T(\beta)/\partial\beta$ close to one are not the only reason that convergence to RE may be very slow. In the present paper slow convergence arises because of the non-linearities of the model out of (but close to) the limit point.

has a tendency to be larger than fundamental growth. Since agents will use this higher-than-fundamental stock price growth to update their beliefs in the next period, β_t will tend to overshoot a , which will reinforce the upward tendency further. It is at this point where the self-referential nature of the learning mechanism makes a difference for the dynamics under learning.³⁶ Conversely, if $\beta_t = a$ in a bearish market ($\Delta\beta_t < 0$), beliefs display downward momentum, i.e., a tendency to undershoot the RE value.

We have argued before that in the limit the mapping from actual to perceived expectations is given by $T(\cdot, 0) = a$ so that actual growth is not affected by perceived growth. During the transition, however, $\Delta\beta_t$ is not equal to zero and the expression for T given in equation (11) highlights that $\Delta\beta_t \neq 0$ imparts substantial non-linearity in the model. These non-linear features are summarized below.

Result 2: For all $\beta \in (0, \delta^{-1})$

- a) For $\Delta\beta > 0$ the map $T(\cdot, \Delta\beta)$ is increasing and convex and converges to $+\infty$ as $\beta \rightarrow \delta^{-1}$,
- b) For $\Delta\beta < 0$ the map $T(\cdot, \Delta\beta)$ is decreasing and concave and converges to $-\infty$ as $\beta \rightarrow \delta^{-1}$.
- c) The level and first and second derivatives of $T(\cdot, \Delta\beta)$ are increasing in $\Delta\beta$.³⁷
- d) Given $\Delta\beta$, the fixed points of $T(\cdot, \Delta\beta)$ are as follows:
 - For $\Delta\beta > 0$ and sufficiently small³⁸, there are two fixed points $a < \bar{\beta} < \overline{\beta} < \delta^{-1}$ (which depend on $\Delta\beta$) such that

$$\begin{aligned} T(\beta, \Delta\beta) &< \beta && \text{if } \beta \in (\bar{\beta}, \overline{\beta}) \\ T(\beta, \Delta\beta) &> \beta && \text{if } \beta \notin (\bar{\beta}, \overline{\beta}) \end{aligned}$$

- For $\Delta\beta > 0$ and large enough, $T(\beta, \Delta\beta) > \beta$ for all $\beta \in (0, \delta^{-1})$ and there are no fixed points.
- If $\Delta\beta < 0$ there is one fixed point $\tilde{\beta} < a$ (which depends on $\Delta\beta$) such that

$$\begin{aligned} T(\beta, \Delta\beta) &> \beta && \text{if } \beta < \tilde{\beta} \\ T(\beta, \Delta\beta) &< \beta && \text{if } \beta > \tilde{\beta} \end{aligned}$$

³⁶It is easy to check that in the model of Timmermann (1996) there is a similar tendency for stock price growth to overshoot, but this has no effect on perceptions of agents. In his model agents' perceptions depend only on exogenous dividends. Therefore, there is not feedback from prices to perceptions and there is no momentum in beliefs.

³⁷To be precise, for $\Delta\beta > \Delta\beta'$ and any $\beta \in (0, \delta^{-1})$ we have $\frac{\partial T(\beta, \Delta\beta)}{\partial \beta} > \frac{\partial T(\beta, \Delta\beta')}{\partial \beta}$ and $\frac{\partial^2 T(\beta, \Delta\beta)}{\partial^2 \beta} > \frac{\partial^2 T(\beta, \Delta\beta')}{\partial^2 \beta}$.

³⁸More precisely, if $\Delta\beta < \frac{(a\delta-1)^2}{4\delta^2 a}$.

These properties can be derived from simple algebra. They are illustrated in Figure 2, which depicts the T -map for each of the three cases described in result 2d) taking into account results 2a)-2c).

The above result can be used to derive the mean dynamics of the model under learning.

Result 3: • If $\Delta\beta_t > 0$ and sufficiently small, letting $\bar{\beta}, \overline{\bar{\beta}}$ be as in Result 2e),

$$\begin{aligned} E_{t-1}\beta_{t+1} &< \beta_t && \text{if } \beta_t \in (\bar{\beta}, \overline{\bar{\beta}}) \\ E_{t-1}\beta_{t+1} &> \beta_t && \text{if } \beta_t \notin (\bar{\beta}, \overline{\bar{\beta}}) \end{aligned}$$

- If $\Delta\beta_t > 0$ and large enough $E_{t-1}\beta_{t+1} > \beta_t$
- If $\Delta\beta_t < 0$, letting $\tilde{\beta}$ be the corresponding value in Result 2e),

$$\begin{aligned} E_{t-1}\beta_{t+1} &> \beta_t && \text{if } \beta_t < \tilde{\beta} \\ E_{t-1}\beta_{t+1} &< \beta_t && \text{if } \beta_t > \tilde{\beta} \end{aligned}$$

We illustrate the mean dynamics in Figure 2 by drawing arrows on the β axis of each graph. An arrow pointing left (right) indicates that the mean dynamics imply a decrease (increase) in β_t .

For the case $\Delta\beta_t > 0$ figure 2 indicates that if $\Delta\beta_t$ is too large (so that the second graph applies) or if β_t is too large (so that we are at the right end of the axis in the first graph), β_t tends to grow, even if it is already much higher than the fundamental RE value a . In the limit, if β_t is close to the upper bound δ^{-1} the change in prices is infinite. Symmetrically, low values of $\Delta\beta_t$ or β_t imply that perceptions have a tendency to move towards $\bar{\beta}$ ($> a$). For beliefs that are high ($\beta_t > \bar{\beta}$) but not too high ($\beta_t < \overline{\bar{\beta}}$) this suggests a stable system, as these beliefs are drawn back towards the fundamental value a .

The previous findings show that the model has the potential to display bubbles: if growth perceptions start to grow (and, say, the second graph of Figure 2 applies), they cross the ‘fundamental’ growth rate a and as long as $\Delta\beta > 0$ there is an upward movement of the expected growth of stock prices β_t . From formula (9) follows that a higher value for β_t implies higher a PD ratio. Therefore, a bubble may occur.

Importantly, when growth perceptions and stock prices are high, a small change in return expectations can generate a very strong price decrease. More precisely, a high β_t combined with a slightly negative $\Delta\beta_t$ may start the downturn of the bubble or even a crash. To make this point we do not need to take a stand on what caused this decrease in growth perceptions: it could either be a low realization of the innovation to dividends ε_t , or simply due to

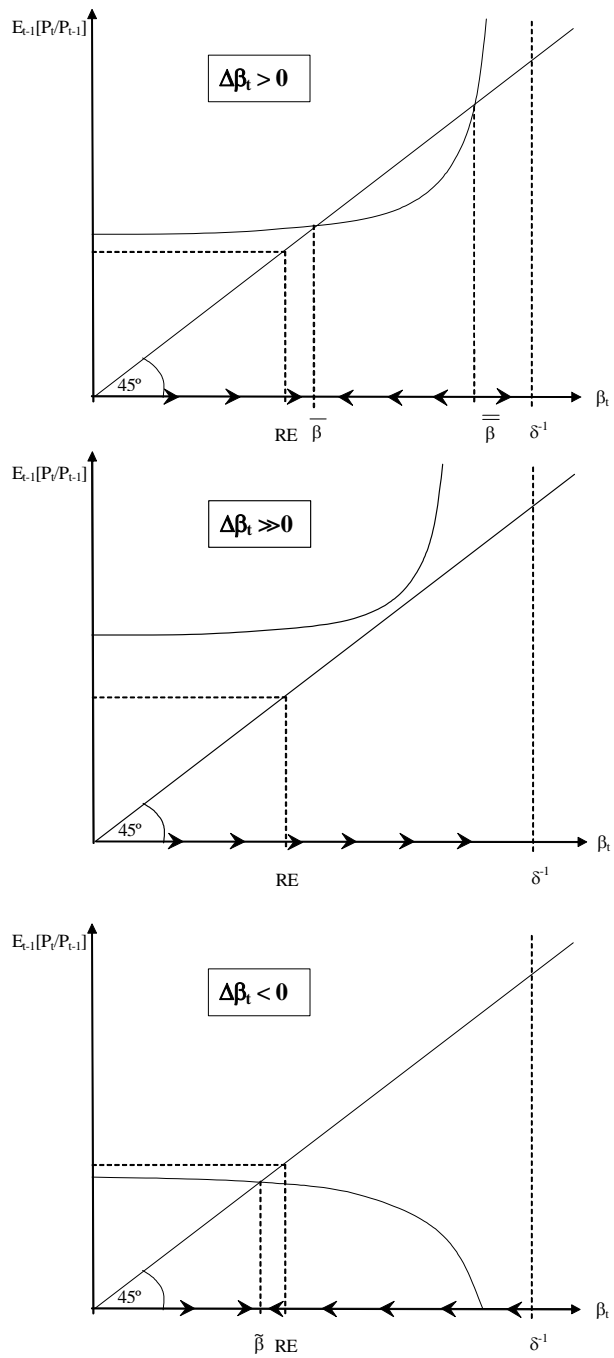


Figure 2: T-map

the learning dynamics, i.e., perceptions entering the interval $(\bar{\beta}, \overline{\beta})$ in the first graph in Figure 2. Going slightly outside of the model in this paper, the drop in expected price growth could also be generated by a Central Bank ‘pricking’ a bubble. Whatever caused the initial downward revision in beliefs, the third graph in figure 2 shows that if a high β_t is combined with $\Delta\beta_t < 0$, $E_{t-1}\beta_{t+1}$ is much lower than the fundamental growth rate a . Therefore, once perceptions have started to fall, they will fall further as the third graph will describe the learning dynamics for many periods. This continued decline in perceptions will cause a fall in the PD ratio, but since $\tilde{\beta} < a$ prices will have a tendency to fall below the fundamental value. Small changes in fundamentals may thus trigger a ‘stock market crash’.

These sudden reversals can not occur for low values of β_t . It is clear that the maps in all graphs in Figure 2 are very similar and when β_t is small, so that the instability discussed in the previous paragraph is only activated at high β 's. The learning model thus implies that a large fall in price may occur when prices are overvalued, but no symmetric price increase for undervalued prices. We summarize the previous findings as follows:

Result 4: If a high β_t is combined with $\Delta\beta_t < 0$ we have

$$E_{t-1} \left(\frac{P_t}{P_{t-1}} \right) \ll \beta_t$$

with the possibility of a ‘market crash’. If β_t is low, $\Delta\beta_t$ does not have a large influence on actual prices.

The analysis of the model’s mean dynamics in this section suggests that the model has the potential of matching all the asset pricing facts mentioned in section 2. Clearly, Results 1, 3 and the possibility of bubbles imply that the learning model generates excess price volatility, matching facts 2 and 3. Occasional market crashes are likely to occur, as in fact 5. Results 1 and 3 imply that learning imparts dynamics into the behavior of prices, causing prices to be very high or very low depending on how β_t combines with β_{t-1} and ε_t and that β_t depends strongly on β_{t-1} . Since the PD ratio is highly related to β_t it is likely that it will be highly serially correlated and that it will help predict stock returns as in facts 3 and 4.

At this writing we have not given too much attention to the equity premium. Simulations show that the model under learning generates a considerable equity premium. This probably occurs for the following reason. While β is growing the first two graphs of Figure 2 show that actual price growth is less different from perceptions than in the third graph of the figure. If actual price growth is more similar to perceived price growth, perceptions change less strongly. This

suggests that if perceived growth is high it tends to have more persistence than if perceived growth is low.³⁹

Finally, we need to introduce a feature that prevents perceived stock price growth from being higher than δ^{-1} so as to insure a positive price in (9). If beliefs are such that $\beta_t > \delta^{-1}$, expected stock return is larger than the inverse of the discount factor and the representative agent will have an infinite demand for stocks at any stock price. The model could be changed in a number of directions to avoid this infinite demand, but in the interest of staying as close as possible to the literature we do not take this route. Instead, we follow Timmermann and Cogley and Sargent and apply the following projection facility: if in some period β_t determined by (7) is larger than some constant $K \leq \delta^{-1}$ then set

$$\beta_t = \beta_{t-1}$$

in that period, otherwise we use (7). The interpretation is that if the observed price growth implies beliefs are too high, agents realize that this would prompt a crazy action (infinite stock demand) and they decide to ignore this observation. The constant K is chosen so that the implied PD is less than a certain upper bound U^{PD} . It turns out that this facility is binding only very rarely and that it does not affect the moments we look at.

3.3.3 Simulation under risk neutrality

To illustrate the previous discussion of the model under learning by reporting simulation results in a calibrated example. 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 mean and standard deviation of US dividends summarized in table 1. Using the log-normality assumption we set

$$a = 1.00346, \quad s = 3.63. \quad (13)$$

The discount factor is

$$\delta = 0.9872$$

and implies that the PD ratio of the RE model matches the observed average ratio in the data.

In the learning model we set

$$\beta_0 = a \quad \text{and} \quad \alpha_1 = 50$$

³⁹This would be a different and complementary mechanism to the transition from an initial pessimistic belief emphasized in Cogley and Sargent (2006).

These starting values are chosen to insure that the agents' expectations will not depart too much from rationality. Agents have high confidence on the RE belief. The initial value for α implies that after twelve years β_t is halfway between a and the observed sample mean. The bounds on β_t are set so that the price dividend ratio will never exceed 500.

Table 4 shows the average moments (across realizations) of each statistic computed by each model with 288 observations, together with the 95% probability interval of the statistic across realizations.⁴⁰

	U.S. Data	RE	Learning
	First and second moments		
$E(r^S)$	2.36	1.30 [0.93,1.64]	1.61 [1.32,1.91]
$E(r^B)$	0.16	1.30 [1.30,1.30]	1.30 [1.30,1.30]
$E(PD)$	105.4	105.4 [105.4,105.4]	77.6 [60.2,100.1]
σ_{r^S}	11.5	3.67 [3.42,3.92]	4.68 [4.19,5.19]
σ_{PD}	35.4	0.00 [0.00,0.00]	19.3 [9.7,35.2]
$\rho(PD_t, PD_{t-1})$	0.95	-	0.991 [0.981,0.997]
	Excess return predictability		
	Coefficient on PD		
1 yr	-0.0017	-	-0.0022 [-0.0049,-0.0007]
5 yrs	-0.0118	-	-0.0106 [-0.0215,-0.0032]
10 yrs	-0.0267	-	-0.0186 [-0.0354,-0.0049]
15 yrs	-0.0580	-	-0.0249 [-0.0476,-0.0049]
	R^2 value:		
1 yr	0.05	0.00	0.08 [0.02,0.17]
5 yrs	0.34	0.00	0.30 [0.05,0.57]
10 yrs	0.46	0.00	0.43 [0.04,0.77]
15 yrs	0.53	0.00	0.50 [0.03,0.84]

Table 4: Data and model under risk neutrality

The column labeled US data reports statistics that have been discussed in section 2. It is clear that the RE model fails to explain key asset pricing moments, see the column labeled RE. Consistent with our discussion the RE equilibrium fails to match the equity premium, the low risk free rate, the variability of stock returns and PD ratio, the serial correlation of the PD ratio, the

⁴⁰To compute these statistics we use 5000 realizations each of 288 periods, which the same length as the available data. Since we abstract from learning about dividends the RE and learning model both imply constant real bond yields. We thus do not report this statistic in the table.

predictability of excess returns.⁴¹

The learning model shows a higher volatility of stock returns, high volatility and high persistence of the PD ratio, and the coefficients and R^2 of the excess predictability regressions all move strongly in the direction of the data. This is consistent with our discussion of the mean dynamics under learning. Some statistics of the learning model do not match exactly the moments in the data⁴², but the purpose of the table is to show that adding learning improves enormously the ability of the model to match observations. This finding is robust to changing α_1 , as long as it is fairly high. It is also robust to changes in the bounds, which are active in very few periods in each simulation.

4 Estimation and testing

For illustrative purposes the previous section used the most simple model with the most standard learning scheme, imposing also the same parameter values in the RE and learning model. In this section we add some elements of generality to the model and disconnect the parameters in each model. All this increases the chances of each model to match the data.

We estimate and test the models with the method of simulated moments and discuss various factors influencing the stability of the stock market under learning.

4.1 Risk aversion

We now introduce risk aversion in both models and habit persistence in consumption in the RE model only. The asset pricing literature under RE shows that these features improve the chances of the RE model to match the equity premium and to generate variability of the PD ratio. Moreover, by allowing for habit persistence we introduce an additional parameter in the utility function under RE. Since the learning model has also one additional free parameter (α_1) both models will have the same number of free model parameters.

Following Abel's (1990) extension of Lucas (1978) we consider a representa-

⁴¹Since PD is constant under RE, the coefficients c_1 of the predictability equation are undefined. This is not the case for the R^2 values.

⁴²The interest rate (which in the learning model we just assume equal to the RE value) does not show any variability, but this model was not set out to do this and in the paper we will not try to explain variability of interest rates. The level of the PD ratio is not matched, but the discount factor was chosen to favor the RE model on this aspect of the model, the estimation section will allow different parameters for each model and then learning will do well. Surprisingly, the model with learning does generate an equity premium (of about 1% per year), even for the risk neutral case. We will not pursue this here, as our focus is on price volatility, but this is an issue that we will take up later on.

tive consumer-investor solving

$$\begin{aligned} \max_{\{S_t, C_t\}} \quad & \tilde{E}_0 \sum_{t=0}^{\infty} \delta^t \frac{\left(\frac{C_t}{C_{t-1}^\kappa}\right)^{1-\sigma} - 1}{1-\sigma} \\ \text{s.t.} \quad & P_t S_t + C_t = (P_t + D_t) S_{t-1} \end{aligned}$$

where C_t denotes consumption, S_t the agent's stock holdings at the end of period t , $\sigma \geq 0$ the coefficient of relative risk aversion and κ the habit parameter. Dividends are as before. The parameter $\kappa \geq 0$ regulates the weight given to the past consumption, the habit is external to the agent.

4.2 Learning

In the model under learning we set $\kappa = 0$. The investor's first-order conditions, and the assumption (as in the previous section) that agents know the conditional expectations of dividends deliver the asset pricing equation

$$P_t = \delta \tilde{E}_t \left(\left(\frac{C_t}{C_{t+1}} \right)^\sigma P_{t+1} \right) + \delta E_t \left(\frac{D_t^\sigma}{D_{t+1}^{\sigma-1}} \right) \quad (14)$$

For the risk-neutral case ($\sigma = 0$) this simplifies to equation (5) studied in the previous section.

We now generalize also the learning scheme in order to give it a chance to be asymptotically rational. For this purpose, we start by analyzing the RE solution. For general risk aversion, and using the market clearing condition $C_t = D_t$ it is easy to see that RE stock prices are given by⁴³

$$\begin{aligned} P_t^{RE} &= \frac{\delta \beta^{RE}}{1 - \delta \beta^{RE}} D_t \\ \beta^{RE} &= a^{1-\sigma} e^{-\sigma(1-\sigma)\frac{\sigma^2}{2}} \end{aligned} \quad (15)$$

From equation (14) follows that agents have to forecast $\left(\frac{C_t}{C_{t+1}}\right)^\sigma P_{t+1}$ and given that the RE solution this implies

$$E_t \left(\left(\frac{C_t}{C_{t+1}} \right)^\sigma P_{t+1}^{RE} \right) = \beta^{RE} P_t^{RE}$$

It is thus natural to specify the learning mechanism with expectation functions

$$\beta_t P_t = \tilde{E}_t \left(\left(\frac{C_t}{C_{t+1}} \right)^\sigma P_{t+1} \right) \quad (16)$$

⁴³To show this, note that

$$\beta^{RE} = E_t \left(\left(\frac{C_t}{C_{t+1}} \right)^\sigma \frac{P_{t+1}^{RE}}{P_t^{RE}} \right) = E_t \left((a\varepsilon_{t+1})^{1-\sigma} \right) = a^{1-\sigma} e^{-\sigma(1-\sigma)\frac{\sigma^2}{2}}$$

where β_t is agents' best estimate of $E\left(\frac{C_t}{C_{t+1}}\right)^\sigma \frac{P_{t+1}}{P_t}$ which is interpreted as *risk-adjusted expected* stock price growth. Therefore, it is natural to write

$$\beta_t = \beta_{t-1} + \frac{1}{\alpha_t} \left[\left(\frac{C_{t-2}}{C_{t-1}} \right)^\sigma \frac{P_{t-1}}{P_{t-2}} - \beta_{t-1} \right] \quad (17)$$

The gain sequence is unchanged from the previous section. Given the form of the RE equilibrium these assumptions give a chance for the learning scheme thus written to be asymptotically rational. Appendix D shows that the learning scheme globally converges to RE, i.e., $\beta_t \rightarrow \beta^{RE}$ a.s.

Using (16), (14) and the fact that $E_t\left(\frac{D_t^\sigma}{D_{t+1}^{\sigma-1}}\right) = \beta^{RE} D_t$ gives

$$P_t = \frac{\delta \beta^{RE}}{1 - \delta \beta_t} D_t \quad (18)$$

$$\frac{P_t}{P_{t-1}} = \left(1 + \frac{\delta \Delta \beta_t}{1 - \delta \beta_t} \right) a \varepsilon_t \quad (19)$$

Now we should study the map T from perceived to actual expectations of the risk-adjusted price growth $\frac{P_{t+1}}{P_t} \left(\frac{C_t}{C_{t+1}} \right)^\sigma$. Using (19) and market clearing $C_t = D_t$ we have:⁴⁴

$$T(\beta_{t+1}, \Delta \beta_{t+1}) \equiv \beta^{RE} + \frac{\beta^{RE} \delta \Delta \beta_{t+1}}{1 - \delta \beta_{t+1}} \quad (20)$$

Clearly, this mapping T maintains all the features discussed in the previous section: we have momentum, non-linear behavior, etc. The only difference is that risk aversion $\sigma > 0$ changes the value of the limit point β^{RE} relative to the asymptote δ^{-1} . It is well known that, for σ sufficiently large, β^{RE} as well as the variance of realized risk-adjusted stock price growth under RE are increasing with σ .⁴⁵ This means that, to the extent that β_t tends to be around β^{RE} and this is closer to δ^{-1} , it is more likely that β_t will be near the asymptote and the instability under learning is even higher.

Another effect of risk aversion is that it is now the term $\left(\frac{D_{t-1}}{D_{t-2}} \right)^{-\sigma} \frac{P_{t-1}}{P_{t-2}}$ which changes the beliefs of agents in each period. This term is likely to have a larger variance than in the risk neutral case, since it also depends on ε_{t-1} . A large variance of $\left(\frac{D_{t-1}}{D_{t-2}} \right)^{-\sigma} \frac{P_{t-1}}{P_{t-2}}$ implies that a small realization of ε has a bigger chance of causing a large change in β_t and to deviate from the limiting

⁴⁴To see this, note that $T(\beta_{t+1}, \Delta \beta_{t+1}) \equiv E_t\left(\frac{P_{t+1}}{P_t} \left(\frac{C_t}{C_{t+1}}\right)^\sigma\right) = E_t\left(\left(1 + \frac{\delta \Delta \beta_{t+1}}{1 - \delta \beta_{t+1}}\right) (a \varepsilon_{t+1})^{1-\sigma}\right) = \beta^{RE} + \frac{\beta^{RE} \delta \Delta \beta_{t+1}}{1 - \delta \beta_{t+1}}$

⁴⁵For the parameter values of this paper, β^{RE} increases with σ as long as $\sigma > \approx 3$.

value. It is well known that, for σ sufficiently large, the variance of realized risk-adjusted stock price growth under RE are increasing with σ .⁴⁶

We conclude that, qualitatively, the main features of the model under learning are likely to remain after risk aversion is introduced.

4.3 RE model with habit persistence

Models of learning are often criticized because they add too many degrees of freedom. Indeed, by introducing learning we have a new free parameter in the model (namely, the precision on the initial prior given by α_1). To give the RE model an equal number of degrees of freedom we allow a free value for the habit parameter κ .

This model is well known to be able to replicate the equity premium and to have a variable PD ratio.

$$\frac{P_t^{RE}}{D_t} = A (a\varepsilon_t)^{\kappa(\sigma-1)} \quad (21)$$

for a certain constant A . Details are given in appendix A. It is clear that now the PD ratio has some variability, although it will not display serial correlation. Clearly, this is not the best model that can be found in the RE literature to match the above mentioned facts. Results for the RE model should thus be understood as an illustration only.

4.4 Method of Simulated Moments

We give a detailed account of the econometric procedure in Appendix C, but give an overview here. We estimate and test both models adapting the method of simulated moments (MSM) to take care of short samples. We find parameter values that match some of the asset price statistics listed in tables 1 and 2 as closely as possible. The measure of ‘closeness’ is a quadratic form with a weighting matrix that estimates the variance covariance matrix of the moments matched. As usual in MSM, the value of this distance at the minimum provides a test of the model.

We deviate from the standard practice in MSM in two ways: first, we match the data to *short sample statistics* generated by the model, as opposed to the usual practice of using the long run moments. More precisely, given a model, we draw many histories of 288 observations from the model, compute the statistic at hand for each history, and we compute the relevant simulated moments from

⁴⁶The formula for the variance is

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

This variance reaches a minimum for $\sigma = 1$.

the distribution of this statistic across realizations. This is computationally more intensive, but the usual practice of looking at long run moments from the data is not appropriate in our case, since the learning model converges to RE so that the asymptotic moments of the model under learning do not allow to distinguish between RE and learning. Also, our procedure has a better chance of capturing any short sample bias that may be present in the calculations of the statistics.

The second adaptation concerns the weighting matrix that is used in the quadratic form that defines the distance of simulated to actual moments. Usually, this matrix consists of the inverse of an estimator of the infinite sum of autocovariances of the moments (the ‘ S_w ’ matrix) and is estimated from the autocovariances in the data. This matrix is very difficult to estimate, mostly because of the presence of an infinite sum that has to be truncated or approximated. Several possible estimates have been designed for this purpose. Instead, we use the autocovariances computed from the distribution across realizations in the short samples generated by the model. This avoids approximations of the infinite sum involved in S_w and, in addition, captures any possible short sample bias.

These two modifications are irrelevant asymptotically. This procedure is thus as well grounded on asymptotic theory as common practice, but they are likely to capture the true short-sample properties of the model much better than the asymptotic moments, they allow to distinguish between RE and learning, and they are likely to give a better estimate of the S_w matrix.

For the learning model the parameter vector to be estimated is $\theta = (\delta, \sigma, \alpha_1, a, s)$, for the RE model $\theta = (\delta, \sigma, a, s, \kappa)$ so for both models the number of parameters is $n = 5$. Note that since we now estimate the parameters of the dividend process (a, s) the estimates will not match exactly the actual observed values as we did in section 3, but the econometric procedure will find the point estimates that help explain the overall observed moments.

We choose to match the following statistics

$$Eh(y_t) = \begin{bmatrix} E(r^s) \\ E(r^B) \\ E(PD) \\ E\left(\frac{\Delta D}{D}\right) \\ \rho(PD_t, PD_{t-1}) \\ \sigma_{rs} \\ \sigma_{PD} \\ \sigma_{\frac{\Delta D}{D}} \\ c_1^{10} \\ R_{10}^2 \end{bmatrix}$$

This is a summary of the statistics that the literature has considered relevant in terms of the facts 1 to 4 described in section 2. It basically includes the statistics reported in table 1 plus the coefficient and R^2 at ten years reported in Table 2 (we do not include all coefficients and R^2 's to economize on computation time).

5 Estimation Results

Table 5 below shows the estimated parameter values that set the simulated moments as close as possible to the actual observed values for these statistics. Parameter estimates appear reasonable on a priori ground for both the RE and the learning model. For the learning model the weight on the initial belief (α_1) reflects the tendency of the data to give a large but finite weight to the initial belief being equal to RE. The risk aversion parameters are relatively high but within the ranges that have been used in many studies. The parameter values for the dividend process change slightly from the case where mean and standard deviations of dividend growth were matched perfectly as in (13). The habit parameter for the RE case is very high compared to other estimates in the literature.

	Learning model	RE model (with habits)
a	0.355	0.380
s	3.65	3.40
δ	0.996	0.993
σ	4.9	6.0
α_1	70	-
κ	-	0.8

Table 5: Estimated model parameters

Table 6 below summarizes the goodness of fit of each model. We report the average and standard deviation for each statistic (with $N = 288$ observations) implied by the model with parameters given by the point estimates in the previous table.

Let us first concentrate on the RE column. The PD ratio now has some variation, but the model clearly fails to match its serial correlation (not surprisingly, given equation (21)). The variance of PD is very small. As is well known this model can match the equity premium, but to do so the variance of stock returns and interest rates has to be very high. Actually, for the above estimation, the equity premium is overpredicted. It appears that the estimation procedure selected a very high value of κ to try and match the high variance of PD, but in so doing it generated a very large variance of returns and an equity premium too large. The model had the potential to show excess return predictability, since

both future returns current PD depend on today 's innovation to the dividend, but it turns out that the model fails to match the predictability

The learning model, however, performs very well. The model with risk aversion maintains the high variability and serial correlation of PD as in section 3, but in addition now it matches the equity premium. The point estimate of some model moments is not exactly like the observed moment, but this tends to occur for moments that, in the short sample, have a large variance. This happens because the estimation procedure optimally gives less importance to matching exactly high variance moments. Still, we see that the observed moment values are always within one standard deviation of the estimated value.

Finally, the last two lines in the table report the results of testing the overidentifying restrictions. This is an overall measure of how well the model matches the selected moments. The RE model has a huge value for this statistic, implying a p-value of zero (almost up to machine precision). On the other hand, the model under learning is accepted at the 5% level and marginally rejected at 10% (one-sided confidence intervals).

	U.S. data	Learning model	RE model (with habits)
$E(r^s)$	2.36	2.47 (0.34)	3.70 (0.34)
$E(r^B)$	0.16	0.21 (0.22)	0.20 (0.83)
$E(PD)$	105.4	98.6 (36.7)	105.4 (0.88)
$E\left(\frac{\Delta D}{D}\right)$	0.346	0.371 (0.213)	0.377 (0.210)
σ_{r^s}	11.5	14.0 (3.7)	22.7 (1.3)
σ_{PD}	35.4	67.9 (29.0)	14.4 (0.65)
$\sigma_{\frac{\Delta D}{D}}$	3.63	3.66 (0.14)	3.41 (0.14)
$\rho(PD_t, PD_{t-1})$	0.95	0.94 (0.02)	-0.00 (0.06)
Excess returns predictability:			
Coefficient on PD (10 yrs)	-0.0267	-0.0142 (0.0079)	-0.0066 (0.0211)
R^2 (10 yrs)	0.46	0.36 (0.16)	0.00 (0.01)
Test statistic overident. restr.	-	9.54	$4.4 \cdot 10^4$
p-value	-	0.09	0.00

Table 6: Data, model moments and goodness of fit

The summary is, clearly, that introducing learning generates an enormous improvement in the fit of the model. This, despite the fact that we used the simplest version of the asset pricing model with the simplest learning mechanism. Notice that the estimation tells the model to use a learning scheme that does not deviate too much from rationality, since the estimated confidence in the initial beliefs (centered at the fundamental RE value) is very high.

This goodness of fit of the learning model is very robust. Changing the parameters considerably does not change the behavior of the model drastically, and from eyeball inspection of the simulations the variables in the model roughly behave in a similar way as the data.

6 Conclusions

The failure of equilibrium asset pricing models under RE to account for basic moments of the data has been well documented. Introducing learning in a simple asset pricing model generates asset pricing dynamics that are much more in line with the empirical behavior of stock prices. Since learning-induced deviations from rational expectations are small, the results of this paper show that even slight non-rationalities in expectations can have large implications for the behavior of asset prices. This has been accomplished with only minor model modifications: we just introduced a simple learning mechanism in a simple asset pricing model. Key to our results is the assumption that agents care about future prices, so that expectations in the model influence price movements and these feed back into expectations.

The magnitude of the improvement achieved by introducing learning is very large. The model is accepted in a formal test under the method of simulated moments; that a dynamic equilibrium model of asset prices survives formal econometric testing when matching so many moments is, to say the least, uncommon in the literature.

This large improvement was not achieved by introducing many degrees of freedom. The model under learning has the same number of parameters as a the basic RE model with habit persistence. The choice of learning scheme is far from arbitrary, since least squares learning is known to have a number of desirable features. In our formulation, this learning scheme can be interpreted as a small departure from RE for two reasons: *i*) initial beliefs are assumed to be at the RE and agents have high confidence in this RE value and *ii*) the learning scheme is asymptotically rational: in the long run agents would realize that their forecasts are as good as those of someone who knew the whole model. Therefore, in the long run agents would have no incentive to deviate from their learning scheme.

The work shown in this paper can be improved in many ways. We wanted our model economy to be as close as possible to the standard literature. In doing this, the model has a number of weak points. One weak point is that the rationality bounds along the transition, as they were formally defined in Marcet and Nicolini (2003), are currently not satisfied. We know of various changes to the model that would deliver these bounds, but this seems an issue to be taken up subsequently.

Also, it turns out that prices in our model are very sensitive to changes in expectations. This is in part what allowed to match the data, but the impression is that prices are ‘too’ sensitive to expectations in the model. Related to this is the fact that if expectations are higher than a certain bound (δ^{-1}) there is no positive price that clears the market and expectations have to be sent back below this bound. In part, the reason for this sensitivity is due to the homogeneous agent assumption: under this assumption no agent ever sells a stock, so the actual price is, in a way, ‘irrelevant’. There are a number of features that can be introduced in the model to make prices adjust less quickly, such as agents that have to sell stocks at some points in time, or financial frictions. We are exploring various alternatives in this direction.

Also to be explored is the relationship to monetary policy. RE models are also not very rich in terms of the interactions predicted between market volatility and various other aspects of the economy such as the conduct of monetary policy, the degree of investors’ risk aversion, or the presence of speculative investors with short investment horizons. Under learning low real interest rates are likely to increase stock price volatility, since the asymptote of the T map will be closer to the long run value of the beliefs. Speculative investors, to the extent that they care less about dividends and more about prices, act in a similar way and they make the asymptote dangerously close to long run beliefs. A model with learning thus suggests a different role for monetary policy and investors’ risk attitude that seems to be consistent with views generally expressed by central bankers, e.g., Papademos (2005).

What does our model say about the long run behavior of stock prices?. It predicts doom: our model is perfectly consistent with stock prices that for many periods have a very high growth rate and PD ratios much higher than RE. But in the long run it converges to RE, so that PD and stock price growth converge to their "fundamental" RE value. Therefore, stock price growth in the long will be much lower than during the transition. If ours is the right model, given that PD is currently so high compared to its historical values, stockholders will do well to stay away from stocks. Of course, the observed behavior may be explained by other alternatives that do not predict doom, for example, a change of trend in dividends. It is of interest, we think, to try and extract as much information as possible from actual data to see the possible evolution of stock prices in the long run by comparing the behavior of these models under learning.

A RE model with risk-aversion and habits

Under RE, in the habits model, the investor's first-order conditions and the market clearing condition $C_t = D_t$ deliver the asset pricing equation

$$P_t = \delta E_t \left(\left(\frac{D_t^{\kappa(\sigma-1)}}{D_{t+1}^\sigma} \right) \left(\frac{D_t^\sigma}{D_{t-1}^{\kappa(\sigma-1)}} \right) (P_{t+1} + D_{t+1}) \right)$$

Together with the process for dividends (1) this implies that under rational expectations

$$E \left(\frac{P_t + D_t}{P_{t-1}} \right) = a \left(E(\varepsilon_t^{1+\kappa(\sigma-1)}) + A^{-1} \right) E(a\varepsilon_t)^{-\kappa(\sigma-1)} \quad (23a)$$

$$E(R_t) = \delta^{-1} \frac{E(a\varepsilon_t)^{-\kappa(\sigma-1)}}{E(a\varepsilon_t)^{-\sigma}} \quad (23b)$$

$$A = \frac{\delta E(a\varepsilon_t)^{1-\sigma}}{1 - \delta E(a\varepsilon_t)^{(\kappa-1)(\sigma-1)}} \quad (23c)$$

B Model with learning about dividends

We now assume that agents learn to forecast future dividends in addition to learning how to forecast future price. We directly consider the general model with risk-aversion from section 4.2. With learning about future dividends and future price equation (14) becomes

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

Under RE one has

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

This justifies that learning agents will forecast future dividends according to

$$\tilde{E}_t \left(\frac{D_{t+1}^{1-\sigma}}{D_t^\sigma} \right) = \gamma_t D_t$$

where γ_t is agents's best estimate of $\tilde{E}_t \left(\left(\frac{D_{t+1}}{D_t} \right)^{1-\sigma} \right)$, 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

$$\gamma_t = \gamma_{t-1} + \frac{1}{\alpha_t} \left(\frac{D_{t-1}^{1-\sigma}}{D_{t-2}^{1-\sigma}} - \gamma_{t-1} \right) \quad (24)$$

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

$$\gamma_0 = \beta^{RE}.$$

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 (17). With these assumptions realized price and price growth are

$$P_t = \frac{\delta \gamma_t}{1 - \delta \beta_t} D_t$$

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

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

$$T(\beta_{t+1}, \Delta \beta_{t+1}) \equiv \frac{\gamma_{t+1}}{\gamma_t} \left(\beta^{RE} + \frac{\beta^{RE} \delta \Delta \beta_{t+1}}{1 - \delta \beta_{t+1}} \right)$$

which differs from (20) only by the factor $\frac{\gamma_{t+1}}{\gamma_t}$. From (24) it is clear that $\frac{\gamma_{t+1}}{\gamma_t}$ evolves exogenously and that $\lim_{t \rightarrow \infty} \frac{\gamma_{t+1}}{\gamma_t} = 1$ since $\lim_{t \rightarrow \infty} \gamma_t = \beta^{RE}$ and $\alpha_t \rightarrow \infty$. 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. Simulating the learning model with dividend learning for the estimated learning model from section 5 reveals that the models with and without learning produce essentially identical asset price statistics.⁴⁷ This is shown in table 7 below.

⁴⁷To compute bond returns in the case with dividend learning we assume (in close analogy to the other learning setups) that

$$\tilde{E}_t \left[\left(\frac{D_{t+1}}{D_t} \right)^{-\sigma} \right] = \phi_t$$

where

$$\phi_t = \phi_{t-1} + \frac{1}{\alpha_t} \left(\frac{D_{t-1}^{-\sigma}}{D_{t-2}^{-\sigma}} - \phi_{t-1} \right)$$

$$\phi_0 = E_t \left[\left(\frac{D_{t+1}}{D_t} \right)^{-\sigma} \right]$$

$$= E \left[(a \varepsilon_t)^{-\sigma} \right] = a^{-\sigma} e^{\sigma(1+\sigma) \frac{\sigma^2}{2}}$$

The gross real bond return from t to $t+1$ is then given by $(\delta \phi_t)^{-1}$.

	Learning model with RE about dividends	Learning model with dividend learning
$E(r^s)$	2.47 (0.34)	x
$E(r^B)$	0.21 (0.22)	x
$E(PD)$	98.6 (36.7)	x
$E\left(\frac{\Delta D}{D}\right)$	0.371 (0.213)	x
σ_{r^s}	14.0 (3.7)	x
σ_{PD}	67.9 (29.0)	x
$\sigma_{\frac{\Delta D}{D}}$	3.66 (0.14)	x
$\rho(PD_t, PD_{t-1})$	0.94 (0.02)	x
Excess returns predictability:		
Coefficient on PD (10 yrs)	-0.0142 (0.0079)	x
R^2 (10 yrs)	0.36 (0.16)	x

Table 7: Learning model with and without dividend learning

C Short Sample MSM

We use the simulated method of moments to estimate models adapting it to match short-sample moments.

Let N be the sample size, and $(\mathbf{y}_1, \dots, \mathbf{y}_N)$ the observed sample, with \mathbf{y}_t containing m variables. Let $h : R^m \rightarrow R^q$ be a moment function, giving the moments to be matched, and let M_N be the sample moments observed from the data:

$$M_N \equiv \frac{1}{N} \sum_{t=1}^N h(\mathbf{y}_t)$$

Let $\theta \in R^n$ denote a vector of possible model parameter values to be estimated. Let ω^s denote a realization of shocks and denote $(y_1(\theta, \omega^s), \dots, y_N(\theta, \omega^s))$ the random variables corresponding to a history of length N generated by the model for a realization ω^s . Define the moments from the model as

$$M_N(\theta) \equiv \widehat{E} \left(\frac{1}{N} \sum_{t=1}^N h(y_t(\theta)) \right)$$

where \widehat{E} is obtained from replicating a large number (S) of histories of length N , computing the moment $\frac{1}{N} \sum_{t=1}^N h(y_t(\theta), \omega^s)$ for each history, and averaging over all replications. Formally,

$$\widehat{E} \left(\frac{1}{N} \sum_{t=1}^N h(y_t(\theta)) \right) \equiv \frac{1}{S} \sum_{s=1}^S \left(\frac{1}{N} \sum_{t=1}^N h(y_t(\theta, \omega^s)) \right)$$

Notice that we deviate from the usual practice in MSM, since the usual practice involves matching observed moments to unconditional moments generated by the model in the long run, so that \widehat{E} is usually computed by averaging over one very long observation. Of course, in this setup, initial conditions have to be specified, either as a constant that has been observed (this would be the case, for example, in a growth model with fixed initial capital where the capital is observed) or as a coefficient to be estimated and, therefore, to be included in θ (this is the learning model of this paper, where the initial values for the constant gain has to be estimated).

The estimator we use is, as usual, in two steps. First, we first use some initial weighting matrix $\widetilde{\Omega}$, which is just required to be positive definite, to find an initial (asymptotically inefficient) estimator $\widetilde{\theta}$

$$\widetilde{\theta} = \arg \min_{\theta} (M_N(\theta) - M_N)' \widetilde{\Omega}^{-1} (M_N(\theta) - M_N) \quad (25)$$

Then, we let $\Omega_N(\theta)$ be the variance covariance matrix of $M_N(\theta)$:

$$\Omega(\theta) \equiv \widehat{E} \left[\left(\frac{1}{N} \sum_{t=1}^N h(y_t(\theta)) - M_N(\theta) \right) \left(\frac{1}{N} \sum_{t=1}^N h(y_t(\theta)) - M_N(\theta) \right)' \right]$$

where, again, \widehat{E} is obtained by averaging $\left[\frac{1}{N} \sum_{t=1}^N h(y_t(\theta)) - M_N(\theta) \right] \left[\frac{1}{N} \sum_{t=1}^N h(y_t(\theta)) - M_N(\theta) \right]'$ over S replications. The inverse of this matrix, evaluated at the initial estimate, gives an optimal weighting matrix. This is the second departure from the usual practice: here we just compute "directly " the variance of the moments implied by the model, instead of first estimating some autocovariances, then adding up over some lags, and weighting each autocovariance as would be done, for example, in the Newey-West procedure.

Finally, our estimator is defined as

$$\widehat{\theta}_N = \arg \min_{\theta} (M_N(\theta) - M_N)' \Omega_N(\widetilde{\theta})^{-1} (M_N(\theta) - M_N) \quad (26)$$

we can be certain we use optimally (asymptotically) the instruments.

Therefore, this differs from standard MSM in two ways:

1. Usually, the simulated moments \widehat{E} are computed from long run averages, intended to estimate the unconditional moment in the long run, i.e., with the steady state distribution. By computing \widehat{E} with (numerical means) of sample averages we are considering the effects of the transition, crucial in our model of learning, and we may take care of some short sample distribution biases that may be present in the estimation.

2. The optimal weighting matrix $\Omega(\theta)$ is not found by averaging autocorrelations at different lags, but by computing the variance (numerically) of the of statistics. This avoids truncating the sum and having to apply some HAC estimator and, again, it takes care of the short sample transition.

Of course, these changes do not affect the asymptotic validity of the estimator:

Using standard argument one can show (I hope) that:

- $\hat{\theta}_N \rightarrow \theta_0$ a.s. as $N \rightarrow \infty$
- $\hat{\theta}_N$ is efficient among all MSM estimators for any initial weighting matrix $\tilde{\Omega}$
- $(M_N(\hat{\theta}_N) - M_N)' \Omega_N(\tilde{\theta})^{-1} (M_N(\hat{\theta}_N) - M_N) (1 + \frac{1}{S})^{-1} \rightarrow \chi_{q-n}^2$ in distribution as $N \rightarrow \infty$, where S is the number of replications used in computing the simulated moments \tilde{E} .

To obtain the minima, we first simulate the learning model on a coarse parameter grid $\theta \in [0 : 0.5 : 5] \times [0.986 : 0.001 : 0.996] \times [50 : 25 : 125, 150 : 50 : 300] \times [0.31 : 0.01 : 0.38] \times [3.4 : 0.1 : 3.8]$ where $\theta = (\sigma, \delta, \alpha_1, E(\frac{\Delta D}{D}), \sigma_{\frac{\Delta D}{D}})$. Using results from the coarse grid we then refine the grid to $[4 : 0.1 : 5] \times [0.990 : 0.001 : 0.998] \times [50 : 10 : 120] \times [0.345 : 0.005 : 0.375] \times [3.6 : 0.05 : 3.8]$. At each gridpoint we compute the mean of the considered moments $M_N(\theta)$ and the moment covariance matrix $\Omega(\theta)$ using $S=1000$ simulations of $N=288$ model periods each, i.e., the length of our empirical sample. The initial weighting matrix is $\tilde{\Omega} = \Omega(\underline{\theta})$ where $\underline{\theta} = \arg \min_{\theta} (M_N - M_N(\theta))' \Omega^{-1}(\theta) (M_N - M_N(\theta))$.

It is a good idea to match average bond returns, since this pins down the discount factor. But since we simplified our model by assuming no variation of interest rates, bond returns in the learning model are constant over time, which implies a singular moment matrix due to the zero variation of interest rates in the model. There are several alternatives to correct for this problem. We assume a small measurement error ME for average bond returns and impose it on the corresponding diagonal entry in the moment matrix. The standard error of ME is set equal to the standard error of the estimated mean bond return in the data,

$$\text{i.e., } std(ME) = std\left(\frac{1}{T} \sum_{j=1}^T r_j^B\right) \approx \sqrt{\frac{1}{T} \left(\sum_{j=-10}^{10} \frac{1}{T} cov(r_t^B, r_{t-10}^B) \right)} = 0.22\%,$$

where T denotes the sample length.

The test for overidentifying restrictions has 5 degrees of freedom (the number of moments 10 minus the number of estimated parameters 5).

When estimating the rational expectations model with habits, we proceed as above, except that now $\theta = (\sigma, \delta, \kappa, E(\frac{\Delta D}{D}), \sigma_{\frac{\Delta D}{D}})$ and the grid is given by $[0 : 0.5 : 6] \times [0.988 : 0.001 : 0.996] \times [0.1 : 0.1 : 0.9] \times [0.34 : 0.01 : 0.38] \times [3.4 : 0.1 : 3.8]$.

D Convergence of least squares to RE

We show convergence directly for the general learning model with risk aversion from section 4.2. To obtain convergence we need bounded shocks. In particular, we assume existence of some $U^\varepsilon < \infty$ such that

$$\begin{aligned}\text{Prob}(\varepsilon_t < U^\varepsilon) &= 1 \\ \text{Prob}(\varepsilon_t^{1-\sigma} < U^\varepsilon) &= 1\end{aligned}$$

Furthermore, we assume that the projection facility is not binding in the RE equilibrium

$$\frac{P_t^{RE}}{D_t} = \frac{\delta\beta^{RE}}{1 - \delta\beta^{RE}} < U^{PD}$$

where $\beta^{RE} = E[(a\varepsilon_t)^{1-\sigma}]$ and P_t^{RE} is the price in the RE equilibrium.

Since price growth in temporary equilibrium is determined by two lags of β , the adaptation of the stochastic control framework of Ljung (1977) by Marcat and Sargent (1989) or Evans and Honkapohja (2001) is not applicable.⁴⁸ Therefore, we provide a separate proof which proceeds in two steps. First, we show that the projection facility will almost surely cease to be binding after some finite time. In a second step, we show that β_t converges to β^{RE} from that time onwards.

The projection facility implies

$$\beta_t = \begin{cases} \beta_{t-1} + \alpha_t^{-1} \left((a\varepsilon_{t-1})^{-\sigma} \frac{P_{t-1}}{P_{t-2}} - \beta_{t-1} \right) & \text{if } \frac{\delta a}{1 - \delta \left(\beta_{t-1} + \alpha_t^{-1} \left((a\varepsilon_{t-1})^{-\sigma} \frac{P_{t-1}}{P_{t-2}} - \beta_{t-1} \right) \right)} < U^{PD} \\ \beta_{t-1} & \text{otherwise} \end{cases} \quad (27)$$

⁴⁸It may be possible to adapt Ljung's theorem to this case, but it is not immediate how this can be done. The technical problem is the following. Since P/P_{-1} depends on two lags of β we would need to study convergence of the parameter $\gamma_t \equiv (\beta_t, \beta_{t-1})$. We then have that the law of motion of observables satisfies

$$\frac{P_t}{P_{t-1}} = T(\gamma_t)\varepsilon_t$$

which is a special case of the laws of motion considered in Ljung (1977). The stochastic control formulation assume the following law of motion for γ_t :

$$\gamma_t = \gamma_{t-1} + \alpha_t^{-1} Q(\gamma_{t-1}, \frac{P_t}{P_{t-1}}, t)$$

This formulation is consistent with the definition of γ if Q in the second row insures that $\gamma_{2,t} = \gamma_{1,t-1}$, which requires

$$Q_2(\gamma_{t-1}, \frac{P_t}{P_{t-1}}, t) \equiv \alpha_t(\gamma_{1,t-1} - \gamma_{2,t-1})$$

Yet, for fixed arbitrary γ we have $\alpha_t(\gamma_1 - \gamma_2) \rightarrow \infty$ violating the key condition in Ljung that this limit has to be well-defined. Therefore, the convergence theorems of Ljung are not directly applicable in this formulation.

If the lower equality applies one has $(a\varepsilon_{t-1})^{-\sigma} \frac{P_{t-1}}{P_{t-2}} \geq \beta_{t-1}$ and this gives rise to the following inequalities

$$\beta_t \leq \beta_{t-1} + \alpha_t^{-1} \left((a\varepsilon_{t-1})^{-\sigma} \frac{P_{t-1}}{P_{t-2}} - \beta_{t-1} \right) \quad (28)$$

$$|\beta_t - \beta_{t-1}| \leq \alpha_t^{-1} \left| \left((a\varepsilon_{t-1})^{-\sigma} \frac{P_{t-1}}{P_{t-2}} - \beta_{t-1} \right) \right| \quad (29)$$

which hold for all t . Substituting recursively in (28) for past β 's delivers

$$\begin{aligned} \beta_t &\leq \frac{1}{t-1+\alpha_1} \left(\sum_{j=0}^{t-1} (a\varepsilon_j)^{-\sigma} \frac{P_j}{P_{j-1}} + (\alpha_1 - 1) \beta_0 \right) \\ &= \underbrace{\frac{t}{t-1+\alpha_1} \left(\frac{1}{t} \sum_{j=0}^{t-1} (a\varepsilon_j)^{1-\sigma} + \frac{\alpha_1 - 1}{t} \beta_0 \right)}_{=T_1} + \underbrace{\frac{1}{t-1+\alpha_1} \left(\sum_{j=0}^{t-1} \frac{\delta \Delta\beta_j}{1-\delta\beta_j} (a\varepsilon_j)^{1-\sigma} \right)}_{=T_2} \end{aligned} \quad (30)$$

where the second line follows from (19). Since $T_1 \rightarrow \beta^{RE}$ for $t \rightarrow \infty$ a.s., β_t will eventually be bounded away from its upper bound if we can establish $|T_2| \rightarrow 0$ a.s.. This is achieved by noting that

$$\begin{aligned} |T_2| &\leq \frac{1}{t-1+\alpha_1} \sum_{j=0}^{t-1} \frac{\delta (a\varepsilon_j)^{1-\sigma}}{1-\delta\beta_j} |\Delta\beta_j| \\ &\leq \frac{U^\varepsilon}{t-1+\alpha_1} \sum_{j=0}^{t-1} \frac{a^{1-\sigma} \delta |\Delta\beta_j|}{1-\delta\beta_j} \\ &\leq \frac{U^\varepsilon}{t-1+\alpha_1} \frac{a^{1-\sigma} U^{PD}}{\beta^{RE}} \sum_{j=0}^{t-1} |\Delta\beta_j| \end{aligned} \quad (31)$$

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 the bound on the price dividend ratio insuring that $\delta\beta^{RE} (1-\delta\beta_j)^{-1} < U^{PD}$. Next, observe that

$$(a\varepsilon_j)^{-\sigma} \frac{P_t}{P_{t-1}} = \frac{1-\delta\beta_{t-1}}{1-\delta\beta_t} (a\varepsilon_t)^{1-\sigma} < \frac{(a\varepsilon_t)^{1-\sigma}}{1-\delta\beta_t} < \frac{a^{1-\sigma} U^\varepsilon U^{PD}}{\delta\beta^{RE}} \quad (32)$$

where the equality follows from (18), the first inequality from $\beta_{t-1} > 0$, and the second inequality from the bounds on ε and PD . Using result (32), equation (29) implies

$$|\beta_t - \beta_{t-1}| \leq \alpha_t^{-1} \left| (a\varepsilon_t)^{-\sigma} \frac{P_{t-1}}{P_{t-2}} - \beta_{t-1} \right| \leq \alpha_t^{-1} \left(\frac{a^{1-\sigma} U^\varepsilon U^{PD}}{\delta\beta^{RE}} + \delta^{-1} \right)$$

where the second inequality follows from the triangle inequality and the fact that $\beta_{t-1} < \delta^{-1}$. Since $\alpha_t \rightarrow \infty$ this establishes that $|\Delta\beta_t| \rightarrow 0$ and, therefore, $\frac{1}{\bar{t}-1+\alpha_1} \sum_{j=0}^{t-1} |\Delta\beta_j| \rightarrow 0$. Then (31) implies that $|T_2| \rightarrow 0$ a.s. as $t \rightarrow \infty$. By taking the lim sup on both sides on (30), it follows from $T_1 \rightarrow \beta^{RE}$ and $|T_2| \rightarrow 0$ that

$$\limsup_{t \rightarrow \infty} \beta_t \leq \beta^{RE}$$

a.s.. The projection facility is thus operative infinitely often with probability zero. Therefore, there exists a set of realizations ω with measure one and a $\bar{t} < \infty$ (which depends on the realization ω) such that the projection facility does not operate for $t > \bar{t}$.

We now proceed with the second step of the proof. Consider, for a given realization ω , a \bar{t} for which the projection facility is not operative after this period. Then the upper equality in (27) holds for all $t > \bar{t}$ and simple algebra gives

$$\begin{aligned} \beta_t &= \frac{1}{t - \bar{t} + \alpha_{\bar{t}}} \left(\sum_{j=\bar{t}}^{t-1} (a\varepsilon_j)^{-\sigma} \frac{P_j}{P_{j-1}} + \alpha_{\bar{t}} \beta_{\bar{t}} \right) \\ &= \frac{t - \bar{t}}{t - \bar{t} + \alpha_{\bar{t}}} \left(\frac{1}{t - \bar{t}} \sum_{j=\bar{t}}^{t-1} (a\varepsilon_j)^{1-\sigma} + \frac{1}{t - \bar{t}} \sum_{j=\bar{t}}^{t-1} \frac{\delta \Delta\beta_j}{1 - \delta\beta_j} (a\varepsilon_j)^{1-\sigma} + \frac{\alpha_{\bar{t}}}{t - \bar{t}} \beta_{\bar{t}} \right) \end{aligned} \quad (33)$$

for $t > \bar{t}$. Equations (28) and (29) now hold with equality for $t > \bar{t}$. Similar operations as before then deliver

$$\frac{1}{t - \bar{t}} \sum_{j=\bar{t}}^{t-1} \frac{\delta \Delta\beta_j}{1 - \delta\beta_j} (a\varepsilon_j)^{1-\sigma} \rightarrow 0$$

a.s. for $t \rightarrow \infty$. Finally, taking the limit on both sides of (33) establishes

$$\beta_t \rightarrow \beta^{RE}$$

a.s. as $t \rightarrow \infty$. ■

References

- ABEL, A. B. (1990): "Asset Prices under Habit Formation and Catching Up with the Joneses," *American Economic Review*, 80, 38–42.
- ADAM, K. (2005): "Experimental Evidence on the Persistence of Output and Inflation," CEPR Working Paper No. 4885, (forthcoming *Economic Journal*).
- 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.
- CAMPBELL, J. Y., AND R. J. SHILLER (1988): “Stock Prices, Earnings, and Expected Dividends,” *Journal of Finance*, 43, 661–676.
- CAMPBELL, J. Y., AND M. YOGO (2005): “Efficient Test of Stock Return Predictability,” *Harvard University mimeo*.
- CARCELES-POVEDA, E., AND C. GIANNITSAROU (2006): “Asset Pricing with Adaptive Learning,” SUNY Stony Brook and Cambridge University mimeo.
- 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.
- EVANS, G. W., AND S. HONKAPOHJA (2001): *Learning and Expectations in Macroeconomics*. Princeton University Press, Princeton.
- FAMA, E. F., AND K. R. FRENCH (1988): “Dividend Yields and Expected Stock Returns,” *Journal of Financial Economics*, 22, 3–25.
- LEROY, S. F., AND R. PORTER (1981): “The Present-Value Relation: Test Based on Implied Variance Bounds,” *Econometrica*, 49, 555–574.
- LJUNG, L. (1977): “Analysis of Recursive Stochastic Algorithms,” *IEEE Transactions on Automatic Control*, 22, 551–575.
- 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.

- MEHRA, R., AND E. C. PRESCOTT (1985): “The Equity Premium: A Puzzle,” *Journal of Monetary Economics*, 15, 145–161.
- MISHKIN, F. S., AND E. N. WHITE (2002): “U.S. Stock Market Crashes and Their Aftermath: Implications for Monetary Policy,” *NBER Working Paper No. 8992*.
- PAPADEMOS, L. (2005): “Interview with the Financial Times on 19 December 2005,” Available at <http://www.ecb.int/press/key/date/2005/html/sp051219.en.html>.
- POTERBA, J. M., AND L. S. SUMMERS (1988): “Mean Reversion on Stock Prices,” *Journal of Financial Economics*, 22, 27–59.
- SANTOS, M. S., AND M. WOODFORD (1997): “Rational Asset Pricing Bubbles,” *Econometrica*, 65, 19–57.
- SHILLER, R. J. (1981): “Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?,” *American Economic Review*, 71, 421–436.
- 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.
- WEIL, P. (1989): “The Equity Premium Puzzle and the Risk-Free Rate Puzzle,” *Journal of Monetary Economics*, 24, 401–421.
- WEITZMAN, M. L. (2005): “Risk, Uncertainty, and Asset-Pricing ‘Puzzles’,” *Harvard University mimeo*.