

# Stock Price Booms and Expected Capital Gains

By KLAUS ADAM, ALBERT MARCET AND JOHANNES BEUTEL\*

*Investors' subjective capital gains expectations are a key element explaining stock price fluctuations. Survey measures of these expectations display excessive optimism (pessimism) at market peaks (troughs). We formally reject the hypothesis that this is compatible with rational expectations. We incorporate subjective price beliefs with such properties into a standard asset-pricing model with rational agents (internal rationality). The model gives rise to boom-bust cycles that temporarily delink stock prices from fundamentals. It quantitatively replicates many asset-pricing moments, including the observed strong positive correlation between the price dividend ratio and survey return expectations, which cannot be matched by rational expectations.*

*JEL: G12, D84*

*Keywords: asset pricing, subjective beliefs, internal rationality*

*'Bull-markets are born on pessimism, grow on skepticism, mature on optimism and die on euphoria'*

Sir John Templeton, Founder of Templeton Mutual Funds

Following the recent boom and bust cycles in a number of asset markets around the globe, there exists renewed interest in understanding better the forces contributing to the emergence of such drastic asset price movements. This paper argues that movements in investor optimism and pessimism, as measured by the movements in investors' subjective

\* Adam: Department of Economics, University of Mannheim, L7, 3-5, 68131 Mannheim, Germany and CEPR, adam@uni-mannheim.de. Marcet: Institut d'Anàlisi Econòmica (CSIC), Universitat Autònoma de Barcelona, Campus Bellaterra, 08193 Barcelona, Spain and ICREA, MOVE, Barcelona GSE and CEPR, marcet.albert@gmail.com. Beutel: Deutsche Bundesbank, Wilhelm-Epstein-Straße 14, 60431 Frankfurt am Main, Germany, johannes.beutel@bundesbank.de. This paper replaces an earlier paper titled 'Booms and Busts in Asset Prices' by Adam and Marcet, which appeared 2010 as Bank of Japan - IMES Discussion Paper No. 2010-E-2. Thanks go to Fernando Alvarez, Michele Boldrin, Chryssi Giannitsarou, Vivien Lewis, Morten Ravn, Ricardo Reis, Mike Woodford, conference participants at the Banque de France and Chicago Fed Conference on Asset Price Bubbles, ESSIM 2010 in Tarragona, 2017 AEA Meetings, MONFISPOL network, and seminar participants at Columbia University, University College London, London School of Economics, Yale University, Harvard University, Northwestern University, Stony Brook, IMF, Mitsui Conference at Michigan University, Northwestern University, University of Chicago, New York University and London Business School for helpful comments and suggestions. Research assistance from Jeanine Baumert, Oriol Carreras, Lena Gerko, Dmitry Matveev, Tongbin Zhang and especially Sebastian Merkel is greatly appreciated. Klaus Adam thanks the Bank of Japan for the hospitality offered during early stages of this project. Albert Marcet acknowledges support from Programa de Excelencia del Banco de España, Plan Nacional (Ministry of Education), SGR (Generalitat de Catalunya). Klaus Adam and Albert Marcet acknowledge support from the European Research Council under the EU 7th Framework Programme (FP/2007-2013), Starting Grant Agreement No. 284262 (Adam) and Advanced Grant Agreement No. 324048 (Marcet). Klaus Adam and Albert Marcet thank Fondation Banque de France for its support on this project. The paper represents the authors' personal opinions and does not necessarily reflect the views of the Deutsche Bundesbank, the Eurosystem, or its staff. All errors remain ours. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

expectations about future capital gains, are a crucial ingredient for understanding these fluctuations.

We present an asset pricing model that incorporates endogenous belief dynamics about expected capital gains. The model gives rise to sustained stock price booms and busts and is consistent with the behavior of investors' capital gains expectations, as measured by survey data. The presented modeling approach differs notably from the standard approach in the consumption-based asset pricing literature, which proceeds by assuming that stock price fluctuations are fully efficient. Campbell and Cochrane 1999 and Bansal and Yaron 2004, for example, present models in which stock price fluctuations reflect the interaction of investor preferences and stochastic driving forces in a setting with optimizing investors who hold rational expectations.

We first present empirical evidence casting considerable doubt on the prevailing view that stock price fluctuations are efficient. Specifically, we show that the rational expectations (RE) hypothesis gives rise to an important counterfactual prediction for the behavior of investors' return or capital gain expectations.<sup>1</sup> This counterfactual prediction is a model-independent implication of the RE hypothesis, but - as we explain below - key for understanding stock price volatility and its efficiency properties.

As previously noted by Fama and French 1988, the empirical behavior of asset prices implies that rational return expectations correlate negatively with the price dividend (PD) ratio. Somewhat counter-intuitively, the RE hypothesis thus predicts that investors have been particularly pessimistic about future stock returns in the early part of the year 2000, when the tech stock boom and the PD ratio of the S&P500 reached their all-time maximum. As we document, the available survey evidence implies precisely the opposite: all quantitative survey measures of investors' expected return (or capital gain) available for the U.S. economy, unambiguously and unanimously correlate positively with the PD ratio; and perhaps not surprisingly, return expectations reached a temporary maximum rather than a minimum in the early part of the year 2000, i.e., precisely at the peak of the tech stock boom, a fact previously shown in Vissing-Jorgensen 2003.

We present formal econometric tests of the null hypothesis that the survey evidence is consistent with RE and demonstrate that the hypothesis of rational return or capital gain

<sup>1</sup>Since most variation in returns is due to the variation in capital gains, we tend to use both terms interchangeably.

expectations is overwhelmingly rejected by the data. Our tests correct for small sample bias, account for autocorrelations in the error structure, are immune to the presence of differential information on the part of agents and to the presence of measurement error in survey data. An appealing feature of the tests is that they also provide clues about why the RE hypothesis fails: the failure arises because survey expectations and RE covary differently with the PD ratio, a finding that is useful for guiding the search for alternative and empirically more plausible expectations models.

The positive comovement of stock prices and survey expectations suggests that price fluctuations are amplified by overly optimistic beliefs at market peaks and by overly pessimistic beliefs at market troughs. Furthermore, it suggests that investors' capital gains expectations are influenced - at least partly - by the capital gains observed in the past, in line with evidence presented by Malmendier and Nagel 2011. Indeed, a simple adaptive updating equation captures the time series behavior of the survey data and its correlation with the PD ratio very well.

Taken together, these observations motivate the construction of an asset pricing model in which agents hold subjective beliefs about price outcomes. We do so using the framework of Internal Rationality (IR), developed in Adam and Marcet 2011, which allows considering maximizing investors that hold subjective price beliefs within an otherwise standard Lucas 1978 asset pricing model.<sup>2</sup> Within this framework, agents optimally update beliefs using Bayes' law.

With this modification, the Lucas model becomes *quantitatively* consistent with important aspects of the data. Using confidence intervals based on the simulated method of moments, we find that the model matches key moments describing the observed volatility of stock prices and the positive correlation between the PD ratio and subjective return expectations. This is obtained even though we use the simplest version of the Lucas model with time separable preferences and standard stochastic driving processes. The same model under RE is very far from explaining the data and produces - amongst other things - far too little price volatility and the wrong sign for the correlation between the PD ratio and investors return expectations.

<sup>2</sup>As is explained in Adam and Marcet 2011, subjective price beliefs are consistent with optimizing behavior in the presence of lack of common knowledge about agents' beliefs and preferences.

The strong improvement in the model's empirical performance arises because agents' attempts to improve their knowledge about price behavior can temporarily delink asset prices from their fundamental (RE) value and give rise to belief-driven boom and bust cycles in stock prices. This occurs because with imperfect information about the price process, optimal behavior prescribes that agents use past capital gain observations to learn about the stochastic process governing the behavior of capital gains; this generates a feedback between capital gain expectations and realized capital gains that can drive booms and busts in stock prices.

Suppose - in line with the empirical evidence - that agents become more optimistic about future capital gains whenever they are positively surprised by past capital gains.<sup>3</sup> A positive surprise for the capital gains observed in the previous period then increases optimism about the capital gains associated with investing in the asset today. If such increased optimism leads to an increase in investors' asset demand and if this demand effect is sufficiently strong, then positive past surprises trigger further positive surprises today, and thus further increases in optimism tomorrow. As we show analytically, stock prices in our model do increase with capital gain optimism whenever the substitution effect of increased optimism dominates the wealth effect of such belief changes. Asset prices in the model then display sustained price booms, similar to those observed in the data.

After a sequence of sustained increases, countervailing forces come into play that endogenously dampen the upward price momentum, eventually halt it and cause a reversal. Specifically, in a situation where increased optimism about capital gains has led to a stock price boom, stock prices make up for a larger share of agents' total wealth.<sup>4</sup> As we show analytically, this causes the wealth effect to become as strong as (or even stronger than) the substitution effect when expectations about stock price appreciation are sufficiently high.<sup>5</sup> Increases in optimism then cease to cause further increases in stock demand and thus stock prices, so that investors' capital gains expectations turn out to be too optimistic relative to the realized outcomes. This induces downward revision in beliefs, which gives

<sup>3</sup>Such positive surprises may be triggered by fundamental shocks, e.g., a high value for realized dividend growth.

<sup>4</sup>This occurs because stock prices are high, but also because agents discount other income streams, e.g., wage income, at a higher rate.

<sup>5</sup>With CRRA utility, this happens whenever the coefficient of relative risk aversion is larger than one.

rise to negative price momentum and an asset price bust.

The previous arguments show how belief dynamics can temporarily delink asset prices from their fundamental value. Clearly, these price dynamics are inefficient as they are not justified by innovations to preferences or other fundamentals.

Since we depart from RE, our model requires introducing an explicit assumption about agents' price beliefs. Various elements guide this modelling choice. First, we choose price beliefs such that there are no 'black swan' like events, i.e., we insure that agents have contingency plans for all prices that they actually encounter along the equilibrium path. Second, we choose the subjective price process such that it gives rise to capital gain expectations that are consistent with the behavior of survey expectations. In particular, agents believe the average growth rate of stock prices to slowly drift over time, which is consistent with the presence of prolonged periods of price booms that are followed by price busts. Given these beliefs, equilibrium prices will indeed display prolonged periods of above average and below average growth.

More generally, the present paper shows how the framework of Internal Rationality allows studying learning about market behavior in a model of intertemporal decision making. It thereby improves on shortcomings present in the learning literature, where agents' belief updating equations and choices are often not derived from individual maximization and are optimal only in the limit once learning converges to the RE outcome. We thus provide explicit microfoundations for settings in which subjective beliefs about market outcomes matter for these outcomes.

The bulk of the paper considers a representative agent model. This is motivated by the desire to derive results analytically and to show how a rather small deviation from the standard paradigm helps reconciling the model with the data. A range of extensions consider - amongst other things - a heterogenous agents version and more elaborate subjective belief structures. These extensions allow replicating additional data features, e.g., the equity premium.

The remainder of the paper is structured as follows. The next section discusses the related literature. Section II then shows that the price dividend ratio (PD) ratio covaries positively with survey measures of investors' return expectations and that this is incompatible with the RE hypothesis. It also shows that the time series of survey expectations

can be captured by a fairly simple belief updating equation. Section III introduces our asset pricing model with subjective beliefs. As a benchmark, section IV determines the RE equilibrium. Section V introduces a specific model of subjective price beliefs, which relaxes agents' prior about price behavior relative to the RE equilibrium beliefs. It also derives the Bayesian updating equations characterizing the evolution of subjective beliefs over time. After imposing market clearing in section VI, we present closed form solutions for the PD ratio in section VII in the special case of vanishing uncertainty. Using the analytical solution, we explain how the interaction between belief updating dynamics and price outcomes can endogenously generate boom and bust dynamics in asset prices. Section VIII estimates the fully stochastic version of the model using a mix of calibration and simulated method of moments estimation. It shows that the model successfully replicates a number of important asset pricing moments, including the positive correlation between expected returns and the PD ratio. It also explains how the model gives rise to a high Sharpe Ratio and to a low volatility for the risk free interest rate. Section IX shows that the estimated model can replicate the low frequency movements in the time series of the US postwar PD ratio, as well as the available time series of survey data. Section X presents a number of robustness checks and extensions of the basic model. A conclusion briefly summarizes and discusses potential avenues for future research. Technical material and proofs can be found in the appendix.

## I. Related Literature

Following Bob Shiller's 1981 seminal observation that stock price volatility cannot be explained by the volatility of rational dividend expectations, the asset pricing literature made considerable progress in explaining stock price behavior. Bansal and Yaron 2004 and Campbell and Cochrane 1999, for example, developed consumption based RE models in which price fluctuations result from large and persistent swings in investors' stochastic discount factor. Section II shows, however, that RE models fail to capture the behavior of investors' return expectations. This strongly suggests that RE models fall short of providing a complete explanation of the sources of stock price volatility.

Attributing stock price fluctuations to 'sentiment' fluctuations or issues of learning has long had an intuitive appeal. A substantial part of the asset pricing literature intro-

duces subjective beliefs to model investor 'sentiment'. The standard approach resorts to Bayesian RE modeling, which allows for subjective beliefs about fundamentals, while keeping the assumption that investors know the equilibrium pricing function linking stock prices to fundamentals. Following early work by Timmermann 1993 and Barberis, Shleifer and Vishny 1998, a substantial literature follows this approach. It finds that the additional stock price volatility generated from learning is overall small compared to the gap that exists relative to the data.

Recent work by Barberis et al. 2014, for example, considers a time-separable utility framework where some investors have rational dividend beliefs while others extrapolate from past dividend observations.<sup>6</sup> While this allows to successfully replicate survey behavior, the standard deviation of the PD ratio falls one order of magnitude short of that observed in the data, so that there are no significant stock price boom and bust episodes.<sup>7</sup>

In ongoing work, Hirshleifer and Yu 2012 and Choi and Mertens 2013 consider Bayesian RE models with time non-separable preferences and investors who extrapolate past fundamentals. They show how extrapolation of fundamentals endogenously generates long-run consumption risk and thereby some increase in asset price volatility. Collin-Dufresne, Michael and Lochstoer 2016 show how learning about fundamentals can give rise to sizable Sharpe ratios, provided agents have a preference for an early resolution of uncertainty. Standard preference parameterizations in this class of models imply, however, that agents are willing to give up a big part of consumption to resolve consumption risk early, see Epstein, Farhi and Strzalecki 2014.

The modeling approach pursued in the present paper differs fundamentally from the one discussed in the previous paragraphs. The Bayesian RE literature assumes that

<sup>6</sup>Barberis et al. 2014 claim in the main text of their paper that agents agree on dividend behavior and that the only difference between rational agents and extrapolators is that the latter extrapolate future prices from past prices (p.4 of their paper). Yet, as it turns out, their model is one where some agents extrapolate past dividends while others have rational dividend expectations and all agents know the equilibrium pricing function, in line with standard Bayesian RE modeling. This can be seen in the proof of their proposition 1, which reverse-engineers a process for extrapolators' dividend beliefs (their equations (A16) and (A18)) that differs from the true dividend process (their equation (1)). The proof of the proposition explicitly acknowledges that these subjective dividend beliefs are chosen such that extrapolators' price beliefs (their equations (3) and (4)) are consistent with the equilibrium pricing function: "*At the same time, in order to compute the values of the derived parameters that govern their consumption and portfolio decisions, extrapolators need to be aware of the price equation (A11).*" (p.19 of their paper). This implies that rational agents and extrapolators disagree about the dividend process but agree about the equilibrium pricing function, in line with standard Bayesian RE modeling. It also implies that in their setting, learning from price behavior is observationally equivalent to learning from dividend behavior, unlike in the setting presented in this paper.

<sup>7</sup>Due to the CARA utility setup the volatility of the PD ratio also asymptotically converges to zero.

agents find it difficult to forecast fundamental shocks (agents hold subjective dividend beliefs), but that agents can predict perfectly price outcomes conditional on the history of observed fundamentals (agents know the equilibrium pricing function mapping the history of dividends into price outcomes). Assuming that agents know the pricing function provides agents with a substantial amount of information about market behavior, suggesting that it is of interest to study the effects of relaxing this informational assumption. Our agents entertain a distribution of prices for given fundamentals, which is non-degenerate and that does not coincide with the model at hand.<sup>8</sup>

To show that a key element for understanding stock price volatility is investor's imperfect knowledge about how prices are formed, we make the distinction to the Bayesian RE literature as stark as possible: we assume that agents find it easy to predict fundamentals, i.e., assume agents hold RE about dividends, but find it difficult to predict price behavior, i.e., agents do not know the equilibrium pricing function. We show that a simple asset pricing model can then replicate survey data and generate sufficient volatility for the PD ratio, including occasional boom and bust episodes. This is achieved in a setting with standard time-separable preferences and obtained because there is a much stronger propagation of economic disturbances when agents learn about the equilibrium pricing function: belief changes then affect stock price behavior and stock prices feed back into belief changes; this allows movements in prices and beliefs to mutually reinforce each other during price boom and bust phases, thereby increasing price volatility. The feedback from market outcomes into beliefs is absent in a Bayesian RE setting.

The literature on robust control and asset prices, e.g. Cogley and Sargent 2008, considers settings where investors are uncertain about the process for fundamentals. In line with Bayesian RE modeling, this literature assumes that investors know the equilibrium pricing function.

The literature on adaptive learning previously considered deviations from rational price expectations using asset pricing models where investors learn about price behavior. Marcet and Sargent 1992, for example, study convergence to RE when agents estimate an

<sup>8</sup>This is related to work by Angeletos and La'O 2013 who consider a setting in which agents are uncertain about the price at which they will be able to trade. They show how sentiment shocks can give rise to perfectly self-fulfilling fluctuations in aggregate outcomes. Sentiment shocks in their setting result from extrinsic uncertainty; in our setting they are triggered from intrinsic sources of uncertainty.



incorrect model of stock prices by least squares learning. A range of papers in the adaptive learning literature argues that learning generates additional stock price volatility. Bullard and Duffy 2001 and Brock and Hommes 1998, for example, show that learning dynamics can converge to complicated attractors that increase asset return volatility, when the RE equilibrium is unstable.<sup>9</sup> Lansing 2010 shows how near-rational bubbles can arise in a model with learning about price behavior. Branch and Evans 2011 present a model where agents learn about risk and return and show how it gives rise to bubbles and crashes. Boswijk, Hommes and Manzan 2007 estimate a model with fundamentalist and chartist traders whose relative shares evolve according to an evolutionary performance criterion, showing that the model can generate a run-up in asset prices and subsequent mean-reversion to fundamental values. DeLong et al. 1990 show how the pricing effects of positive feedback trading survives or even get amplified by the introduction of rational speculators.

The approach used in the present paper differs along several dimensions from the contributions mentioned in the previous paragraph. First, we compare quantitatively the implications of our model with the data, i.e., we match a standard set of asset pricing moments capturing stock price volatility and use formal asymptotic distribution to evaluate the goodness of fit. Second, we compare the model to evidence obtained from survey data. Third, we present a model that derives investors' consumption and stockholding plans from properly specified microfoundations. In particular, we consider agents that solve an infinite horizon decision problem and hold a consistent set of beliefs, we discuss conditions for existence and uniqueness of optimal plans, as well as conditions insuring that the optimal plan has a recursive representation. The adaptive learning literature often relies on shortcuts that amount to introducing additional behavioral elements into decision making and postulates beliefs that become well specified only in the limit, if convergence to RE occurs.<sup>10</sup>

In prior work, Adam, Marcet and Nicolini 2016 present a model in which investors learn about risk-adjusted price growth and show how such a model can quantitatively replicate a set of standard asset pricing moments describing stock price volatility. While

<sup>9</sup>Stability under learning dynamics is defined in Marcet and Sargent 1989.

<sup>10</sup>See section 2 in Adam and Marcet 2011 for a detailed discussion.

replicating stock price volatility and postulating beliefs that are hard to reject in the light of the existing asset price data and the outcomes generated by the model, their setup falls short of explaining survey evidence. Specifically, it counterfactually implies that stock return expectations are constant over time. Adam, Marcet and Nicolini 2016 also solve for equilibrium prices under the assumption that dividend and trading income are a negligible part of total income. We solve the model without this assumption and show that it can play an important role for the model solution, for example, it gives rise to an endogenous upper bound for equilibrium prices.<sup>11</sup>

The experimental and behavioral literature provides further evidence supporting the presence of subjective price beliefs. Hirota and Sunder 2007 and Asparouhova, Bossaerts, Roy and Zame 2013, for example, implement the Lucas asset pricing model in the experimental laboratory and document that there is excess volatility in prices that is unaccounted for by the rational expectations equilibrium and that likely arises from participants' expectations about future prices. Furthermore, the type of learning employed in the present model is in line with evidence presented in Malmendier and Nagel 2011 who show that experienced returns affect beliefs about future asset returns.<sup>12</sup>

## II. Stock Prices & Stock Price Expectations

This section shows that survey expectations of future stock prices are inconsistent with the notion that agents hold rational stock price expectations. Indeed, our formal econometric tests, presented in sections II.B and II.C, show that stock market investors display undue optimism about future stock prices when the PD ratio is high and undue pessimism when the PD ratio is low. Section II.D then illustrates how simple adaptive price predictions, in line with Malmendier and Nagel 2011; 2013, quantitatively capture the relationship between survey expectations and the PD ratio.

<sup>11</sup>In line with the approach in the Bayesian RE literature, Adam Marcet and Nicolini 2016 impose an exogenous upper bound on agents' beliefs, a so-called 'projection facility', so as to insure existence of finite equilibrium prices.

<sup>12</sup>Nagel and Greenwood 2009 show that - in line with this hypothesis - young mutual fund managers displayed trend chasing behavior over the tech stock boom and bust around the year 2000.

*A. Survey Expectations and the PD Ratio*

As pointed out by Vissing-Jorgensen 2003, Greenwood and Shleifer 2014 and a previous version of the current paper, survey expectations of future returns (or capital gains) display a positive correlation with the PD ratio, while actual returns (or capital gains) display a negative correlation.<sup>13</sup>

Appendix A.A2 documents this fact for a range of surveys and figure 2 illustrates it using our preferred survey, the UBS Gallup Survey, which is based on a representative sample of approximately 1.000 U.S. investors that own at least 10.000 US\$ in financial wealth.<sup>14</sup> Figure 2 graphs the US PD ratio (the black line) together with measures of the cross-sectional average of investors' one-year ahead expected real return.<sup>15</sup> Return expectations are expressed in terms of quarterly real growth rates and the figure depicts two expectations measures: investors' expectations about the one year ahead stock market return, as well as their expectations about the one year ahead returns on their own stock portfolio. These measures behave very similarly over the period for which they overlap, but the latter is available for a longer time period. Figure 2 reveals that there is a strong positive correlation between the PD ratio and expected returns. The correlation between the expected own portfolio returns and the PD ratio is +0.70 and even higher for expected stock market returns (+0.79). Moreover, investors' return expectations were highest at the beginning of the year 2000, which is precisely the year the PD ratio reached its peak during the tech stock boom. Investors then expected annualized real returns of around 13% from stock investments, while the subsequently realized returns turned out to be particularly dismal. Conversely, investors were most pessimistic in the year 2003 when the PD ratio reached its bottom, expecting then annualized real returns of below 4%.

This evidence suggests that survey data is incompatible with rational expectations and that stock prices seem to play a role in the formation of expectations about stock returns.

<sup>13</sup>A related observation is that return forecast errors implied by survey data can be predicted using the PD ratio, see Bacchetta, Mertens, and Wincoop 2009.

<sup>14</sup>About 40% of respondents own more than 100.000 US\$ in financial wealth. As is documented in appendix A.A2, this subgroup does not behave differently.

<sup>15</sup>To be consistent with the asset pricing model presented in later sections we report expectations of real returns. The nominal return expectations from the survey have been transformed into real returns using inflation forecasts from the Survey of Professional Forecasters. Results are robust to using instead the Michigan Survey inflation forecast.

Yet, evidence based on comparing two correlations can only be suggestive, as it is subject to several econometric shortcomings. For example, if investors possess information that is not observed by the econometrician, as might be considered likely, then the correlation between the fully rational return forecasts and the PD ratio will differ from the correlation between realized returns and the PD ratio. The same holds true if survey expectations are measured with error, as one can reasonably expect. Furthermore, results in Stambaugh 1999 imply that with the PD ratio being such a persistent process, there is considerable small sample bias in these correlations, given the relatively short time spans over which investor expectations can be tracked. Finally, a highly serially correlated predictor variable (PD ratio), whose innovations are correlated with the variable that is to be predicted (future returns), gives rise to spurious regression and thus spurious correlation problems, see Ferson et al. 2003 and Campbell and Yogo 2006. There exists also no standard approach allowing to correct for these small sample issues when comparing correlations.<sup>16</sup> Comparisons involving correlations are thus insufficient for rejecting the hypothesis that survey expectations are consistent with RE. To deal with these concerns, the next sections construct formal econometric tests that take these concerns fully into account.

### B. RE Test with Small Sample Adjustments

This section develops a RE test that takes into account the concerns expressed in the previous section. While the present section emphasizes the derivation of analytical results, section II.C provides further tests that rely entirely on Monte Carlo simulation.

Let  $E_t^{\mathcal{P}}$  denote agents' subjective expectations operator based on information up to time  $t$ , which can differ from the rational expectations operator  $E_t$ . Let  $R_{t,t+N}$  denote the real cumulative stock returns between period  $t$  and  $t+N$  and let  $\mathcal{E}_t^N = E_t^{\mathcal{P}} R_{t,t+N} + \mu_t^N$  denote the (potentially noisy) observation of expected returns, as obtained - for example - from survey data, where  $\mu_t^N$  is measurement error.<sup>17</sup>

<sup>16</sup>Any test must take into account the joint distribution of the correlation estimates in order to make statistically valid statements.

<sup>17</sup>Since the Shiller survey reports expectations about capital gains instead of returns,  $R_{t,t+N}$  denotes the real growth rate of stock prices between periods  $t$  and  $t+N$  when using the Shiller survey.

Let us linearly project the random variable  $E_t^{\mathcal{P}} R_{t,t+N}$  on  $\frac{P_t}{D_t}$  to define

$$(1) \quad E_t^{\mathcal{P}} R_{t,t+N} = a^N + c^N \frac{P_t}{D_t} + u_t^N,$$

where

$$(2) \quad E(x_t u_t^N) = 0,$$

for  $x_t' = (1, P_t/D_t)$ . The operator  $E$  denotes the objective expectation for the *true* data generating process, whatever is the process for agents' expectations. The projection residual  $u_t^N$  captures variations in agents' actual expectations that cannot be linearly attributed to the price-dividend ratio.<sup>18</sup> It summarizes all other information that agents believe to be useful in predicting  $R_{t,t+N}$ .<sup>19</sup>

Due to the potential presence of measurement error, one cannot directly estimate equation (1), but given the observed return expectations  $\mathcal{E}_t^N$ , one can write the following *regression* equation

$$(3) \quad \mathcal{E}_t^N = a^N + c^N \frac{P_t}{D_t} + u_t^N + \mu_t^N.$$

Assuming that the measurement error  $\mu_t^N$  is orthogonal to the current PD ratio<sup>20</sup>, we have the orthogonality condition

$$(4) \quad E[x_t(u_t^N + \mu_t^N)] = 0.$$

Let  $\widehat{c}_T^N$  denote the OLS estimator of  $c^N$  in equation (3), given a sample of length  $T$ .<sup>21</sup>

We can specify an additional *regression* equation like equation (3), but with actual

<sup>18</sup>The residual  $u_t^N$  is likely to be correlated with current and past observables (other than the PD ratio) and thus serially correlated.

<sup>19</sup>The projection in equation (1) and the error are well-defined as long as agents' expectations  $E_t^{\mathcal{P}} R_{t,t+N}$  and  $P_t/D_t$  are stationary and have bounded second moments.

<sup>20</sup>We allow  $\mu_t^N$  to be serially correlated and correlated with equilibrium variables other than  $P/D_t$ .

<sup>21</sup>Although the residuals  $u_t^N$  and the measurement errors  $\mu_t^N$  are likely to be serially correlated, the OLS estimate is consistent.

returns as dependent variable

$$(5) \quad R_{t,t+N} = \mathbf{a}^N + \mathbf{c}^N \frac{P_t}{D_t} + \mathbf{u}_t^N,$$

where

$$(6) \quad E[x_t \mathbf{u}_t^N] = 0.$$

Let  $\widehat{\mathbf{c}}_T^N$  denote the OLS estimate of  $\mathbf{c}^N$  with  $T$  observations.

The reader can probably guess that the regression estimates are useful here because under the hypothesis of RE we have  $c^N = \mathbf{c}^N$ , so that the estimates  $\widehat{\mathbf{c}}_T^N$  and  $\widehat{\mathbf{c}}_T^N$  are both consistent estimates of the *same* parameter. This gives rise to the following test:

**PROPOSITION 1:** *Assume the process  $\{R_{t,t+N}, \mathcal{E}_t^N, P_t/D_t\}$  is stationary and ergodic, all moments are such that asymptotic distributions exist<sup>22</sup>,  $E(\mu_t) = E(\mu_t P_t/D_t) = 0$ , and  $P_t/D_t$  is part of agents' time  $t$  information set. Then*

*a) Under the null hypothesis of rational expectations*

$$(7) \quad \sqrt{T} \frac{\widehat{\mathbf{c}}_T^N - \widehat{\mathbf{c}}_T^N}{\widehat{\sigma}_{c-c}} \rightarrow N(0, 1) \text{ in distribution as } T \rightarrow \infty,$$

where  $\widehat{\sigma}_{c-c}^2$  is a consistent estimate of  $\text{var}(\widehat{\mathbf{c}}^N - \widehat{\mathbf{c}}^N)$ .<sup>23</sup>

*b) Suppose in addition that*

$$(8) \quad \frac{P_t}{D_t} = \overline{PD}(1 - \rho) + \rho \frac{P_{t-1}}{D_{t-1}} + \varepsilon_t^{PD}$$

for  $\rho \in (-1, 1)$ , where  $(u_t^N + \mu_t^N, \mathbf{u}_t^N, \varepsilon_{t+1}^{PD})$  is normally distributed, i.i.d., with mean zero, and  $E(\mu_t^N \varepsilon_{t+1}^{PD}) = 0$ . Under the null hypothesis of rational expectations, the small

<sup>22</sup>More precisely, we assume: i) bounded second moments of  $(R_{t,t+N}, \mathcal{E}_t^N, P_t/D_t)$ , ii)  $\text{var}(P_t/D_t) > 0$ , and iii) letting the prediction error  $\varepsilon_t^N = R_{t,t+N} - E_t(R_{t,t+N})$  where  $E_t$  is taken with respect to the objective distribution and the information available to investors at  $t$ , we have

$$S_w = \sum_{k=-\infty}^{\infty} E \left( \begin{bmatrix} u_t^N + \mu_t^N \\ u_t^N + \varepsilon_t^N \end{bmatrix} \begin{bmatrix} u_{t-k}^N + \mu_{t-k}^N \\ u_{t-k}^N + \varepsilon_{t-k}^N \end{bmatrix}' \otimes x_t x_{t-k}' \right) < \infty.$$

<sup>23</sup>Equation (A5) in the proof of proposition 1 provides an explicit expression for a consistent variance estimator.

sample bias of  $\widehat{\mathbf{c}}_T^N - \mathbf{c}_T^N$  in the test-statistic (7) is

$$(9) \quad E(\widehat{\mathbf{c}}_T^N - \mathbf{c}_T^N) = \frac{\text{cov}(\varepsilon_{t+1}^{PD}, \varepsilon_t^N)}{\text{var}(\varepsilon_t^{PD})} E(\widehat{\rho}_T - \rho),$$

where  $\varepsilon_t^N \equiv R_{t,t+N} - E_t R_{t,t+N}$  denotes the rational prediction error and  $E(\widehat{\rho}_T - \rho)$  the small sample bias in the estimation of  $\rho$  for a sample of length  $T$ .

The proof of proposition 1 can be found in appendix A.A3. It treats equations (3)-(6) as a seemingly unrelated regression system and uses the fact that under rational expectations one has  $\mathbf{c}^N = \mathbf{c}^N$ . Part b) of the proposition follows from results in Stambaugh 1999.

Part a) of proposition 1 uses minimal assumptions to obtain an asymptotically valid result. Essentially, all that is needed is stationarity of the observables and orthogonality of the measurement error. The test is asymptotically robust to serial-correlation and heteroskedasticity of the error terms.

Part b) of proposition 1 deals with small sample bias in the test statistic. Since  $\rho$  is close to one, we have  $E(\widehat{\rho}_T - \rho) < 0$ ; and since future stock price increases are likely to be correlated with future surprises to returns, i.e.,  $\text{cov}(\varepsilon_{t+1}^{PD}, \varepsilon_t^N) > 0$ , we tend to get  $E(\widehat{\mathbf{c}}_T^N - \mathbf{c}_T^N) < 0$  in small samples, even if in fact  $\mathbf{c}^N = \mathbf{c}^N$ .

The fraction on the right-hand side of the bias expression in equation (9) can be estimated from observables using the calculated errors from equations (3), (5) and (8) and the fact that under the null

$$\text{cov}(\varepsilon_{t+1}^{PD}, \varepsilon_t^N) = \text{cov}(\varepsilon_{t+1}^{PD}, \mathbf{u}_t^N) - \text{cov}(\varepsilon_{t+1}^{PD}, u_t^N + \mu_t^N).$$

The bias  $E(\widehat{\rho}_T - \rho)$  in equation (9) is approximately given by  $E(\widehat{\rho}_T - \rho) \simeq -\frac{1+3\rho}{T}$ , see Marriott and Pope 1954. Since the true bias is a non-linear function of  $\rho$  and the analytical linear approximation less precise in the relevant range of  $\rho$  close to one<sup>24</sup>, we compute the bias  $E(\widehat{\rho}_T - \rho)$  using Monte-Carlo integration.<sup>25</sup>

Interestingly, our RE test  $\mathbf{c}^N = \mathbf{c}^N$  is less prone to small sample bias than tests for the significance of the individual regression coefficients ( $\mathbf{c}^N = 0$  and  $c^N = 0$ ). This follows

<sup>24</sup>See, for instance, figure 1 in MacKinnon and Smith 1998.

<sup>25</sup>Given the estimated values of  $\overline{PD}$ ,  $\rho$ ,  $\sigma_{\varepsilon^{PD}}^2$  we simulate 10.000 realizations of  $PD$  of length  $T$ , compute  $\widehat{\rho}$  for each realization, average over realizations to obtain an approximation for  $E(\widehat{\rho}_T)$  and compute the bias correction accordingly.

from the proof of proposition 1, which shows that

$$E(\widehat{\mathbf{c}}_T^N - \mathbf{c}^N) = E(\widehat{c}_T^N - c^N) + \frac{\text{cov}(\varepsilon_{t+1}^{PD}, \varepsilon_t^N)}{\text{var}(\varepsilon_t^{PD})} E(\widehat{\rho}_T - \rho).$$

The previous equation implies that the small sample bias present in the individual estimate of  $c^N$ , i.e.,  $E(\widehat{c}_T^N - c^N)$ , cancels when testing for  $\mathbf{c}^N = c^N$ .

The test outcomes associated with proposition 1 are reported in table 1a.<sup>26</sup> The table reports the bias-corrected point estimates of  $c^N$  and  $\mathbf{c}^N$ , as well as the bias-corrected p-values for the test based on proposition 1.<sup>27</sup> We use the UBS, CFO and Shiller surveys and consider various ways for extracting expectations from these surveys.<sup>28</sup> The point estimates always satisfy  $\widehat{c}^N > 0$  and  $\widehat{\mathbf{c}}^N < 0$ . The difference between the two estimates is statistically significant at the 1% level in all cases, except when using the survey median from the CFO survey, where p-values are around 3% to 5%.<sup>29</sup>

Overall, the test results in table 1a provide strong evidence against the notion that survey expectations are compatible with rational expectations. Table 1a also shows that agents are overly optimistic when the  $PD$  ratio is high and overly pessimistic when the  $PD$  ratio is low. This suggests that current prices have an ‘excessive role’ in influencing current return expectations. Clearly, if the asset price and survey data were generated by a rational expectations model, say the models of Campbell and Cochrane 1999 or Bansal and Yaron 2004, the tests in table 1a would have been accepted.

### C. Additional Small Sample Adjustments

The closed-form expressions for the small sample bias derived in the previous section are useful for understanding the nature of the bias. At same time, they fall short of completely addressing small sample issues. In particular, the result stated in part b) of proposition 1 relies on assuming that the regression residuals  $u_t^N + \mu_t^N$  and  $\mathbf{u}_t^N$  in

<sup>26</sup>We used 4 lags for Newey West estimator and we checked that results are robust to increasing the lag length up to 12 lags. For each considered survey, we use data on actual returns (or excess returns, or price growth) for the same time period for which survey data is available when computing the p-values.

<sup>27</sup>The p-values are computed using the bias corrected test statistic  $\sqrt{T} \frac{\widehat{c}_T^N - c^N - E(\widehat{c}_T^N - c^N)}{\widehat{\sigma}_{c-c}}$ .

<sup>28</sup>See appendix A.A1 for information on the data sources.

<sup>29</sup>We conjecture that the CFO provides less significant results because the sample starts in Q3:2000, thus does not include the upswing of the tech boom period, unlike the UBS sample. As a result, the CFO sample period displays less mean reversion in prices, which accounts for the fact that the estimates of  $\mathbf{c}$  are less negative and less significant.



equations (3) and (5) are i.i.d. Yet, this is unlikely to hold in our application.

Consider first the residual  $\mathbf{u}_t^N$ . Under the null hypothesis of RE, we have  $\mathbf{u}_t^N = u_t^N + \varepsilon_t^N$ , where  $\varepsilon_t^N \equiv R_{t,t+N} - E_t R_{t,t+N}$  is a prediction error from the *true* data-generating process. For short prediction horizons ( $N = 1$ ),  $\varepsilon_t^N$  is indeed serially uncorrelated, but for longer horizons ( $N > 1$ ), the residuals  $\varepsilon_t^N$  denote forecast errors for overlapping prediction horizons. Since we use quarterly data and prediction horizons between one and ten years, serial correlation of  $\mathbf{u}_t^N$  is a relevant concern.

Next, consider the residual  $u_t^N$ . Under the null hypothesis, it captures information that agents use to predict future returns over and above the PD ratio. The variables capturing such information can be expected to be themselves serially correlated.<sup>30</sup> Indeed, once we take into account serial correlation below, we find a quarterly persistence for  $u_t^N$  between 0.77 and 0.93.

An additional concern with the small sample result in proposition 1 is that the variance  $\widehat{\sigma}_{c-c}$  is usually underestimated in regressions involving highly serially correlated variables. Ferson, Sarkissian and Simin 2003 and Campbell and Yogo 2006 show that this leads to spurious regression problems that cause the null hypothesis to be rejected too often in the kind of regressions we are dealing with.

To address all these issues, this section constructs p-values using a Monte-Carlo procedure to find the actual distribution of the statistic  $\sqrt{T}(\widehat{\mathbf{c}}^N - \widehat{c}^N) / \widehat{\sigma}_{c-c}$ . The aim is to build a model for returns that is as close as possible to the one used in Stambaugh 1999, but that allows for serial correlation of the error terms. To this end, we consider in addition to equation (8) an equation for returns of the form

$$(10) \quad R_{t,t+N} = \mathbf{A}^N + \mathbf{C}^N \frac{P_t}{D_t} + \mathbf{U}_{t+N}^N,$$

with given constants  $\mathbf{A}^N$  and  $\mathbf{C}^N$ . We allow  $\mathbf{U}_t^N$  to be serially correlated by specifying it as an AR(1) process. Since the innovations  $\varepsilon_t^{PD}$  to the PD ratio in equation (8) are likely also a key component of the innovation to returns, we allow the innovations to  $\mathbf{U}_t^N$  to be

<sup>30</sup>For example, in a model with capital accumulation,  $u_t^N$  is a function of the capital stock, which is highly serially correlated.

correlated with  $\varepsilon_t^{PD}$  and consider

$$(11) \quad \mathbf{U}_t^N = \chi \mathbf{U}_{t-1}^N + \eta_t + \lambda \varepsilon_t^{PD}$$

for  $\eta_t \sim N(0, \sigma_\eta^2)$ , independent of  $\varepsilon_j^{PD}$  at all dates  $j$ , and for given constants  $\chi$  and  $\lambda$  satisfying  $|\chi| < 1$ .

Notice that equation (10) is not a special case of equation (5) in section II.B. The reason is that  $\mathbf{U}_{t+N}^N$  is correlated with  $\varepsilon_{t+N-j}^{PD}$  for  $j \geq 0$  and thus correlated with  $PD_t$ :

$$E(\mathbf{U}_{t+N}^N PD_t) = \frac{(\chi)^N \lambda \sigma_{\varepsilon^{PD}}^2}{1 - (\chi\rho)^2}.$$

As a result, the regression coefficients ( $\mathbf{a}^N, \mathbf{c}^N$ ) in equation (5) do not satisfy  $\mathbf{a}^N = \mathbf{A}^N$ ,  $\mathbf{c}^N = \mathbf{C}^N$ , whenever  $\chi\lambda \neq 0$ . It would thus be incorrect to plug our estimates of  $\mathbf{a}^N, \mathbf{c}^N$  into equation (10) for the purpose of running the Monte Carlo simulations.

To estimate the parameters in equations (10) and (11) we proceed as follows. We lag equation (10) by one period, multiply by  $\chi$  and subtract it from equation (10). This delivers

$$(12) \quad R_{t,t+N} = \mathbf{A}^N(1 - \chi) + \chi R_{t-1,t+N-1} + \mathbf{C}^N \frac{P_t}{D_t} - \chi \mathbf{C}^N \frac{P_{t-1}}{D_{t-1}} + \lambda \varepsilon_{t+N}^{PD} + \eta_{t+N},$$

which can be estimated using non-linear least squares and the observed explanatory variables  $(R_{t-1,t+N-1}, \frac{P_t}{D_t}, \frac{P_{t-1}}{D_{t-1}}, \widehat{\varepsilon}_{t+N}^{PD})$ , because these explanatory variables are orthogonal to  $\eta_{t+N}$ . We thus have consistent and efficient estimates for  $\chi, \lambda, \sigma_\eta^2, \mathbf{A}^N$  and  $\mathbf{C}^N$ . We plug these estimates into equations (10) and (11) to simulate  $R_{t,t+N}$ .

To compute expected returns under the null hypothesis of RE, we compute the true expectation of returns, which are given by

$$E_t(R_{t,t+N}) = \mathbf{A}^N + \mathbf{C}^N \frac{P_t}{D_t} + (\chi)^N \mathbf{U}_t^N.$$

Using these results, we can simulate all the variables involved in equations (3) and (5), compute the statistic  $\sqrt{T}(\widehat{\mathbf{c}}^N - \widehat{\mathbf{c}}^N) / \widehat{\sigma}_{c-c}$  for each simulation and study its small sample distribution, using the sample sizes of the considered survey source. We then compute

the probability that  $\sqrt{T}(\widehat{\mathbf{c}}^N - \widehat{\mathbf{c}}^N) / \widehat{\sigma}_{c-c}$  in the Monte-Carlo simulations is smaller than the corresponding value we find for the data. This provides a p-value for the one-sided test of RE, when the alternative hypothesis is  $\mathbf{c}^N > \mathbf{c}^N$ , i.e., that survey returns respond more strongly to the PD ratio than actual returns.

Table 1b reports the outcomes of this procedure. The second column in the table reports the estimated value for  $\chi$ . It shows that the residuals  $\mathbf{U}_t^N$  in equation (11) are indeed serially correlated. We find that this leads to considerable spurious regression problems, as the standard deviation of the test statistic  $\sqrt{T}(\widehat{\mathbf{c}}^N - \widehat{\mathbf{c}}^N) / \widehat{\sigma}_{c-c}$  is indeed around 2 to 3 times larger than its asymptotic value of 1.

Table 1b also reports the bias corrected estimates  $\widehat{\mathbf{c}}^N$  and  $\widehat{\mathbf{c}}^N$ .<sup>31</sup> Compared to the results in table 1a, the estimates for  $\widehat{\mathbf{c}}^N$  are considerably less negative; the ones for the CFO and Shiller 1 year sample even become positive. Yet, the bias corrected estimates for  $\widehat{\mathbf{c}}^N$  in table 1b also become more positive when compared to the ones reported in table 1a. Therefore, despite the spurious regression problems, which cause an increase in the true variance of the test statistic, the RE hypothesis is soundly rejected.<sup>32</sup> The level of the rejection is now considerably lower than the one reported in table 1a, but still highly significant. Since the involved sample lengths are not very large, this is a remarkable result.

Table 1c repeats the analysis when letting  $R_{t,t+N}$  denote excess stock returns rather than stock returns.<sup>33</sup> We construct excess return expectations following Bacchetta et al. 2009, i.e., we assume that the  $N$  period ahead risk-free interest rate is part of agents' information set and subtract it from the (expected) stock return.<sup>34</sup> We find that the strength of the rejection of RE is then somewhat lower when compared to table 1b. Still, for most survey sources one obtains p-values near or below 5%. The somewhat lower p-values show that the approach based on plain stock returns, as reported in table 1b, offers a slightly more powerful test of the RE hypothesis. Interestingly, this is in line with the

<sup>31</sup>The point estimates correct for small sample bias using the mean of the estimator found in the Monte-Carlo simulations with serially correlated errors and using the fact that  $\mathbf{c}^N = \mathbf{C}^N + (\chi)^N \lambda(1 - \rho^2)/(1 - (\chi\rho)^2)$ .

<sup>32</sup>As in table 1a, the rejection is less strong for the CFO survey, especially when using the survey median. See the discussion in footnote 29 for a discussion for the potential reasons behind this result.

<sup>33</sup>For the Shiller survey, which reports price growth expectations,  $R_{t,t+N}$  now denotes excess prices growth.

<sup>34</sup>Following Bacchetta et al. 2009, we use the constant maturity interest rates available from the FRED database at the St. Louis Federal Reserve Bank.

main hypothesis of this paper, namely, that agents form their expectations about stock prices by extrapolating the behavior of past price growth. Under this hypothesis, subtracting the risk-free interest rate adds noise to the independent variables on the l.h.s. of regression equations (3) and (5), which is consistent with the observed decrease in the significance levels.

#### D. How Models of Learning May Help

This section illustrates that a simple ‘adaptive’ approach to forecasting stock prices is a promising alternative to explain the joint behavior of survey expectations and stock price data.

Figure 2 shows that the peaks and troughs of the PD ratio are located very closely to the peaks and troughs of investors’ return expectations. This suggests that agents become optimistic about future capital gains whenever they have observed high capital gains in the past. Such behavior can be captured by models where agents’ expectations are influenced by past experience, prompting us to temporarily explore the assumption that the log of agents’ subjective conditional capital gain expectations  $\ln \tilde{E}_t [P_{t+1}/P_t]$  evolves according to the following adaptive prediction model

$$(13) \quad \ln \tilde{E}_t [P_{t+1}/P_t] = (1 - g) \ln \tilde{E}_{t-1} [P_t/P_{t-1}] + g \ln P_{t-1}/P_{t-2},$$

where  $g > 0$  indicates how strongly capital gain expectations are updated in the direction of past price growth observations. While equation (13) may appear ad-hoc, we show in section V how a very similar equation can be derived from Bayesian belief updating in a setting where agents estimate the persistent component of price growth from the data. Note that equation (13) incorporates price growth observations only with a lag, in line with the theoretical model that we consider later on.

One can feed into equation (13) the historical price growth data of the S&P 500 over the postwar period. Together with an assumption about capital gain expectations at the start of the sample this delivers a time series of implied capital gain expectations  $\ln \tilde{E}_t [P_{t+1}/P_t]$  that can be compared to the expectations from the UBS survey.<sup>35</sup> Set-

<sup>35</sup>We transform the UBS survey measures of return expectations into a measure of price growth expectations using

ting the unobserved initial price growth expectations in Q1:1946 equal to 0% , figure 3 reports the outcome of this procedure, when estimating the gain parameter  $g$  using nonlinear least squares to minimize the distance between the expectations implied by equation (13) and the observed survey expectations.<sup>36</sup> The resulting point estimate is given by  $g = 0.0264$  and has a standard error equal to 0.00168. Figure 3 shows that the adaptive prediction model captures the behavior of UBS expectations extremely well: the correlation between the two series is equal to +0.91.

There also exists a strong positive relationship between the PD ratio and the capital gains expectations implied by equation (13). Figure 4 documents this relationship for the entire postwar period by plotting the joint distribution of the capital gains expectations (as implied by equation (13)) and the PD ratio in the data.<sup>37</sup> When regressing the PD ratio on a constant and the expectations of the adaptive prediction model, one obtains an  $R^2$  coefficient of 0.43; using also the square of the expectations, the  $R^2$  rises further to 0.48.<sup>38</sup>

Interestingly, the relationship between implied price growth expectations and the PD ratio depicted in figure 4 seems to have shifted upwards after the year 2000, as indicated by the squared icons in the figure. Indeed running the previous regressions on the linear and squared expectations separately before and after the year 2000, one obtains much higher  $R^2$  values, namely 0.72 and 0.77, respectively. We will come back to this issue in section IX.

Overall, it emerges from figure 4 that variations in expected capital gains  $\tilde{E}_t [P_{t+1}/P_t]$  can account - in a purely statistical sense - for a large share of the variability in the postwar PD ratio. This suggests that an asset pricing model consistent with equation (13), which additionally predicts a positive relationship between the PD ratio and subjective expectations about future capital gains, has a good chance of replicating the observed positive comovement between price growth expectations and the PD ratio. The next sections spell out the microfoundations of such a model.

the identity  $R_{t+1} = \frac{P_{t+1}}{P_t} + \frac{D_{t+1}}{P_t} = \frac{P_{t+1}}{P_t} + \beta^D \frac{D_t}{P_t}$  where  $\beta^D$  denotes the expected quarterly growth rate of dividends that we set equal to the sample average of dividend growth over Q1:1946-Q1:2012, i.e.  $\beta^D = 1.0048$ . Results regarding implied price growth are very robust towards changing  $\beta^D$  to alternative empirically plausible values.

<sup>36</sup>The figure reports log quarterly expected growth rates for real stock prices.

<sup>37</sup>As before, we set the unobserved initial price growth expectations in Q1:1946 equal to 0%.

<sup>38</sup>The p-values for the coefficients on the linear and squared expectations are all statistically significant at the 1% level.

### III. A Simple Asset Pricing Model

Consider an endowment economy populated by a unit mass of infinitely lived agents  $i \in [0, 1]$  with time-separable preferences. Agents trade one unit of a stock in a competitive stock market. They earn each period an exogenous non-dividend income  $W_t \geq 0$  that we refer to as ‘wages’ for simplicity. Stocks deliver the exogenous dividend  $D_t \geq 0$ . Dividend and wage incomes take the form of perishable consumption goods.

**The Investment Problem.** Investor  $i$  solves

$$(14a) \quad \max_{\{C_t^i \geq 0, S_t^i\}_{t=0}^{\infty}} E_0^{\mathcal{P}^i} \sum_{t=0}^{\infty} \delta^t u(C_t^i)$$

$$\text{s.t.} \quad S_t^i P_t + C_t^i = S_{t-1}^i (P_t + D_t) + W_t \quad \text{for all } t \geq 0$$

$$(14b) \quad \underline{S} \leq S_t^i \leq \bar{S}$$

$$S_{-1}^i = 1,$$

where  $C^i$  denotes consumption,  $u$  the instantaneous utility of the consumer, assumed to be continuous, differentiable, increasing and strictly concave,  $S^i$  the agent’s stockholdings, which are subject to upper and lower limits such that  $-\infty < \underline{S} < 1 < \bar{S} < \infty$ , and  $P \geq 0$  the (ex-dividend) price of the stock.  $\mathcal{P}^i$  denotes the agent’s subjective probability measure, which may or may not satisfy the rational expectations hypothesis. Further details of  $\mathcal{P}^i$  will be specified below.

To simplify the exposition, we do not explicitly consider trade in risk free bonds ( $B_t^i$ ) in problem (14), instead impose from the outset that actual bond prices and agents’ subjective beliefs about future bond prices are such that agents hold and plan to hold zero bonds at all times ( $B_t^i \equiv 0$ ). When bonds are in zero net supply, these beliefs are in line with the equilibrium outcomes, when all agents are identical or when agents face the borrowing constraint  $B_t^i \geq 0$ .<sup>39</sup>

**Dividend and Wage Income.** As is standard in the literature, we assume that divi-

<sup>39</sup>In the latter case, the ‘most patient agent’ prices the bond, i.e., letting  $1 + r_t^b$  denote the gross real return on a one period risk free bond, we have  $(1 + r_t^b)^{-1} = \sup_{i \in [0, 1]} \delta E_t^{\mathcal{P}^i} u'(C_{t+1}^i) / u'(C_t^i)$ .

dends grow at a constant rate and that dividend growth innovations are unpredictable

$$(15) \quad \ln D_t = \ln \beta^D + \ln D_{t-1} + \ln \varepsilon_t^D,$$

where  $\beta^D \geq 1$  denotes gross mean dividend growth and  $\ln \varepsilon_t^D$  an i.i.d. growth innovation described further below.

We also specify an exogenous wage income process  $W_t$ , which is chosen such that the resulting aggregate consumption process  $C_t = W_t + D_t$  is empirically plausible: first, in line with Campbell and Cochrane 1999, we set the standard deviation of consumption growth to be 1/7 of the standard deviation of dividend growth; second, again following these authors, we set the correlation between consumption and dividend growth equal to 0.2; third, we choose a wage process in the model such that the average consumption-dividend ratio in the model ( $E[C_t/D_t]$ ) equals the average ratio of personal consumption expenditure to net dividend income in U.S. postwar data and displays persistence similar to that observed in the data. All this can be parsimoniously achieved using the following wage income process.

$$(16) \quad \ln \left( 1 + \frac{W_t}{D_t} \right) = (1 - p) \ln(1 + \rho) + p \ln \left( 1 + \frac{W_{t-1}}{D_{t-1}} \right) + \ln \varepsilon_t^W,$$

where  $1 + \rho$  is the average consumption-dividend ratio and  $p \in [0, 1)$  its quarterly persistence. The innovations are given by

$$(17) \quad \begin{pmatrix} \ln \varepsilon_t^D \\ \ln \varepsilon_t^W \end{pmatrix} \sim iiN \left( -\frac{1}{2} \begin{pmatrix} \sigma_D^2 \\ \sigma_W^2 \end{pmatrix}, \begin{pmatrix} \sigma_D^2 & \sigma_{DW} \\ \sigma_{DW} & \sigma_W^2 \end{pmatrix} \right),$$

with  $E\varepsilon_t^D = E\varepsilon_t^W = 1$ . Given the variance of dividend growth  $\sigma_D^2$ , which can be estimated from dividend data, one can use  $\sigma_{DW}$  and  $\sigma_W^2$  to impose the desired volatility of consumption growth and the desired correlation with dividend growth. Choosing  $\rho = 22$ , one obtains the observed average consumption-dividend ratio in the data and setting  $p$  to a value close to one, one replicates the observed persistence of the consumption-dividend ratio in the data. Appendix A.A4 provides further details.

**The Agents' Underlying Probability Space.** Agents hold a set of subjective proba-

bility beliefs about all payoff-relevant variables that are beyond their control. In addition to fundamental variables (dividends and wages), agents perceive competitive stock prices as beyond their control. Therefore, the belief system also specifies probabilities about prices. Formally, letting  $\Omega$  denote the space of possible realizations for infinite sequences, a typical element  $\omega \in \Omega$  is given by  $\omega = \{P_t, D_t, W_t\}_{t=0}^{\infty}$ . As usual,  $\Omega^t$  then denotes the set of all (nonnegative) price, dividend and wage histories from period zero up to period  $t$  and  $\omega^t$  its typical element. The underlying probability space for agents' beliefs is then given by  $(\Omega, \mathcal{B}, \mathcal{P}^i)$  with  $\mathcal{B}$  denoting the corresponding  $\sigma$ -Algebra of Borel subsets of  $\Omega$ , and  $\mathcal{P}^i$  a probability measure over  $(\Omega, \mathcal{B})$ .

The agents' plans will be contingent on the history  $\omega^t$ , i.e., the agent chooses state-contingent consumption and stockholding functions

$$(18) \quad C_t^i : \Omega^t \rightarrow \mathcal{R}_+$$

$$(19) \quad S_t^i : \Omega^t \rightarrow [\underline{S}, \overline{S}]$$

The fact that  $C^i$  and  $S^i$  depend on price realizations is a consequence of optimal choice under uncertainty, given that agents consider prices to be exogenous random variables.

The previous setup is general enough to accommodate situations where agents learn about the stochastic processes governing the evolution of prices, dividends, and wages. For example,  $\mathcal{P}^i$  may arise from a stochastic process describing the evolution of these variables that contains unknown parameters or hidden variables about which agents hold prior beliefs. The presence of unknown parameters or hidden variables then implies that agents update their beliefs using the observed realizations of prices, dividends and wages. A particular example of this kind will be presented in section V when we discuss learning about stock price behavior.

The probability space defined above is more general than that specified in a RE analysis of the model, where  $\Omega$  contains usually only the variables that are exogenous to the model (in this case  $D_t$  and  $W_t$ ), but not variables that are endogenous to the model and exogenous to the agent only (in this case  $P_t$ ). In the RE literature, the Bayesian RE literature or the literature modeling robustness concerns, agents are assumed to know the equilibrium pricing function  $P_t((D, W)^t)$  mapping histories of dividends and wages



into a market prices. Prices then carry only redundant information and can be excluded from the probability space without loss of generality. The more general formulation we entertain here allows us to consider agents who do not know exactly which price materializes given a particular history of dividends and wages; our agents do have a view about the distribution of  $P_t$  conditional on  $(D, W)^t$ , but in their minds this is a proper distribution, not a point mass as in the RE case. Much akin to academic economists, investors in our model have not converged on a single asset pricing model that associates one market price with a given history of exogenous fundamentals.

**Parametric Utility Function.** To obtain closed-form solutions, we consider in the remaining part of the paper the utility function

$$(20) \quad u(C_t) = \frac{C_t^{1-\gamma}}{1-\gamma} \quad \text{with } \gamma > 1.$$

We furthermore assume that

$$(21) \quad \delta\beta^{RE} < 1,$$

where  $\beta^{RE} = (\beta^D)^{1-\gamma} E[(\varepsilon_t^W)^{-\gamma} (\varepsilon_t^D)^{1-\gamma}]$ . This insures existence of an equilibrium under rational price expectations.

**Existence of a Recursive Solution.** Since solving the optimization problem (14) for general (potentially non-rational) price beliefs is non-standard, it is worth pointing out that this problem is well defined. Existence of a maximum is guaranteed by the stock limits (14b), which ensure that the choice of stocks is made over a compact set, in combination with a bounded continuous objective function (20), satisfying  $u(C_t) \leq 0$ . Sufficiency of first order conditions is guaranteed because the agents' problem is concave. Given the subjective price beliefs introduced in the remaining part of the paper, agents' posterior will follow a recursive form. Using a standard normalization of the utility function and of the budget constraint, one can then guarantee that the optimal solution to (14)

takes the form

$$(22) \quad S_t^i = S^i \left( S_{t-1}^i, \frac{P_t}{D_t}, \frac{W_t}{D_t}, m_t^i \right),$$

where  $m_t^i$  is a sufficient statistic characterizing the subjective distributions about future values of  $\left( \frac{D_{t+j}}{D_{t+j-1}}, \frac{P_{t+j}}{D_{t+j}}, \frac{W_{t+j}}{D_{t+j}} \right)$  for  $j > 0$ .<sup>40</sup> See appendix A.A5 for details.

#### IV. Rational Expectations (RE) Equilibrium

As a point of reference, we determine the stock price under the assumption of RE. For the limiting cases with an i.i.d. consumption-dividend ratio ( $p = 0$ ) or with a unit root in this ratio ( $p \rightarrow 1$ ), one can derive simple closed form expressions for the RE stock price, see Appendix A.A6:

**PROPOSITION 2:** *If agents hold rational expectations and if price expectations satisfy the usual transversality condition (stated explicitly in appendix A.A6), then the RE equilibrium price dividend ratio is*

$$(23) \quad \frac{P_t^{RE}}{D_t} = \frac{\delta \beta^{RE}}{1 - \delta \beta^{RE}} \quad \text{for } p \rightarrow 1$$

$$(24) \quad \frac{P_t^{RE}}{D_t} = b \frac{\delta \tilde{\beta}^{RE}}{1 - \delta \tilde{\beta}^{RE}} (\varepsilon_t^W)^\gamma \quad \text{for } p = 0,$$

where  $\beta^{RE} \equiv (\beta^D)^{1-\gamma} E[(\varepsilon_t^W)^{-\gamma} (\varepsilon_t^D)^{1-\gamma}]$ ,  $b \equiv E[(\varepsilon_t^W)^{-\gamma} (\varepsilon_t^D)^{1-\gamma}] e^{\gamma(1-\gamma)\frac{\sigma_D^2}{2}}$  and  $\tilde{\beta}^{RE} \equiv (\beta^D)^{1-\gamma} e^{\gamma(\gamma-1)\sigma_D^2/2}$ .

The previous result shows that under RE the equilibrium PD ratio inherits the persistence properties of the consumption-dividend ratio process. Specifically, for  $p = 0$ , the PD ratio is an i.i.d. process, thus fails to match the persistence of the PD ratio observed in the data. For the empirically more plausible case with a persistent consumption-dividend ratio ( $p$  close to 1), the PD ratio also becomes persistent but volatility then strongly falls. In the limit  $p \rightarrow 1$ , which is the case considered in Campbell and Cochrane 1999, the

<sup>40</sup>The solution for optimal consumption plans follows from (22) and the flow budget constraint.

RE PD ratio becomes a constant. Price growth is then given by

$$(25) \quad \ln P_{t+1}^{RE} - \ln P_t^{RE} = \ln \beta^D + \ln \varepsilon_{t+1}^D,$$

so that one-step-ahead price growth and return expectations are constant over time. Moreover, as long as  $p$  is bounded below 1, the RE equilibrium implies a *negative* correlation between the PD ratio and expected future returns, contrary to what is evidenced by survey data.<sup>41</sup> The RE equilibrium thus has difficulties in jointly matching the persistence and volatility of the PD ratio and the behavior of survey returns.

## V. Learning about Capital Gains and Internal Rationality

This section specifies agents' beliefs. Once one departs from rational expectations, these beliefs become part of the micro-foundations of the model. The benchmark beliefs presented in the next section aim at specifying price beliefs in a parsimonious way such that they have a chance of replicating the behavior of survey expectations. We study alternative and more elaborate belief formulations in sections X.B and X.C.

### A. Agents' Belief System: Benchmark Specification

To focus on the effects of learning about price behavior, we assume that agents know the processes (15)-(17), i.e., agents hold rational dividend and wage expectations. This sets us apart from the Bayesian RE literature, which focuses on learning about the behavior of exogenous fundamentals.

We also have to specify price beliefs. Our specification is motivated by the following observations. Under RE, price growth is approximately equal to dividend growth, whenever the consumption-dividend ratio is persistent ( $p$  close to 1), see proposition 2. Price growth in the data, however, can persistently outstrip dividend growth, thereby giving rise to a persistent increase in the PD ratio and an asset price boom, see figure 1; conversely, it can fall persistently short of dividend growth and give rise to a price bust. This behavior of actual asset prices suggests that it is of interest to relax agents' beliefs about price growth behavior. Indeed, in view of the actual behavior of asset prices, agents may

<sup>41</sup>Appendix A.A6 derives the general expression for the RE price for the case  $p \in (0, 1)$ .

entertain a more general model of price behavior, incorporating the possibility that the growth rate of prices persistently exceeds/falls short of the growth rate of dividends. To the extent that the equilibrium asset prices implied by these beliefs display such data-like behavior, agents' beliefs will be generically validated

In line with the previous discussion, our benchmark assumption is that agents perceive prices to evolve according to

$$(26) \quad \ln P_{t+1} - \ln P_t = \ln \beta_{t+1} + \ln \varepsilon_{t+1},$$

where  $\varepsilon_{t+1}$  denotes a transitory shock to price growth and  $\beta_{t+1}$  a persistent price growth component that drifts slowly over time according to

$$(27) \quad \ln \beta_{t+1} = \ln \beta_t + \ln \nu_{t+1} .$$

For simplicity, we assume that agents perceive the innovations  $\ln \varepsilon_t$  and  $\ln \nu_t$  to be jointly normally distributed according to

$$(28) \quad \begin{pmatrix} \ln \varepsilon_t \\ \ln \nu_t \end{pmatrix} \sim iiN \left( \begin{pmatrix} -\frac{\sigma_\varepsilon^2}{2} \\ -\frac{\sigma_\nu^2}{2} \end{pmatrix}, \begin{pmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & \sigma_\nu^2 \end{pmatrix} \right).$$

Since agents observe stock price growth, but do not separately observe the persistent and transitory subcomponents driving it, the previous setup defines a filtering problem in which agents need to decompose observed price growth into its persistent and transitory elements, so to forecast optimally.<sup>42</sup>

The presented belief setup has a number of appealing features. First, it allows capturing the fact that stock price growth displays large transitory fluctuations by imposing  $\sigma_\nu^2 \ll \sigma_\varepsilon^2$ . In the limiting case where  $\sigma_\nu^2$  is close to zero, the persistent price growth component behaves almost like a constant, as in the RE solution with  $p \rightarrow 1$ . The sub-

<sup>42</sup>Note that we do not incorporate mean-reversion into price growth beliefs in our benchmark setting. This is for simplicity, as we wish to consider the most parsimonious way to include perceived booms and busts. In addition, we seek to determine the model-endogenous forces that lead to a reversal of boom and bust dynamics, i.e., we do not want to obtain reversals because they are hard-wired into beliefs. We extend the setup to one with mean reversion in section X.B.

jective price beliefs specified above can thus be interpreted as a small deviation from RE equilibrium beliefs. Second, the benchmark setup can capture periods with sustained increases in the PD ratio ( $\beta_{t+1} > \beta^D$ ) and sustained decreases ( $\beta_{t+1} < \beta^D$ ), in line with the behavior of the PD ratio in the data. To the extent that the PD ratio in the model also displays such behavior, agents' beliefs about the presence of a persistent price growth component will also be generically validated by the model outcome. Third, as we shall show in the next section, the belief specification implies that agents perceived price growth expectations will approximately evolve according to equation (13), which captures the time series behavior of survey expectations.

As with any assumption about model primitives, one can entertain other plausible alternatives. Section X.B explores alternative ways to specify the belief system, which incorporate mean reversion in the PD ratio.

### *B. Efficiency of Stock Prices and Internal Rationality*

Among academics there appears to exist a widespread belief that rational behavior and knowledge of the fundamental processes (dividends and wages in our case) jointly dictate a certain process for stock prices and thus the price beliefs agents can rationally entertain.<sup>43</sup> This view stipulates that rational behavior implies knowledge that current stock prices must equal a discounted sum of dividends. Individual rationality and rational expectations about fundamentals would then provide investors with knowledge of the equilibrium pricing function (as is assumed under RE or Bayesian RE), so that postulating subjective price beliefs, e.g., those specified in the previous section, would be inconsistent with the assumption of optimal behavior on the part of agents.

This view is correct only in special cases. Considering the case with risk neutral agents, Adam and Marcet 2011 show that it is correct only if agents are homogeneous and do not face trading constraints. In a setting with heterogeneous risk neutral agents and trading constraints, it fails to be correct. As we show below, agents in our model are 'internally rational' even in the homogeneous agent case: their behavior is optimal given an internally consistent system of subjective beliefs about variables that are beyond their

<sup>43</sup>We often received this reaction during seminar presentations.

control, which includes prices. All that is needed is a concave utility function.<sup>44</sup>

To illustrate this point, consider first risk neutral agents with rational dividend expectations and ignore limits to stock holdings. Forward-iteration on the agents' optimality condition

$$(29) \quad u'(C_t^i) = \delta E_t^{\mathcal{P}^i} \left[ u'(C_{t+1}^i) \frac{P_{t+1} + D_{t+1}}{P_t} \right]$$

then delivers the present value relationship

$$P_t = E_t \left[ \sum_{i=1}^T \delta^i D_{t+i} \right] + \delta^T E_t^{\mathcal{P}^i} [P_{t+T}],$$

which is independent of the agent's own choices. Provided agents' price beliefs satisfy a standard transversality condition ( $\lim_{T \rightarrow \infty} \delta^T E_t^{\mathcal{P}^i} [P_{t+T}] = 0$  for all  $i$ ), each rational agent would conclude that there must be a degenerate joint distribution for prices and dividends given by

$$(30) \quad P_t = E_t \left[ \sum_{i=1}^{\infty} \delta^i D_{t+i} \right] \text{ a.s.}$$

Since the r.h.s. of the previous equation is fully determined by dividend expectations, the beliefs about the dividend process deliver the price process compatible with optimal behavior. In such a setting, it would be plainly inconsistent with optimal behavior to assume the subjective price beliefs (26)-(27).<sup>45</sup>

Next, consider a concave utility function  $u(\cdot)$  satisfying standard Inada conditions. Forward iteration on (29) and assuming an appropriate transversality condition then delivers

$$(31) \quad P_t u'(C_t) = E_t^{\mathcal{P}} \left[ \sum_{j=1}^{\infty} \delta^j u'(C_{t+j}) D_{t+j} \right] \text{ a.s.}$$

<sup>44</sup>Obviously, the subsequent discussion takes for granted that the fact that investors are homogenous is not common knowledge.

<sup>45</sup>See Adam and Marcet 2011 for a discussion of how in the presence of trading constraints, this conclusion breaks down with linear consumption preferences.

Unlike in equation (30), the right side of the previous equation depends on the agent's subjective consumption plans. From equation (18) follows that future consumption plans are of the form  $C_{t+j}((P, D, W)^{t+j})$ . Expected future consumption thus depends on the agent's price expectations, say those implied by equation (26). Indeed, whatever are the agent's price expectations, subjective consumption plans will adjust, so as to satisfy the first order condition (29). As a result, equation (31) will also hold. Equation (31) thus fails to impose any restriction on what optimizing agents can possibly believe about the price process, given their knowledge about the processes for  $W_t$  and  $D_t$ . With the considered non-linear utility function, we can thus simultaneously assume that agents maximize utility, hold the subjective price beliefs (26)-(27) and have rational expectations about dividends and wages.

The fact that equilibrium prices can be written as a discounted sum of the form (31) differs notably from the standard pricing formula studied in modern asset pricing theory, which determines equilibrium prices using

$$P_t u'(W_t + D_t) = E_t^{\mathcal{P}} \left[ \sum_{j=1}^{\infty} \delta^j u'(W_{t+j} + D_{t+j}) D_{t+j} \right] \text{ a.s.}$$

The r.h.s. of this equation uses equilibrium values of future consumption instead of the subjective future consumption plans showing up in equation (31). In this sense, stock prices satisfying (31) are not efficient, because they do not discount dividends with the 'objective' discount factor.

### C. Learning about Capital Gains

Under internal rationality, the specification of the perceived price process (26) dictates the way agents learn from observed prices. The presence of an unobserved permanent component in equation (26) gives rise to an optimal filtering problem. To obtain a parsimonious description of this problem, we specify conjugate prior beliefs about the unobserved persistent component  $\ln \beta_t$  at  $t = 0$ . Specifically, agent  $i$ 's prior is

$$(32) \quad \ln \beta_0 \sim N(\ln m_0^i, \sigma^2),$$

where prior uncertainty  $\sigma^2$  is assumed to be equal to its Kalman filter steady state value, i.e.,

$$(33) \quad \sigma^2 \equiv \frac{-\sigma_v^2 + \sqrt{(\sigma_v^2)^2 + 4\sigma_v^2\sigma_\varepsilon^2}}{2}.$$

Equations (26), (27) and (32), and knowledge of the dividend and wage income processes (17) then jointly specify agents' probability beliefs  $\mathcal{P}^i$ .

The optimal Bayesian filter then implies that the posterior beliefs following some history  $\omega^t$  are given by<sup>46</sup>

$$(34) \quad \ln \beta_t | \omega^t \sim N(\ln m_t^i, \sigma^2),$$

with

$$(35) \quad \ln m_t^i = \ln m_{t-1}^i - \frac{\sigma_v^2}{2} + g \left( \ln P_t - \ln P_{t-1} + \frac{\sigma_\varepsilon^2 + \sigma_v^2}{2} - \ln m_{t-1}^i \right)$$

$$(36) \quad g = \frac{\sigma^2}{\sigma^2 + \sigma_\varepsilon^2}.$$

Agents' beliefs can thus be parsimoniously summarized by a single state variable ( $m_t^i$ ) describing agents' degree of optimism about future capital gains. These beliefs evolve recursively according to equation (35) and imply that

$$(37) \quad \ln m_t^i = \ln E_t^{\mathcal{P}^i} \left[ \frac{P_{t+1}}{P_t} \right] - \sigma^2/2,$$

so that equation (35) is - up to some (small) variance correction and the presence of a time lag - identical to the adaptive prediction model (13) considered in section II.D.

The subjective price beliefs (26),(27) and (32) generate perfect foresight equilibrium price expectations in the special case in which prior beliefs are centered at the growth

<sup>46</sup>See theorem 3.1 in West and Harrison 1997. Choosing a value for  $\sigma^2$  different from the steady state value (33) would only add a deterministically evolving variance component  $\sigma_t^2$  to posterior beliefs with the property  $\lim_{t \rightarrow \infty} \sigma_t^2 = \sigma^2$ , i.e., it would converge to the steady state value.



rate of dividends, i.e.,

$$\ln m_0^i = \ln \beta^D,$$

and when considering the limiting case with vanishing uncertainty, where  $(\sigma_\varepsilon^2, \sigma_v^2, \sigma_D^2, \sigma_W^2) \rightarrow 0$ . Agents' prior beliefs at  $t = 0$  about price growth in  $t \geq 1$  then increasingly concentrates at the perfect foresight outcome  $\ln \beta^D$ , see equations (26) and (27). With price and dividend expectations being at their perfect foresight value, the perfect foresight equilibrium prices implied by proposition 2 for the vanishing noise limit become the equilibrium outcome at  $t = 0$ . Importantly, it continues to be possible to study learning dynamics in the limit with vanishing risk: keeping the limiting ratio  $\sigma_v^2/\sigma_\varepsilon^2$  finite and bounded from zero as uncertainty vanishes, the Kalman gain parameter  $g$ , defined in equation (36), remains well-specified and satisfies  $\lim \frac{\sigma_v^2}{\sigma_\varepsilon^2} = \lim \frac{g^2}{1-g}$ . We will exploit this fact in section VII when presenting analytical results.

## VI. Dynamics under Learning

This section explains how equilibrium prices are determined under the subjective beliefs introduced in the previous section and how they evolve over time.

Agents' stock demand is given by equation (22). Stock demand depends on the belief  $m_t^i$ , which characterizes agents' capital gains expectations. These beliefs evolve according to (35). As a benchmark, we shall now assume that all agents hold identical beliefs ( $m_t^i = m_t$  for all  $i$ ). In doing so, we assume that the marginal investor has return expectations in line with those documented for the survey data. We consider the homogenous agent case first, so as to stay as close as possible to the representative agent settings typically considered under RE. Section VIII.B considers extensions to settings with heterogenous beliefs.

Using the representative agent assumption, the fact that stocks are in unit supply and imposing market clearing in periods  $t$  and  $t - 1$  in equation (22), one obtains that the equilibrium price in any period  $t \geq 0$  solves

$$(38) \quad 1 = S \left( 1, \frac{P_t}{D_t}, \frac{W_t}{D_t}, m_t \right).$$

The beliefs  $m_t$  and the price dividend ratio  $P_t/D_t$  are now simultaneously determined via

equations (35) and (38). Due to a complementarity between realized capital gains and expected future capital gains<sup>47</sup>, this simultaneity can give rise to multiple market clearing price and belief pairs. In the learning literature, the standard approach to resolve this issue consists of assuming that agents use only lagged endogenous variables to update model estimates (Eusepi and Preston 2011, Evans and Honkapohja 2001, Marcet and Nicolini 2003, Sargent, Williams and Zha 2009). In our benchmark specification we shall follow this approach, but we provide below Bayesian foundations for lagged updating in a setting where agents can observe current prices. Furthermore, as we show in section X.C, lagged belief updating is objectively optimal because agents who update beliefs using current price information experience lower utility in equilibrium.

Introducing an updating lag into equation (35) can be justified as internally rational by slightly modifying the information structure. The modification is relatively straightforward and consists of assuming that agents observe at any time  $t$  information about the lagged transitory price growth component  $\varepsilon_{t-1}$  entering equation (26). In such a setting, it is optimal to give less weight to the last available price growth observation, when estimating the permanent component  $\beta_t$ . This is so because the last observation is more noisy. Formally, appendix A.A7 shows that in the limit where agents learn all transitory factors with a lag, Bayesian updating implies

$$(39) \quad \ln m_t = \ln m_{t-1} + g (\ln P_{t-1} - \ln P_{t-2} - \ln m_{t-1}) + g \ln \varepsilon_t^1,$$

where updating now occurs using only lagged price growth (even though agents do observe current prices) and where  $\ln \varepsilon_t^1 \sim iiN(\frac{-\sigma_\varepsilon^2}{2}, \sigma_\varepsilon^2)$  is a time  $t$  innovation to agent's information set (unpredictable using information available to agents up to period  $t - 1$ ). The shock  $\ln \varepsilon_t^1$  thereby captures the information that agents receive in period  $t$  about about the transitory price growth component  $\ln \varepsilon_{t-1}$ .

With this slight modification, agents' beliefs  $m_t$  are now pre-determined at time  $t$ , so that the economy evolves according to a uniquely determined recursive process: equation (38) determines the market clearing price for period  $t$  given the beliefs  $m_t$  and equation

<sup>47</sup>Intuitively, a higher PD ratio implies higher realized capital gains and thus higher expectations of future gains via equation (35). Higher expected future gains may in turn induce a higher willingness to pay for the asset, thereby justifying the higher initial PD ratio.

(39) determines how time  $t$  beliefs are updated following the observation of the new market clearing price.<sup>48</sup>

## VII. Equilibrium: Analytic Findings

This section derives a closed form solution for the equilibrium asset price for the special case where all agents hold the same subjective beliefs  $\mathcal{P}$  and where these beliefs imply no (or vanishing) uncertainty about future prices, dividends and wages. While the absence of uncertainty is unrealistic from an empirical standpoint, it helps us in deriving key insights into how the equilibrium price depends on agents' beliefs, as well as on how prices and beliefs evolve over time.<sup>49</sup> The empirically more relevant case with uncertainty will be considered in section IX using numerical solutions.

We present a series of results that increasingly adds assumptions on agents' beliefs system  $\mathcal{P}$ . The next section provides a closed form expression for the equilibrium PD ratio as a function of agents' subjective expectations about future stock market returns for any belief system  $\mathcal{P}$  without uncertainty. Section VII.B then discusses the pricing implications of this result for the subjective capital gains beliefs presented in section V. Finally, section VII.C shows how the interaction between asset price behavior and subjective belief revisions can temporarily de-link asset prices from their fundamental value, i.e., give rise to a self-reinforcing boom and bust cycle in asset prices along which subjective expected returns rise and fall.

### A. Main Result

Letting  $R_{t+1} \equiv (P_{t+1} + D_{t+1})/P_t$  denote the stock return, we have following main result:<sup>50</sup>

**PROPOSITION 3:** *Suppose  $u(C) = C^{1-\gamma}/(1-\gamma)$ , agents' beliefs  $\mathcal{P}$  imply no uncertainty about future prices, dividends and wages. Assume given posterior beliefs about*

<sup>48</sup>There could still be an indeterminacy arising from the fact that  $S(\cdot)$  is non-linear, so that equation (38) may not have a unique solution. We have not encountered such problems in our analytical solution or when numerically solving the model.

<sup>49</sup>An analytic solution can be found because in the absence of uncertainty one can evaluate more easily the expectations of nonlinear functions of future variables showing up in agents' FOCs.

<sup>50</sup>The proof can be found in appendix A.A8.

prices, dividends and wages that satisfy

$$(40) \quad \lim_{T \rightarrow \infty} E_t^{\mathcal{P}} R_T > 1 \text{ and } \lim_{T \rightarrow \infty} E_t^{\mathcal{P}} \left( \sum_{j=1}^T \left( \prod_{i=1}^j \frac{1}{R_{t+i}} \right) W_{t+j} \right) < \infty,$$

then, if  $|\underline{S}|$ ,  $|\bar{S}|$  are sufficiently large, the equilibrium PD ratio in period  $t$  is given by

$$(41) \quad \begin{aligned} \frac{P_t}{D_t} &= \left( 1 + \frac{W_t}{D_t} \right) \sum_{j=1}^{\infty} \left( \left( \delta^{\frac{1}{\gamma}} \right)^j \left( E_t^{\mathcal{P}} \prod_{i=1}^j \frac{1}{R_{t+i}} \right)^{1 - \frac{1}{\gamma}} \right) \\ &\quad - \frac{1}{D_t} E_t^{\mathcal{P}} \left( \sum_{j=1}^{\infty} \left( \prod_{i=1}^j \frac{1}{R_{t+i}} \right) W_{t+j} \right) \end{aligned}$$

Conditions (40) insure that the infinite sums in the pricing equation (41) converge.<sup>51</sup> Under the additional assumption that agents hold rational wage and dividend expectations and that  $W_t/D_t = \rho$ , equation (41) simplifies further to<sup>52</sup>

$$(42) \quad \begin{aligned} \frac{P_t}{D_t} &= (1 + \rho) \sum_{j=1}^{\infty} \left( \left( \delta^{\frac{1}{\gamma}} \right)^j \left( E_t^{\mathcal{P}} \prod_{i=1}^j \frac{1}{R_{t+i}} \right)^{\frac{\gamma-1}{\gamma}} \right) \\ &\quad - \rho \left( \sum_{j=1}^{\infty} (\beta^D)^j \left( E_t^{\mathcal{P}} \prod_{i=1}^j \frac{1}{R_{t+i}} \right) \right). \end{aligned}$$

We now discuss the implications of equation (42), focusing on the empirically relevant case where  $\rho > 0$  and  $\gamma > 1$ .

Consider first the upper term on the r.h.s. of equation (42), which is decreasing in the expected asset returns. This emerges because for  $\gamma > 1$  the wealth effect of a change in return expectations then dominates the substitution effect, so that expected asset demand and therefore the asset price has a tendency to decrease as return expectations increase. The negative wealth effect thereby increases in strength if the ratio of wage to dividend income ( $\rho$ ) increases. This is the case because higher return expectations also reduce the

<sup>51</sup>These are satisfied, for example, for the expectations associated with the perfect foresight RE solution. Equation (41) then implies for  $p = 1$  that the PD ratio equals the perfect foresight PD ratio, i.e., the value given by equation (24) with  $\varepsilon_t^W = 1$ , as is easily verified. Conditions (40) are equally satisfied for the subjective beliefs defined in section V, when considering the case with vanishing uncertainty ( $\sigma_\varepsilon^2, \sigma_v^2, \sigma_D^2, \sigma_W^2 \rightarrow 0$ ).

<sup>52</sup>In deriving equation (42) we abstract from transitional dynamics in  $W_t/D_t$  and set  $W_t/D_t = \rho$ .

present value of wage income.

Next, consider the lower term on the r.h.s. of equation (42), including the negative sign pre-multiplying it. This term depends positively on the expected returns and captures a substitution effect that is associated with increased return expectations. This substitution effect only exists if  $\rho > 0$ , i.e., only in the presence of non-dividend income, and it is increasing in  $\rho$ . It implies that increased return expectations are associated with increased stock demand and thus with a higher PD ratio in equilibrium. It is this term that allows the model to match the positive correlation between expected returns and the PD ratio.

This substitution effect is present even in the limiting case with log consumption utility ( $\gamma \rightarrow 1$ ). The upper term on the r.h.s. of equation (42) then vanishes because the substitution and wealth effects associated with changes in expected returns cancel each other, but the lower term still induces a positive relationship between prices and return expectations. The substitution effect is also present for  $\gamma > 1$  and can then dominate the negative wealth effect arising from the upper term on the r.h.s. of (42). Consider, for example, the opposite limit with  $\gamma \rightarrow \infty$ . Equation (42) then delivers

$$\frac{P_t}{D_t} = \sum_{j=1}^{\infty} \left( 1 + \rho \sum_{j=1}^{\infty} (1 - (\beta^D)^j) \right) \left( E_t^{\mathcal{P}} \prod_{i=1}^j \frac{1}{R_{t+i}} \right).$$

Since  $\beta^D > 1$ , there is a positive relationship between prices and expected asset returns, whenever  $\rho$  is sufficiently large. The two limiting results ( $\gamma \rightarrow 1$  and  $\gamma \rightarrow \infty$ ) thus suggest that for sufficiently large  $\rho$  the model can generate a positive relationship between return expectations and the PD ratio, in line with the evidence obtained from survey data.

### B. PD Ratio and Expected Capital Gains

Since future stock returns depend on the belief system and on current PD ratio, equation (42) does not yet give an explicit solution for the PD ratio. For the belief system introduced in section V<sup>53</sup>, we have  $E_t^{\mathcal{P}} [P_{t+i}] = (m_t)^i P_t$  and  $E_t^{\mathcal{P}} D_{t+i} = (\beta^D)^i D_t$ , so

<sup>53</sup>Appendix A.A9 proves that condition (40) is satisfied for all beliefs  $m_t > 0$ .

that without uncertainty

$$E_t^{\mathcal{P}} R_{t+i}^{-1} = \frac{E_t^{\mathcal{P}} P_{t+i-1}}{E_t^{\mathcal{P}} P_{t+i} + E_t^{\mathcal{P}} D_{t+i}} = \left( m_t + (\beta^D / m_t)^{i-1} \beta^D \frac{D_t}{P_t} \right)^{-1}.$$

Substituting this into (42) gives a non-linear relationship between the PD ratio and the subjective capital gain expectations  $m_t$ . Since a closed-form solution for the PD ratio is unavailable we show numerical solutions of this equation.

Figure 5 depicts the relationship between the PD ratio and  $m_t$  using the parameterization employed in our quantitative application later on, but abstracting from future uncertainty.<sup>54</sup> Figure 5 shows that there is a range of price growth beliefs around the perfect foresight value ( $m_t = \beta^D$ ) over which the PD ratio depends positively on expected price growth, similar to the positive relationship between expected returns and the PD ratio derived analytically in the previous section. Over this range, the substitution effect dominates the wealth effect because our calibration implies that dividend income finances only a small share of total consumption (approximately 4.3%). As a result, stock market wealth is only a small share of the total present value of household wealth (the same 4.3%) when beliefs assume their perfect foresight value ( $m_t = \beta^D$ ).

Figure 5 also reveals that there exists a capital gains belief beyond which the PD ratio starts to decrease for higher  $m$ . Mathematically, this occurs because if  $m_t \rightarrow \infty$ , expected returns also increase without bound<sup>55</sup>, so that  $E_t^{\mathcal{P}} \prod_{i=1}^j \frac{1}{R_{t+i}} \rightarrow 0$ . From equation (42) one then obtains  $\frac{P_t}{D_t} \rightarrow 0$ .

The economic intuition for the existence of a maximum PD ratio is as follows: for higher  $m_t$  the present value of wage income is declining, as increased price growth optimism implies higher expected returns<sup>56</sup> and therefore a lower discount factor. This can be seen by noting that the FOC (A15) can alternatively be written as

$$1 = \delta E_t^{\mathcal{P}} \left[ \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} R_{t+1} \right],$$

<sup>54</sup>The parameter values are given by those listed in table 2 and by the estimated parameters for the model with diagonal matrix from table 3, i.e.,  $g = 0.262$ ,  $\gamma = 2.03$ ,  $\delta = 0.99514$ ,  $p = 0.95$ .

<sup>55</sup>This follows from  $E_t^{\mathcal{P}} R_{t+i+1} = E_t^{\mathcal{P}} \frac{P_{t+i+1} + D_{t+i+1}}{P_{t+i}} > E_t^{\mathcal{P}} \frac{P_{t+i+1}}{P_{t+i}} = m_t$ .

<sup>56</sup>This is shown in appendix A.A11, which depicts the relationship between expected capital gains and expected returns at various forecast horizons.

which implies that increased return expectations  $E_t^P R_{t+1}$  imply a lower discount factor  $\delta E_t^P [(C_{t+1}/C_t)^{-\gamma}]$ .<sup>57</sup> With increased optimism, the present value of wage income thus falls. At the same time, stock market wealth initially increases strongly. Indeed, at the maximum PD ratio, stock market wealth amounts to approximately 4.5 times the value it assumes in the perfect foresight solution, see figure 5. This relative wealth shift has the same effect as a decrease in the wage to non-wage income ratio  $\rho$ . As argued in section VII.A, for sufficiently small values of  $\rho$  the income effect starts to dominate the substitution effect, so that prices start to react negatively to increased return optimism.

### C. Endogenous Boom and Bust Dynamics

We now explain how the interplay between price realizations and belief updating can temporarily de-link asset prices from their fundamental values. This process emerges endogenously and takes the form of a sustained asset price boom along which expected returns rise and that ultimately results in a price bust along which expected returns fall. This feature allows the model to generate volatile asset prices and to capture the positive correlation between expected returns and the PD ratio.

Consider figure 5 and a situation in which agents become optimistic, in the sense that their capital gains expectations  $m_t$  increase slightly above the perfect foresight value  $m_{t-1} = \beta^D$  entertained in the previous period.<sup>58</sup> Figure 5 shows that this increase in expectations leads to an increase in the PD ratio, i.e.,  $P_t/D_t > P_{t-1}/D_{t-1}$ . Moreover, due to the relatively steep slope of the PD function, realized capital gains will strongly exceed the initial increase in expected capital gains. The belief updating equation (39) then implies further upward revisions in price growth expectations and thus further capital gains, leading to a sustained asset price boom in which the PD ratio and return expectations jointly move upward.

The price boom comes to an end when expected price growth reaches a level close to where the PD function in figure 5 reaches its maximum.<sup>59</sup> At this point, stock prices

<sup>57</sup>This holds true under the maintained assumption of no or vanishing uncertainty.

<sup>58</sup>In the model with uncertainty, such upward revisions can be triggered by fundamentals, e.g., by an exceptionally high dividend growth realization in the previous period, which is associated with an exceptionally high price growth realization.

<sup>59</sup>In the model with noise, fundamental shocks, e.g., a low dividend growth realization, can cause the process to end well before reaching this point.

grow at most at the rate of dividends ( $\beta^D$ ), but agents hold considerably more optimistic expectations about future capital gains ( $m_t > \beta^D$ ). Investors' high expectations will thus be disappointed, which subsequently leads to a reversal. During a price boom, expected price growth and actual price growth thus mutually reinforce each other. The presence of an upper bound in prices implies, however, that the boom must come to an end. Price growth must thus eventually become very low, sending actual and expected stock price growth eventually down.

The previous dynamics are also present in the stochastic model considered in the next sections. They introduce low frequency movements in the  $PD$  ratio, allowing the model to replicate boom and bust dynamics and thereby empirically plausible amounts of asset price volatility, despite assuming standard consumption preferences. These dynamics also generate a positive correlation between the  $PD$  ratio and expected returns.<sup>60</sup>

In the deterministic model, however, the dynamics for the  $PD$  ratio tend to be temporary phenomena because beliefs tend to converge to the perfect foresight equilibrium:<sup>61</sup>

**LEMMA 1:** *Consider the limiting case without uncertainty and suppose investors hold rational dividend and wage beliefs.*

- 1) For any  $m_t > 0$ , we have  $\lim_{t \rightarrow \infty} m_t = \beta^D$ , whenever  $\lim_{t \rightarrow \infty} m_t$  exists.
- 2) For  $m_t$  sufficiently close to  $\beta^D$  and  $g < \frac{1}{2}$ , we have  $\lim_{t \rightarrow \infty} m_t = \beta^D$  if

$$(43) \quad -1 < \frac{\beta^D}{PD(\beta^D)} \left. \frac{\partial PD(m)}{\partial m} \right|_{m=\beta^D} < 1,$$

where  $PD(m)$  is the equilibrium  $PD$  ratio associated with beliefs  $m$ , as implied by equation (42).

The first result in the lemma provides a global convergence result. It shows that if beliefs settle down in this economy, they must settle down on the perfect foresight equilibrium value. When this is the case, equilibrium prices also converge to the perfect

<sup>60</sup>While the arguments above only show that expected capital gains correlate positively with the  $PD$  ratio, Appendix A.A11 shows that expected capital gains and expected returns comove positively, so that expected returns also comove positively with the  $PD$  ratio.

<sup>61</sup>The proof of lemma 1 can be found in appendix A.A10.



foresight value. While technically one cannot rule out convergence to deterministic or chaotic cycles, the second result in the lemma shows that locally beliefs do converge to the perfect foresight equilibrium, whenever the elasticity of the PD ratio with respect to price growth beliefs is below one in absolute value and the gain parameter  $g$  not too large. Condition (43) is satisfied, for example, for the parameterization of the estimated models reported in table 3.

To illustrate the global belief dynamics further, figure 6 depicts how beliefs evolve over time using the parameterization of the estimated model from table 3.<sup>62</sup> The arrows in the figure indicate, starting from any point  $(m_t, m_{t-1})$  in the plane, the direction in which the belief pair  $(m_t, m_{t-1})$  evolves.<sup>63</sup> The black dot indicates the position of the perfect foresight equilibrium  $(m_t = m_{t-1} = \beta^D)$ , which is a rest point of the dynamics. In line with what we find when simulating the model, figure 6 strongly suggests that beliefs globally converge to the perfect foresight equilibrium in the absence of stochastic disturbances.

Figure 6 also shows that it takes time for beliefs to settle down at the perfect foresight equilibrium and that agents will make persistent forecast errors along the transition path. In particular, when agents are optimistic and stock prices high ( $m_t > \beta^D$ ), there is a tendency for prices and beliefs to return towards their perfect foresight values, so that realized capital gains are low. Similarly, when agents are pessimistic and stock prices low ( $m_t < \beta^D$ ), the tendency of prices and beliefs to return to the perfect foresight values implies that realized capital gains are high. Note that this pattern of forecast errors is consistent with the one documented for the survey data in section II.

### VIII. Matching Asset Pricing Moments

This section evaluates the ability of the model to replicate key asset pricing moments when using dividend and wage shocks as fundamental driving forces.

The set of data moments that we seek to replicate is listed in the second column of table 3. The first eight asset pricing moments listed in the table are standardly used

<sup>62</sup>The vector field uses the parameters from the estimated model with diagonal matrix and estimated  $p$  reported in table 3. The vector fields for the other models from table 3 look almost identical.

<sup>63</sup>To increase readability of the graph, the length of the arrows  $l$  is nonlinearly rescaled by dividing by  $l^{3/4}$  and then linearly adjusted, so as to fit into the picture.

in the asset pricing literature to summarize the main features of stock price volatility.<sup>64</sup> They include the mean, standard deviation and autocorrelation of the quarterly PD ratio ( $E[PD]$ ,  $Std[PD]$  and  $Corr[PD_t, PD_{t-1}]$ , respectively), the mean and standard deviation of quarterly real stock returns ( $E[r^s]$  and  $Std[r^s]$ ), the risk-free interest rate ( $E[r^b]$ ) and the regression coefficient ( $\mathbf{c}$ ) and R-square value ( $R^2$ ) obtained in equation (5) when  $R_{t,t+N}$  is the five year ahead excess return.<sup>65</sup> As is well known, it is difficult for RE models with time separable utility functions to jointly match these moments.

We augment the standard set of moments in table 3 by the correlation between the PD ratio and expected stock returns from the UBS survey, denoted  $Corr[PD_t, E_t^P R_{t+1}]$ , so as to capture the behavior of survey data.

Numerically solving the non-linear asset pricing model with subjective beliefs turns out to be computationally time-consuming, despite the fact that we extensively rely on parallelization in the solution algorithms.<sup>66</sup> For this reason, we match the model to the data by calibrating most parameters and estimate only a key subset using the simulated method of moments (SMM).

Table 2 reports the calibrated parameters and the calibration targets.<sup>67</sup> The mean and standard deviation of dividend growth ( $\beta^D$  and  $\sigma_D$ ) are chosen to match the corresponding empirical moments of the U.S. dividend process. The ratio of non-dividend to dividend income ( $\rho$ ) is chosen to match the average dividend-consumption ratio in the U.S. for 1946-2011.<sup>68</sup> The standard deviation of wage innovations ( $\sigma_W$ ) and the covariance between wage and dividend innovations ( $\sigma_{DW}$ ) are chosen, in line with Campbell and Cochrane 1999, such that the correlation between consumption and dividend growth is 0.2 and the standard deviation of consumption growth is one seventh of the the standard

<sup>64</sup>These moments are also considered in Adam, Marcet, and Nicolini 2016.

<sup>65</sup>The regression also contains a constant whose value is statistically insignificant and not reported in the table.

<sup>66</sup>The numerical solution is obtained by numerically determining the stock demand function (22) solving the FOC (29) under the subjectively perceived dividend, wage and price dynamics, where agents understand that their beliefs evolve according to (39). We verify that in the limiting case without uncertainty, our numerical solution algorithm recovers the analytical solution derived in proposition 3. Furthermore, in the case with uncertainty, we insure the accuracy of the numerical solution by verifying that the Euler equation errors are in the order of  $10^{-5}$  over the relevant area of the state space. Insuring this requires a considerable amount of adjustment by hand of the grid points and grid size used for spanning the model's state space. Further details of the solution approach are described in appendix A.A12. The MatLab code used for solving the model is available upon request.

<sup>67</sup>The targets are chosen to match features of the fundamental processes emphasized in the asset pricing literature.

<sup>68</sup>See appendix A.A4 for further details.

deviation of dividend growth.<sup>69</sup> The perceived uncertainty in stock price growth ( $\sigma_\varepsilon$ ) is set equal to the empirical standard deviation of stock price growth.<sup>70</sup>

This leaves us with four remaining parameters: the updating parameter  $g$ , the time discount factor  $\delta$ , the risk aversion parameter  $\gamma$  and the persistence parameter for the wage-dividend ratio  $p$ . Letting  $\theta = (g, \delta, \gamma, p)$  denote this set of parameters and  $\Theta$  the set of admissible values, the SMM estimate  $\hat{\theta}$  is given by

$$(44) \quad \hat{\theta} = \arg \min_{\theta \in \Theta} [\hat{\mathcal{S}} - \mathcal{S}(\theta)]' \hat{\Sigma} [\hat{\mathcal{S}} - \mathcal{S}(\theta)],$$

where  $\hat{\mathcal{S}}$  is the set of moments in the data to be matched (the ones listed in table 3),  $\mathcal{S}(\theta)$  the corresponding moment from the model for parameter values  $\theta$  and  $\hat{\Sigma}$  a weighting matrix.

We pursue two estimation approaches, one that implements efficient SMM and an alternative one that emphasizes more directly the replication of the data moments. Both approaches have advantages and disadvantages, as we discuss below, but ultimately deliver similar estimates and model moments. All estimations exclude the risk free rate from the set of moments to be matched, as the model has a hard time in fully replicating the equity premium. We thus only report the risk-free interest rate implied by the estimated models.

Our first estimation approach chooses the weighting matrix  $\hat{\Sigma}$  in equation (44) to be equal to the inverse of the estimated covariance matrix of the data moments  $\hat{\mathcal{S}}$ , as required for efficient SMM estimation. While an efficient weighting matrix is desirable for estimation, it also causes some difficulties. First, the weighting matrix turns out to be approximately singular; following Adam, Marcet and Nicolini 2016 we thus exclude some moments from the estimation that are nearly redundant, namely we exclude the excess return regression coefficient ( $\mathbf{c}$ ) from the set of estimated moments.<sup>71</sup> Second, we also have to exclude the correlation between the PD ratio and surveyed expected returns, since the short sample for surveys provides a less reliable estimate of  $\hat{\Sigma}$ .

<sup>69</sup>For details on how this can be achieved, see appendix A.A4.

<sup>70</sup>Since the gain parameter  $g$  will be small, the contribution of  $\sigma_v^2$  in (26) is negligible.

<sup>71</sup>See section 5 and the online appendix in Adam, Marcet and Nicolini 2016 for details on asymptotic distribution of SMM, how to use a systematic criterion for excluding moments, details on how to estimate  $\hat{\Sigma}$  and how to compute  $\mathcal{S}(\theta)$  using small samples.

The second estimation approach uses a diagonal weighting matrix  $\widehat{\Sigma}$  in equation (44), with the diagonal entries consisting of the inverse of the individually estimated variances of the corresponding data moments in  $\widehat{\mathcal{S}}$ . Although this alternative is less efficient from an econometric point of view, the matrix  $\widehat{\Sigma}$  is now guaranteed to be invertible, therefore we can use the full set of moments in the estimation.<sup>72</sup> Furthermore, this criterion just minimizes the sum of t-statistics of all the moments, therefore emphasizes more directly matching the moments, as typically pursued in the asset pricing literature.

An unconstrained minimization of the objective function (44) over  $\theta = (g, \delta, \gamma, p)$  turns out to be numerically unstable and computationally too costly. For this reason, we impose additional restrictions on the parameter space  $\Theta$ . These restrictions can only constrain the empirical performance of the model, so that the goodness of fit results presented below constitute a lower bound on what the model can potentially achieve. Specifically, we impose  $\delta(\beta^D)^{-\gamma} = 0.995(\beta^D)^{-2}$ , where  $\beta^D$  assumes the value from table 3. This additional restriction is inspired by the fact that - according to our experience - the model can perform reasonably well for  $(\delta, \gamma) = (0.995, 2)$  and helps resolving numerical instabilities in our solution routines. We furthermore restrict the persistence parameter to  $p \in \{0.95, 0.999\}$ , which is inspired by the fact that the sample autocorrelation of  $\log(1 + W_t/D_t)$  is very high in US postwar data (about 0.99).

Table 3 reports the estimation outcome in terms of implied model moments, estimated parameters and t-ratios.<sup>73</sup>

In terms of estimated parameters, the discount factor is estimated to be close to one and relative risk aversion is slightly above 2. The estimated gains ( $\widehat{g}$ ) are very close to the values obtained in the empirical section II.D.

For both estimation approaches, the model matches the data moments rather well. In particular, the model easily replicates the positive correlation between the PD ratio and expected stock returns ( $\text{Corr}[\text{PD}_t, E_t^P R_{t+1}]$ ), in line with the value found in the survey data. This is achieved, even though this moment was not used in the estimation using the

<sup>72</sup> As mentioned before, the risk free rate is always excluded from estimation.

<sup>73</sup> The t-ratio is the ratio of the gap between the model and the data moment over the standard deviation of the moment in the data, as implied by the weighting matrix. For excluded moments we use the individually estimated standard deviations in the numerator. For the case of an efficient weighting matrix, the t-ratio for moments included is computed according to the proper covariance matrix that delivers a standard normal distribution, see the online appendix in Adam Marcat and Nicolini 2016.

efficient weighting matrix. The model also performs well in terms of producing sufficient volatility for the PD ratio ( $\text{Std}[\text{PD}]$ ) and stock returns ( $\text{Std}[r^s]$ ). If anything, the model tends to produce too much volatility. The model also succeeds in replicating the mean and autocorrelation of the PD ratio ( $E[\text{PD}]$  and  $\text{Corr}[\text{PD}_t, \text{PD}_{t-1}]$ ) and the evidence on excess return predictability ( $\mathbf{c}$  and  $R^2$ ), even though the regression coefficient  $\mathbf{c}$  was not included in the set of moments to be matched.

The most significant shortcoming of the model concerns the equity premium. While it matches the average stock return ( $E[r^s]$ ), it predicts a too high value for the risk free rate ( $E[r^b]$ ). Nevertheless, the model can explain about half of the equity premium observed in the data. Given the low estimated value for the degree of relative risk aversion ( $\hat{\gamma}$ ), this is remarkable and we explore the mechanism behind this outcome further in section VIII.B below.

Abstracting from the mean risk free rate, the estimation using the diagonal weighting matrix generates t-ratios below two for all model moments and often even t-ratios below one. The estimation using the full matrix performs similarly well, but implies a slightly high volatility for stock returns and the PD ratio. The empirical performance of the subjective belief model is overall very good, especially when compared to the performance of the rational expectations version of the model, as reported in table 4 using the same parameters as for the estimated model (with diagonal matrix) in table 3. Under RE, the t-ratios all increase in absolute terms, with some of the increases being quite dramatic. With objective price beliefs, the model produces insufficient asset price volatility (too low values for  $\text{Std}[\text{PD}]$  and  $\text{Std}[r^s]$ ) and the wrong sign for the correlation between the PD ratio and expected stock returns ( $\text{Corr}[\text{PD}_t, E_t^P R_{t+1}]$ ). It also gives rise to a small negative equity premium.<sup>74</sup> These features are rather robust across alternative parameterizations of the RE model and highlight the strong quantitative improvement obtained by incorporating subjective belief dynamics.

Table 5 shows that the performance of the subjective belief model is also rather robust across different values for the persistence parameter  $p$  of the wage-dividend process. The table reports the estimation outcomes when repeating the estimation from table 3

<sup>74</sup>For the considered parameterization, the equity premium is slightly negative because stocks are a hedge against wage income risk.

but imposing  $p = 0$  (no persistence) and  $p = 0.999$  (near unit root behavior). The ability of the subjective belief model to match the empirical asset pricing moments is hardly affected by the persistence parameter.

Figure 7 illustrates how the subjective belief model improves empirical performance. The figure depicts the equilibrium PD ratio (y-axis) as a function of agents' capital gain beliefs (x-axis). It graphs this relationship once for the model with uncertainty (black line) and once for the vanishing noise limit analyzed in the previous section (red line).<sup>75</sup> While the presence of price, dividend and wage risk lowers the equilibrium PD ratio compared to a setting without risk, the functional form of the relationship remains qualitatively unchanged. The intuition from the vanishing noise limit thus carries over to the model with noise: the model continues to give rise to occasional boom and bust dynamics in asset prices.

#### A. Further Evidence on Model Performance

Table 6 presents further asset pricing moments that have not been used in the estimation of the model. It reports data moments, the corresponding model moments for the estimated subjective belief model from table 3, for the RE version of model and for an augmented subjective belief model featuring stock supply shocks, as introduced below.

Table 6 shows that the subjective belief model successfully replicates the low volatility of the risk free interest rate present in the data. The subjective belief model even slightly underpredicts volatility. This shows that the ability to generate highly volatile stock returns does not rely on counterfactually making the risk free interest rate very volatile. In terms of matching the low volatility of the risk free rate, the performance of the subjective belief model is approximately as good as that of the RE model.

Table 6 also presents evidence on the quarterly autocorrelations of stock returns and excess stock returns. It shows that the subjective belief model overpredicts both autocorrelations relative to the ones present in the data. As table 6 shows, this feature arises not solely due to the presence of subjective beliefs: the rational expectations versions of the model also overpredicts the autocorrelation of excess stock returns.

<sup>75</sup>For the stochastic solution, the equilibrium PD in figure 7 is determined from the market clearing condition (38) assuming  $W_t/D_t = \rho$ , to be comparable with the value this variable assumes in the vanishing risk limit. The figure assumes the parameters implied by the estimated model with diagonal matrix from table 3.

The failure of the subjective belief model to deliver a near zero high-frequency autocorrelation for returns is related to the fact that the low frequency returns generated by learning also show up at high frequency because the model features few transitory shocks to stock prices. As we show below, the subjective belief model becomes consistent with the observed value of the quarterly autocorrelations of stock returns and excess stock returns, when adding small iid disturbances to the model.

To illustrate this point in a simple way, we consider in appendix A.A13 a model where equilibrium stock supply in period  $t$  satisfies

$$S_t = e^{\varepsilon_t^s},$$

where  $\varepsilon_t^s \sim iiN(-\sigma_{\varepsilon^s}^2/2, \sigma_{\varepsilon^s}^2)$  denotes an exogenous shock to the supply of stocks. This shock can represent extraneous demand for stocks, e.g., the one arising from noise or liquidity traders.<sup>76</sup> This generalized setup nests the model without noise traders for the special case  $\sigma_{\varepsilon^s} = 0$ .

As it turns out, small positive values for the standard deviation  $\sigma_{\varepsilon^s}$  bring the model fully in line with the observed autocorrelations. The last two columns in table 6 illustrate this fact and report the model implied autocorrelations, when setting  $\sigma_{\varepsilon^s} = 1.25 \cdot 10^{-3}$ , which implies that the standard deviation of aggregate stock supply is 0.125% of the average outstanding amount of stocks (which is equal to one). Table 6 shows that the autocorrelations are now fully consistent with those observed in the data, while the standard deviation of the risk free rate is hardly affected. Moreover, as shown in appendix A.A13, the presence of stock supply shocks has only small effects on the other asset pricing moments reported in table 3.

### B. Subjective Consumption Plans and the Sharpe Ratio

A somewhat surprising feature of the estimated models in Table 3 is that they give rise to a fairly high Sharpe ratio, despite the fact that the estimated risk aversion is relatively low. Specifically, the estimated models imply that the quarterly unconditional Sharpe

<sup>76</sup>Alternatively, the shocks  $\varepsilon_t^S$  may capture changes to asset float, as discussed in Ofek and Richardson 2003 and Hong, Scheinkman, and Xiong 2006. In any case, they capture (exogenous) stock demand or supply that is not coming from the consumers described in equation (14).

ratio,  $(E[r^s] - E[r^b]) / Std[r^s]$ , is close to 0.1. Since the model implied equity premium is too low, this value falls significantly short of the value of Sharpe ratio in the data (0.22). Yet, it significantly exceeds the value it would assume if agents held rational price expectations. Under RE, the Sharpe ratio would be tiny and approximately be given by<sup>77</sup>

$$(45) \quad \frac{E[r^s] - E[r^b]}{Std[r^s]} \approx \gamma \cdot E [Std_t(C_{t+1}/C_t)] = 0.00576.$$

In the presence of subjective price beliefs, however, the Sharpe ratio is approximately equal to

$$(46) \quad \frac{E[r_{t+1}^s] - E[r_t^b]}{Std[r^s]} \approx \underbrace{\frac{E[r_{t+1}^s] - E[E_t^P[r_{t+1}^s]]}{Std[r^s]}}_{\text{Subj. return pessimism}} + \gamma \underbrace{\frac{E[Std_t^P[C_{t+1}/C_t]]}{E[Std_t[C_{t+1}/C_t]]}}_{\text{Ratio of subj. to obj. consumption volatility}} E[Std_t[C_{t+1}/C_t]],$$

which illustrates the existence of two additional factors affecting the Sharpe ratio: subjective return pessimism, as present in models where agents entertain robustness concerns, e.g., Cogley and Sargent 2008, and the relative volatility of subjective consumption plans compared to the ones emerging under rational expectations. Subjective return pessimism contributes to the Sharpe ratio because pessimism depresses stock prices and thereby generates a higher equity premium ex-post. Subjective consumption volatility affects the equity premium for standard reasons.

For the estimated models from table 3, we find that about one third of the model's Sharpe ratio is due to subjective return pessimism and about two thirds due to the second term on the r.h.s of equation (46).<sup>78</sup> In both estimated models, the second component is so large because the average subjective standard deviation of consumption growth is

<sup>77</sup> See appendix A.A14 for a derivation of equations (45) and (46).

<sup>78</sup> The approximate Sharpe ratio decomposition (46) works quantitatively well for the estimated models from table 3: the actual Sharpe ratio are 10.7% and 10.2% for the model estimated using a diagonal and an efficient matrix, respectively, while the terms on the r.h.s. in equation (46) sum to 10.5% and 10.9%.



about 12.5 times larger than the average objective standard deviation.<sup>79</sup>

Analyzing which future shocks contribute in agents' mind to this high level of subjectively expected consumption volatility, we find that transitory price growth shocks are the main culprit, i.e., the perceived shocks  $\varepsilon_{t+1}$  in the subjective price evolution equation (26). These shocks move prices permanently up (or down), without an associated change in fundamentals, and prompt agents to make contingent plans to sell (or buy) stocks in the future in response to these shocks. These expectations of future trade increase subjective expected consumption volatility, while the representative agent assumption insures that agents will never actually buy or sell stocks in equilibrium. As a result, subjective and objective consumption volatility diverge considerably.

The reason for this discrepancy is largely due the fact that the homogeneous agent model imposes zero trade in equilibrium. Equilibrium consumption is then not affected by stock prices. As we now illustrate, one can resolve this discrepancy by introducing heterogeneity among agents.

We show this using two simple heterogeneous agent extensions of our baseline model. The first extension studies an ad-hoc modification in which agents' beliefs at time  $t$  are given by

$$(47) \quad \ln m_t^i = \ln m_t + \ln \mu_t^i,$$

where  $\ln \mu_t^i \sim iiN(0, \sigma_\mu^2)$  is an idiosyncratic and *transitory* shock to agent  $i$ 's capital gain optimism and where  $m_t$  evolves as in the baseline model. The second extension considers a fully microfounded extension, where - in line with the belief updating equation (39) - agents' expectations evolve according to

$$(48) \quad \ln m_t^i = \ln m_{t-1}^i + g (\ln P_{t-1} - \ln P_{t-2} - \ln m_{t-1}^i) + g \ln \varepsilon_t^{1,i},$$

where  $(g \ln \varepsilon_t^{1,i}) \sim iiN(0, \sigma_\varepsilon^2)$  is an idiosyncratic with rather *persistent* effects on beliefs.<sup>80</sup> In both extensions, belief heterogeneity induces trade in equilibrium, which in-

<sup>79</sup>The precise ratios are 12.87 (12.64) for the model from table 3 estimated using a diagonal (efficient) matrix. For both models, we have  $E[Std_t[C_{t+1}/C_t]] = 0.00273$ .

<sup>80</sup>As is clear from appendix A.A7, the shock  $\ln \varepsilon_t^{1,i}$  represents information that the agent receives about the perceived

creases the objective volatility of individual consumption growth relative to the representative agent setting. We then calibrate the variances  $\sigma_\mu^2$  and  $\sigma_\varepsilon^2$  such that the objective and subjective standard deviations are approximately aligned ( $E [Std_t^P [C_{t+1}^i/C_t^i]] \approx E [Std_t [C_{t+1}^i/C_t^i]]$ ).

Table 7 reports the asset pricing moments and t-ratios for both heterogeneous agent extensions.<sup>81</sup> As becomes clear from table 7, the extended models can replicate the asset pricing moments equally well as the baseline model, but also align the subjective and objective standard deviations of consumption growth. The ability of the model to generate a low risk free rate is actually improved relative to the representative agent model, with the average risk free rate even becoming slightly negative for the model with persistent shocks.<sup>82</sup> Overall, table 7 illustrates the robustness of our quantitative results towards allowing for investor heterogeneity. Adam et al. 2015 explore in greater detail a heterogeneous agent version of the present model and show that the model successfully matches important stylized facts regarding trading volume.

### C. Expectational Errors

Since our model matches stock price behavior and the comovement pattern between survey expectations and stock prices, agents' expectations in the model fail to satisfy the rational expectations hypothesis, in line with what has been documented for the survey evidence. To document the quantitative importance of this fact, this section applies the RE tests from section II.B to the subjectively expected one-quarter-ahead capital gain expectations from our estimated models. In particular, we test the acceptance rates of the rational expectations hypothesis ( $H_0 : c = \mathbf{c}$ ) at various significance intervals and various sample lengths, using the test statistic described in proposition 1.<sup>83</sup> Table 8 reports the outcomes.<sup>84</sup> It shows that at conventional significance levels the rational expectations

transitory price shock. Since these shocks exist only in subjective terms, the baseline model sets  $\ln \varepsilon_t^i = \ln \varepsilon_t = 0$  for all  $t$ .

<sup>81</sup>The table uses the model parameterization of the estimated model from table 3 (diagonal matrix). For the case with persistent belief shocks, we slightly adjust the updating gain  $g$  to improve the match with the asset pricing moments. The simulations use 51 different agents.

<sup>82</sup>The strong effect on the risk free rate is due to the tight borrowing limits ( $B_t^i \geq 0$ ) that we impose, see the discussion in section III. The effect would likely be less strong for less strict borrowing limits, but this would require fundamental changes to the solution approach, including an increase in the number of state variables.

<sup>83</sup>We include the small sample bias correction reported in proposition 1.

<sup>84</sup>Rejection frequencies have been computed from 1000 random samples of the specified sample length that were randomly drawn from a time series of 100,000 simulated model periods, with the first 10,000 periods being

hypothesis can be rejected with high likelihood, even for relatively short sample lengths. This is in line with the empirical evidence documented in section II.

### IX. Historical PD Ratio and Survey Evidence

This section shows that the estimated subjective belief model successfully replicates the low-frequency movements of the postwar U.S. PD ratio, as well as the available time series of survey expectations. We illustrate this point using the model from table 5, which has been estimated under the restriction that the wage-dividend ratio displays close to unit root behavior ( $p = 0.999$ ). This is motivated by the fact that estimated models that imply more mean reversion in the wage-dividend ratio, say the ones from table 3, cannot capture the time series behavior of the PD ratio after the year 1990. After 1990, the wage-dividend ratio started to persistently diverge from a value close to its sample mean of 22 to a value close to 12 at the end of the sample period in 2012.

To determine the model fit, we feed the historically observed price growth observations into the model's belief updating equation (39), so as to obtain a model-implied belief process.<sup>85</sup> We then combine this belief series with the historically observed wage-dividend ratio to obtain from the model's equilibrium pricing function the model implied prediction for the postwar PD series. Formally, this is done by plugging implied expectations and wage-dividend ratio in equation (38) and backing out the corresponding  $PD$  ratio. In doing so, we choose the unobserved price growth belief at the start of the sample period ( $m_{Q1:1946}$ ) so as to minimize the sum of absolute deviations of the model-implied PD ratio from the PD ratio in the data.<sup>86</sup>

The resulting belief sequence is depicted in figure 8. Figure 9 graphs the model implied PD ratio together with the PD ratio from the data. It reveals that the model captures a lot of the low-frequency variation in the historically observed PD ratio. It captures particularly well the variations before the year 2000, including the strong run-up in the PD ratio from the mid 1990's to the year 2000. The model also predicts a strong decline of the PD ratio after the year 2000, but overpredicts the decline relative to the data.

discarded as burn-in.

<sup>85</sup>Since we cannot observe the shocks  $\ln \varepsilon_t^1$  in (39), we set them equal to zero. Gaps between the model predicted and actual PD ratio may thus partly be due to these shocks.

<sup>86</sup>This delivers initial price growth expectations in Q1:1946 of -1.7%.

The gap after the year 2000 emerging in figure 9 is hardly surprising, given the empirical evidence presented in figure 4, which shows that the relationship between the PD ratio and the expectations implied by equation (13) has shifted upward in the data following the year 2000. While we can only speculate about potential reasons causing this shift, the exceptionally low real interest rates implemented by the Federal Reserve following the reversal of the tech stock boom and following the collapse of the subsequent housing boom may partly contribute to the observed discrepancy. This suggests that incorporating the asset pricing effects of monetary policy decisions might improve the model fit. This is, however, beyond the scope of the present paper.

Figure 10 depicts the model-implied price growth expectations and the price growth expectations from the UBS survey.<sup>87</sup> While the model fits the survey data rather well, it predicts after the year 2003 considerably lower capital gains expectations, which partly explains why the model underpredicts the PD ratio in figure 9 towards the end of the sample period. While the expectations gap in figure 10 narrows considerably after the year 2004, this fails to be the case for the PD ratio in figure 9. Underprediction of expected price growth thus explains only partly the deterioration of the fit of the PD ratio towards the end of the sample period.

## X. Robustness Analysis

This section explores the robustness of our main findings to changing key model parameters and to using more general price belief systems. The next section studies the pricing effects of alternative model parameters. Section X.B discusses the implications of more general price belief systems that incorporate mean-reversion of the PD ratio. Finally, section X.C considers the pricing and welfare implications of introducing agents who use current price information for updating beliefs.

### A. Model Parameterization

This section studies how the model's ability to generate boom and bust dynamics in stock prices depends on key model parameters. The equilibrium pricing function for our baseline parameterization, depicted in figure 5, allows for self-reinforcing stock price

<sup>87</sup>See footnote 35 for how to compute price growth expectations from the UBS survey.

boom and bust dynamics because the price dividend ratio increases initially strongly with capital gain optimism. This section explores the robustness of this feature by studying how the equilibrium pricing is affected by the coefficient of relative risk aversion ( $\gamma$ ), the discount factor ( $\delta$ ), and the average wage to dividend income ratio ( $\rho$ ).

Figure 11 depicts the equilibrium pricing function for alternative parameter choices. Each panel plots the pricing function from our baseline parameterization, as well as the those generated by increasing or decreasing the values for  $\gamma$ ,  $\delta$  and  $\rho$ . The top panel, for example, shows that lowering (increasing) the coefficient of relative risk aversion increases (reduces) the hump in the PD function and moves it to the left (right), thereby causing asset price booms to become more (less) likely and larger in size. Similar effects are associated with increasing (decreasing) the wage to dividend income ratio ( $\rho$ ) and with increasing (decreasing) the discount factor ( $\delta$ ). Overall, figure 11 shows that the model can produce hump-shaped equilibrium PD functions over a fairly wide set of parameter specifications.

### B. Generalized Belief System

This section considers a generalized price belief system, which implies that investors expect the PD ratio to eventually mean-revert over time. This is motivated by the fact that the price belief system (26)-(27) employed in the main part of the paper, along with knowledge of the dividend process, does not imply that agents expect the log of the  $PD$  ratio to mean revert over time.

We find that the pricing implications of the model do not depend on the expected long run behavior for the PD ratio and that very similar equilibrium pricing functions can be generated by belief systems that imply persistent changes in the PD ratio but where the PD ratio is ultimately mean reverting.

It is important to note that the survey data provide little support for mean reverting price growth expectations over the available forecast horizons and thus for mean reversion in the expected PD ratio. In particular, the last four rows in table 1a show that the regression coefficient  $\hat{c}$  obtained from regressing the expected 10 year ahead capital gain forecasts from the Shiller survey on the PD ratio is more than 10 times larger than the corresponding regression coefficient obtained from regressing the one year ahead fore-

cast on the PD ratio. The expected annualized ten year price growth thus reacts stronger to movements in the PD ratio than the expected one year price growth. This shows that one should incorporate only a mild degree of mean reversion into subjective price growth beliefs, as the belief system would otherwise become inconsistent with the survey evidence.

To analyze the effects of mean reversion in price beliefs, we now assume that investors perceive prices to evolve according to

$$(49) \quad \ln P_{t+1} = \ln \beta_{t+1} + \ln P_t + (1 - \eta_{PD}) (\ln PD - \ln P_t/D_t) + \ln \varepsilon_{t+1}$$

$$(50) \quad \ln \beta_{t+1} = (1 - \eta_\beta) \ln \beta^D + \eta_\beta \ln \beta_t + \ln v_{t+1},$$

where  $\ln PD$  denotes the perceived long run mean of the log PD ratio and  $\eta_{PD} \in [0, 1]$ ,  $\eta_\beta \in [0, 1]$  are given parameters. For  $\eta_{PD} = \eta_\beta = 1$  these equations deliver the benchmark price belief system (26)-(27) studied in the main part of the paper.<sup>88</sup> For  $\eta_\beta < 1$ , equation (50) implies that agents expect mean reversion in the persistent price growth component  $\ln \beta_t$  towards the mean growth rate of dividends ( $\ln \beta^D$ ); if in addition  $\eta_{PD} < 1$ , equation (49) implies that agents expect  $\ln P_t/D_t$  to eventually return to its long run mean  $\ln PD$ .<sup>89</sup>

Suppose that  $\eta_\beta < 1$  and  $\eta_{PD} < 1$ , that agents are optimistic about future capital gains, i.e., they believe  $\beta_t$  to be above  $\beta^D$ , and that agents observe a PD ratio above its long-run mean ( $P_t/D_t > PD$ ). Provided  $\eta_\beta$  and  $\eta_{PD}$  are sufficiently close to one, equations (49)-(50) imply that agents expect a fairly persistent boom in the PD ratio, as is the case with the benchmark belief system (26)-(27). Yet, unlike with the benchmark, they also expect a price bust further down the road, because the PD ratio is expected to eventually return to its long run value.

Equations (49) and (50) jointly imply that optimal belief updating about the unobserved persistent stock price growth component  $\ln \beta_t$  is described by a generalized version of equation (39), which states that the posterior mean  $\ln m_t \equiv E^{\mathcal{P}}(\ln \beta_t | P^t)$  in

<sup>88</sup>As before, we assume that agents have rational expectations about the dividend and wage income processes.

<sup>89</sup>This can be seen by subtracting (15) from (49).

steady state evolves according to<sup>90</sup>

$$(51) \quad \ln m_t = (1 - \eta_\beta) \ln \beta^D + \eta_\beta \ln m_{t-1} \\ + g \left( \begin{array}{c} \ln P_{t-1} - \ln P_{t-2} - \ln m_{t-1} \\ -(1 - \eta_{PD}) (\ln PD - \ln P_{t-2}/D_{t-2}) \end{array} \right) + g \ln \varepsilon_t^1.$$

Using the generalized belief and updating equations, figure 12 depicts the impact of different perceived values for  $\eta_{PD}$  and  $\eta_\beta$  on the equilibrium pricing function when setting the perceived long-run mean  $\ln PD$  in (49) equal to the perfect foresight value of the  $PD$  ratio.<sup>91</sup> The figure's top panel plots the effects of decreasing  $\eta_{PD}$  below one, while keeping  $\eta_\beta = 0.9999$ . For the considered values of  $\eta_{PD}$ , agents expect the log PD ratio to mean revert by 1%, 2%, or 3% per year towards its long run mean ( $\ln PD$ ). The figure also plots the outcome when there is virtually no mean reversion ( $\eta_{PD} = 0.9999$ ). The top panel of figure 12 shows that by introducing mean reversion, one pushes the peak of the equilibrium PD ratio to the right and also lowers its height. The shift to the right occurs because agents only expect a persistent boom in the PD ratio if the increase in the PD ratio implied by the persistent growth component ( $\ln \beta_t$ ) outweighs the mean reversion generated by the negative feedback from the deviation of the PD from its long run value in equation (49). The downward shift in the PD ratio occurs because mean reversion causes investors to expect lower and eventually negative returns sooner.

The second panel in figure 12 depicts the effects of decreasing  $\eta_\beta$  below one, while keeping  $\eta_{PD} = 0.9999$ . As before, we consider values for  $\eta_\beta$  that imply virtually no mean reversion and mean reversion by 1%, 2% and 3% per year towards the long run value ( $\ln \beta^D$ ). The panel shows that the pricing implications are very similar to those of decreasing  $\eta_{PD}$ .

Finally, the bottom panel depicts the effects of jointly decreasing  $\eta_\beta$  and  $\eta_{PD}$ . The pricing implication of such simultaneous changes turn out to be considerably stronger,

<sup>90</sup>Since equation (49) only introduces an additional observable variable into (26) and equation (50) only adds a known constant and known mean reversion coefficient relative to (27), the arguments delivering equation (39) as optimal Bayesian updating directly generalize to equation (51).

<sup>91</sup>We assume that  $\sigma_\varepsilon^2$  and  $\sigma_v^2$  assume the same values as in the baseline specification, so that the gain parameter also remains unchanged at  $g = 0.02515$ . This parameterization makes sure that - for the values of  $\eta_\beta$  and  $\eta_{PD}$  considered below - the perceived standard deviation for stock price growth implied by equations (50)-(49) approximately matches the standard deviation of stock price growth in the data.

compared to the case where only one of the persistence parameters decreases. Figure 13 illustrates why this sharp difference occurs. It graphs the expected path of the PD ratio for different parameter combinations  $(\eta_{PD}, \eta_{\beta})$ .<sup>92</sup> If either  $\eta_{PD}$  or  $\eta_{\beta}$  are close to 1, agents expect a rather prolonged stock price boom that is expected to revert only in the distant future. Expected stock returns are thus high for many periods before they turn negative. This differs notably from the case where both persistence parameters fall significantly below one ( $\eta_{PD} = \eta_{\beta} = 0.97^{\frac{1}{4}}$ ). The expected stock price boom is then much smaller and considerably more short-lived, so that returns are lower and expected to become negative earlier. From the discussion following equation (42), it should be clear that the implied path for expected returns then cannot sustain a high PD ratio as an equilibrium outcome.

Summing up, the model continues to give rise to hump-shaped equilibrium pricing functions, even if agents ultimately expect mean reversion in the PD ratio, provided the generalized belief system implies that agents expect elevated capital gains to be sufficiently persistent.

### C. Current Price Information for Belief Updating

The baseline model specification postulates a belief structure that makes it subjectively optimal for agents to update beliefs  $m_t$  based on price information up to period  $t - 1$  only, see equation (39). The present section demonstrates that lagged belief updating is also objectively optimal for agents: belief updating with current price observations robustly generates lower experienced utility for agents.<sup>93</sup> In this sense, the baseline specification with lagged belief updating can be interpreted as the limiting outcome of a setting in which agents choose - based on experienced utility - whether or not to use current prices for updating beliefs.

To show that lagged updating generates objectively higher utility, we consider an extended model setup where a share  $\alpha \in [0, 1]$  of investors updates beliefs using current price growth information, with the remaining share  $1 - \alpha$  using lagged price growth infor-

<sup>92</sup>The figure assumes that the equilibrium PD ratio initially equals  $PD_0 = 150$ , i.e., it is above its long run value, and that agents are mildly optimistic about future capital gains with  $\ln m_0 = 1\% > \ln \beta^D$ .

<sup>93</sup>This finding is in line with results reported in Adam et al. 2015, who show that agents whose beliefs are more reactive to price growth observations tend to do worse than agents whose beliefs display less sensitivity.



mation.<sup>94</sup> While the beliefs of lagged updaters are predetermined, the beliefs of current updaters vary simultaneously with the stock price  $P_t$ . The latter creates a potential for multiple market clearing equilibrium price and belief pairs.

To assess the implications of simultaneous belief updating, we consider the estimated benchmark parameters from table 3 (diagonal matrix) and different values for  $\alpha \in [0, 1]$ . To select between equilibrium prices in periods where multiple equilibrium price exist, we consider two alternative selection rules.<sup>95</sup> The first rule selects the equilibrium price that is closest to the previous period's market clearing price. The second selection rule selects the price that is furthest away from the previous market clearing price.

Table 9 reports the expected discounted utility of agents that use current and lagged belief updating for different values of  $\alpha$  and for the two considered selection rules.<sup>96</sup> It shows that independently of the selection rule and independently of the share of current updaters  $\alpha$ , utility of lagged updaters always exceeds that of current updaters. Current updaters thus have an incentive to switch to lagged updating, i.e., to the setting considered in our baseline specification.

## XI. Conclusions

We present a model with rationally investing agents that gives rise to market failures in the sense that the equilibrium stock price deviates from its fundamental value. These deviations take the form of asset price boom and bust cycles that are fueled by the belief-updating dynamics of investors who behave optimally given their imperfect knowledge about the behavior of stock prices. Optimal belief updating also causes investors' subjective capital gain expectations to comove positively with the price-dividend ratio, consistent with the evidence available from investor surveys.

As we argue, these features cannot be replicated within asset pricing models that impose rational price expectations. Moreover, the developed statistical tests show that the behavior of survey return expectations is incompatible with the rational expectations hy-

<sup>94</sup>Updaters using current price information update beliefs according to equation (39), but replace  $\ln P_{t-1} / \ln P_{t-2}$  by  $\ln P_t / \ln P_{t-1}$  on the right-hand side.

<sup>95</sup>In the vast majority of cases, we find three market clearing equilibrium prices, conditional on there being multiplicity. In less than 0.1% of the periods with multiplicities we find five market clearing prices.

<sup>96</sup>The table reports the unconditional expectation of discounted consumption utility using the objective distribution for consumption, as realized in equilibrium. Appendix A.A15 reports the associated asset pricing moments.

pothesis.

Taken together, this suggests that asset price dynamics are to a large extent influenced by investors' subjective optimism and pessimism, i.e., the asset price fluctuations observed in the data are to a large extent inefficient. The inefficiency arises because equilibrium stock prices are determined by the sum of dividends that is discounted using the stochastic discount factor implied by investors' subjective consumption plans, which are influenced by investors' subjective price beliefs. This differs from the standard setup under RE, where stock prices equal the sum of dividends discounted by the 'objective' stochastic discount factor.

Due to the simplicity of the setup, these inefficient price fluctuations do not yield adverse welfare implications in our baseline model.<sup>97</sup> For models incorporating investor heterogeneity, e.g., the extension presented in section VIII.B, or models featuring endogenous output or stock supply processes, stock price fluctuations may have significant effects on welfare. Exploring these within a setting that generates quantitatively credible amounts of asset price fluctuations appears to be an interesting avenue for further research. Such research will in turn lead to further important questions, e.g., whether policy can and should intervene with the objective to stabilize asset prices.

In deriving our results, we assumed that all agents in the economy become more (or less) optimistic when observing capital gains above (or below) their expectations. While the quantitative model predictions survive when investors are heterogeneous in the degree to which they respond to observed capital gains, see Adam et al. 2015, it appears of interest to assess the potential price impact generated by speculators with rational price expectations. While rational speculators can contribute to price destabilization, as in DeLong et al. 1990, they may also help with price stabilization, as in Barberis et al. 2014. Exploring this issue further, especially in connection with the limits to arbitrage emphasized in Shleifer and Vishny 1997, appears to be a fruitful avenue for further research.

<sup>97</sup>This is true if one evaluates welfare using ex-post realized consumption.

	$\hat{c}$ $\cdot 10^3$	$\hat{\mathbf{c}}$ $\cdot 10^3$	bias $\cdot 10^3$ $-E(\hat{\mathbf{c}} - \hat{c})$	$p$ -value $H_0: c = \mathbf{c}$	$\hat{c}$ $\cdot 10^3$	$\hat{\mathbf{c}}$ $\cdot 10^3$	bias $\cdot 10^3$ $E(\hat{\mathbf{c}} - \hat{c})$	$p$ -value $H_0: c = \mathbf{c}$
Survey measure	S&P 500, real returns							
	Survey Average				Survey Median			
UBS, >100k, 1 yr, SPF	0.58	-2.46	0.432	0.0000	0.48	-2.49	0.415	0.0000
UBS, >100k, 1 yr, Michigan	0.57	-2.46	0.452	0.0000	0.47	-2.49	0.413	0.0000
UBS, all, 1 yr, SPF	0.57	-2.46	0.424	0.0000	0.49	-2.49	0.401	0.0000
UBS, all, 1 yr, Michigan	0.56	-2.46	0.442	0.0000	0.48	-2.49	0.433	0.0000
CFO, 1 yr, SPF	0.20	-1.67	0.222	0.0011	0.25	-1.37	0.325	0.0471
CFO, 1 yr, Michigan	0.27	-1.67	0.200	0.0006	0.34	-1.37	0.313	0.0362
	Dow Jones, real price growth							
	Survey Average				Survey Median			
Shiller, 1 yr, SPF	0.26	-1.22	0.235	0.0011	0.24	-1.20	0.265	0.0015
Shiller, 1 yr, Michigan	0.33	-1.22	0.232	0.0006	0.31	-1.20	0.238	0.0007
Shiller, 10 yrs, SPF	4.73	-7.25	-1.367	0.0000	6.15	-7.24	-1.440	0.0000
Shiller, 10 yrs, Michigan	4.24	-7.25	-1.423	0.0000	5.65	-7.24	-1.462	0.0000

Table 1a: Rational Expectations test,  $H_0 : c = \mathbf{c}$   
(stock returns, small sample bias correction from proposition 1)

The table reports  $p$ -values for the rational expectations test from section II.B for different survey sources and different ways to extract expectations from the surveys. Estimates and  $p$ -values are bias corrected as described in section II.B and proposition 1.  $\hat{c}$  is the estimate of  $c^N$  in equation (3) and  $\hat{\mathbf{c}}$  the estimate of  $\mathbf{c}^N$  in equation (5). The column labeled 'bias' reports the small sample bias of  $\hat{\mathbf{c}} - \hat{c}$  as implied by proposition 1. The UBS and CFO surveys report return expectations for the S&P 500, the Shiller surveys report capital gain expectations for the Dow Jones Index. In the regressions,  $R_{t,t+N}$  denotes returns, except for the Shiller survey, where it denotes capital gains. The columns labeled 'Survey Average' compute expectations using the cross-sectional average of return expectations, the columns labeled 'Survey Median' uses the median expectation. In the first column, SPF and Michigan refer to different approaches to compute real expected returns, with the former using inflation expectations from the Survey of Professional Forecasters (SPF) and the latter using the Michigan survey; 1 yr and 10 yr refer to forecast horizons of 1 and 10 years, respectively; UBS, >100k indicates a restricted sample using only UBS survey participants with more than 100.000 USD in financial wealth; UBS, all indicates the use of all survey participants. The UBS expectations refer to expectations of the future stock market returns.

	$\chi$	$\hat{c} \cdot 10^3$	$\hat{c} \cdot 10^3$	bias $\cdot 10^3$ $-E(\hat{c} - \hat{c})$	$p$ -value $H_0: c \leq \mathbf{c}$	$\hat{c} \cdot 10^3$	$\hat{c} \cdot 10^3$	bias $\cdot 10^3$ $E(\hat{c} - \hat{c})$	$p$ -value $H_0: c \leq \mathbf{c}$
Survey measure		S&P 500, real returns							
		Survey Average				Survey Median			
UBS*, > 100k, 1 yr, SPF	0.82	2.80	-0.64	0.054	0.014	2.68	-0.62	0.062	0.013
UBS*, > 100k, 1 yr, Michigan	0.82	2.79	-0.64	0.054	0.016	2.67	-0.62	0.071	0.014
UBS*, all, 1 yr, SPF	0.82	2.78	-0.64	0.072	0.013	2.69	-0.62	0.063	0.013
UBS*, all, 1 yr, Michigan	0.82	2.77	-0.64	0.072	0.015	2.68	-0.62	0.055	0.014
CFO, 1 yr, SPF	0.94	5.83	3.19	-0.568	0.046	6.59	4.07	-0.550	0.111
CFO, 1 yr, Michigan	0.94	5.90	3.19	-0.568	0.044	6.67	4.07	-0.520	0.107
		Dow Jones, real price growth							
		Survey Average				Survey Median			
Shiller, 1 yr, SPF	0.92	4.68	2.34	-0.621	0.020	4.68	2.38	-0.627	0.018
Shiller, 1 yr, Michigan	0.92	4.74	2.34	-0.624	0.016	4.74	2.38	-0.627	0.017
Shiller, 10 yrs, SPF	0.76	8.07	-2.65	-0.130	0.016	9.45	-2.66	-0.200	0.014
Shiller, 10 yrs, Michigan	0.76	7.48	-2.65	0.200	0.023	8.85	-2.66	-0.200	0.019

Table 1b: Rational Expectations test,  $H_0 : c \leq \mathbf{c}$   
(stock returns, one-sided test, small sample corrections with serially correlated errors and non-linear least squares estimates from section II.C)

The column labeled  $\chi$  reports the estimated persistence parameter of the return residuals  $\mathbf{U}_t^N$  in equation (11). See the caption of table 1a for further information.

	$\chi$	$\hat{c} \cdot 10^3$	$\hat{c} \cdot 10^3$	bias $\cdot 10^3$ $-E(\hat{c} - \hat{c})$	p-value $H_0: c \leq \mathbf{c}$	$\hat{c} \cdot 10^3$	$\hat{c} \cdot 10^3$	bias $\cdot 10^3$ $E(\hat{c} - \hat{c})$	p-value $H_0: c \leq \mathbf{c}$
Survey measure	S&P 500, real excess returns								
	Survey Average					Survey Median			
UBS, > 100k, 1 yr, SPF	0.77	2.00	-1.11	0.153	0.051	1.89	-1.10	0.154	0.058
UBS, > 100k, 1 yr, Michigan	0.77	2.00	-1.11	0.157	0.056	1.89	-1.10	0.158	0.054
UBS, all, 1 yr, SPF	0.77	1.98	-1.11	0.161	0.056	1.90	-1.10	0.154	0.057
UBS, all, 1 yr, Michigan	0.77	1.98	-1.11	0.154	0.054	1.90	-1.10	0.160	0.062
CFO, 1 yr, SPF	0.93	5.58	3.00	-0.532	0.067	6.44	4.12	-0.517	0.155
CFO, 1 yr, Michigan	0.93	5.59	3.00	-0.540	0.065	6.45	4.12	-0.522	0.151
	Dow Jones, real excess price growth								
	Survey Average					Survey Median			
Shiller, 1 yr, SPF	0.91	4.01	1.85	-0.562	0.028	4.01	1.84	-0.548	0.028
Shiller, 1 yr, Michigan	0.91	4.00	1.85	-0.547	0.030	4.01	1.84	-0.540	0.032
Shiller, 10 yrs, SPF	0.78	6.30	-5.16	0.151	0.030	7.68	-5.15	0.175	0.023
Shiller, 10 yrs, Michigan	0.78	6.14	-5.16	0.383	0.033	7.52	-5.15	-0.214	0.026

Table 1c: Rational Expectations test,  $H_0 : c \leq \mathbf{c}$   
(excess stock returns, one-sided test, small sample corrections with serially correlated errors and non-linear least squares estimates from section II.C.

The column labeled  $\chi$  reports the estimated persistence parameter of the return residuals  $\mathbf{U}_t^N$  in equation (11). See the caption of table 1a for further information.

Parameter	Value	Calibration Target
$\beta^D$	1.0048	average quarterly real dividend growth
$\sigma_D$	0.0192	std. deviation quarterly real dividend growth
$\rho$	22	average consumption-dividend ratio
$\sigma_{DW}$	$-3.51 \cdot 10^{-4}$	jointly chosen s.t. $\text{corr}_t(C_{t+1}/C_t, D_{t+1}/D_t) = 0.2$
$\sigma_W$	0.0186	and $\text{std}_t(C_{t+1}/C_t) = \frac{1}{7} \text{std}_t(D_{t+1}/D_t)$
$\sigma_\varepsilon$	0.0816	std. deviation of quarterly real stock price growth

Table 2: Model calibration

	<b>U.S. Data</b> 1946:1-2012:1 (quart. real values)	<b>Subj. Belief Model</b> (efficient $\Sigma$ matrix)		<b>Subj. Belief Model</b> (diagonal $\Sigma$ matrix)	
	Moment	Moment	t-ratio	Moment	t-ratio
<b>E[PD]</b>	139.8	120.1	-0.79	115.2	-0.98
<b>Std[PD]</b>	65.2	95.6	2.06	88.2	1.56
<b>Corr[PD<sub>t</sub>, PD<sub>t-1</sub>]</b>	0.98	0.98	0.65	0.98	0.90
<b>Std[r<sup>s</sup>]</b>	8.00	8.88	2.20	7.74	-0.64
<i>c</i>	-0.0041	-0.0055	-1.11	-0.0050	-0.77
<i>R</i> <sup>2</sup>	0.25	0.23	-0.13	0.20	-0.34
<b>E[r<sup>s</sup>]</b>	1.89	1.91	0.04	1.82	-0.15
<b>E[r<sup>b</sup>]</b>	0.13	1.01	5.16	0.99	5.08
<b>UBS Survey Data:</b>					
<b>Corr[PD<sub>t</sub>, E<sub>t</sub><sup>P</sup>R<sub>t+1</sub>]</b>	0.79	0.80	0.07	0.79	0.01
<b>Estimates:</b>					
$\widehat{g}$		0.0282		0.0262	
$\widehat{\delta}$		0.99524		0.99514	
$\widehat{\gamma}$		2.05		2.03	
<i>p</i>		0.95		0.95	

Table 3: Asset pricing moments, data and estimated models

The table reports U.S. asset pricing moments (second column) using the data sources described in Appendix A.A1, the moments of the estimated models (columns three and five) and t-ratios for the model moments (columns four and six) from the SMM estimation. T-ratios are defined as (data moment - model moment)/(estimated standard deviation of the model moment). The reported moments are as follows: E[PD], Std(PD) and Corr[PD<sub>t</sub>, PD<sub>t-1</sub>] denote the mean, standard deviation and autocorrelation of the quarterly price dividend ratio, respectively; E[r<sup>s</sup>] and Std[r<sup>s</sup>] denote the mean and standard deviation of the real quarterly stock return, expressed in percentage points, respectively; E[r<sup>b</sup>] is the mean risk free interest rate, expressed in percentage points; *c* and *R*<sup>2</sup> denote, respectively, the regression coefficient and *R*<sup>2</sup> value obtained in equation (5) when *R*<sub>*t*,*t*+*N*</sub> is excess stock returns and *N* = 5 years; Corr[PD<sub>t</sub>, E<sub>t</sub><sup>P</sup>R<sub>t+1</sub>] denotes the correlation between the PD ratio and the subjective return expectations, with these being measured in the data column using the mean of the UBS survey (own portfolio), deflated by the mean inflation expectations from the SPF survey. The estimated parameters are the updating gain *g* from equation (39), the time discount factor  $\widehat{\delta}$ , the coefficient of relative risk aversion  $\widehat{\gamma}$  and the persistence *p* of the wage-dividend ratio from equation (16).

	<b>U.S. Data</b>	<b>RE Model</b>	
	Moment	Moment	t-ratio
<b>E[PD]</b>	139.8	108.3	-1.26
<b>Std[PD]</b>	65.2	11.3	-3.64
<b>Corr[PD<sub>t</sub>, PD<sub>t-1</sub>]</b>	0.98	0.95	-11.20
<b>Std[r<sup>s</sup>]</b>	8.00	1.49	-16.22
<i>c</i>	-0.0041	-0.0181	-11.09
<i>R</i> <sup>2</sup>	0.25	0.72	3.66
<b>E[r<sup>s</sup>]</b>	1.89	1.41	-1.04
<b>E[r<sup>b</sup>]</b>	0.13	1.52	8.15
<b>UBS Survey Data:</b>			
<b>Corr[PD<sub>t</sub>, E<sub>t</sub><sup>P</sup>R<sub>t+1</sub>]</b>	0.79	-1.00	-26.36

Table 4: RE asset pricing moments

The table reports U.S. asset pricing moments (second column) using the data sources described in Appendix A.A1, the moments of the rational expectations (RE) model (column three) and the t-ratios of the RE model (column four). The RE model uses the parameterization of the estimated subjective belief model from table 3 (diagonal matrix). See table 3 for a description of the moment labels in the first column.

Restriction imposed:	<b>U.S. Data</b> 1946:1-2012:1 (quart. real values)	<b>Subj. Belief Model</b> (diagonal matrix) <b>p=0</b>		<b>Subj. Belief Model</b> (diagonal matrix) <b>p=0.999</b>	
	Moment	Moment	t-ratio	Moment	t-ratio
<b>E[PD]</b>	139.8	112.9	-1.07	116.8	-0.92
<b>Std[PD]</b>	65.2	82.1	1.14	98.5	2.25
<b>Corr[PD<sub>t</sub>, PD<sub>t-1</sub>]</b>	0.98	0.98	-0.22	0.99	1.78
<b>Std[r<sup>s</sup>]</b>	8.00	7.91	-0.22	7.73	-0.68
<i>c</i>	-0.0041	-0.0045	-0.36	-0.0037	0.27
<i>R</i> <sup>2</sup>	0.25	0.15	-0.74	0.14	-0.79
<b>E[r<sup>s</sup>]</b>	1.89	1.83	-0.14	1.81	-0.17
<b>E[r<sup>b</sup>]</b>	0.13	0.97	4.92	1.00	5.12
<b>UBS Survey Data:</b>					
<b>Corr[PD<sub>t</sub>, E<sub>t</sub><sup>P</sup>R<sub>t+1</sub>]</b>	0.79	0.85	0.82	0.75	-0.65
<b>Estimates:</b>					
$\hat{g}$		0.0222		0.0272	
$\hat{\delta}$		0.99510		0.99519	
$\hat{\gamma}$		2.02		2.04	

Table 5: Asset pricing moments (restricted estimation)

The table reports U.S. asset pricing moments (second column) using the data sources described in Appendix A.A1, the moments and t-ratios of the estimated subjective belief model when restricting the persistence parameter  $p$  from equation (16) to zero (third and fourth column) and when restricting  $p$  to 0.999 (fifth and sixth column). See table 3 for a description of the labels used in the first column.



	<b>Data</b>	<b>Std. Dev.</b>	<b>Estimated Model</b> <small>(table 3, diagonal matrix)</small>	<b>RE Version</b> <small>(table 3, diagonal matrix)</small>	<b>Estimated Model</b> <b>with supply shocks</b>
	<b>Moment</b>	<b>Dev.</b>	<b>Moment t-ratio</b>	<b>Moment t-ratio</b>	<b>Moment t-ratio</b>
<b>Std. deviation</b>					
risk free rate	1.11	0.24	0.83	0.62	0.62
<b>Autocorrelations:</b>					
Excess stock return	0.12	0.046	0.53	0.33	0.12
Stock return	0.13	0.045	0.54	0.04	0.14
					-2.05
					0.18
					0.37

Table 6: Further data and model moments

	<b>U.S. Data</b> 1946:1-2012:1 (quart. real values)	<b>Subj. Belief Model</b> (transitory belief shocks)		<b>Subj. Belief Model</b> (persistent belief shocks)	
	Moment	Moment	t-ratio	Moment	t-ratio
<b>E[PD]</b>	139.8	113.5	-1.05	120.6	-0.77
<b>Std[PD]</b>	65.2	85.9	1.40	90.8	1.73
<b>Corr[PD<sub>t</sub>, PD<sub>t-1</sub>]</b>	0.98	0.98	0.63	0.98	0.51
<b>Std[r<sup>s</sup>]</b>	8.00	7.37	-1.58	7.93	-0.16
<i>c</i>	-0.0041	-0.0047	-0.53	-0.0064	-1.81
<i>R</i> <sup>2</sup>	0.25	0.18	-0.47	0.25	0.05
<b>E[r<sup>s</sup>]</b>	1.89	1.80	-0.19	1.87	-0.06
<b>E[r<sup>b</sup>]</b>	0.13	0.34	1.26	-0.65	-4.58
<b>UBS Survey Data:</b>					
<b>Corr[PD<sub>t</sub>, E<sup>P</sup><sub>t</sub>R<sub>t+1</sub>]</b>	0.79	0.79	-0.08	0.85	0.78
<b>Consumption volatility:</b>					
<b>E[Std<sup>P</sup><sub>t</sub>C<sub>t+1</sub><sup>i</sup>/C<sub>t</sub><sup>i</sup>]</b>		3.38		4.55	
<b>E[Std<sub>t</sub>C<sub>t+1</sub><sup>i</sup>/C<sub>t</sub><sup>i</sup>]</b>		3.23		4.99	
<b>Parameters:</b>					
<i>g</i>		0.0262		0.0282	
<i>δ</i>		0.99514		0.99514	
<i>γ</i>		2.03		2.03	
<i>σ</i> <sub>μ</sub>		0.0030		-	
<i>σ</i> <sub>ε</sub>		-		0.0018	

Table 7: Heterogeneous Agent Model with Equilibrium Trade

The table reports U.S. asset pricing moments (second column) using the data sources described in Appendix A.A1, the moments and t-ratios of the subjective belief model with idiosyncratic transitory belief shocks (third and fourth column), see equation (47), and the moments and t-ratios of the subjective belief model with idiosyncratic persistent belief shocks (fifth and sixth column), see equation (48).  $E[\text{Std}_t^P C_{t+1}^i / C_t^i]$  denotes the unconditional expectation of the subjective standard deviation of consumption growth and  $E[\text{Std}_t C_{t+1}^i / C_t^i]$  the unconditional expectation of the objective standard deviation of consumption growth. The standard deviation of the idiosyncratic belief shocks  $\sigma_\mu$  and  $\sigma_\varepsilon$  has been chosen so as to approximately equate subjective and objective standard deviation of consumption growth. See table 3 for a description of the remaining labels used in the first column.

	Significance level	Sample length in quarters					
		25	50	75	100	150	200
Model from Table 7 (diagonal matrix)	1%	0.828	0.944	0.971	0.981	0.984	0.987
	5%	0.900	0.970	0.981	0.987	0.988	0.990
Model from Table 7 (efficient matrix)	1%	0.779	0.938	0.968	0.975	0.975	0.989
	5%	0.892	0.974	0.986	0.981	0.980	0.991

Table 8: Rejection frequencies for the RE hypothesis ( $H_0 : c = \mathbf{c}$ ) for the estimated asset pricing models

The reported rejection frequencies are based on the rational expectations test developed in section II.B of the paper and include the small sample bias correction stated in proposition 1.

Selection rule : Share of current updaters ( $\alpha$ )	Closest price		Furthest price	
	Lagged updaters	Current updaters	Lagged updaters	Current Updaters
0.0	-4.02	-4.07	-4.02	-4.07
0.1	-4.02	-4.06	-4.02	-4.06
0.2	-4.01	-4.06	-4.01	-4.06
0.3	-4.01	-4.05	-4.01	-4.06
0.4	-4.01	-4.05	-4.00	-4.06
0.5	-4.00	-4.05	-3.99	-4.06
0.6	-4.00	-4.04	-3.98	-4.06
0.7	-3.99	-4.04	-3.97	-4.06
0.8	-3.99	-4.03	-3.89	-4.11
0.9	-3.97	-4.03	-3.39	-4.24
1.0	-3.95	-4.02	-2.46	-4.02

Table 9: Expected utility of current and lagged belief updaters

The table reports the objectively expected utility for agents that use current and one period lagged prices to update beliefs for different equilibrium selection rules, as described in section X.C, and different shares  $\alpha$  of current updaters. Since  $u(C) < 0$ , expected utility is always negative. The model is parameterized using the estimated parameters from table 3 (diagonal matrix).

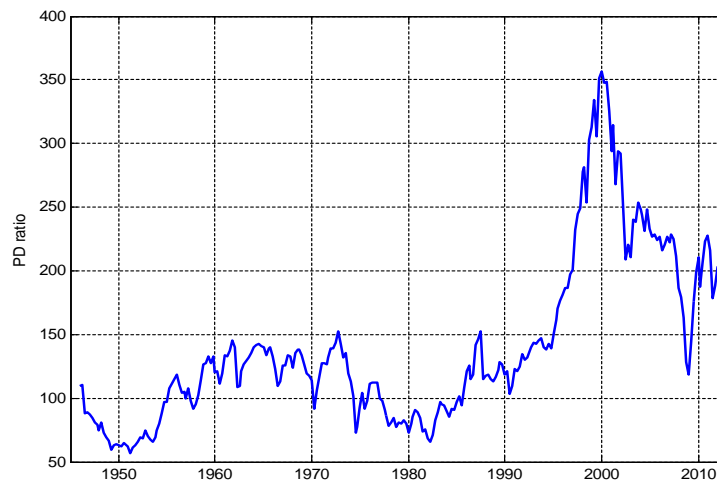


FIGURE 1. QUARTERLY PD RATIO OF THE S&P 500, Q1:1946-Q1:2012

(PD RATIO IS DEFINED AS STOCK PRICES OVER QUARTERLY DIVIDENDS; DIVIDENDS ARE DESEASONALIZED BY AVERAGING DIVIDEND PAYMENTS OVER THE CURRENT AND PREVIOUS THREE QUARTERS)

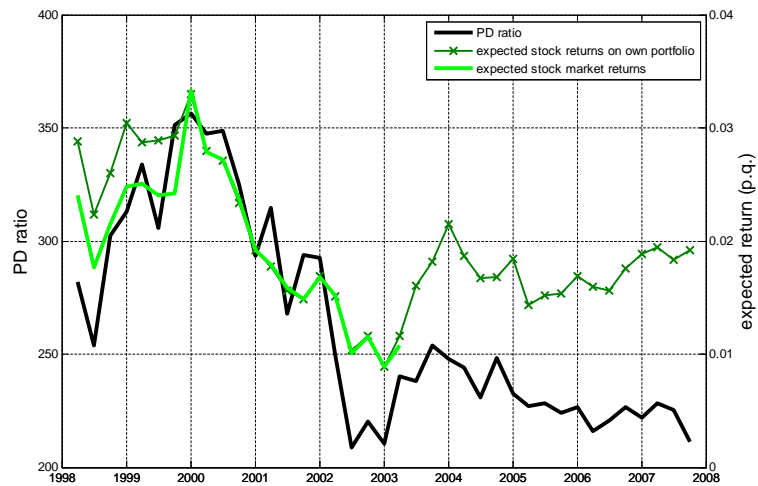


FIGURE 2. QUARTERLY PD RATIO OF THE S&P 500 (L.H.S. SCALE) AND INVESTORS' EXPECTED QUARTERLY REAL RETURNS FROM THE UBS GALLUP SURVEY (R.H.S. SCALE), Q2:1998-Q4:2007

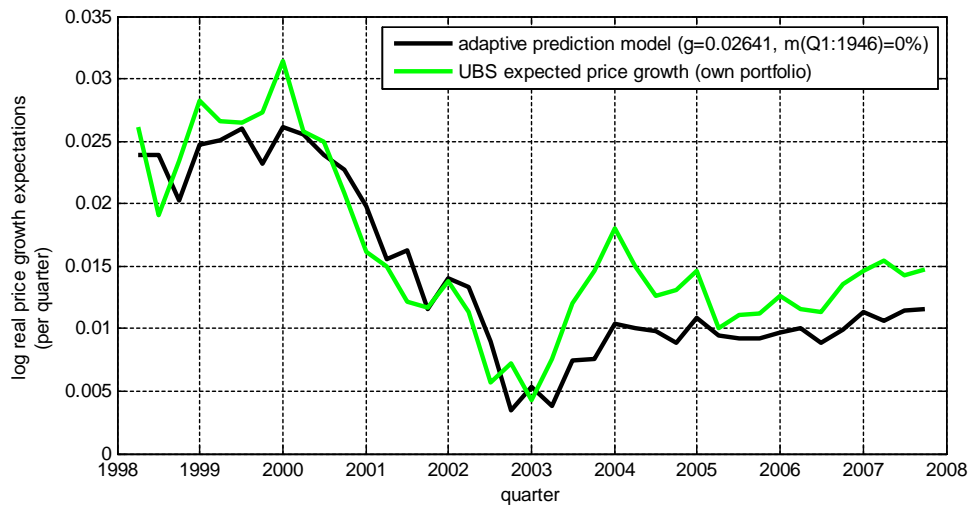


FIGURE 3. REAL PRICE GROWTH EXPECTATIONS IMPLIED BY THE UBS SURVEY EXPECTATIONS AND BY THE ADAP- TIVE PREDICTION MODEL (EQUATION (13)), Q2:1998-Q4:2007

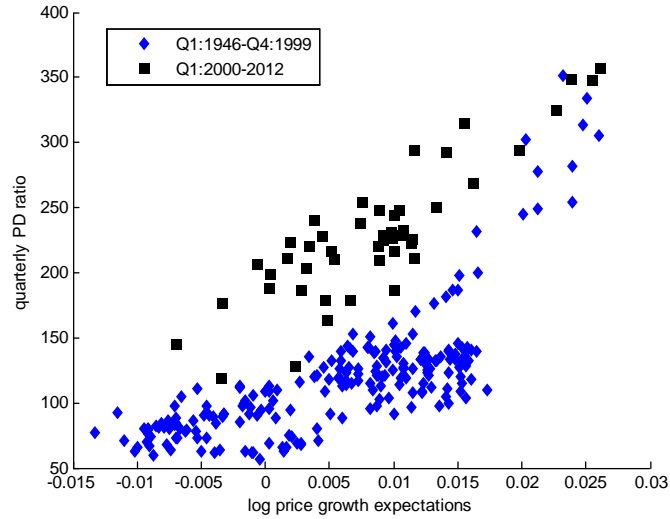


FIGURE 4. QUARTERLY PD RATIO OF THE S&P 500 GRAPHED AGAINST THE QUARTERLY REAL PRICE GROWTH EXPECTATIONS IMPLIED BY THE ADAP TIVE PREDICTION MODEL (EQUATION (13)) FOR THE SAME QUARTER, Q1:1946-Q1:2012

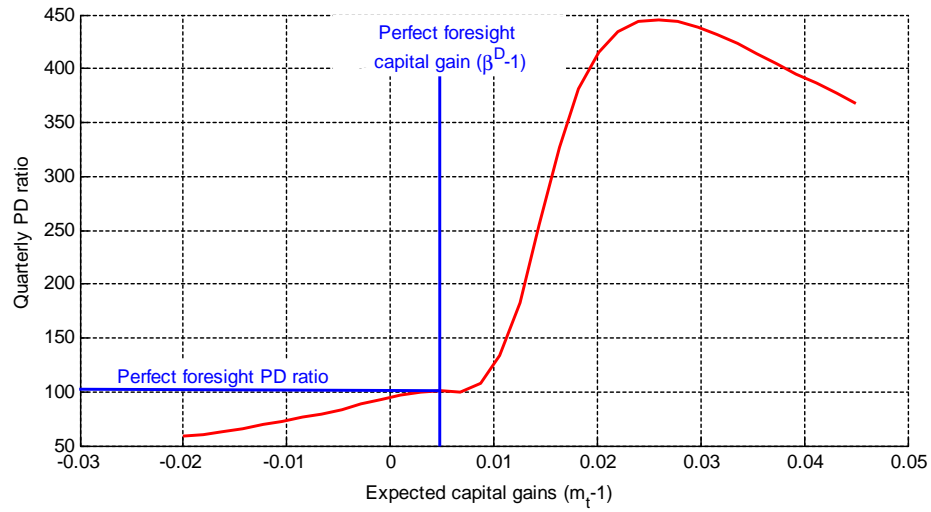


FIGURE 5. QUARTERLY PD RATIO AND EXPECTED CAPITAL GAINS (VANISHING NOISE LIMIT, ESTIMATED MODEL FROM TABLE 3, DIAGONAL MATRIX)

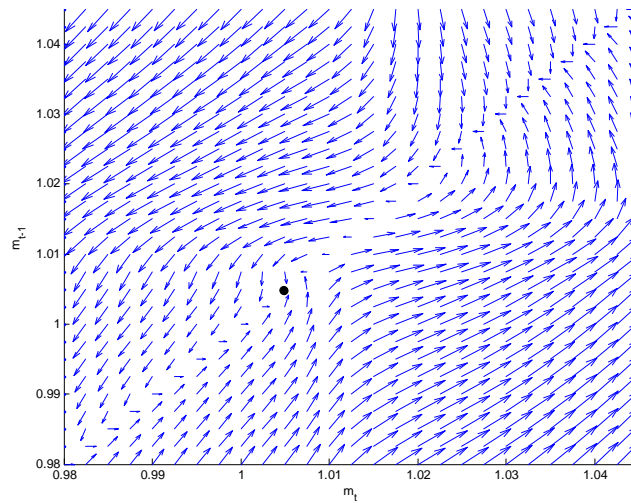


FIGURE 6. ILLUSTRATION OF GLOBAL BELIEF DYNAMICS UNDER LEARNING IN THE ABSENCE OF SHOCKS (ESTIMATED MODEL FROM TABLE 3, DIAGONAL MATRIX). THE BLACK DOT INDICATES THE PERFECT FORESIGHT EQUILIBRIUM BELIEF.

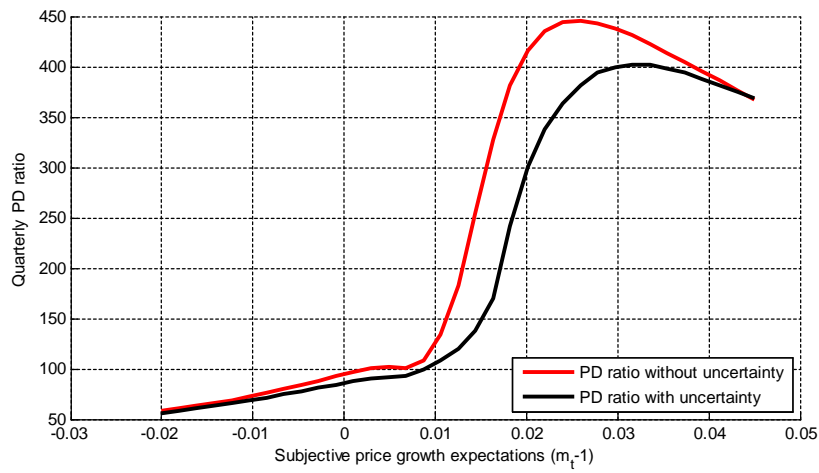


FIGURE 7. THE EFFECTS OF UNCERTAINTY ON THE EQUILIBRIUM PD RATIO (ESTIMATED MODEL FROM TABLE 3, DIAGONAL MATRIX)

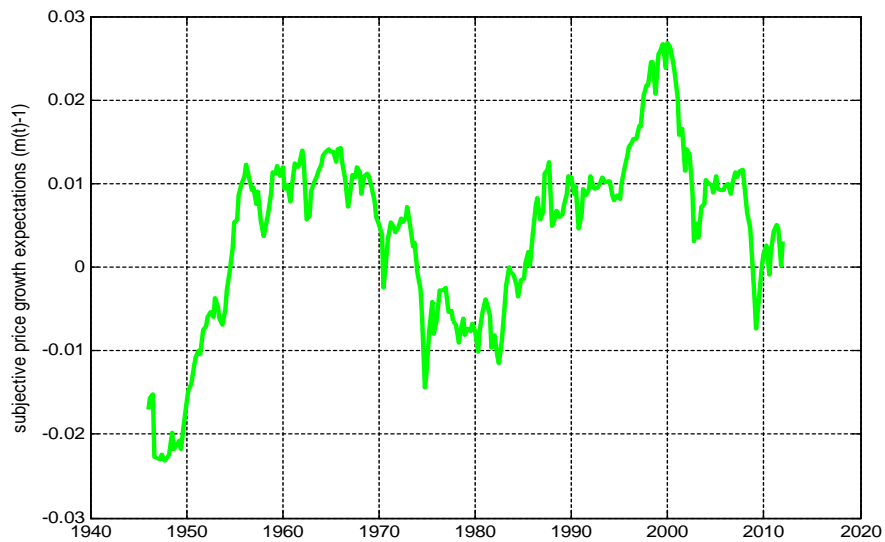


FIGURE 8. PRICE GROWTH EXPECTATIONS IMPLIED BY BAYESIAN UPDATING AND HISTORICAL STOCK PRICE GROWTH (ESTIMATED MODEL FROM TABLE 5,  $P=0.999$ ), Q1:1946-Q1:2012

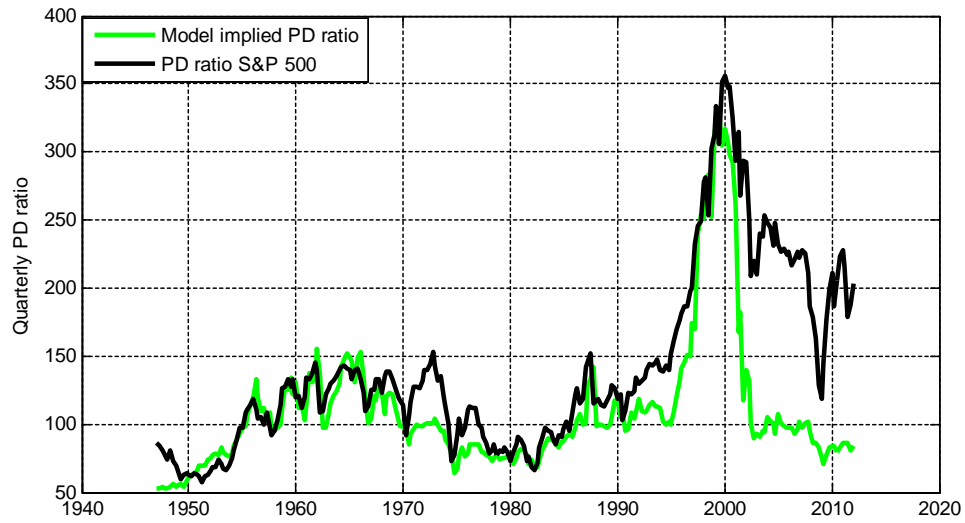


FIGURE 9. HISTORICAL PD RATIO: DATA VERSUS MODEL, Q1:1946-Q1:2012 (ESTIMATED MODEL FROM TABLE 5,  $P=0.999$ )

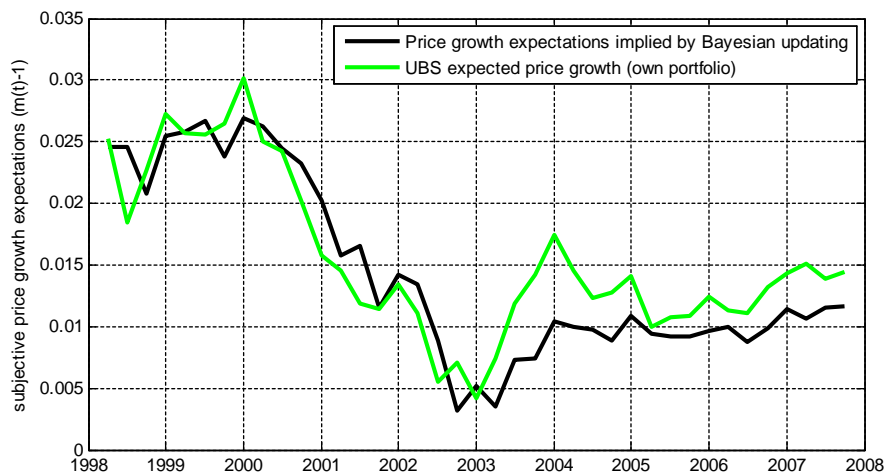


FIGURE 10. PRICE GROWTH EXPECTATIONS: UBS SURVEY VS. BAYESIAN UPDATING MODEL (ESTIMATED MODEL FROM TABLE 5,  $P=0.999$ ), Q2:1998-Q4:2007



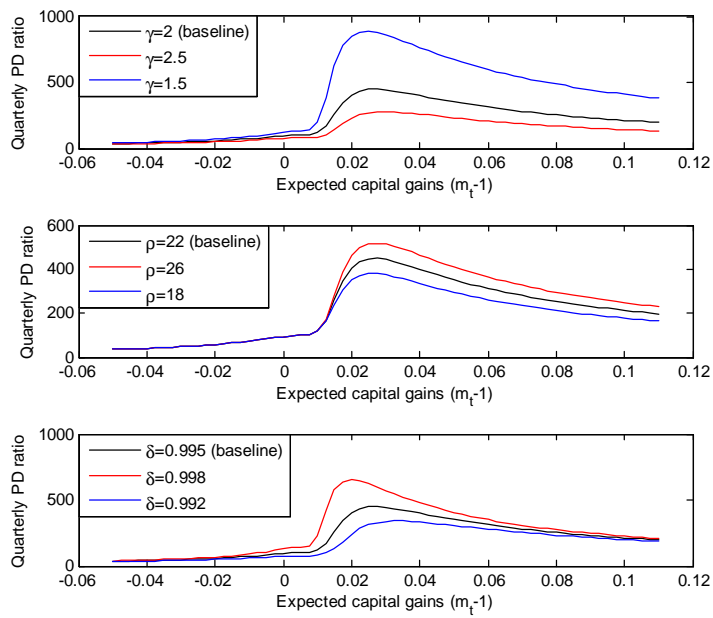


FIGURE 11. EFFECTS OF  $(\gamma, \rho, \delta)$  ON THE EQUILIBRIUM PD FUNCTION

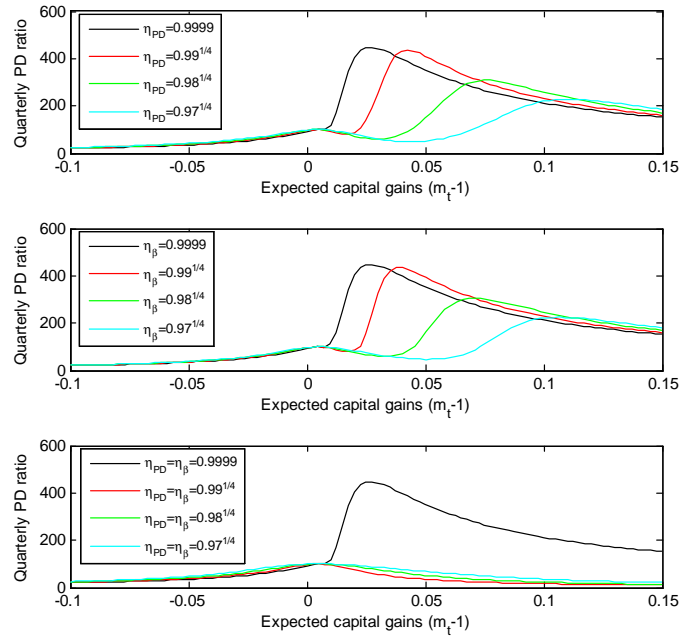


FIGURE 12. EQUILIBRIUM PRICING FUNCTION (GENERALIZED BELIEF SYSTEM)

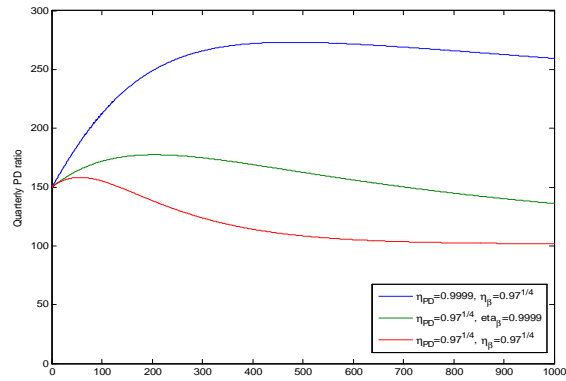


FIGURE 13. PERSISTENCE PARAMETERS AND EXPECTED PD PATH (GENERALIZED BELIEF SYSTEM)

## REFERENCES

- Adam, Klaus, Albert Marcet, and Juan Pablo Nicolini.** 2016. "Stock Market Volatility and Learning." *Journal of Finance*, 71(1): 33–82.
- Adam, Klaus, and Albert Marcet.** 2011. "Internal Rationality, Imperfect Market Knowledge and Asset Prices." *Journal of Economic Theory*, 146: 1224–1252.
- Adam, Klaus, Johannes Beutel, Albert Marcet, and Sebastian Merkel.** 2015. "Can a Financial Transaction Tax Prevent Stock Price Booms?" *Journal of Monetary Economics*, 76: S90–S109.
- Angeletos, George-Marios, and Jennifer La'O.** 2013. "Sentiments." *Econometrica*, 81(2): 739–779.
- Aruoba, Boragan, Jesus Fernandez-Villaverde, and Juan Rubio-Ramirez.** 2006. "Comparing solution methods for dynamic equilibrium economies." *Journal of Economic Dynamics and Control*, 30.
- Asparouhova, Elena, Peter Bossaerts, Nilanjan Roy, and William Zame.** 2013. "Experiments on the Lucas Asset Pricing Model." *NBER Macroeconomics Annual 2013 (forthcoming)*.
- Bacchetta, Philippe, Elmar Mertens, and Eric Van Wincoop.** 2009. "Predictability in Financial Markets: What Do Survey Expectations Tell Us?" *Journal of International Money and Finance*, 28: 406–426.
- Bansal, Ravi, and Amir Yaron.** 2004. "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles." *Journal of Finance*, 59: 1481–1509.
- Barberis, Nicholas, Andrei Shleifer, and Robert Vishny.** 1998. "A Model of Investor Sentiment." *Journal of Financial Economics*, 49: 307–343.
- Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer.** 2014. "X-CAPM: An Extrapolative Capital Asset Pricing Model." *Journal of Financial Economics (forthcoming)*.
- Boswijk, Peter, Cars H. Hommes, and Sebastiano Manzan.** 2007. "Behavioral Heterogeneity in Stock Prices." *Journal of Economic Dynamics and Control*, 31: 1938–1970.
- Branch, William A., and George Evans.** 2011. "Learning about Risk and Return: A Simple Model of Bubbles and Crashes." *American Economic Journal: Macroeconomics*, 3: 159–191.
- Brock, William A., and Cars 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, James, and John Duffy.** 2001. "Learning and Excess Volatility." *Macroeconomic Dynamics*, 5: 272–302.

- Campbell, John Y.** 2003. "Consumption-Based Asset Pricing." In *Handbook of Economics and Finance*, ed. G. M. Constantinides, M. Harris and R. Stulz, 803–887. Amsterdam:Elsevier.
- Campbell, John Y., and John H. Cochrane.** 1999. "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." *Journal of Political Economy*, 107: 205–251.
- Campbell, John Y., and Motohiro Yogo.** 2006. "Efficient Test of Stock Return Predictability." *Journal of Financial Economics*, 81: 27–60.
- Carroll, Christopher D.** 2006. "The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems." *Economics Letters*, 91: 312–320.
- Choi, James J., and Thomas M. Mertens.** 2013. "Extrapolative Expectations and the Equity Premium." *mimeo*.
- Cogley, Timothy, and Thomas J. Sargent.** 2008. "The Market Price of Risk and the Equity Premium: A Legacy of the Great Depression?" *Journal of Monetary Economics*, 55: 454–476.
- Collin-Dufresne, Pierre, Michael Johannes, and Lars A. Lochstoer.** 2016. "Parameter Learning in General Equilibrium: The Asset Pricing Implications." *American Economic Review*, 106: 664–698.
- DeLong, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann.** 1990. "Positive Feedback Investment Strategies and Destabilizing Rational Speculation." *Journal of Finance*, 45: 379–395.
- Epstein, Larry G., Emmanuel Farhi, and Tomasz Strzalecki.** 2014. "How Much Would You Pay to Resolve Long-Run Risk?" *American Economic Review*, 104: 2680–2697.
- Eusepi, Stefano, and Bruce Preston.** 2011. "Expectations, Learning, and Business Cycle Fluctuations." *American Economic Review*, 101: 2844–2872.
- Evans, George W., and Seppo Honkapohja.** 2001. *Learning and Expectations in Macroeconomics*. Princeton:Princeton University Press.
- Fama, Eugene F., and Kenneth R. French.** 1988. "Dividend Yields and Expected Stock Returns." *Journal of Financial Economics*, 22: 3–25.
- Ferson, Wayne E., Sergei Sarkissian, and Timothy T. Simin.** 2003. "Spurious Regressions in Financial Economics?" *The Journal of Finance*, 58(4): 1393–1413.
- F. H. C. Marriott, J. A. Pope.** 1954. "Bias in the Estimation of Autocorrelations." *Biometrika*, 41(3/4): 390–402.
- Greenwood, Robin, and Andrei Shleifer.** 2014. "Expectations of Returns and Expected Returns." *Review of Financial Studies*, 27 (3).
- Hirota, Shinichi, and Shyam Sunder.** 2007. "Price Bubbles Sans Dividend Anchors: Evidence from Laboratory Stock Markets." *Journal of Economic Dynamics and Control*, 31: 1875–1909.

- Hirshleifer, David, and Jianfeng Yu.** 2012. "Asset Pricing in Production Economies with Extrapolative Expectations." *mimeo*.
- Hong, Harrison, Jose Scheinkman, and Wei Xiong.** 2006. "Asset Float and Speculative Bubbles." *Journal of Finance*, 61: 1073–1117.
- Lansing, Kevin J.** 2010. "Rational and Near-Rational Bubbles Without Drift." *Economic Journal*, 120: 1149–1174.
- Lucas, Robert E.** 1978. "Asset Prices in an Exchange Economy." *Econometrica*, 46: 1426–1445.
- MacKinnon, James G., and Antony A. Smith.** 1998. "Approximate Bias Correction in Econometrics." *Journal of Econometrics*, 85: 205–230.
- Malmendier, Ulrike, and Stefan Nagel.** 2011. "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking." *Quarterly Journal of Economics*, 126: 373–416.
- Malmendier, Ulrike, and Stefan Nagel.** 2013. "Learning from Inflation Experiences." Berkeley University Working Paper.
- Marcet, Albert, and Juan Pablo Nicolini.** 2003. "Recurrent Hyperinflations and Learning." *American Economic Review*, 93: 1476–1498.
- Marcet, Albert, and Thomas J. Sargent.** 1989. "Convergence of Least Squares Learning Mechanisms in Self Referential Linear Stochastic Models." *Journal of Economic Theory*, 48: 337–368.
- Marcet, Albert, and Thomas J. Sargent.** 1992. "The Convergence of Vector Autoregressions to Rational Expectations Equilibria." In *Macroeconomics: A Survey of Research Strategies*. , ed. Alessandro Vercelli and Nicola Dimitri, 139–164. Oxford:Oxford University Press.
- Nagel, Stefan, and Robin Greenwood.** 2009. "Inexperienced Investors and Bubbles." *Journal of Financial Economics*, 93(2): 239–258.
- Ofek, Eli, and Matthew Richardson.** 2003. "Dotcom Mania: The Rise and Fall of Internet Stock Prices." *Journal of Finance*, 58: 1113–1137.
- Rendahl, Pontus.** 2013. "Inequality Constraint and Euler Equation Based Solution Methods." *Economic Journal (forthcoming)*.
- Roy, Roch.** 1989. "Asymptotic covariance structure of serial correlations in multivariate time series." *Biometrika*, 76: 824–7.
- Sargent, Thomas J., Noah Williams, and Tao Zha.** 2009. "The Conquest of South American Inflation." *Journal of Political Economy*, 117(2): 211–256.
- Shiller, Robert J.** 1981. "Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?" *American Economic Review*, 71: 421–436.
- Shleifer, Andrei, and Robert W. Vishny.** 1997. "The limits of Arbitrage." *Journal of Finance*, 52: 35–55.

- Stambaugh, Robert F.** 1999. "Predictive Regressions." *Journal of Financial Economics*, 54: 375–421.
- Timmermann, Allan.** 1993. "How Learning in Financial Markets Generates Excess Volatility and Predictability in Stock Prices." *Quarterly Journal of Economics*, 108: 1135–1145.
- Vissing-Jorgensen, Annette.** 2003. "Perspectives on Behavioral Finance: Does "Irrationality" Disappear with Wealth? Evidence from Expectations and Actions." Boston:NBER.
- West, Mike, and Jeff Harrison.** 1997. *Bayesian Forecasting and Dynamic Models*. Berlin:Springer.

## APPENDIX (NOT FOR PUBLICATION)

## A1. Data Sources

**Stock price data:** our stock price data is for the United States and has been downloaded from ‘The Global Financial Database’ (<http://www.globalfinancialdata.com>). The period covered is Q1:1946-Q1:2012. The nominal stock price series is the ‘SP 500 Composite Price Index (w/GFD extension)’ (Global Fin code ‘\_SPXD’). The daily series has been transformed into quarterly data by taking the index value of the last day of the considered quarter. To obtain real values, nominal variables have been deflated using the ‘USA BLS Consumer Price Index’ (Global Fin code ‘CPUSAM’). The monthly price series has been transformed into a quarterly series by taking the index value of the last month of the considered quarter. Nominal dividends have been computed as follows

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

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

**Interest rate data:** As nominal interest rate we use the 90 Days T-Bills Secondary Market (Global Fin code ITUSA3SD). The weekly (to the end of 1953) and daily (after 1953) series has been transformed into a quarterly series using the interest rate corresponding to the last week or day of the considered quarter and is expressed in quarterly rates (not annualized). To obtain real values, nominal variables have been deflated using the ‘USA BLS Consumer Price Index’ (Global Fin code ‘CPUSAM’).

**Stock market survey data:** The UBS survey is the UBS Index of Investor Optimism, which is available (against a fee) at

[http://www.ropercenter.uconn.edu/data\\_access/data/datasets/ubs\\_investor.html](http://www.ropercenter.uconn.edu/data_access/data/datasets/ubs_investor.html).

The quantitative question on stock market expectations has been surveyed over the period Q2:1998-Q4:2007 with 702 responses per month on average and has thereafter been suspended. For each quarter we have data from three monthly surveys, except for the first four quarters and the last quarter of the survey period where we have only one monthly survey per quarter. The Shiller survey data covers individual investors over the period Q1:1999Q1-Q4:2012 and has been kindly made available to us by Robert Shiller at Yale University. On average 73 responses per quarter have been recorded for the question on stock price growth. Since the Shiller data refers to the Dow Jones, we used the PD ratio for the Dow Jones, which is available at <http://www.djaverages.com/>, to compute correlations. The CFO survey is collected by Duke University and CFO magazine and collects responses from U.S. based CFOs over the period Q3:2000-Q4:2012 with on average 390 responses per quarter, available at <http://www.cfosurvey.org/>.

**Inflation expectations data:** The Survey of Professional Forecasters (SPF) is available from the Federal Reserve Bank of Philadelphia at <http://www.phil.frb.org/research-and-data/real-time-center/survey-of-professional-forecasters/>. The Michigan Surveys of Consumers are collected by Thomson Reuters/University of Michigan

(<http://www.sca.isr.umich.edu/>).

#### A2. Correlations between PD Ratio and Actual/Survey Returns

This appendix documents for the UBS Survey, the CFO survey and the Shiller individual investor survey that there exists a positive correlation between the PD ratio and survey expected returns (or capital gains), but a negative correlation between the PD ratio and actual returns (or capital gains).

Table A1 documents the positive correlation for the UBS survey. Results are independent of how one extracts expectations from the survey (using the median or the mean expectation, using inflation expectations from the Michigan survey or the Survey of Professional Forecasters (SPF) to obtain real return expectations, using plain nominal returns instead of real returns, or when restricting attention to investors with more than 100,000 US\$ in financial wealth). The numbers reported in brackets in table A1 (and in subsequent tables) are autocorrelation robust p-values for the hypothesis that the correlation is smaller or equal to zero.<sup>98</sup> These p-values are *not adjusted for small sample bias*, as there is no generally accepted approach for how to perform such adjustments. This said, the p-values for the null hypothesis are all below the 5% significance level and in many cases below the 1% level.

A positive correlation is equally obtained when considering other survey data. Table A2 reports the correlations between the PD ratio and the stock price growth expectations from Bob Shiller's Individual Investors' Survey.<sup>99</sup> The table shows that price growth expectations are also strongly positively correlated with the PD ratio, suggesting that the variation in expected returns observed in the UBS survey is due to variations in expected capital gains. Table A2 also shows that correlations seem to become stronger for longer prediction horizons.

Table A3 reports the correlations for the stock return expectations reported in the Chief Financial Officer (CFO) survey which surveys chief financial officers from large U.S. corporations. Again, one finds a strong positive correlation.

Table A4 reports the correlations between the PD ratio and the realized real returns (or capital gains) in the data, using the same sample periods as are available for the surveys considered in tables A1 to A3, respectively. The point estimate for the correlation is negative in all cases, although the correlations fall short of being significant the 5% level due to the short sample length for which the survey data is available.

<sup>98</sup>The sampling width is four quarters, as is standard for quarterly data, and the test allows for contemporaneous correlation, as well as for cross-correlations at leads and lags. The p-values are computed using the result in Roy 1989.

<sup>99</sup>Shiller's price growth data refers to the Dow Jones Index. The table thus reports the correlation of the survey measure with the PD ratio of the Dow Jones.



UBS Gallup	Nominal		Real Ret. Exp.		Real Ret. Exp.	
	Return Exp.		(SPF)		(Michigan)	
	Average	Median	Average	Median	Average	Median
<b>Own portfolio,</b> > 100k US\$	0.80 (0.01)	0.78 (0.01)	0.79 (0.01)	0.77 (0.01)	0.84 (0.01)	0.83 (0.01)
<b>Own portfolio,</b> all investors	0.80 (0.01)	0.76 (0.02)	0.79 (0.01)	0.75 (0.02)	0.84 (0.01)	0.80 (0.01)
<b>Stock market,</b> > 100k US\$	0.90 (0.03)	0.89 (0.04)	0.90 (0.03)	0.88 (0.03)	0.91 (0.03)	0.88 (0.03)
<b>Stock market,</b> all investors	0.90 (0.03)	0.87 (0.04)	0.90 (0.03)	0.87 (0.04)	0.91 (0.03)	0.88 (0.03)

Table A1: Correlation between PD ratio and 1-year ahead expected return measures (UBS Gallup Survey, robust p-values in parentheses, without small sample correction for p-values )

Shiller Survey	Nominal Capital Gain Exp.		Real Capital Gain. Exp. (SPF)		Real Capital Gain Exp. (Michigan)	
	Average	Median	Average	Median	Average	Median
Horizon						
<b>1 month</b>	0.46 (0.01)	0.48 (0.01)	0.45 (0.01)	0.47 (0.01)	0.46 (0.01)	0.49 (0.01)
<b>3 months</b>	0.57 (0.01)	0.64 (0.00)	0.54 (0.01)	0.61 (0.00)	0.56 (0.01)	0.62 (0.01)
<b>6 months</b>	0.58 (0.01)	0.75 (0.01)	0.54 (0.02)	0.70 (0.01)	0.56 (0.02)	0.71 (0.01)
<b>1 year</b>	0.43 (0.03)	0.69 (0.01)	0.38 (0.05)	0.62 (0.01)	0.42 (0.04)	0.64 (0.02)
<b>10 years</b>	0.74 (0.01)	0.75 (0.01)	0.66 (0.02)	0.71 (0.01)	0.71 (0.02)	0.75 (0.01)

Table A2: Correlation between PD ratio and expected stock price growth (Shiller's Individual Investors' Survey, robust p-values in parentheses, without small sample correction for p-values )

CFO Survey	Nominal Return Exp.		Real Return Exp. (SPF)		Real Return Exp. (Michigan)	
	Average	Median	Average	Median	Average	Median
<b>1 year</b>	0.71 (0.00)	0.75 (0.00)	0.62 (0.00)	0.69 (0.00)	0.67 (0.00)	0.72 (0.00)

Table A3: Correlation between PD ratio and 1-year ahead expected stock return measures (CFO Survey, robust p-values in parentheses, without small sample correction for p-values )

Variables	Time Period	Stock Index	Correlation
PD, 1 year-ahead real return	UBS Gallup sample (stock market exp.)	S&P 500	-0.66 (0.08)
PD, 1 year-ahead real price growth	Shiller 1 year sample	Dow Jones	-0.42 (0.06)
PD, 10 year-ahead real price growth	Shiller 10 year sample	Dow Jones	-0.88 (0.16)
PD, 1 year-ahead real return	CFO sample	S&P 500	-0.46 (0.06)

Table A4: Correlation between PD and actual real returns/capital gains (robust p-value in parentheses, without small sample correction for p-values)

### A3. Proof of Proposition 1

#### Proof of part a)

Under the null hypothesis of rational expectations ( $E_t^P = E_t$ ) equation (1) implies

$$(A1) \quad R_{t,t+N} = a^N + c^N \frac{P_t}{D_t} + u_t^N + \varepsilon_t^N,$$

where  $\varepsilon_t^N$  is the prediction error  $R_{t,t+N} - E_t R_{t,t+N}$  from the *true* data-generating process, the conditional expectation is taken with respect to investors' information at  $t$ . Since  $P_t/D_t$  is in this information set under RE and given (2) we have

$$(A2) \quad E [x_t (u_t^N + \varepsilon_t^N)] = 0.$$

Therefore,  $u_t^N + \varepsilon_t^N = \mathbf{u}_t^N$  and the null hypothesis of rational expectations implies

$$(A3) \quad c^N = \mathbf{c}^N.$$

Equations (3) (5) define a SUR system of equations with dependent variables  $\mathcal{E}_t^N$  and  $R_{t,t+N}$ , and explanatory variables in both equations  $x_t = (1, \frac{P_t}{D_t})'$ . Under the null hypothesis the error terms satisfy the orthogonality conditions (4) and (A2).

For part a) of Proposition 1 we define the OLS estimator equation by equation  $\widehat{\beta}_T$  as

$$\widehat{\beta}_T \equiv \begin{bmatrix} \widehat{a}_T^N \\ \widehat{c}_T^N \\ \widehat{\mathbf{a}}_T^N \\ \widehat{\mathbf{c}}_T^N \end{bmatrix} = \left( I_2 \otimes \sum_{t=1}^T x_t x_t' \right)^{-1} \sum_{t=1}^T \begin{bmatrix} \mathcal{E}_t^N \\ R_{t,t+N} \end{bmatrix} \otimes x_t,$$

where  $I_2$  is a  $2 \times 2$  identity matrix. A standard result ensures that under the assumptions OLS equation by equation is consistent and efficient among the set of estimators that use orthogonality conditions (4) and (A2).

As is well known, as  $T \rightarrow \infty$  we have

$$(A4) \quad \sqrt{T} (\hat{\beta}_T - \beta_0) \rightarrow N \left( 0, [I_2 \otimes E(x_t x_t')]^{-1} S_w [I_2 \otimes E(x_t x_t')]^{-1} \right),$$

in distribution, where

$$S_w = \Gamma_0 + \sum_{k=1}^{\infty} (\Gamma_k + \Gamma_k')$$

$$\Gamma_k = E \left( \begin{bmatrix} u \mu_t \\ u \varepsilon_t \end{bmatrix} [u \mu_{t-k}, u \varepsilon_{t-k}] \otimes x_t x_{t-k}' \right),$$

where  $u \mu_t \equiv u_t^N + \mu_t^N$  and  $u \varepsilon_t \equiv u_t^N + \varepsilon_t^N$ . The footnote of the proposition contains all boundedness conditions required to ensure validity of asymptotic distribution,  $E(x_t x_t')$  is invertible because  $\text{var}(P_t/D_t) > 0$ .

To build the test-statistic, we only need to find an estimator for var-cov matrix in (A4). We estimate  $E(x_t x_t')$  by  $\frac{1}{T} \sum_{t=1}^T x_t x_t'$ . Since  $u \mu_t$  and  $u \varepsilon_t$  are not forecasting errors, there is no reason why  $\Gamma_k$  should be zero for any  $k$ , so we use a Newey-West estimator  $\hat{S}_w$ . Therefore, postmultiplying  $(\hat{\beta}_T - \beta_0)$  by  $[0, 1, 0, -1]'$  in (A4), and letting

$$(A5) \quad \hat{\sigma}_{c-c}^2 \equiv \frac{1}{T} [0, 1, 0, -1] \left[ I_2 \otimes \frac{1}{T} \sum_{t=1}^T x_t x_t' \right]^{-1} \hat{S}_w \left[ I_2 \otimes \frac{1}{T} \sum_{t=1}^T x_t x_t' \right]^{-1} \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix}$$

we have that under the null hypothesis

$$\sqrt{T} \frac{\hat{c}_T^N - \hat{c}_T^N}{\hat{\sigma}_{c-c}} \rightarrow N(0, 1) \text{ in distribution.}$$

### Proof of part b)

Equations (3) (5) in the current paper are each of the form of equation (1) in Stambaugh 1999. Focusing first on (3), our  $(\mathcal{E}_t^N, u_t + \mu_t, P D_t, \varepsilon_{t+1}^{PD})$  play the role of  $(y_t, u_t, x_{t-1}, v_t)$  in Stambaugh. Note, in particular, that to match his framework we need to have  $P D_t$  play the role of  $x_{t-1}$ , implying that our  $\varepsilon_{t+1}^{PD}$  plays the role of Stambaugh's  $v_t$ . Therefore, assumption at the bottom of page 378 in Stambaugh requires that  $(u_t^N + \mu_t^N, \varepsilon_{t+1}^{PD})$  is jointly normal. Under normality, using orthogonality of measurement error, it follows from proposition 4 in Stambaugh 1999 that

$$(A6) \quad E(\hat{c}_T^N - c^N) = \frac{\text{cov}(\varepsilon_{t+1}^{PD}, u_t^N)}{\text{var}(\varepsilon_{t+1}^{PD})} E(\hat{\rho}_T - \rho)$$

where  $\hat{\rho}_T$  is the OLS estimator of  $\rho$ . Since  $u_t$  contains information that under the null is useful for predicting future returns we could expect  $\frac{\text{cov}(\varepsilon_{t+1}^{PD}, u_t^N)}{\text{var}(\varepsilon_{t+1}^{PD})} \neq 0$  and the bias to be

non-zero. Similarly, we have,

$$(A7) \quad E(\widehat{c}_T^N - c^N) = \frac{\text{cov}(\varepsilon_{t+1}^{PD}, u_t^N + \varepsilon_t^N)}{\text{var}(\varepsilon_t^{PD})} E(\widehat{\rho}_T - \rho)$$

$$(A8) \quad = E(\widehat{c}_T^N - c^N) + \frac{\text{cov}(\varepsilon_{t+1}^{PD}, \varepsilon_t^N)}{\text{var}(\varepsilon_t^{PD})} E(\widehat{\rho}_T - \rho).$$

#### A4. Parameterization of the Wage Process

We set  $1 + \rho$  equal to the average consumption-dividend ratio in the U.S. over the period 1946-2011, using the ‘Personal Consumption Expenditures’ and ‘Net Corporate Dividends’ series from the Bureau of Economic Analysis. This delivers  $\rho = 22$ . The consumption-dividend ratio fluctuates considerably over time and displays a close to unit root behavior, with the quarterly sample autocorrelation being equal to 0.99, prompting us to consider only values close to one for the persistence parameter  $p$ .

Following Campbell and Cochrane 1999, our remaining calibration targets are

$$(A9) \quad \sigma_{c,t} = \frac{1}{7} \sigma_D$$

and

$$(A10) \quad \rho_{c,d,t} = 0.2,$$

where  $\sigma_{c,t}$  denotes the conditional standard deviation of log consumption growth and  $\rho_{c,d,t}$  the conditional correlation between log consumption growth and log dividend growth. Aggregate consumption is given by  $C_t = D_t + W_t$  so that

$$\begin{aligned} \log \frac{C_t}{C_{t-1}} &= \log \frac{D_t \left(1 + \frac{W_t}{D_t}\right)}{D_{t-1} \left(1 + \frac{W_{t-1}}{D_{t-1}}\right)} \\ &= c + \log \varepsilon_t^D + \log \varepsilon_t^W - (1 - p) \log \left(1 + \frac{W_{t-1}}{D_{t-1}}\right), \end{aligned}$$

where  $c$  summarizes constant terms. Conditional variance of log consumption growth is thus equal to

$$(A11) \quad \sigma_{c,t}^2 = \sigma_D^2 + \sigma_W^2 + 2\sigma_{DW}$$

and conditional covariance between log consumption and dividend growth given by

$$\begin{aligned} \text{cov}_{t-1} \left( \log \frac{C_t}{C_{t-1}}, \log \frac{D_t}{D_{t-1}} \right) &= \text{cov}_{t-1} (\log \varepsilon_t^D + \log \varepsilon_t^W, \log \varepsilon_t^D) \\ &= \text{cov} (\log \varepsilon_t^D + \log \varepsilon_t^W, \log \varepsilon_t^D) \\ &= \sigma_D^2 + \sigma_{DW}. \end{aligned}$$

The conditional correlation is

$$\begin{aligned}\rho_{c,d,t} &= \text{corr}_{t-1} \left( \log \frac{C_t}{C_{t-1}}, \log \frac{D_t}{D_{t-1}} \right) \\ &= \frac{\sigma_D^2 + \sigma_{DW}}{\sqrt{(\sigma_D^2 + \sigma_W^2 + 2\sigma_{DW})\sigma_D}} \\ &= 7 \frac{\sigma_D^2 + \sigma_{DW}}{\sigma_D^2},\end{aligned}$$

where the last line uses (A9) and (A11). Targeting a correlation of 0.2 thus delivers

$$(A12) \quad \sigma_{DW} = \left( \frac{0.2}{7} - 1 \right) \sigma_D^2.$$

From (A11) we then get

$$(A13) \quad \sigma_W^2 = \left( \frac{1}{49} - 1 - 2 \left( \frac{0.2}{7} - 1 \right) \right) \sigma_D^2.$$

Equations (A12) and (A13) deliver our calibration targets. As is easily verified, the implied covariance matrix for the innovations  $(\log \varepsilon_t^D, \log \varepsilon_t^W)$  has a positive determinant.

We now check what the calibration implies for the variance of unconditional log consumption growth  $\sigma_C^2$ . Using the fact that time  $t$  shocks are independent of  $\log \left( 1 + \frac{W_{t-1}}{D_{t-1}} \right)$  and letting  $\sigma_C$  denote the standard deviation of unconditional log consumption growth, we get

$$\sigma_C^2 = \sigma_D^2 + 2\sigma_{DW} + \sigma_W^2 + (1-p)^2 \text{var} \left( \log \left( 1 + \frac{W}{D} \right) \right).$$

The unconditional variance of  $\log \left( 1 + \frac{W}{D} \right)$  is given by

$$\text{var} \left( \log \left( 1 + \frac{W}{D} \right) \right) = \frac{\sigma_W^2}{1-p^2},$$

hence

$$\begin{aligned}\sigma_C^2 &= \sigma_D^2 + 2\sigma_{DW} + \frac{2(1-p)}{1-p^2} \sigma_W^2 \\ &= \sigma_D^2 + 2 \left( \frac{0.2}{7} - 1 \right) \sigma_D^2 + \frac{2(1-p)}{1-p^2} \left( \frac{1}{49} - 1 - 2 \left( \frac{0.2}{7} - 1 \right) \right) \sigma_D^2\end{aligned}$$

For  $p \rightarrow 1$ , we thus have the result that unconditional and conditional consumption growth volatility are identical ( $\sigma_C = \sigma_{c,t} = \sigma_D/7$ ). For lower values of the persistence parameter  $p$ , unconditional consumption volatility decreases somewhat relative to conditional consumption volatility. For the parameter values considered in the main text, we

have

$$\left(\frac{\sigma_D}{\sigma_C}\right) \approx \begin{cases} 7.0 & \text{for } p = 1.0 \\ 4.7 & \text{for } p = 0.95 \end{cases} .$$

A5. *Existence of Optimum, Sufficiency of FOCs, Recursive Solution*

Substituting out  $C_t$  using the budget constraint (14a) in problem (14), one obtains a problem that involves consumption choices only. Given the stock holding limits, the choice set is compact. It is also non-empty since  $S_t = 1$  for all  $t$  is feasible. The following condition then insures existence of optimal plans:

**Condition 1** The utility function  $u(\cdot)$  is bounded above and for all  $i \in [0, 1]$

$$(A14) \quad E_0^{\mathcal{P}^i} \sum_{t=0}^{\infty} \delta^t u(W_t + D_t) > -\infty.$$

The expression on the left-hand side of condition (A14) is the utility associated with never trading stocks ( $S_t^i = 1$  for all  $t$ ). Since this policy is always feasible, condition (A14) guarantees that the objective function in (14a) is also bounded from below, even if the flow utility function  $u(\cdot)$  is itself unbounded below.

From  $\gamma > 1$ , see equation (20), we have  $u(C_t) = \frac{C_t^{1-\gamma}}{1-\gamma} \leq 0$  and thus a utility function that is bounded above. Provided (A14) holds, the optimization problem (14a) thus maximizes a bounded continuous utility function over a compact set, which guarantees existence of a maximum. Under the assumptions made in the main text (utility function given by (20), knowledge of (17) and (21)), condition (A14) indeed holds, as can be seen from the following derivations:

$$\begin{aligned} & E_0^{\mathcal{P}^i} \sum_{t=0}^{\infty} \delta^t u(W_t + D_t) \\ &= E_0 \sum_{t=0}^{\infty} \delta^t u(W_t + D_t) \\ &= \frac{1}{1-\gamma} E_0 \sum_{t=0}^{\infty} \delta^t \left(1 + \frac{W_t}{D_t}\right) D_t^{1-\gamma} \\ &= \frac{1}{1-\gamma} \left(\left(1 + \frac{W_0}{D_0}\right) D_0\right)^{1-\gamma} E_0 \sum_{t=0}^{\infty} \delta^t \left(\frac{1 + \frac{W_t}{D_t}}{1 + \frac{W_0}{D_0}} \frac{D_t}{D_0}\right)^{1-\gamma} . \end{aligned}$$

Using

$$\begin{aligned} \frac{1 + \frac{W_t}{D_t}}{1 + \frac{W_0}{D_0}} &= \left(\frac{1 + \rho}{1 + \frac{W_0}{D_0}}\right)^{1-p^t} \prod_{j=0}^{t-1} (\varepsilon_{t-j}^W)^{p^j} \\ \frac{D_t}{D_0} &= (\beta^D)^t \prod_{j=0}^{t-1} \varepsilon_{t-j}^D \end{aligned}$$

we get

$$\begin{aligned} & E_0^{\mathcal{P}^i} \sum_{t=0}^{\infty} \delta^t u(W_t + D_t) \\ &= \frac{1}{1-\gamma} \left( \left(1 + \frac{W_0}{D_0}\right) D_0 \right)^{1-\gamma} E_0 \sum_{t=0}^{\infty} \left( \frac{1+\rho}{1+\frac{W_0}{D_0}} \right)^{(1-p^t)(1-\gamma)} (\delta\beta^D)^{t(1-\gamma)} \left( \prod_{j=0}^{t-1} \varepsilon_{t-j}^D (\varepsilon_{t-j}^W)^{p^j} \right)^{1-\gamma} \end{aligned}$$

The infinite sum is bounded if  $\delta (\beta^D)^{1-\gamma} E[(\varepsilon^W)^{p^j} \varepsilon^D]^{1-\gamma}$  is bounded below one for all  $j > 0$ . The following derivations establish this fact:

$$\begin{aligned} \delta (\beta^D)^{1-\gamma} E \left[ \left( (\varepsilon^W)^{p^j} \varepsilon^D \right)^{1-\gamma} \right] &= \delta (\beta^D)^{1-\gamma} \left( E \left[ (\varepsilon^W)^{p^j(1-\gamma)} \right] + E \left[ (\varepsilon^D)^{1-\gamma} \right] \right) \\ &= \delta (\beta^D)^{1-\gamma} \left( e^{p^j(\gamma-1)(1+p^j(\gamma-1))\frac{\sigma_W^2}{2}} + e^{\gamma(\gamma-1)\frac{\sigma_D^2}{2}} \right) \\ &\leq \delta (\beta^D)^{1-\gamma} \left( e^{(\gamma-1)(1+(\gamma-1))\frac{\sigma_W^2}{2}} + e^{\gamma(\gamma-1)\frac{\sigma_D^2}{2}} \right) \\ &= \delta (\beta^D)^{1-\gamma} E \left[ (\varepsilon^W)^{-\gamma} (\varepsilon^D)^{1-\gamma} \right] < 1, \end{aligned}$$

where the weak inequality follows from  $\gamma > 1$  and  $p \in [0, 1]$  and the strict inequality from (21). This establishes existence of optimal plans.

Since (14a) is a strictly concave maximization problem the maximum is unique. With the utility function being differentiable, the first order conditions

$$(A15) \quad u'(C_t^i) = \delta E_t^{\mathcal{P}^i} \left[ u'(C_{t+1}^i) \frac{P_{t+1} + D_{t+1}}{P_t} \right]$$

plus a standard transversality condition are necessary and sufficient for the optimum.

**Recursive Formulation.** We have a recursive solution whenever the optimal stock-holding policy can be written as a time-invariant function  $S_t^i = S^i(x_t)$  of some state variables  $x_t$ . We seek a recursive solution where  $x_t$  contains appropriately rescaled variables that do not grow to infinity. With this in mind, we impose the following condition:

**Condition 2** The flow utility function  $u(\cdot)$  is homogeneous of degree  $\eta \geq 0$ . Furthermore, the beliefs  $\mathcal{P}^i$  imply that  $\theta_t \equiv \left( \frac{D_t}{D_{t-1}}, \frac{P_t}{D_t}, \frac{W_t}{D_t} \right)$  has a state space representation, i.e., the conditional distribution  $\mathcal{P}^i(\theta_{t+1}|\omega^t)$  can be written as

$$(A16) \quad \mathcal{P}^i(\theta_{t+1}|\omega^t) = \mathcal{F}^i(m_t^i)$$

$$(A17) \quad m_t^i = \mathcal{R}^i(m_{t-1}^i, \theta_t)$$

for some finite-dimensional state vector  $m_t^i$  and some time-invariant functions  $\mathcal{F}^i$  and  $\mathcal{R}^i$ .

Under Condition 2, problem (14a) can then be re-expressed as

$$(A18) \quad \max_{\{S_t^i \in [\underline{S}, \bar{S}]\}_{t=0}^{\infty}} E_0^{\mathcal{P}^i} \sum_{t=0}^{\infty} \delta^t \mathcal{D}_t u \left( S_{t-1}^i \left( \frac{P_t}{D_t} + 1 \right) - S_t^i \frac{P_t}{D_t} + \frac{W_t}{D_t} \right),$$

given  $S_{-1}^i = 1$ , where  $\mathcal{D}_t$  is a time-varying discount factor satisfying  $\mathcal{D}_{-1} = 1$  and

$$\mathcal{D}_t = \mathcal{D}_{t-1} (\beta^D \varepsilon_t^D)^{\eta}.$$

The return function in (A18) depends only on the exogenous variables contained in the vector  $\theta_t$ . Since the beliefs  $\mathcal{P}^i$  are assumed to be recursive in  $\theta_t$ , standard arguments in dynamic programming guarantee that the optimal solution to (A18) takes the form (22). This formulation of the recursive solution is useful, because scaling  $P_t$  and  $W_t$  by the level of dividends eliminates the trend in these variables, as desired. This will be useful when computing numerical approximations to  $S^i(\cdot)$ . The belief systems  $\mathcal{P}^i$  introduced in section V will satisfy the requirements stated in condition 2.

#### A6. Proof of Proposition 2

For general  $p$  we have

$$\begin{aligned} \left(1 + \frac{W_t}{D_t}\right) &= (1 + \rho)^{(1-p)} \left(1 + \frac{W_{t-1}}{D_{t-1}}\right)^p \ln \varepsilon_t^W \\ D_t &= \beta^D D_{t-1} \varepsilon_t^D \end{aligned}$$

so that for  $S_t^i = 1$  for all  $t \geq 0$ , the budget constraint implies  $C_t^i = D_t + W_t = D_t(1 + \frac{W_t}{D_t})$ .

Substituting this into the agent's first order condition delivers

$$(A19) \quad P_t = \delta E_t \left[ \left( \frac{\left(1 + \frac{W_{t+1}}{D_{t+1}}\right) D_{t+1}}{\left(1 + \frac{W_t}{D_t}\right) D_t} \right)^{-\gamma} (P_{t+1} + D_{t+1}) \right].$$

Assuming that the following transversality condition holds

$$(A20) \quad \lim_{j \rightarrow \infty} E_t \left[ \delta^j \left\{ \left( \frac{1 + \frac{W_{t+j}}{D_{t+j}}}{1 + \frac{W_t}{D_t}} \right) \frac{D_{t+j}}{D_t} \right\}^{-\gamma} P_{t+j} \right] = 0,$$

one can iterate forward on (A19) to obtain

$$\frac{P_t}{D_t} = E_t \left[ \sum_{j=1}^{\infty} \delta^j \left( \frac{1 + \frac{W_{t+j}}{D_{t+j}}}{1 + \frac{W_t}{D_t}} \right)^{-\gamma} \left( \frac{D_{t+j}}{D_t} \right)^{1-\gamma} \right].$$



Using  $D_{t+j}/D_t = (\beta^D)^j \prod_{k=1}^j \varepsilon_{t+k}^D$  and

$$\frac{1 + \frac{W_{t+j}}{D_{t+j}}}{1 + \frac{W_t}{D_t}} = \left( \frac{1 + \rho}{1 + \frac{W_t}{D_t}} \right)^{1-p^j} \prod_{i=0}^{t-1} (\varepsilon_{t+j-i}^W)^{p^i},$$

one has

$$\begin{aligned} \frac{P_t}{D_t} &= E_t \left[ \sum_{j=1}^{\infty} \delta^j \left( \frac{1 + \frac{W_{t+j}}{D_{t+j}}}{1 + \frac{W_t}{D_t}} \right)^{-\gamma} \left( \frac{D_{t+j}}{D_t} \right)^{1-\gamma} \right] \\ (A21) \quad &= \sum_{j=1}^{\infty} (\delta(\beta^D)^{1-\gamma})^j E_t \left[ \left( \frac{1 + \rho}{1 + \frac{W_t}{D_t}} \right)^{-\gamma(1-p^j)} \prod_{k=1}^j (\varepsilon_{t+k}^D)^{1-\gamma} (\varepsilon_{t+k}^W)^{-\gamma p^k} \right] \end{aligned}$$

The infinite sum in the previous expression is bounded, if  $E_t \left[ (\varepsilon_{t+1}^W)^{-\gamma p^{j-1}} (\varepsilon_{t+1}^D)^{1-\gamma} \right]$  remains bounded away from one for all  $j > 1$ . This follows from the following derivations:

$$\begin{aligned} &E_t \left[ (\varepsilon_{t+1}^W)^{-\gamma p^{j-1}} (\varepsilon_{t+1}^D)^{1-\gamma} \right] = \\ &= E_t \left[ (\varepsilon_{t+1}^W)^{-\gamma p^{j-1}} \right] + E_t \left[ (\varepsilon_{t+1}^D)^{1-\gamma} \right] \\ &= e^{-\gamma p^{j-1} \frac{-\sigma_W^2}{2} + (\gamma p^{j-1})^2 \frac{\sigma_W^2}{2}} + e^{(1-\gamma) \frac{-\sigma_D^2}{2} + (1-\gamma)^2 \frac{\sigma_D^2}{2}} \\ &\leq e^{\gamma(1+\gamma) \frac{\sigma_W^2}{2}} + e^{(\gamma-1)\gamma \frac{\sigma_D^2}{2}} = E_t \left[ (\varepsilon_{t+1}^W)^{-\gamma} (\varepsilon_{t+1}^D)^{1-\gamma} \right] < 1, \end{aligned}$$

where the inequality in the third to last line follows from  $\gamma > 1$  and  $p \in [0, 1]$  and the last inequality from assumption (21). For the special cases  $p = 1$  and  $p = 0$ , equation (A21) delivers the expressions stated in proposition 2.

#### A7. Bayesian Foundations for Lagged Belief Updating

We now present a slightly modified information structure for which Bayesian updating gives rise to the lagged belief updating equation (39). Specifically, we generalize the perceived price process (26) by splitting the temporary return innovation  $\ln \varepsilon_{t+1}$  into two independent subcomponents:

$$\ln P_{t+1} - \ln P_t = \ln \beta_{t+1} + \ln \varepsilon_{t+2}^1 + \ln \varepsilon_{t+1}^2$$

with  $\ln \varepsilon_{t+2}^1 \sim iiN(-\frac{\sigma_{\varepsilon_1}^2}{2}, \sigma_{\varepsilon_1}^2)$ ,  $\ln \varepsilon_{t+1}^2 \sim iiN(-\frac{\sigma_{\varepsilon_2}^2}{2}, \sigma_{\varepsilon_2}^2)$  and

$$\sigma_{\varepsilon}^2 = \sigma_{\varepsilon_1}^2 + \sigma_{\varepsilon_2}^2.$$

We then assume that in any period  $t$  agents observe the prices, dividends and wages up to period  $t$ , as well as the innovations  $\varepsilon_t^1$  up to period  $t$ . Agents' time  $t$  information set thus consists of  $I_t = \{P_t, D_t, W_t, \varepsilon_t^1, P_{t-1}, D_{t-1}, W_{t-1}, \varepsilon_{t-1}^1, \dots\}$ . By observing the innovations  $\varepsilon_t^1$ , agents learn - with a one period lag - something about the temporary components of price growth. The process for the persistent price growth component  $\ln \beta_t$  remains as stated in equation (27), but we now denote the innovation variance by  $\sigma_v^2$  instead of  $\sigma_v^2$ . As before,  $\ln m_t$  denotes the posterior mean of  $\ln \beta_t$  given the information available at time  $t$ . We prove below the following result:

**PROPOSITION 4:** Fix  $\sigma_\varepsilon^2 > 0$  and consider the limit  $\sigma_{\varepsilon 2}^2 \rightarrow 0$  with  $\sigma_v^2 = \sigma_{\varepsilon 2}^2 g^2 / (1 - g)$ . Bayesian updating then implies

$$(A22) \quad \ln m_t = \ln m_{t-1} + g (\ln P_{t-1} - \ln P_{t-2} - \ln m_{t-1}) - g \ln \varepsilon_t^1$$

The modified information structure thus implies that only lagged price growth rates enter the current state estimate, so that beliefs are predetermined, precisely as assumed in equation (39). Intuitively, this is so because lagged returns become infinitely more informative relative to current returns as  $\sigma_{\varepsilon 2}^2 \rightarrow 0$ , which eliminates the simultaneity problem. For non-vanishing uncertainty  $\sigma_{\varepsilon 2}^2$  the weight of the last observation actually remains positive but would still be lower than that given to the lagged return observation, see equation (A25) in the proof below and the subsequent discussion for details.

We now provide the proof of the previous proposition. Let us define the following augmented information set  $\tilde{I}_{t-1} = I_{t-1} \cup \{\varepsilon_t^1\}$ . The posterior mean for  $\beta_t$  given  $\tilde{I}_{t-1}$ , denoted  $\ln m_{t|\tilde{I}_{t-1}}$  is readily recursively determined via

$$(A23) \quad \ln m_{t|\tilde{I}_{t-1}} = \ln m_{t-1|\tilde{I}_{t-2}} - \frac{\sigma_v^2}{2} + \tilde{g} \left( \ln P_{t-1} - \ln P_{t-2} - \ln \varepsilon_t^1 + \frac{\sigma_v^2 + \sigma_{\varepsilon 2}^2}{2} - \ln m_{t-1|\tilde{I}_{t-2}} \right)$$

and the steady state posterior uncertainty and the Kalman gain by

$$(A24) \quad \begin{aligned} \sigma^2 &= \frac{-\sigma_v^2 + \sqrt{(\sigma_v^2)^2 + 4\sigma_v^2\sigma_{\varepsilon 2}^2}}{2} \\ \tilde{g} &= \frac{\sigma^2}{\sigma_{\varepsilon 2}^2} \end{aligned}$$

We furthermore have

$$E[\ln P_t - \ln P_{t-1} | \tilde{I}_{t-1}] = \ln m_{t|\tilde{I}_{t-1}} - \frac{\sigma_{\varepsilon 1}^2 + \sigma_{\varepsilon 2}^2}{2}$$

and

$$(\ln P_t - \ln P_{t-1}) - E[\ln P_t - \ln P_{t-1} | \tilde{I}_{t-1}] = \ln \beta_t - \ln m_{t|\tilde{I}_{t-1}} + \ln \varepsilon_{t+1}^1 + \ln \varepsilon_t^2 + \frac{\sigma_{\varepsilon 1}^2 + \sigma_{\varepsilon 2}^2}{2}$$

so that

$$\begin{pmatrix} \ln \beta_t \\ \ln P_t - \ln P_{t-1} \end{pmatrix} | \tilde{I}_{t-1} \sim N \left( \begin{pmatrix} \ln m_{t|\tilde{I}_{t-1}} \\ \ln m_{t|\tilde{I}_{t-1}} - \frac{\sigma_{\varepsilon_1}^2 + \sigma_{\varepsilon_2}^2}{2} \end{pmatrix}, \begin{pmatrix} \sigma^2 & \sigma^2 \\ \sigma^2 & \sigma^2 + \sigma_{\varepsilon_1}^2 + \sigma_{\varepsilon_2}^2 \end{pmatrix} \right),$$

where the covariance between  $\ln \beta_t | \tilde{I}_{t-1}$  and  $\ln P_t - \ln P_{t-1} | \tilde{I}_{t-1}$  can be computed by exploiting the fact that  $\ln \beta_t - \ln m_{t|\tilde{I}_{t-1}}$  and  $\ln \varepsilon_{t+1}^1 + \ln \varepsilon_t^2$  are independent and using  $\ln \beta_t - E_t[\ln \beta_t | \tilde{I}_{t-1}] = \ln \beta_t - \ln m_{t|\tilde{I}_{t-1}}$ . Using standard normal updating formulae, we can thus compute

$$(A25) \quad \ln m_{t|I_t} = E[\ln \beta_t | I_t] = E[\ln \beta_t | \tilde{I}_{t-1}, \ln P_t - \ln P_{t-1}] \\ = \ln m_{t|\tilde{I}_{t-1}} + \frac{\sigma^2}{\sigma^2 + \sigma_{\varepsilon_1}^2 + \sigma_{\varepsilon_2}^2} \left( \ln P_t - \ln P_{t-1} + \frac{\sigma_{\varepsilon_1}^2 + \sigma_{\varepsilon_2}^2}{2} - \ln m_{t|\tilde{I}_{t-1}} \right),$$

where the second equality exploits the fact that  $D_t, W_t$  contain no information about  $\ln \beta_t$ , and the second inequality follows from the fact that  $\tilde{I}_{t-1}$  contains information up to  $t-1$  and including  $\varepsilon_t^1$ , since the latter is independent of  $(\ln P_t - \ln P_{t-1})$ .

Since  $\frac{\sigma^2}{\sigma^2 + \sigma_{\varepsilon_1}^2 + \sigma_{\varepsilon_2}^2} < \frac{\sigma^2}{\sigma_{\varepsilon_2}^2} = \tilde{g}$ , the weight of the price observation dated  $t$  is reduced relative to the earlier observation dated  $t-1$  because it is ‘noisier’. Now consider the limit  $\sigma_{\varepsilon_2}^2 \rightarrow 0$  and along the limit choose  $\sigma_{\varepsilon_1}^2 = \sigma_\varepsilon^2 - \sigma_{\varepsilon_2}^2$  and  $\sigma_v^2 = \frac{g^2}{1-g} \sigma_{\varepsilon_2}^2$ , as assumed in the proposition. From  $\sigma^2 \rightarrow 0$  and equation (A25) it then follows that  $\ln m_{t|I_t} = \ln m_{t|\tilde{I}_{t-1}}$ , i.e., the weight of the last observation price converges to zero. Moreover, from  $\sigma_v^2 = \frac{g^2}{1-g} \sigma_{\varepsilon_2}^2$  and (A24) we get  $\tilde{g} = g$ . Using these results, equation (A23) implies equation (A22).

#### A8. Proof of Proposition 3

The proof uses the assumption of no uncertainty so that for any function  $f$  we have  $E_t^P f(X_{t+j}, Y_{t+j}) = f(E_t^P X_{t+j}, E_t^P Y_{t+j})$ . Simplifying notation (and slightly abusing it) in this appendix we let  $X_{t+j} = E_t^P X_{t+j}$  for all  $j \geq 1$ , so that  $X_{t+j}$  below denotes the subjective expectation conditional on information at time  $t$  of the variable  $X$  at time  $t+j$ . The first order conditions (A15) can then be written as

$$(A26) \quad C_t^{-\gamma} = C_{t+1}^{-\gamma} \delta R_{t+1} \implies C_t = \delta^{-\frac{j}{\gamma}} \prod_{\tau=1}^j R_{t+\tau}^{-\frac{1}{\gamma}} C_{t+j}$$

for all  $t, j \geq 0$ , assuming the stock limits are not binding in periods  $t, t+1, \dots, t+j-1$ . Iterating forward  $N$  periods on the budget constraint of the agent and using the fact that

either  $\prod_{\tau=1}^N R_{t+\tau}^{-1} \rightarrow 0$  or  $S_{t+N} \rightarrow 0$  as  $N \rightarrow \infty$  we have

$$(P_t + D_t) S_{t-1} = \sum_{j=0}^{\infty} \left( \prod_{\tau=1}^j R_{t+\tau}^{-1} \right) (C_{t+j} - W_{t+j})$$

Using equation (A26) to substitute out  $C_{t+j}$  gives

$$(P_t + D_t) S_{t-1} = \sum_{j=0}^{\infty} \left( \prod_{\tau=1}^j R_{t+\tau}^{-1} \right) \left[ \left( \frac{W_t}{D_t} + 1 \right) \delta^{\frac{j}{\gamma}} \prod_{\tau=1}^j R_{t+\tau}^{\frac{1}{\gamma}} - W_{t+j} \right]$$

assuming the stock limits are large enough not to be binding.

Imposing on the previous equation  $S_{t-1} = 1$  (the market clearing condition for  $S_{t-1}$  if  $t \geq 0$ , or the initial condition for period  $t = 0$ ) and  $C_t = D_t + W_t$  (the market clearing condition for consumption) one obtains

$$\frac{P_t}{D_t} + 1 = \sum_{j=0}^N \left( \prod_{\tau=1}^j R_{t+\tau}^{-1} \right) \left[ \left( \frac{W_t}{D_t} + 1 \right) \delta^{\frac{j}{\gamma}} \prod_{\tau=1}^j R_{t+\tau}^{\frac{1}{\gamma}} - \frac{W_{t+j}}{D_t} \right]$$

Cancelling the terms for  $j = 0$  in each summation gives for the market-clearing price

$$\frac{P_t}{D_t} = \sum_{j=1}^N \left( \prod_{\tau=1}^j R_{t+\tau}^{-1} \right) \left[ \left( \frac{W_t}{D_t} + 1 \right) \delta^{\frac{j}{\gamma}} \prod_{\tau=1}^j R_{t+\tau}^{\frac{1}{\gamma}} - \frac{W_{t+j}}{D_t} \right]$$

Using (40) gives (41).

#### A9. Verification of Conditions (40)

For the vanishing noise limit of the beliefs specified in section V we have

$$\begin{aligned} E_t^{\mathcal{P}}[P_{t+j}] &= (m_t)^j P_t \\ E_t^{\mathcal{P}}[D_{t+j}] &= (\beta^D)^j D_t \\ E_t^{\mathcal{P}}[W_{t+j}] &= (\beta^D)^j W_t \end{aligned}$$

where we have abstracted from transitional dynamics in the  $W_t/D_t$  ratio and assume  $W_t/D_t = \rho$ , as transitional dynamics do not affect the limit results. We first verify the inequality on the l.h.s. of equation (40). We have

$$\lim_{T \rightarrow \infty} E_t^{\mathcal{P}}[R_T] = m_t + \lim_{T \rightarrow \infty} \left( \frac{\beta^D}{m_t} \right)^{T-1} \beta^D \frac{D_t}{P_t},$$

so that for  $m_t > 1$  the limit clearly satisfies  $\lim_{T \rightarrow \infty} E_t^{\mathcal{P}}[R_T] > 1$  due to the first term on the r.h.s.; for  $m_t < 1$  the second term on the r.h.s. increases without bound, due to  $\beta^D > 1$ , so that  $\lim_{T \rightarrow \infty} E_t^{\mathcal{P}}[R_T] > 1$  also holds.

In a second step we verify that the inequality condition on the r.h.s. of equation (40) holds for all subjective beliefs  $m_t > 0$ . We have

$$\begin{aligned} \lim_{T \rightarrow \infty} E_t^{\mathcal{P}} \left( \sum_{j=1}^T \left( \prod_{i=1}^j \frac{1}{R_{t+i}} \right) W_{t+j} \right) &= \lim_{T \rightarrow \infty} W_t E_t^{\mathcal{P}} \left( \sum_{j=1}^T (\beta^D)^j \left( \prod_{i=1}^j \frac{1}{R_{t+i}} \right) \right) \\ (A27) \qquad \qquad \qquad &= \lim_{T \rightarrow \infty} W_t \sum_{j=1}^T X_j \end{aligned}$$

where

$$(A28) \qquad X_j = \frac{(\beta^D)^j}{\prod_{i=1}^j \left( m_t + \left( \frac{\beta^D}{m_t} \right)^{i-1} \beta^D \frac{D_t}{P_t} \right)} \geq 0$$

A sufficient condition for the infinite sum in (A27) to converge is that the terms  $X_j$  are bounded by some exponentially decaying function. The denominator in (A28) satisfies

$$\begin{aligned} &\prod_{i=1}^j \left( m_t + \left( \frac{\beta^D}{m_t} \right)^{i-1} \beta^D \frac{D_t}{P_t} \right) \\ (A29) \qquad &\geq (m_t)^j + \left( \frac{\beta^D}{m_t} \right)^{j \left( \frac{i-1}{2} \right)} \beta^D \frac{D_t}{P_t}, \end{aligned}$$

where the first term captures the the pure products in  $m_t$ , the second term the pure products in  $\left( \frac{\beta^D}{m_t} \right)^{i-1} \beta^D \frac{D_t}{P_t}$ , and all cross terms have been dropped. We then have

$$\begin{aligned} X_j &= \frac{(\beta^D)^j}{\prod_{i=1}^j \left( m_t + \left( \frac{\beta^D}{m_t} \right)^{i-1} \beta^D \frac{D_t}{P_t} \right)} \\ &\leq \frac{(\beta^D)^j}{(m_t)^j + \left( \frac{\beta^D}{m_t} \right)^{j \left( \frac{i-1}{2} \right)} \beta^D \frac{D_t}{P_t}} \\ &= \frac{1}{\left( \frac{m_t}{\beta^D} \right)^j + \left( \frac{\beta^D}{m_t} \right)^{j \left( \frac{i-1}{2} \right)} \frac{1}{(\beta^D)^{j-1}} \frac{D_t}{P_t}}, \end{aligned}$$

where all terms in the denominator are positive. For  $m_t \geq \beta^D > 1$  we can use the first term in the denominator to exponentially bound  $X_j$ , as  $X_j \leq \left( \frac{\beta^D}{m_t} \right)^j$ ; for  $m_t < \beta^D$  we

can use the second term:

$$X_j \leq \frac{1}{\left(\frac{\beta^D}{m_t}\right)^{j\left(\frac{j-1}{2}\right)} \frac{1}{(\beta^D)^{j-1}} \frac{D_t}{P_t}} = \frac{1}{\left(\left(\frac{\beta^D}{m_t}\right)^{\frac{j}{2}} \frac{1}{\beta^D}\right)^{j-1} \frac{D_t}{P_t}}$$

Since  $m_t < \beta^D$  there must be a  $J < \infty$  such that

$$\left(\frac{\beta^D}{m_t}\right)^{\frac{j}{2}} \frac{1}{\beta^D} \geq \frac{\beta^D}{m_t} > 1$$

for all  $j \geq J$ , so that the  $X_j$  are exponentially bounded for all  $j \geq J$ .

#### A10. Proof of Lemma 1

Proof of lemma 1: We start by proving the first point in the lemma. The price, dividend and belief dynamics in the deterministic model are described by the following equations

$$\begin{aligned} \text{(A30)} \quad \ln m_t &= \ln m_{t-1} + g(\ln P_{t-1} - \ln P_{t-2} - \ln m_{t-1}) \\ \ln P_t - \ln D_t &= f(\ln m_t) \\ \ln D_t - \ln D_{t-1} &= \beta^D, \end{aligned}$$

where  $f(\cdot)$  is a continuous function, implicitly defined by the log of the  $P_t/D_t$  solution to equation (42).<sup>100</sup> Substituting the latter two equations into the first delivers

$$\ln m_t - \ln m_{t-1} = g(f(\ln m_{t-1}) - f(\ln m_{t-2}) + \ln \beta^D - \ln m_{t-1}).$$

If  $\ln m_t$  converges, then the l.h.s. of the previous equation must converge to zero. Since  $f(\cdot)$  is continuous, this means that  $m_{t-1}$  must converge to  $\beta^D$ , as claimed.

We now prove the second point in the lemma. The belief dynamics implied by the second order difference equation (A30) can be expressed as a two-dimensional first order difference equation using the mapping  $F : R^2 \rightarrow R^2$ , defined as

$$F(x) = \begin{pmatrix} x_1 + g(\ln PD(e^{x_1}) - \ln PD(e^{x_2}) + \ln \beta^D - x_1) \\ x_1 \end{pmatrix},$$

so that

$$\begin{pmatrix} \ln m_t \\ \ln m_{t-1} \end{pmatrix} = F \begin{pmatrix} \ln m_{t-1} \\ \ln m_{t-2} \end{pmatrix}.$$

Clearly,  $F$  has a fixed point at the RE solution, i.e.,  $(\ln \beta^D, \ln \beta^D)' = F(\ln \beta^D, \ln \beta^D)'$ . Moreover,  $m_t$  locally converges to the RE beliefs if and only if

$$\text{(A31)} \quad \frac{\partial F(\ln \beta^D, \ln \beta^D)}{\partial x'} = \begin{pmatrix} 1 + g(\zeta - 1) & -g\zeta \\ 1 & 0 \end{pmatrix}$$

<sup>100</sup>Since we are interested in asymptotic results and since  $W_t/D_t \rightarrow \rho$ , pricing is asymptotically given by equation (42).

has all eigenvalues less than one in absolute value, where  $\zeta \equiv \left. \frac{\partial \ln PD(e^{\ln m})}{\partial \ln m} \right|_{m=\beta^D} = \frac{\beta^D}{PD(\beta^D)} \left. \frac{\partial PD(m)}{\partial m} \right|_{m=\beta^D}$  with  $|\zeta| < 1$ . The eigenvalues of the matrix in equation (A31) are

$$\lambda = \frac{1 + g(\zeta - 1) \pm \sqrt{(1 + g(\zeta - 1))^2 - 4g\zeta}}{2}.$$

From  $|\zeta| < 1$  and  $g < \frac{1}{2}$  follows that  $(1 + g(\zeta - 1))^2 - 4g\zeta > 1 - 2g\zeta - 2g \geq 0$ , so that all eigenvalues are real. As is easily verified, we have  $\lambda^+ < 1$  because

$$\begin{aligned} 1 + g(\zeta - 1) &< 2 - \sqrt{(1 + g(\zeta - 1))^2 - 4g\zeta} \Leftrightarrow \\ \sqrt{(1 + g(\zeta - 1))^2 - 4g\zeta} &< 1 - g(\zeta - 1) \Leftrightarrow \\ (1 + g(\zeta - 1))^2 - 4g\zeta &< (1 - g(\zeta - 1))^2 \Leftrightarrow \\ 2g(\zeta - 1) - 4g\zeta &< -2g(\zeta - 1) \Leftrightarrow \\ -g &< 0 \end{aligned}$$

and  $\lambda^- < 1$  because

$$-1 + g(\zeta - 1) < \sqrt{(1 + g(\zeta - 1))^2 - 4g\zeta}$$

where the l.h.s. is negative and the r.h.s. positive. We have  $\lambda^+ > -1$  if and only if

$$1 + g(\zeta - 1) > -2 - \sqrt{(1 + g(\zeta - 1))^2 - 4g\zeta}$$

From  $|\zeta| < 1$  and  $g < \frac{1}{2}$  the l.h.s. is weakly positive, while the r.h.s. is strictly negative. We have  $\lambda^- > -1$  if and only if

$$\begin{aligned} 3 + g(\zeta - 1) &> \sqrt{(1 + g(\zeta - 1))^2 - 4g\zeta} \Leftrightarrow \\ (2 + (1 + g(\zeta - 1)))^2 &> (1 + g(\zeta - 1))^2 - 4g\zeta \Leftrightarrow \\ 1 + (1 + g(\zeta - 1)) &> -g\zeta \Leftrightarrow \\ 2 + g(2\zeta - 1) &> 0 \end{aligned}$$

The last equation holds since  $|\zeta| < 1$  and  $g < \frac{1}{2}$ . This shows that the eigenvalues of (A31) are all inside the unit circle.

#### A11. Capital Gains Expectations and Expected Returns: Further Details

Figure A1 depicts how expected returns at various horizons depend on agent's expected price growth expectations using the same parameterization as used in figure 5. It shows that expected returns covary positively with capital gains expectations for  $m_t \geq \beta^D$ , as has been claimed in the main text. The flat part at around  $m_t - 1 \approx 0.01$  arises because in that area the PD ratio increases strongly, so that the dividend yield falls. Only

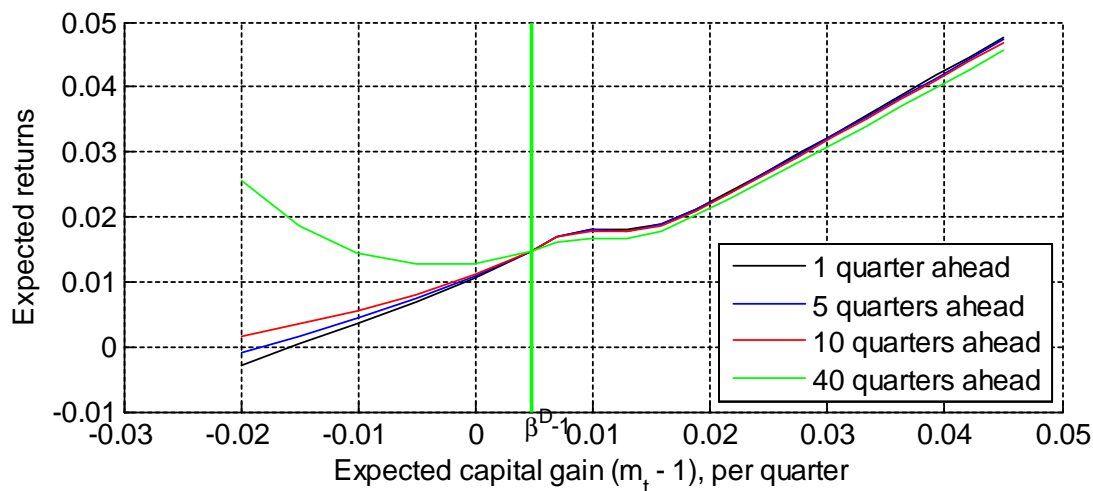


FIGURE A1. EXPECTED RETURN AS A FUNCTION OF EXPECTED CAPITAL GAIN

for pessimistic price growth expectations ( $m_t < \beta^D$ ) and long horizons of expected returns we find a negative relationship. The latter emerges because with prices expected to fall, the dividend yield will rise and eventually result in high return expectations.

#### A12. Numerical Solution Algorithm

**Algorithm:** We solve for agents' state-contingent, time-invariant stockholdings (and consumption) policy (22) using time iteration in combination with the method of endogenous grid points. Time iteration is a computationally efficient, e.g., Aruoba et al. 2006, and convergent solution algorithm, see Rendahl 2013. The method of endogenous grid points, see Carroll 2006, economizes on a costly root finding step which speeds up computations further.

**Evaluations of Expectations:** Importantly, agents evaluate the expectations in the first order condition (A15) according to their subjective beliefs about future price growth and their (objective) beliefs about the exogenous dividend and wage processes. Expectations are approximated via Hermite Gaussian quadrature using three interpolation nodes for the exogenous innovations.

**Approximation of Optimal Policy Functions:** The consumption/stockholding policy is approximated by piecewise linear splines, which preserves the nonlinearities arising in particular in the PD dimension of the state space. Once the state-contingent consumption policy has been found, we use the market clearing condition for consumption goods to determine the market clearing PD ratio for each price-growth belief  $m_t$ .

**Accuracy:** Carefully choosing appropriate grids for each belief is crucial for the accuracy of the numerical solution. We achieve maximum (relative) Euler errors on the order of  $10^{-3}$  and median Euler errors on the order of  $10^{-5}$  (average:  $10^{-4}$ ).

Using our analytical solution for the case with vanishing noise, we can assess the accuracy of our solution algorithm more directly. Setting the standard deviations of



exogenous disturbances to  $10^{-16}$  the algorithm almost perfectly recovers the equilibrium PD ratio of the analytical solution: the error for the numerically computed equilibrium PD ratio for any price growth belief  $m_t$  on our grid is within 0.5 % of the analytical solution.

### A13. Model with Stock Supply Shocks

With the supply shocks specified in the main text, the individual optimization problem remains unchanged. The only point where the model changes is when we compute market clearing prices. These now have to satisfy the relation

$$S(e^{\varepsilon_t^s}, \frac{P_t}{D_t}, \frac{W_t}{D_t}, m_t) = e^{\varepsilon_t^s},$$

where

$$\varepsilon_t^s \sim iiN(-\frac{\sigma_{\varepsilon^s}^2}{2}, \sigma_{\varepsilon^s}^2).$$

To illustrate model performance in the presence of supply shocks and to compare it to the model without supply shocks, we continue to keep the parameter values for  $(\gamma, \delta, p)$  equal to the estimated ones for the model from table 3 (diagonal matrix) and only reestimate the gain value  $g$ , so as to fit the asset pricing moments reported in table 3, using a diagonal weighting matrix. Clearly, an even better match with the data moments can be achieved by reestimating all parameters. For  $\sigma_{\varepsilon^s}^2 = 1.25 \cdot 10^{-3}$ , the new gain estimate is  $\hat{g} = 0.02315$ .

Table A5 below reports the moments and t-ratios for the model with supply shocks. The model implied standard deviation for the risk free rate and the autocorrelations of the excess stock returns and stock returns are reported in table 6 in the main text. In terms of the t-ratios in table A1, the model performs equally well as the baseline models from table 3, except for the autocorrelation of the PD ratio which is now somewhat lower. The fit with the latter moment could be improved further by relaxing the assumption that  $\varepsilon_t^s$

is iid.

	<b>Subj. Belief Model with Supply Shocks</b>	
	Moment	t-ratio
<b>E[PD]</b>	107.3	-1.30
<b>Std[PD]</b>	74.1	0.60
<b>Corr[PD<sub>t</sub>, PD<sub>t-1</sub>]</b>	0.96	-6.66
<b>Std[r<sup>s</sup>]</b>	7.45	-1.36
<i>c</i>	-0.0044	-0.28
<i>R</i> <sup>2</sup>	0.16	-0.63
<b>E[r<sup>s</sup>]</b>	1.81	-0.18
<b>E[r<sup>b</sup>]</b>	0.98	4.99
<b>UBS Survey Data:</b>		
<b>Corr[PD<sub>t</sub>, E<sub>t</sub><sup>P</sup>R<sub>t+1</sub>]</b>	0.75	-0.69

Table A5: Asset pricing moments, estimated model with supply shocks from table 6 in the main text

#### A14. Derivation of Approximate Sharpe Ratios

Under rational and subjective price beliefs, the following two first-order-conditions hold, namely the first order condition for stocks

$$(A32) \quad 1 = E_t^{\mathcal{P}} \left[ \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} (1 + r_{t+1}^s) \right],$$

where  $1 + r_{t+1}^s = \frac{P_{t+1} + D_{t+1}}{P_t}$  is the gross real stock return, and the first order condition for bonds

$$(A33) \quad 1 + r_t^b = \frac{1}{E_t^{\mathcal{P}} \left[ \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \right]}.$$

Equation (A32) can be written as

$$(A34) \quad E_t^{\mathcal{P}} \left[ \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \right] \cdot E_t^{\mathcal{P}} [r_{t+1}^s] + cov_t^{\mathcal{P}} \left[ \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma}, r_{t+1}^s \right] = 1.$$

Dividing the previous equation by  $E_t^{\mathcal{P}} \left[ \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \right]$  and using (A33) one obtains

$$(A35) \quad E_t^{\mathcal{P}} [r_{t+1}^s] - r_t^b = -(1 + r_t^b) \cdot \text{cov}_t^{\mathcal{P}} \left[ \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma}, 1 + r_{t+1}^s \right].$$

Dividing by  $Std_t^{\mathcal{P}} [r_{t+1}^s]$  delivers

$$(A36) \quad \frac{E_t^{\mathcal{P}} [r_{t+1}^s] - r_t^b}{Std_t^{\mathcal{P}} [r_{t+1}^s]} = -(1 + r_t^b) \cdot Std_t^{\mathcal{P}} [\delta (C_{t+1}/C_t)^{-\gamma}] \text{corr}_t^{\mathcal{P}} \left[ \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma}, 1 + r_{t+1}^s \right],$$

where the left-hand side is the (subjective) conditional Sharpe ratio and  $\text{corr}_t^{\mathcal{P}} [\cdot]$  the (subjective) conditional correlation. Using the fact that

$$\begin{aligned} (1 + r_t^b) \delta &\approx 1 \\ \text{corr}_t^{\mathcal{P}} \left[ \delta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma}, 1 + r_{t+1}^s \right] &\approx -1, \end{aligned}$$

where the latter follows from the first order condition (A32), we have under the additional assumption of log-normal consumption growth (which holds exactly in the case with rational price expectations):

$$(A37) \quad \frac{E_t^{\mathcal{P}} [r_{t+1}^s] - r_t^b}{Std_t^{\mathcal{P}} [r_{t+1}^s]} \approx \gamma Std_t^{\mathcal{P}} [C_{t+1}/C_t].$$

For the case with rational price expectations ( $E_t^{\mathcal{P}}[\cdot] = E_t[\cdot]$ ,  $Std_t^{\mathcal{P}}[\cdot] = Std_t[\cdot]$ ), it then follows from (24) that  $Std_t [r_{t+1}^s] \approx Std [r_{t+1}^s]$ , so that by using this relationship to substitute  $Std_t [r_{t+1}^s]$  in (A37) and by applying the unconditional rational expectations operator on both sides of the equation, one obtains equation (45) in the main text.

For the case with subjective price expectations, we have

$$(A38) \quad \begin{aligned} Std_t^{\mathcal{P}} [r_{t+1}^s] &\approx Std^{\mathcal{P}} [r_{t+1}^s] \\ &\approx Std [r^s], \end{aligned}$$

where the first approximation is a feature of the subjective price belief system<sup>101</sup> and

<sup>101</sup> According to agents' subjective beliefs

$$(A39) \quad \begin{aligned} 1 + r_{t+1}^s &= \frac{P_{t+1} + D_{t+1}}{P_t} \\ &= \beta_{t+1} v_{t+1} \varepsilon_{t+1} + \frac{D_t}{P_t} \varepsilon_t^D. \end{aligned}$$

Since  $\frac{D_t}{P_t}$  is small,  $\beta_{t+1} v_{t+1} \approx 1$ , and the standard deviation of  $\varepsilon_t^D$  is small relative to the standard deviation of  $\varepsilon_{t+1}$ , the last term in (A39) contributes little to the standard deviation of  $r_{t+1}^s$ . It then follows from  $\beta_{t+1} v_{t+1} \approx 1$  that

the second approximation due to the way we calibrated the standard deviation  $\sigma_\varepsilon$  of the transitory price shock  $\varepsilon_t$  in table 2. Using (A38) to substitute  $Std_t^{\mathcal{P}} [r_{t+1}^s]$  in (A37) and applying the unconditional expectations operator on both sides of the equation delivers

$$\frac{E [E_t^{\mathcal{P}} [r_{t+1}^s]] - E [r_t^b]}{Std [r^s]} \approx \gamma E [Std_t^{\mathcal{P}} [C_{t+1}/C_t]],$$

which implies (46).

#### A15. Simultaneous Belief Updating

This appendix provides further information about the extended model with a share of simultaneous belief updaters considered in section X.C.

Table A6 reports for different values of  $\alpha$  the model moments when the selection rule chooses the closest market clearing price.<sup>102</sup> The table shows that most asset pricing implications of the subjective belief model are rather robust to allowing for contemporaneous belief updating. The main quantitative effects of introducing current updaters consists of increasing the volatility of stock returns, the volatility of the PD ratio and the equity premium. These effects become more pronounced as the share of current updaters  $\alpha$  increases, as this leads to an increase in the percentage share of periods with multiple market clearing prices. Importantly, however, the objectively expected discounted utility of agents that use lagged belief updating exceeds - for all values of  $\alpha$  - that of agents who use current belief updating, see the last two rows in table A6.<sup>103</sup> Current updaters would thus have an incentive to switch to lagged updating, i.e., to the setting considered in our baseline specification.

Table A7 reports the model moments when the equilibrium selection rule chooses instead the market clearing price that is furthest away from the previous period's price. For  $\alpha \leq 0.7$  the same phenomena occur as for the alternative selection rule considered before, i.e., volatility and risk premia increase, with the quantitative effects being now more pronounced. For  $\alpha \geq 0.8$ , the share of periods with multiple equilibria increases substantially, so that there are more often two consecutive periods with multiple equilibria. The selection rule then creates an oscillating pattern between high and low market clearing prices. This manifests itself in a reduced autocorrelation for the PD ratio, which becomes even negative for  $\alpha \geq 0.9$ . As before, current forecasters' (objectively) expected utility always falls short of that experienced by forecasters using only lagged price growth. For large values of  $\alpha$ , the utility gap widens significantly because the forecast quality of forecasters using current price information deteriorates significantly in the presence of oscillating price patterns.

Overall, we find that - in line with the postulated subjective belief structure in our baseline setting - agents will find it optimal to use only lagged price observations to update beliefs. Even if some agents would use current price information for updating beliefs,

$$Std_t^{\mathcal{P}} [r_{t+1}^s] \approx Std^{\mathcal{P}} [\varepsilon_{t+1}], \text{ which is time invariant, as claimed.}$$

<sup>102</sup>The parameters in tables A6 and A7 are those given by the estimated model from table 3 (diagonal matrix). To compute  $Corr[PD_t, E_t^{\mathcal{P}}[R_{t+1}]]$  we let  $E_t^{\mathcal{P}}[R_{t+1}]$  denote the average return expectations across agents, in line with how the return expectations are computed in the survey data.

<sup>103</sup>The table reports the unconditional expectation of discounted consumption utility using the objective distribution for consumption, as realized in equilibrium.

the model continues to produce high amounts of stock price volatility and also tends to deliver a positive correlation between the PD ratio and subjective expected returns.

	U.S. Data 1946:1-2012:1	Share of current updaters ( $\alpha$ )										
		0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
<b>E[PD]</b>	139.8	115.2	115.7	116.2	117.4	118.9	119.6	123.2	127.1	132.4	135.2	137.1
<b>Std[PD]</b>	65.2	88.2	89.3	90.5	92.9	95.9	97.2	103.2	108.7	115.8	118.8	121.7
<b>Corr[PD, PD<sub>t-1</sub>]</b>	0.98	0.98	0.98	0.98	0.98	0.98	0.97	0.97	0.96	0.96	0.96	0.94
<b>Std[r<sup>s</sup>]</b>	8.00	7.74	8.11	8.59	9.25	9.92	10.46	11.38	12.64	15.41	20.79	28.71
<i>c</i>	-0.0041	-0.0050	-0.0050	-0.0051	-0.0052	-0.0052	-0.0053	-0.0055	-0.0059	-0.0062	-0.0066	-0.0067
<i>R</i> <sup>2</sup>	0.25	0.20	0.19	0.19	0.19	0.19	0.19	0.20	0.21	0.22	0.23	0.24
<b>E[r<sup>s</sup>]</b>	1.89	1.82	1.85	1.89	1.94	2.01	2.08	2.22	2.41	2.73	3.22	3.96
<b>E[r<sup>b</sup>]</b>	0.13	0.99	0.96	0.96	0.96	0.96	0.95	0.95	0.95	0.95	0.94	1.03
<b>UBS Survey Data:</b>												
<b>Corr[PD<sub>t</sub>, E<sub>t</sub><sup>p</sup>R<sub>t+1</sub>]</b>	0.79	0.79	0.80	0.81	0.81	0.82	0.82	0.83	0.85	0.86	0.87	0.87
<b>Percent of periods w multiple equilibria:</b>		0.00	0.00	0.41	1.88	3.74	5.50	8.22	24.22	66.38	87.35	97.76
<b>Expected utility:</b>												
<b>Lagged updating</b>		-4.02	-4.02	-4.01	-4.01	-4.01	-4.00	-4.00	-3.99	-3.99	-3.97	-3.95
<b>Current updating</b>		-4.07	-4.06	-4.06	-4.05	-4.05	-4.05	-4.04	-4.04	-4.03	-4.03	-4.02

Table A6: Asset pricing moments with simultaneous belief updating  
(market clearing price closest to previous period price)

The table reports U.S. asset pricing moments (second column) using the data sources described in Appendix A.A1, see table 3 for a description of the labels used in the first column, and the asset pricing moments for the estimated model from table 3 (diagonal matrix), considering different values for the share  $\alpha$  of belief updaters using current price growth observations (columns 3 to 13). The table also reports the percentage share of periods in which multiple equilibria are encountered along the equilibrium (fifth to last line) and the unconditional expected utility of agents using current or lagged price growth observations for updating beliefs (last two lines). In periods with multiple equilibria, the equilibrium price closest to the previous market clearing price is selected.

	U.S. Data 1946:1-2012:1	Share of current updaters ( $\alpha$ )										
		0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
<b>E[PD]</b>	139.8	115.2	115.7	115.3	112.7	109.1	106.3	107.8	117.7	136.0	167.5	225.3
<b>Std[PD]</b>	65.2	88.2	89.3	89.2	84.9	78.0	71.5	75.5	92.2	111.9	124.2	177.3
<b>Corr[PD<sub>t</sub>, PD<sub>t-1</sub>]</b>	0.98	0.98	0.98	0.98	0.97	0.96	0.94	0.91	0.78	0.31	-0.51	-0.91
<b>Std[r<sup>a</sup>]</b>	8.00	7.74	8.11	8.66	9.37	10.16	11.23	14.72	27.80	64.04	186.20	362.66
$c$	-0.0041	-0.0050	-0.0050	-0.0051	-0.0055	-0.0057	-0.0056	-0.0053	-0.0055	-0.0064	-0.0094	-0.0033
$R^2$	0.25	0.20	0.19	0.19	0.18	0.16	0.13	0.11	0.11	0.15	0.28	0.07
<b>E[r<sup>a</sup>]</b>	1.89	1.82	1.85	1.90	1.96	2.02	2.10	2.46	4.89	18.70	97.08	271.19
<b>E[r<sup>b</sup>]</b>	0.13	0.99	0.96	0.96	0.95	0.95	0.94	0.93	0.86	0.48	-0.59	1.11
<b>UBS Survey Data:</b>												
<b>Corr[PD<sub>t</sub>, E<sub>t</sub><sup>p</sup>R<sub>t+1</sub>]</b>	0.79	0.79	0.80	0.81	0.84	0.85	0.87	0.88	0.91	0.90	0.92	0.96
<b>Percent of periods w multiple equilibria:</b>												
<b>Expected utility:</b>												
<b>Lagged updating</b>		-4.02	-4.02	-4.01	-4.01	-4.00	-3.99	-3.98	-3.97	-3.89	-3.39	-2.46
<b>Current updating</b>		-4.07	-4.06	-4.06	-4.06	-4.06	-4.06	-4.06	-4.06	-4.11	-4.24	-4.02

Table A7: Asset pricing moments with simultaneous belief updating  
(market clearing price furthest away from previous period price)

The table reports U.S. asset pricing moments (second column) using the data sources described in Appendix A.A1, see table 3 for a description of the labels used in the first column, and the asset pricing moments for the estimated model from table 3 (diagonal matrix), considering different values for the share  $\alpha$  of belief updaters using current price growth observations (columns 3 to 13). The table also reports the percentage share of periods in which multiple equilibria are encountered along the equilibrium (fifth to last line) and the unconditional expected utility of agents using current or lagged price growth observations for updating beliefs (last two lines). In periods with multiple equilibria, the equilibrium price furthest away from the previous market clearing price is selected.