Heterogeneous Investors and Stock Market Fluctuations*

Sebastian Hillenbrand†1 and Odhrain McCarthy‡1

[1]New York University

This Version: October 2021

Abstract

We introduce a heterogeneous agent model which features extrapolative beliefs and time-varying risk aversion. The model leads to an empirical framework which we estimate with stock prices, survey data and risk aversion measures. We find that extrapolative beliefs and risk aversion are important drivers of stock prices together explaining 86% of movements in the S&P500 index: extrapolative cash flow expectations explain 34%, extrapolative return expectations explain 23% and time-variation in risk aversion explains 29%. We also find that stock prices would vary by roughly 70% less if all investors were to hold rational beliefs. Our work highlights that investor heterogeneity and the use of survey data to measure their beliefs are key to understanding asset prices.

JEL Classification: G11,G12,G4

Keywords: Heterogeneous investors, extrapolative expectations, stock market valuation, asset pricing, survey expectations, price decomposition

*We would like to thank Jaroslav Borovička, Tim Christensen, Quirin Fleckenstein, Simon Gilchrist, Max Huber, Sydney Ludvigson, Alexi Savov, Johannes Stroebel and Nicholas Zarra for helpful comments. We also thank the CV Starr Center for Applied Economics at New York University for generously providing funding for data purchases.

†shillenb@stern.nyu.edu

‡otm210@nyu.edu

1 Introduction

What causes fluctuations in stock prices? There is still a surprising lack of consensus on what is one of the most fundamental questions in asset pricing. The only resounding consensus in the literature thus far is that prices move too much relative to what is explained by future cash flows (Shiller, 1981). This seminal insight means that if we impose rational expectations, then most of the variation in stock prices must come from time-varying discount rates (Cochrane, 2011). Accordingly, a large "risk-based" literature has assumed the existence of a representative rational investor and proposed different mechanisms to generate time-varying risk premia (e.g., Campbell and Cochrane, 1999; Bansal and Yaron, 2004; Barro, 2006, etc.). At the same time, this has spurred a behavioral literature that stresses the influence of investors' irrational beliefs on stock market fluctuations (e.g., De Bondt and Thaler, 1985; Shiller, 2015, etc.).[1]

The latter literature takes issue with two assumptions implicit in the risk-based literature – namely, that investors form expectations of cash flows rationally and that investors value assets by performing a discounted cash flow analysis, i.e. respecting the Campbell and Shiller (1988) identity when forming expectations. On the first point, researchers have shown that some investors extrapolate past fundamentals when making cash flow forecasts (La Porta, 1996; Bordalo, Gennaioli, Porta, and Shleifer, 2020). On the second point, researchers have shown that some investors rely on past price changes when they form their return expectations (Greenwood and Shleifer, 2014; Cassella and Gulen, 2018).[2]

Despite the list of seemingly contradictory explanations for stock market fluctuations, little progress has been made to reconcile them.[3] This paper aims to bridge the gap by building and testing a heterogeneous investor framework which features extrapolating beliefs, but also accounts for time-variation in investors' risk aversion. To our knowledge, we are the first to estimate a heterogenous investor framework with aggregate stock prices, survey expectations and risk premium measures. Quantifying the root causes of stock market fluctuations, we find that all prior explanations for stock market fluctuations have merit. However, we find that the majority of price fluctuations can be ascribed to belief extrapolation.

We start by developing a stylized model that highlights how aggregate stock prices are formed when various investors are present in the stock market. The model contains rational investors, investors

[1] Although not the focus of this paper, there is also a large behavioral literature that focuses on investors' preferences (e.g., Barberis, Huang, and Santos, 2001, etc.)

[2] The recent Bitcoin, Gamestop and AMC sagas only add to the weight of evidence that these extrapolative investors exist.

[3] There are some models that include extrapolating investors in addition to rational investors (e.g., De Long, Shleifer, Summers, and Waldmann, 1990; Cutler, Poterba, and Summers, 1990; Barberis, Shleifer, and Vishny, 2005; Hirshleifer, Li, and Yu, 2015; Barberis, Greenwood, Jin, and Shleifer, 2015, 2018, etc.). The focus of these paper is often to show that extrapolating investors have an effect on stock prices, i.e. that there are limits to arbitrage.

extrapolating cash flows and investors extrapolating prices. Crucially, we assume that the two former types of investors are "fundamental investors" in that they perform a fundamental analysis when valuing assets by discounting projected cash flows back at their required return. The difference between their valuation and the current stock price determines their perceived attractiveness of stock market investments. Price extrapolators behave differently – they make investments purely on the basis of their short-term return expectations (which are exogenously given in our model) – without considering future cash flows. For example, price extrapolators might form their return expectations by looking only at past price changes (Greenwood and Shleifer, 2014; Cassella and Gulen, 2018). In the model, all investors maximize expected utility, share the same beliefs about the stock market volatility and are subject to the same risk aversion. To ensure that all investors matter for stock prices, we assume a short investment horizon for all investors.[4] In equilibrium, the aggregate stock price is positively related to fundamental investors' cash flow expectations, negatively related to investors' risk aversion and positively related to price extrapolators' return expectations. Importantly, because price extrapolating investors have a "stubborn" demand that does not depend on the current price – but only on their exogenously given return expectation – a greater presence of price extrapolating investors increases the sensitivity of stock prices to movements in risk aversion. In particular, each remaining fundamental investor has to absorb a greater amount of the risky asset when price extrapolators' demand changes in order to clear the market.

To isolate the effect of irrational beliefs on stock prices, we also construct a measure of stock market mispricing defined as the deviation of the equilibrium stock price from the "rational price", i.e. the price that would prevail if only rational investors were active in the market. In the model, mispricing depends positively on the wedge between cash flow extrapolators' and rational investors' expected cash flows ("cash flow wedge") and the wedge between price extrapolators' expected returns and rationally required returns ("return wedge"), which we define as the return compensation that rational investors would demand in a world with only rational investors.

Our stylized model motivates an empirical framework. Accordingly, we collect data to obtain empirical proxies that are closely tied to the theoretical components. For the construction of our empirical proxies we rely on the prior literature which has provided evidence for the presence of these extrapolating investors. To proxy for cash flow extrapolators' expected cash flows, we use equity analysts'

[4] Of course, there are potentially lots of reasons for limits to arbitrage, in addition to the fact that it is almost impossible to "arbitrage" the aggregate stock market. We do not want to take a strong stance here, but view a short-term investment horizon as a simple way to impose limits to arbitrage (De Long, Shleifer, Summers, and Waldmann, 1990). However, we want to note that the empirical results presented below, i.e. that a large fraction of stock price movements can be explained by extrapolative beliefs, are strong evidence that limits to arbitrage exist.

one-year ahead and long-term earnings growth ("LTG") forecasts from I/B/E/S.[5] To obtain a measure for price extrapolators' expected short-term returns, we use the same three primary survey sources of investor expectations used in Greenwood and Shleifer (2014) and follow their methodology to form a composite measure.[6] To proxy for rational investors' expected cash flows, we use predictive regressions for future realized earnings growth.[7] Finally, we draw on the literature to obtain a range of proxies for time-varying risk aversion with our preferred measure being the surplus consumption ratio (Campbell and Cochrane, 1999; Cochrane, 2017).

Before estimating our empirical framework, we first provide both support for our model and suggestive evidence that these heterogeneous investors exist in the market. To do so, we gather novel survey evidence from Bank of America's Fund Manager Survey that asks investors about their current perception of the stock market valuation. We complement this with similar survey data from the Yale SOM International Center of Finance on institutional and on wealthy individual investors. From these surveys, we obtain the fraction of surveyed investors that perceive the stock market to be overvalued. Across all three surveys, this fraction is tightly positively linked to the price-to-earnings ratio of the U.S. stock market. This suggests that the surveyed investors are fairly rational, or in order words, that there is a large fraction of rational investors among them. Theoretically, our model predicts that the fraction of rational investors that perceive the market to be overvalued should be positively related to the cash flow and the return wedge. This is indeed what we find in the data. In our collected surveys, more surveyed investors perceive the market to be overvalued exactly when other investors have overly optimistic cash flow growth and excessively high short-term return expectations.

Next, we estimate our empirical framework using aggregate stock prices. We find that stock prices are highly sensitive to beliefs as measured by survey data. In particular, a 1 percentage point increase in cash flow extrapolators long-term earnings growth expectations raises stock prices by 5.1%. Similarly, if price extrapolators have a 1 percentage point higher return expectation then stock prices increase by 4.5%. As we show in the paper, these sensitivities are economically sensible. Perhaps unsurprisingly, we also find a strong positive relationship between the surplus consumption ratio and stock prices.

[5]We show in Appendix E that these one-year and long-term growth expectations are indeed extrapolative. For LTG forecasts, our results are consistent with Bordalo, Gennaioli, Porta, and Shleifer (2020) who show overreaction in these forecasts. In contrast to Bordalo, Gennaioli, Porta, and Shleifer (2020), we also find the one-year growth forecasts are extrapolative. The reason for this difference lies in the use of a different earnings measures, as further explained in Appendix Section D.5 and E. For further studies that have used I/B/E/S earnings forecast on the aggregate level, see e.g. Chen, Da, and Zhao (2013); Nagel and Xu (2019); De La O and Myers (2021).

[6]As shown in Greenwood and Shleifer (2014) these survey return expectations are extrapolative and heavily load on past returns. We find the same when including the recent sample period.

[7]In future versions, we want to build on recent work (e.g., Bianchi, Ludvigson, and Ma, 2020; de Silva and Thesmar, 2021; van Binsbergen, Han, and Lopez-Lira, 2020, etc.) and use machine learnings methods to construct unbiased out-of-sample one-year and long-term earnings growth forecasts.

Thus, when consumption relative to habit is low – and risk aversion is high – then stock prices are also low. Together, our empirical proxies have an R-squared of more than 80% for explaining S&P500 price movements.

Employing a variance decomposition, we find that extrapolative expectations can together explain 57% of the price movements in the S&P500 index: the price impact of price extrapolators explains 23% of price movements, whilst extrapolative cash flow expectations explain 34%. In particular, the latter explanatory power comes almost exclusively from expectations about long-term cash flows, while expectations about short-term cash flows have an almost negligible impact on prices. In addition, time-variation in consumption-based risk aversion – as proxied for by the surplus consumption ratio – explains 29% of price movements. By contrast, rational investors' expectations of future cash flows explain none of the variation in line with the prior literature (Shiller, 1981; Cochrane, 2011). As an alternative to the surplus consumption ratio, we also use various other proxies for time-varying risk compensation, such as the health of financial intermediaries, but we find that these proxies do not explain much of the stock price variation echoing the findings in Haddad and Muir (2021).

Based on our empirical framework, we find that roughly 70% of aggregate stock market fluctuations can be ascribed as irrational or as "mispricing". To put this result differently, stock prices would vary by 70% less if all investors were rational. This highlights that investors with extrapolating beliefs are a powerful force in the stock market. Our model provides intuition for why this mispricing is so large. Firstly, rational investors' expectations of future cashflows do not vary much and so most of the variation in cash flow extrapolators' expectations of future cash flows (which explains 34% of the variation in prices) can be attributed to mispricing. Second, price extrapolators investors' stock market demand does not necessarily increase (decrease) in response to a decrease (increase) in today's stock price and so having a large fraction of these investors in the market increases the sensitivity of prices to risk aversion. For example, in the case of an increase in risk aversion which causes stock prices to fall, price extrapolators become natural sellers in the stock market. To make the remaining fraction of the fundamental investors willing to absorb this quantity, the price then has to drop more than if there were no price extrapolators in the economy amplifying the sensitivity of stock prices to risk aversion.

How do our results compare to the prior literature, especially the literature who has used aggregate earnings forecast in the Campbell and Shiller (1988) framework? In stark contrast to De La O and Myers (2021), we find that most of the explanatory power of extrapolative cash flow expectations comes from expected long-term earnings growth while almost none comes from expected one-year earnings growth. We show that this contradiction arises, because a recent literature (De La O and Myers, 2021; Bordalo, Gennaioli, Porta, and Shleifer, 2020) uses a measure of the price-to-earnings ratio that is mostly driven by variation in earnings. By contrast, our measure for the price-to-earnings ratio better captures price fluctuations.[8] On this point, it is worth stressing that our work focuses on explaining price fluctuations relative to normalized earnings. This stands in contrast to recent important work by Greenwald, Lettau, and Ludvigson (2014, 2019) who focus on understanding fluctuations in the non-stationary price level.

Turning to long-term earnings expectations, our finding that these are key drivers of price movements qualitatively echo the results of Bordalo, Gennaioli, Porta, and Shleifer (2020). This is despite the different approaches taken to reach this conclusion: we estimate the sensitivity of stock prices to survey expectations of long-term earnings growth (which theoretically is a function of the fraction of investor holding these expectations and the term structure of their earning expectations), while Bordalo, Gennaioli, Porta, and Shleifer (2020) assume that all investors hold these expectations and impose the term structure persistence.

We also relate our findings to recent studies estimating the micro elasticities on the equity share response to changes in expected returns (Giglio, Maggiori, Stroebel, and Utkus, 2021) and the macro elasticities on the change in the stock market value to changes in equity flows (Gabaix and Koijen, 2021). By drawing on these prior studies, we are able to obtain an estimate for the fraction of price extrapolators in the stock market. We find this fraction to be 15.3%. In other words, only a small fraction of investors need to be price extrapolators for them to have a sizeable impact on stock prices.

Overall, we draw two main conclusions from our work. First, our results imply that it is all the more important to understand the formation of beliefs as unveiled by survey data (e.g. Malmendier and Nagel, 2011; Bhandari, Borovička, and Ho, 2019; Nagel and Xu, 2019). In particular, a number of studies document that survey respondents act in accordance with their stated beliefs (e.g. Greenwood and Shleifer, 2014; Gennaioli, Ma, and Shleifer, 2016; Giglio, Maggiori, Stroebel, and Utkus, 2021). We take this as motivation to link economic outcomes, i.e. stock market prices, to survey expectations. Under the assumption that we have an accurate framework for stock market prices and valid empirical proxies, we have recovered estimates for the actual expectations of stock market investors and shown how critical they are to understand stock market fluctuations. Second, our findings emphasize the need to take into account not only investor heterogeneity with regards to their beliefs, but also their investment decisions. For example, it is important to recognize that some investors might not value stocks according to corporate finance principles (e.g. discounted cash flow analysis) and therefore have return expectations that

[8]We do so by using a three-year moving average of actual earnings based on "Street earnings" as reported by I/B/E/S. Street earnings exclude various one-time transitory items relative to Compustat's 'net earnings excluding extraordinary items' which is used in both De La O and Myers (2021) and Bordalo, Gennaioli, Porta, and Shleifer (2020). Using last year's Compustat earnings to scale prices leads, for example, to a large counterintuitive rise in the price-to-earnings ratio during the financial crisis.

contradict classical theories (e.g., Greenwood and Shleifer, 2014, etc.). However, as our results illustrate, this does not mean that we should dismiss consumption-based asset pricing theories.

The remainder of this paper is organized as follows. Section 2 introduces the heterogeneous investor model which forms the basis of our empirical framework. Section 3 describes the construction of the data. Section 4 presents the empirical results including a variance decomposition of stock prices. Section 5 relates our results to the literature and discusses the magnitudes of our estimates. Section 6 concludes. Additional results are included in the Appendix.

2 Theory – Heterogeneous Investors and Stock Prices

The goal of this section is to write down a stylized model that shows how aggregate stock prices are formed when heterogeneous investors are present in the stock market. In addition to providing intuition, the model guides the empirical analyses that follow.

The key feature of our model is that it incorporates different types of investors. First, it distinguishes between fundamental investors and price extrapolators. Simply put, fundamental investors evaluate stocks based on a discounted cash flow analysis. That is, they project cash flows far into the future and discount them back at a required return to obtain a value for a stock. On the other hand, price extrapolators are short-sighted and buy stocks only because they expect prices to go up (further) in the short-run. Second, the model divides the fundamental investors into rational investors and cash flow extrapolators. While rational investors form expectations about future cash flows rationally, cash flow extrapolators can have extrapolative cash flow expectations – that is their expectations might overreact to past cash flow growth. Importantly, we do not model how exactly these beliefs are formed, but focus instead solely on how they influence stock prices.

2.1 Two-period Model

For simplicity, we start with an economy that features only two periods. In period t, each agent i gets to invest her initial wealth w_t (which is the same across agents) into the stock market and the risk-free asset. In period $t+1$, asset payoffs are realized. The risk-free asset yields a net return of zero – it pays back the initial investment – and is in perfectly elastic supply. The stock market is in fixed supply of 1 unit. Its initial price is p_t – which will be determined endogenously in equilibrium – and its stochastic payoff is given by the sum of the future price p_{t+1} and the dividend d_{t+1}.[9] There is a continuum of investors with total mass 1. We distinguish between three main types of investors in the economy: (a) price extrapolators with mass $\mu_p \in (0,1)$, (b) cash flow extrapolators with mass $\mu_c \in (0,1)$, and (c)

[9]We later endogenize the price in period $t+1$ when we go to the infinite-horizon economy.

rational investors with mass $1-\mu_p-\mu_c \in (0,1)$.

2.1.1 Fundamental Investors

We classify cash flow extrapolators and rational investors as the "fundamental investors" that form return expectation based on future payoffs and the current stock price.

Beliefs. Any fundamental investor i believes that the stock market payoff in period $t+1$ follows

$$p_{t+1}+d_{t+1} \sim \mathcal{N}(\mathbb{E}_t^i[p_{t+1}+d_{t+1}], \sigma^2).$$

For now, we assume that all rational investors form the same expectation (denoted by $\mathbb{E}_t^r$). Similarly, all cash flow extrapolators hold the same expectation (denoted by $\mathbb{E}_t^c$). Importantly, the dividend expectations of the cash flow extrapolating investors deviate from the rational expectation, i.e. $\mathbb{E}_t^c[d_{t+1}] \neq \mathbb{E}_t^r[d_{t+1}]$, while both investors agree on the expected future price.[10] Finally, all fundamental investors perceive the volatility of the total payout to be σ.

Asset demand. Agent i's utility is a constant absolute risk aversion function of her wealth at time $t+1$:

$$U^i = -e^{-\gamma_t w_{t+1}^i},$$

where γ_t is the risk aversion coefficient for all agents in period t. Since agent i believes that the stock market payoff follows a normal distribution, maximizing the expectation of utility U_i is equivalent to maximizing

$$\mathbb{E}_t^i[w_{t+1}^i] - \frac{1}{2}\gamma_t \sigma^2 (w_{t+1}^i).$$

Thus, each agent i chooses the number of shares in the stock market λ_t^i that she holds from time t to time $t+1$ to maximize

$$\max_{\lambda_t^i} w_t + \lambda_t^i \left(\mathbb{E}_t^i[p_{t+1}+d_{t+1}] - p_t\right) - \frac{1}{2}\gamma_t (\lambda_t^i)^2 \sigma^2,$$

where $\mathbb{E}_t^i[p_{t+1}+d_{t+1}] - p_t$ is the excess return of the stock market (in dollar terms).
Maximization yields a stock market demand of

$$\lambda_t^i = \frac{\mathbb{E}_t^i[p_{t+1}+d_{t+1}] - p_t}{\gamma_t \sigma^2}. \tag{1}$$

Thus, when determining their demand for the stock market, fundamental investors form expectations

[10] Note that in the infinite horizon version presented below, the price p_{t+1} depends on the dividends d_{t+k} for $k \geq 2$. Then, fundamental investors might have different expectations for the price p_{t+1} depending on their cash flow expectations.

about future dividends (and prices) and compare them to the current price. When their expectations about future dividends are high relative to the current price resulting in a high return expectation, then they want to buy lots of stocks. On the other hand, their demand decreases with higher risk aversion and higher perceived payoff volatility. Thus, fundamental investor behave like they discount cash flows.[11]

2.1.2 Price Extrapolators

Price extrapolators also maximize the same utility function, have the same risk aversion and hold the same beliefs about the payoff volatility as the fundamental investors. However, when determining their stock market demand, they do not compare the current price to next period's expected cash flows, but instead form exogenous return expectations $\mathbb{E}_t^p[r_{t+1}]$.[12] As for the other investor types, we assume that the return expectations are homogeneous across all price extrapolators. Accordingly, price extrapolating investors choose their stock market demand λ_t^p to maximize

$$\max_{\lambda_t^p} w_t + \lambda_t^p \mathbb{E}_t^p[r_{t+1}] - \frac{1}{2}\gamma_t(\lambda_t^p)^2\sigma^2.$$

The stock market demand of any price extrapolator is therefore

$$\lambda_t^p = \frac{\mathbb{E}_t^p[r_{t+1}]}{\gamma_t \sigma^2}. \tag{2}$$

Importantly, the stock market demand of the price extrapolators does not depend on the current price p_t. Thus, while the demand of the fundamental investors is negatively related to the price p_t as highlighted in equation (1), the demand of the price extrapolators might actually increase as a result of a higher price – this has been called "feedback trading" in the literature. Here, we do not need to take a stance on how these return expectations are formed – we take them as exogenously given – as we are only interested in how price extrapolators' beliefs affect prices.

2.1.3 Equilibrium Stock Price

The market clearing equation for the stock market equates supply and demand

$$1 = \int_i \lambda_t^i di = (1 - \mu_c - \mu_p)\frac{\mathbb{E}_t^r[p_{t+1} + d_{t+1}] - p_t}{\gamma_t \sigma^2} + \mu_c \frac{\mathbb{E}_t^c[p_{t+1} + d_{t+1}] - p_t}{\gamma_t \sigma^2} + \mu_p \frac{\mathbb{E}_t^p[r_{t+1}]}{\gamma_t \sigma^2}.$$

[11] To see this more explicitly, suppose that there are only fundamental investors in the economy. This means that the total demand of fundamental investors has to equal the supply of 1 (market clearing condition). Using equation (1), this implies for the stock market price $p_t = \int_i \mathbb{E}_t^i[p_{t+1} + d_{t+1}]di - \gamma_t \sigma^2$. Thus, the price is equal to the average expectation about future cash flows minus a risk adjustment that reflects the quantity of risk (σ^2) times the risk aversion (γ_t). This will also become clearer in the infinite-horizon model.

[12] Note that we again define the return to be in dollar terms, i.e. $r_{t+1} = p_{t+1} + d_{t+1} - p_t$. We use this terminology throughout the theoretical analysis.

Therefore, the stock market price at time t, p_t, equals

$$p_t = \underbrace{\frac{1-\mu_c-\mu_p}{1-\mu_p}\mathbb{E}^r_t[p_{t+1}+d_{t+1}]}_{\text{Rational investors}} + \underbrace{\frac{\mu_c}{1-\mu_p}\mathbb{E}^c_t[p_{t+1}+d_{t+1}]}_{\text{Cash flow extrapolators}} + \underbrace{\frac{\mu_p}{1-\mu_p}\mathbb{E}^p_t[r_{t+1}]}_{\text{Price extrapolators}} - \underbrace{\frac{1}{1-\mu_p}\gamma_t\sigma^2}_{\text{Required returns}} . \tag{3}$$

The price depends on four terms with different signs: It depends (i) positively on the rational investors' expectations of future cash flows (multiplied with the fraction of fundamental investors that are rational), (ii) positively on the cash flow extrapolators' expectations of future cash flows (multiplied with the fraction of fundamental investors that are extrapolating cash flows), (iii) positively on the return expectations of the price extrapolators (multiplied with the ratio of the mass of price extrapolators relative to the mass of fundamental investors), and (iii) negatively on the risk aversion (multiplied with the inverse fraction of fundamental investors in the stock market).

One crucial aspect of this equation is that the sensitivity of the stock price to risk aversion depends positively on the fraction of price extrapolating investors in the economy. Thus, if more price extrapolators are in the stock market, then prices are more sensitive to movements in risk aversion. To see why this is the case, note that, as illustrated in Equation 2, the stock market demand of price extrapolators goes down as risk aversion increases (and as their return expectations are typically positive) and this is independent of the current stock price. In that sense, their stock market demand is "stubborn". Therefore, as the fraction of price extrapolators in the economy μ_p increases, the total amount that price extrapolators want to sell rises. At the same time, this quantity has to be purchased by a smaller fraction of fundamental investors in the economy $1-\mu_p$ to clear markets. In response, fundamental investors require that the price drops more to make them willing to absorb the sell orders than in an economy where the fraction of price extrapolators is small.

2.1.4 Stock Market Mispricing

Our model also allows us to obtain an intuitive definition of stock market mispricing. Namely, we define mispricing as the deviation from the price that would prevail if there were only rational investors in the economy. This hypothetical price, which we denote by p^r_t, is equal to $\mathbb{E}^r_t[p_{t+1}+d_{t+1}]-\gamma_t\sigma^2$. In addition, in this hypothetical world, the required or expected return from a stock market investment would be $\gamma_t\sigma^2 = \mathbb{E}^r_t[p_{t+1}+d_{t+1}] - p^r_t = \mathbb{E}^r_t[r_{t+1}]$. Accordingly, we refer to the term $\gamma_t\sigma^2$ as the *rationally required return* in the following.

We further define expectation wedges as the difference between the expectations of the cash flow and price extrapolators and their rational counterpart. First, we define $\tilde{\mathbb{E}}^c_t[d_{t+1}] = \mathbb{E}^c[d_{t+1}] - \mathbb{E}^r_t[d_{t+1}]$ to be the wedge between cash flow extrapolators' and rational investors' expectations about future cash flows. Second, we define $\tilde{\mathbb{E}}_t^p[r_{t+1}] = \mathbb{E}_t^p[r_{t+1}] - \gamma_t \sigma^2$ to be the wedge between price extrapolators' expected return and the *rationally required return*.

We can then write stock market mispricing as

$$p_t - p_t^r = \frac{\mu_c}{1-\mu_p} \underbrace{\tilde{\mathbb{E}}_t^c[d_{t+1}]}_{\text{Cash flow wedge}} + \frac{\mu_p}{1-\mu_p} \underbrace{\tilde{\mathbb{E}}_t^p[r_{t+1}]}_{\text{Return wedge}} . \tag{4}$$

Thus, stock market mispricing can have two roots in our model: Cash flow extrapolation and price extrapolation.

Subjective Mispricing from the Perspective of the Rational Investors. Finally, we establish a theoretical link between the subjective mispricing of the rational investors and the beliefs of extrapolating investors. In accordance with the survey evidence we present below, we interpret subjective mispricing as the fraction of rational investors that perceive the stock market to be overvalued.

In order to do so, we expand the model by introducing belief dispersion within the group of rational investors. In particular, we assume that rational investor i's expected value of the future payoff is $\mathbb{E}_t^r[p_{t+1} + d_{t+1}] + \epsilon_i$, where $\epsilon_i \sim \mathcal{N}(0, \psi)$. Essentially, the term ϵ_i introduces belief dispersion about the mean payout without changing the perceived payout variance σ^2 of any investor i. Under some assumptions, the pricing equation (4) still holds.[13] From the perspective of rational investor i, the value of the stock market is then given by $p_t^i = p_t^r + \epsilon_i$ and the market is overvalued whenever the actual stock price p_t is higher than p_t^i.

When asking the entire group of rational investors, the fraction that will perceive the market to be overvalued at time t, $\%OV_t$, follows[14]

$$\%OV_t = \Phi\left(\frac{p_t - p_t^r}{\psi}\right) = \Phi\left(\frac{1}{\psi}\left[\frac{\mu_c}{1-\mu_p} \underbrace{\tilde{\mathbb{E}}_t^c[d_{t+1}]}_{\text{Cash flow wedge}} + \frac{\mu_p}{1-\mu_p} \underbrace{\tilde{\mathbb{E}}_t^p[r_{t+1}]}_{\text{Return wedge}} \right]\right), \tag{5}$$

where Φ is the cumulative distribution function of the standard normal distribution.

Equation (5) implies that when the cash flow and the return wedge are both zero – and therefore $p_t = p_t^r$, then exactly 50% of the rational investors will perceive the market as overvalued (as well as undervalued). On the other hand, whenever the expectations of the cash flow extrapolators and return expectations of the price extrapolators are excessively high – and the cash flow and the return

[13] Formally, this requires $\int_{i \in \text{rational}} \left(\mathbb{E}_t^i[p_{t+1} + d_{t+1}] + \epsilon^i\right) di = \mathbb{E}_t^r[p_{t+1} + d_{t+1}]$.

[14] To see this, note that the distribution of $p_t^i = p_t^r + \epsilon^i$ follows $\mathcal{N}(p_t^r, \psi)$. Then $\frac{p_t^i - p_t^r}{\psi} \sim \mathcal{N}(0,1)$. The fraction $\%OV_t$ is then equal to $\mathbb{P}r(p_t \geq p_t^i) = \mathbb{P}r(p_t^i \leq p_t) = \Phi\left(\frac{p_t^i - p_t^r}{\psi}\right)$.

wedge are positive – then more than 50% of the rational investors will say that the market is overvalued. Conversely, less than 50% of rational investors will perceive the market to be overvalued when cash flow and return expectations of the extrapolative investors are irrationally low.

2.2 Infinite-horizon Model

After building the main intuition with a two-period model, we show that our conclusions also hold in an infinite horizon economy. We focus here on the main assumptions and results – the derivations are shown in Appendix Section A.

We assume that investors live for only one period and then pass on their wealth to the next generation upon death. Newborn investors have a total mass of 1 and inherit wealth according to the same shares $(\mu_p, \mu_c, 1-\mu_p-\mu_c)$ in every period.[15] Under these assumptions, investors still maximize their one-period utility as in the static model. We additionally introduce a risk-free rate such that the risk-free asset's gross return is R. For convenience, we further assume that the total payoff volatility σ is the same in every period.

In the infinite-horizon economy, the future stock price p_{t+1} is an endogeneous object. We therefore solve for the equilibrium price by iterating forward on the fundamental investors' expectation about future prices. This again illustrates the key difference between price extrapolators and fundamental investors in our model: while price extrapolators form exogenous short-term return expectations, fundamental investors think about the forces that drive tomorrow's price. These forces are dividends, the risk aversion coefficient and the price impact of price extrapolating investors in future periods.

In order to be able to say something about these forces, we make two further simplifying assumptions. First, we assume that risk aversion and the price extrapolators' return expectation both follow AR(1) processes with autocorrelation parameter ρ. Second, we assume that fundamental investors believe that any fundamental investor born in a future period will adopt their own beliefs.[16] The equilib-

[15] These assumptions ensure that rational investors do not arbitrage away any "mispricing" by pursuing long-term buy-and-hold strategies. Furthermore, they prevent the rational investors from "taking over" the economy and ensure that irrational investors do not become more rational (possibly by learning from their errors) over time. In this regard, our model is thus similar to the noise trader risk model (De Long, Shleifer, Summers, and Waldmann (1990)).

[16] Formally, we assume $\mathbb{E}_t^i[\mathbb{E}_{t+k}^j[x_{t+k+1}]] = \mathbb{E}_t^i[x_{t+k+1}]$ for any fundamental trader j born in period $t+k$, where x_{t+k+1} can be the price or the dividend in period $t+k+1$.

rium price is then

$$p_t = \kappa_0 + \underbrace{\frac{1-\mu_c-\mu_p}{1-\mu_p}\sum_{j=1}^{\infty} R^{-j}\mathbb{E}_t^r[d_{t+j}]}_{\text{Rational investors}} + \underbrace{\frac{\mu_c}{1-\mu_p}\sum_{j=1}^{\infty} R^{-j}\mathbb{E}_t^c[d_{t+j}]}_{\text{Cash flow extrapolators}} + \underbrace{\frac{\mu_p}{1-\mu_p}\frac{1}{1-R^{-1}\rho}\mathbb{E}_t^p[r_{t+1}]}_{\text{Price extrapolators}}$$
$$- \underbrace{\frac{1}{1-\mu_p}\frac{1}{1-R^{-1}\rho}\gamma_t\sigma^2}_{\text{Required returns}}, \tag{6}$$

where κ_0 is a constant derived in the appendix. The intuition is similar to the static model: The price is determined by (i) rational investors' expectation about cash flows, (ii) cash flow extrapolators' expectations about cash flows, (iii) the (current and future) price impact of the price extrapolators, and (iv) risk aversion.

As in the two-period model, we define the cash flow expectation wedge to be $\tilde{\mathbb{E}}_t^c[d_{t+j}] = \mathbb{E}_t^c[d_{t+j}] - \mathbb{E}_t^r[d_{t+j}]$ and the return expectation wedge as $\tilde{\mathbb{E}}_t^p[r_{t+1}] = \mathbb{E}_t^p[r_{t+1}] - \gamma_t\sigma^2$. The hypothetical rational price in the infinite-horizon economy is given by $p_t^r = \kappa_1 + \sum_{j=1}^{\infty} R^{-j}\mathbb{E}_t^r[d_{t+j}] - \frac{1}{1-R^{-1}\rho}\gamma_t\sigma^2$ where κ_1 is a constant. With this at hand, we then define stock market mispricing as the difference between the actual stock market price and the rational price:

$$p_t - p_t^r = \kappa_2 + \frac{\mu_c}{1-\mu_p}\sum_{j=1}^{\infty} R^{-j} \underbrace{\tilde{\mathbb{E}}_t^c[d_{t+j}]}_{\text{Cash flow wedge}} + \frac{\mu_p}{1-\mu_p}\frac{1}{1-R^{-1}\rho} \underbrace{\tilde{\mathbb{E}}_t^p[r_{t+1}]}_{\text{Return wedge}}, \tag{7}$$

where κ_2 is again a contant. As in the two-period economy, stock market mispricing is driven by extrapolative cash flow and return expectations.

To obtain the subjective mispricing from the perspective of the rational investors, we again expand the model by introducing a belief dispersion term ϵ_i within the group of rational investors. We assume that rational investor i's expectation about future dividends follow $\sum_{j=1}^{\infty} R^{-j}\mathbb{E}_t^r[d_{t+j}] + \epsilon_i$, where $\epsilon_i \sim \mathcal{N}(0,\psi)$. The fraction of rational investors that perceive the market to be overvalued at time t, $\%OV_t$, is then given by

$$\%OV_t = \Phi\left(\frac{p_t - p_t^r}{\psi}\right) = \Phi\left(\frac{1}{\psi}\left[\kappa_2 + \frac{\mu_c}{1-\mu_p}\sum_{j=1}^{\infty} R^{-j} \underbrace{\tilde{\mathbb{E}}_t^c[d_{t+j}]}_{\text{Cash flow wedge}} + \frac{\mu_p}{1-\mu_p}\frac{1}{1-R^{-1}\rho} \underbrace{\tilde{\mathbb{E}}_t^p[r_{t+1}]}_{\text{Return wedge}}\right]\right). \tag{8}$$

We reach the same conclusion as in the two-period economy: more rational investors perceive the market to be overvalued when cash flow extrapolators have overly optimistic cash flow expectations and price extrapolators have excessively high return expectations.

3 Data and Variable Construction

The model links stock prices to heterogenous investors' expectations. In this section, we discuss the construction of the empirical proxies that we use to map our model to the data. Writing a stylized model in a simple linear form comes at a cost: the objects formulated in the model are most often not stationary in the data. Thus, we allow for some deviations when we construct our empirical proxies. On the other hand, this flexibility enables us to closely follow the prior empirical literature.

In total, we construct 5 empirical objects: (i) the price-to-earnings ratio (as a stationary measure for the price), (ii) earnings growth expectations as observed in survey data (as a measure of cash flow extrapolators' cash flow expectations), (iii) earnings growth expectations constructed from an econometrician's point of view (as a measure of rational investors' cash flow expectations), (iv) return expectations as observed in survey data (as a measure of price extrapolators' expected return) and (v) the habit factor as a proxy for risk aversion. Additionally, we collect novel survey data on investors' perception of the current stock market valuation.

The main sample is constructed on a quarterly frequency from 1987Q3 to 2019Q4. Only the survey data on stock market valuation start later in 1989Q4. Summary statistics of all variables are reported in Table 1.

3.1 Stock Market Valuation

PE ratio. We use a measure of the price-to-earnings (PE) ratio to value stocks. Specifically, we scale the S&P500 index (obtained from CRSP) by the average of the past three years of earnings, where we use the actual "Street" earnings at the S&P500 level as reported by I/B/E/S.

Appendix Figure B.1 shows that our valuation ratio is similar to two other commonly used valuation ratios – the price-to-dividend ratio and the cyclically adjusted price-to-earnings (CAPE) ratio from Robert Shiller's website. While we prefer our valuation measure (see Appendix Section D.1 for a discussion), our results are robust to using these alternative valuation ratios.

3.2 Cash Flow Expectations of Cash Flow Extrapolators

We use equity analysts' forecasts of earnings growth as reported by Thomson Reuters I/B/E/S to proxy for the expectations of the cash flow extrapolators. We show in Appendix Section E that our constructed expectation proxies are indeed extrapolative using the empirical framework in Coibion and Gorodnichenko (2015). In other words, the surveyed respondents are indeed cash flow extrapolators. For long-term earnings forecasts, our results are consistent with Bordalo, Gennaioli, Porta, and Shleifer

(2020) who show overreaction in these forecasts.[17]

We use forecasts of earnings instead of dividends for two reasons: First, dividend forecasts are only available for a much shorter time period and are sparsely populated in the I/B/E/S data. Second, the attention of the financial media and investors seems to largely lie on earnings estimates. This is consistent with the fact that firms increasingly rely on share repurchases to distribute profits to investors instead of dividends.

Cash flow extrapolators' expectations of short-term earnings growth – $\mathbb{E}_t^c[\Delta e_{t+1}]$. We obtain earnings estimates and reported actuals on the firm-level for S&P500 firms using the I/B/E/S Unadjusted Summary Statistics and the I/B/E/S Unadjusted Actuals files.[18] The Unadjusted Summary Statistics file contains the median analyst's earnings-per-share forecasts spanning different horizons including the one year-ahead and two-year ahead fiscal periods (FY1-FY2).

We proceed in two steps in order to obtain firm-level one-year earnings forecasts. First, as the forecasts are on a per-share basis, we adjust them to the company level by multiplying the forecasts with a firm's outstanding shares. Second, because fiscal-year end periods differ by company, we interpolate between the fiscal year forecasts to obtain a measure of earnings forecasts over the next year (De La O and Myers, 2021). For example, if we are currently in Q4 and a company's fiscal year ends in Q2, then the FY1 and FY2 forecasts correspond to the forecasted one year-earnings in 6 months and 18 months time, respectively. We interpolate the FY1 and FY2 measures to obtain a one-year ahead forecast.[19]

Next, we aggregate the firm-level earnings forecasts to the S&P 500 level by following De La O and Myers (2021). Importantly, despite a high coverage of S&P 500 firms (See Appendix Figure B.3), we do not have forecasts for all companies in the S&P500. To deal with this, we calculate a proxy for the S&P500 earnings by dividing the aggregate earnings expectations of those companies for which we have a 1-year forecast available by the ratio of their aggregate market capitalization to the S&P500 index level. We follow a similar process when aggregating actuals using the Unadjusted Actuals file.[20] As noted by De La O and Myers (2021) the assumption behind the normalization herein is that our sample

[17] In contrast to Bordalo, Gennaioli, Porta, and Shleifer (2020), we also find that short-term earnings forecasts are extrapolative or "overreactive" in the Coibion and Gorodnichenko (2015) framework. These different results can be explained by a different measure for realized earnings: While we use "Street" earnings as reported by I/B/E/S, Bordalo, Gennaioli, Porta, and Shleifer (2020) use GAAP earnings as reported by Compustat. We discuss these differences extensively in Appendix Section D.5.

[18] We merge the I/B/E/S data with the CRSP and Compustat data to obtain information on the number of shares outstanding at the company level (from CRSP) and on the actual earnings according to the GAAP accounting standard (from Compustat). Further detailed information on the matching of I/B/E/S data to CRSP is given in Appendix Section D.3.

[19] Concretely, in this case the 1-year ahead forecast $= \frac{FY1}{2} + \frac{FY2}{2}$.

[20] We first adjust the per-share-level actuals to the company level by multiplying them by a firm's outstanding shares. We then divide the aggregate actuals of those companies in the S&P500 for which we have non-missing actuals by the ratio of their market capitalization to the S&P500 index level.

is representative of the S&P500. In Appendix Section D.4, we show that this assumption holds quite well.

Finally, we compute short-term earnings growth expectations as the log difference between S&P500 1-year ahead forecasts earnings (on a S&P 500 per share basis) and the trailing three-year average I/B/E/S actual earnings (also on a S&P 500 per share basis).

Cash flow extrapolators' expectations of long-term earnings growth – $\mathbb{E}_t^c[LTG]$. The Unadjusted Summary Statistics file also provides the median analyst's forecasts for long-term earnings growth (LTG), defined as the expected annual increase in operating earnings over the company's next business cycle. To obtain LTG estimates at the aggregate S&P500 level we follow Bordalo, Gennaioli, Porta, and Shleifer (2020) by value-weighting firm-level LTG forecasts according to the firms' market capitalization.[21]

3.3 Rational Cash Flow Expectations

Next, we construct the earnings growth expectation that a rational investors would might have held in our sample. For this, we run predictive regressions predicting future earnings growth with contemporaneous variables. As robustness, we also report results where we use realized earnings growth as the rational benchmark following a large literature on rational expectations.

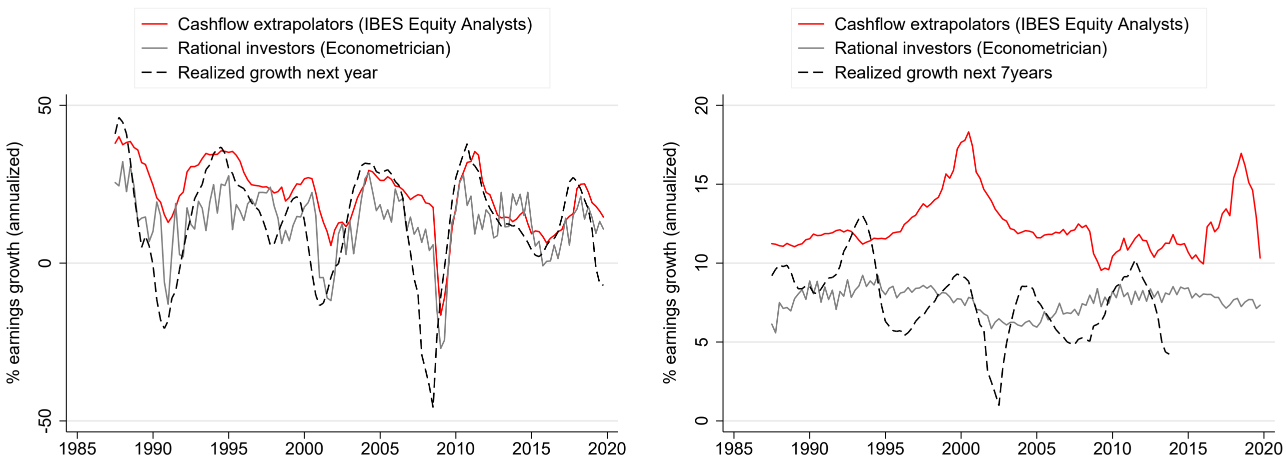
Rational expectations of short-term earnings growth – $\mathbb{E}_t^r[\Delta e_{t+1}]$. We define next year's realized earnings growth computed as the log difference between next year's I/B/E/S actuals and the trailing three-year average I/B/E/S actuals. Consistent with the price-to-earnings ratio and the short-term earnings growth expectations of the cash flow extrapolator, we scale by past three years' earnings to smooth out the volatility in realized earnings. We then run a predictive regression of next year's realized earnings growth on the log change in industrial production (series: INDPRO), real consumption (series: RCON) and real GDP (series: ROUTPUT) over the past year where the independent variables are obtained from FRED. (See Appendix Section D.6 for more information on these regression results). Finally, we use the fitted values as our measure of rational investors' expectation of short-term earnings growth.

Rational expectations of long-term earnings growth – $\mathbb{E}_t^r[LTG]$. The rational expectations of long-term earnings growth are computed analogously. First, we compute realized earnings growth over the next 7 years as the log difference between I/B/E/S actuals in seven years and the trailing three-year average I/B/E/S actuals. Second, we use the predicted values from a regression of realized earnings over the next 7 years on the log change in industrial production, real consumtpion and real GDP over the past 7 years. Figure 1 plots the earnings growth expectation of the cash flow extrapolators and the

[21] As discussed by Bordalo, Gennaioli, Porta, and Shleifer (2020) the correlation between this and an earnings-weighted measure of LTG is 95.4%. Although we prefer the value-weighted measure since stocks with high LTG often have low earnings our results are robust to either specification.

Figure 1: Cash Flow Expectations

(A) Expected 1-year earnings growth – $\mathbb{E}_t(\Delta e_{t+1})$ (B) Expected long-term earnings growth – $\mathbb{E}_t(LTG)$



Note: This figure plots fundamental investors' (i.e., extrapolators and rational investors) expected earnings growth against realized earnings growth. Panel A plots fundamental investors' expected one year earnings growth against realized growth over the next year. Panel B plots fundamental investors' expected long term earnings growth expectations against realized earnings growth over the next five years. Further details on the construction of extrapolators' and rational investors' expected earnings growth are given in Section 3.2 and 3.3, respectively. The period of analysis is 1987:Q3 to 2019:Q4.

rational investors, respectively. As can be seen in Panel A, short-term earnings forecasts of I/B/E/S equity analysts – our proxy group for the cash flow extrapolators – are quite accurate despite exhibiting slight overreaction in the Coibion and Gorodnichenko (2015) framework. However, this conclusion cannot be drawn when it comes to long-term earnings forecasts as shown in Panel B. In particular, analysts' expectations of long-term earnings growth are often high when the rational forecast and future realized earnings growth are low and vice versa. This might come from the fact that long-term earnings expectations are very extrapolative (see Appendix Section E for more discussion on this).

3.4 Return Expectations of Price Extrapolators

To form a measure of price extrapolators' expected returns we follow Greenwood and Shleifer (2014) and use the same three primary survey sources used in that paper - namely (i) the American Association of Individual Investors, (ii) Gallup and (iii) Investors' Intelligence newsletter expectations.

Greenwood and Shleifer (2014) show that these surveyed investors' expectations are extrapolative: they are high whenever past returns have been high and current stock prices are high. In Appendix Section E, we extend the Greenwood and Shleifer (2014) analysis and show that this conclusion also holds when including the recent time period.[22] Thus, the surveyed investors indeed behave like price extrapolators.

[22]The data used in Greenwood and Shleifer (2014) ends in 2011.

American Association of Individual Investors. We obtain the AAII Survey from the American Association of Individual Investors website.[23] This survey measures the percentage of individual investors – the survey is administered to members of the American Association of Individual Investors – who are bullish, neutral, or bearish on the stock market for the next six months. The survey is conducted on a weekly basis since July 1987. Following Greenwood and Shleifer (2014) we compute AAII sentiment as the bullish minus bearish spread. As most of our other data is available on a quarterly average, we take a quarterly average of the data.

Investors' Intelligence newsletter expectations. We purchase the Investors Intelligence newsletter expectations from the Investors Intelligence website which is available on a weekly basis since 1969.[24] To form this newsletter expectations data, the survey editors analyze the outlook of over 120 financial market newsletters and label each newsletter as having "bullish", "bearish" or "neutral" forecasts for stock returns over the near term. Following Greenwood and Shleifer (2014), we compute Investors' Intelligence sentiment as the bullish-bearish spread and use the quarterly average of this data over 1987:Q3-2019:Q4.

Gallup / Wells Fargo. In this survey, Gallup asks participants whether they are "very optimistic", "optimistic", "neutral", "pessimistic" or "very pessimistic" about stock returns over the next year. We follow Greenwood and Shleifer (2014) in computing sentiment as the bullish - bearish spread where bullish (bearish) corresponds to the percentage of responds who are very optimistic or optimistic (pessimistic or very pessimistic) about stock returns over the next year, respectively. The Gallup bullish-bearish spread is available between 1996:Q4-2007:Q3[25] and we extend it to 2019:Q4 by regressing it on the Wells Fargo / Gallup Investor Optimism Index (for which we have data over the entire period) and using the predicted values from this regression for the period 2007:Q4–2019:Q4.[26] We obtain the Wells Fargo / Gallup Investor Optimism Index from two sources. The first is Roper iPoll from which we gather end-of-quarter data for the period 1996:Q4-2007:Q3.[27] The second comes from analyzing Business Wire news releases to obtain quarterly data on the index for the periods 1996:Q4–2019:Q4.[28]

Panel A of Figure 2 plots the three surveys side-by-side. As can be seen, all surveys co-vary posi-

[23] https://www.aaii.com/sentimentsurvey/sent_results.

[24] https://www.investorsintelligence.com/x/us_advisors_sentiment.html

[25] We obtain this data from Roper iPoll.

[26] See Table C.4 for regression results.

[27] Note that the survey was formerly called the "UBS Index of Investor Optimism" during this period and was conducted jointly by Gallup and UBS.

[28] The survey was discontinued between November 2009 through February 2011. Additionally, it appears the survey was not conducted throughout 2008:Q3. During these periods, we replace any missing quarterly value with the last reported non-missing data point.

tively (and significantly so).[29]

Price extrapolators' return expectation – $\mathbb{E}_t^p[r_{t+1}]$. Ultimately, we want a measure for expected returns whilst the three surveys are measures of investors sentiment. To translate them into expected return, we follow Greenwood and Shleifer (2014) by taking advantage of the fact, that between 1998 and 2007, Gallup also asked survey participants about the return they expected to achieve on their portfolio over the next 12 months.[30] We transform our sentiment indices to a composite measure for the return expectations of the price extrapolators via three steps. First, we standardize all three surveys and compute the first principal component (which explains 70.7% of their variation) to form a measure of price exptrapolators expectations index. Second, we use the fitted values from a regression of the Gallup survey on expected portfolio returns on price extrapolators' expectations index. This gives a measure of survey participants' expected returns over the next year between 1996:Q3 and 2019:Q4. Finally, we extend the series back to 1987:Q3 by using fitted value from a regression of the portfolio return survey onto the standardized values AAII Investor Sentiment and the Investors' Intelligence Sentiment, which are both available since 1987:Q3.

3.5 Risk Aversion

Surplus consumption ratio – Habit. The literature has proposed many measures to quantify time-varying risk premium under rational expectations. Our preferred measure is the surplus consumption ratio proposed by Campbell and Cochrane (1999). We follow Cochrane (2017) in constructing a simplified version of the surplus consumption ratio defined as the log difference between log total real per capita consumption C_t and its moving average (the "habit") $X_t = \phi X_{t-1} + (1-\phi)C_t$, where $\phi = 0.9$.[31] The data on personal consumption expenditure (series: "PCE"), on inflation (series: "PCEPI") and on the total population (series: "POP") are obtained from FRED.

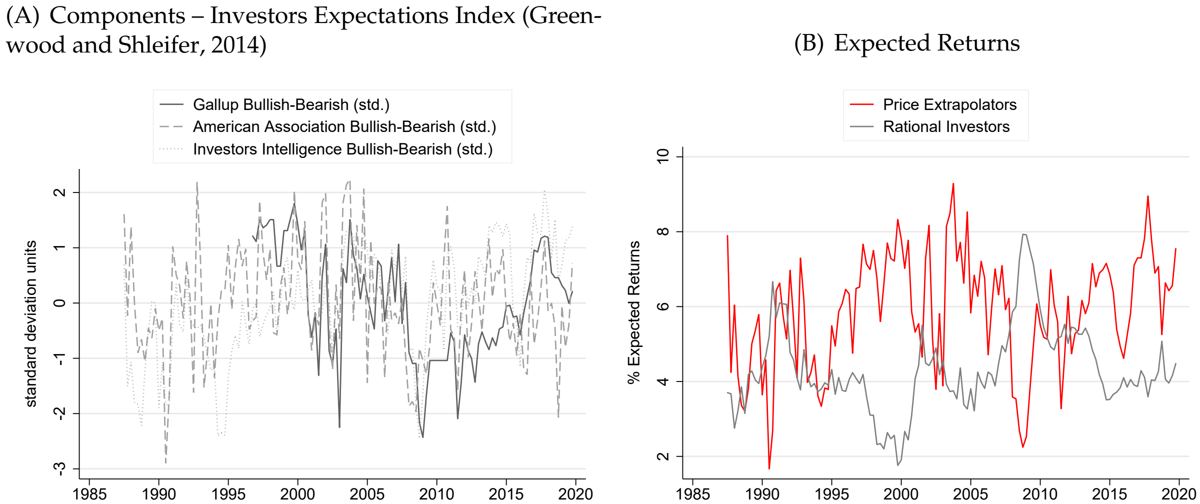
Rationally required return – $\mathbb{E}_t^r[r_{t+1}]$. While we use the consumption surplus ratio directly in our empirical analysis, we also translate this into rationally required returns via two steps. First, we regress future realized annual excess returns on the surplus consumption ratio (controlling for extrapolators' expectations to account for mispricing). From this we obtain the predicted values in future returns due to movements in surplus consumption to form a measure for rational investors' required risk-premia (See Appendix Table C.7 for the regression results). Second, we add back the 1-year T-Bill rate thereby obtaining a measure of rationally required 1yr returns.

[29] Refer to Table C.2 which gives the pairwise correlation coefficients between the three surveys.

[30] We collect this data also from Roper iPoll.

[31] We use monthly data starting in 1959 to construct the surplus consumption ratio. Thus, the initialization $X_0 = C_0$ does not affect our analysis.

Figure 2: Return Expectations



Note: Panel A plots the standardized value of the Gallup Bullish-Bearish spread, the American Association Bullish-Bearish spread, and the Investors Intelligence Bullish-Bearish spread. All of these surveys are used to form price extrapolators' expected returns following Greenwood and Shleifer (2014). Panel B plots price extrapolators' expected returns and rationally required returns. To form price extrapolators' expected one year returns we use fitted values from a regression of Gallup's Survey Investors % return expectations on their portfolio in the next twelve months on the price extrapolators' expectations index (i.e., the first prinicpal component of our optimism indices in Panel A). To form rationally required returns we use fitted values from a regression of future realized excess returns on the market on the surplus consumption ratio (controlling for extrapolators' expectations) to form a measure of rationally required returns. Further details on the construction are given in Sections 3.4 and 3.5, respectively.

In Panel B of Figure 2 we compare the price extrapolators' one-year return expectations to the rationally required 1yr returns implied by the surplus consumption ratio and find that the two series co-vary negatively with a -0.31 correlation coefficient. During the the dot-com boom in the late 1990s and the credit boom in the early 2000s, price extrapolators' return expectation were at near all time-highs before decreasing markedly during both the dot-com bust and financial crisis. Conversely, rational investors' required returns were much lower during the boom periods and increase substantially in the ensuing busts. These results show that the surveyed investors behave indeed like price extrapolators in that they are optimistic (pessimistic) when past have returns have been high (low) in contrast to rational investors' required returns which are counter-cyclical.

3.6 Long-term Real Rate

We also include a proxy for the long-term real rate in our empirical analysis to control for the secular decline in long-term real interest rates that occurred during our sample. In theory, lower real risk-free discount rates should lead to a higher stock market valuation. Our proxy for long-term real rates is the 10-year nominal Treasury yield (obtained from FRED, series: "DGS10") minus survey forecasts for long-term inflation. Specifically, we use the average of the 10-year ahead inflation forecasts from the Bluechip

Survey, the Livingston Survey and the Survey of Professional Forecasters. The data is available on the website of the Federal Reserve Bank of Philadelphia.[32]

3.7 Surveys on the Valuation of the Stock Market

We obtain data on three different surveys that ask investors about their perceived valuation of the U.S. stock market.

Bank of America Fund Manager Survey. Bank of America (BoFA) has surveyed fund managers on a monthly basis on the valuation of stock markets since 1997Q4. They report the fraction of fund managers who believe the stock market to be overvalued. We obtain this data by digitizing the graphical survey results reported in public newspaper articles.[33] This digitization approach (see Appendix Section D.2 for more information) allows us to obtain a fairly accurate proxy of the underlying data, something that we verify by comparing our obtained numbers to numerical results reported in newspaper articles.[34] As we run our analyses on a quarterly frequency, any approximation error from the digitization procedure is unlikely to impact our results. It is worth noting that investors surveyed by BoFA are likely to be sophisticated and important given the sheer size of assets they manage. According to the November 2020 BoFA Fund Manager Survey, the 216 investors surveyed during that month cumulatively managed $583bn in assets.[35]

Yale U.S. Valuation Confidence Indices. Since 1989 the Yale SOM International Center of Finance has surveyed (through third-party survey companies) institutional and wealthy individual investors on the valuation of U.S. stock market and publicly reported the survey results separately for both sets of investors.[36] In particular, this survey asks investors whether *"Stock Prices in the United States, when compared with measures of true fundamental value or sensible investment value, are 1. "Too Low", 2. "Too High", 3. "About Right", 4. "Don't Know"*. The fraction of people who report 2 (*too high*) as a percentage of those who choose either 1, 2, or 3 is the fraction of survey respondents that perceive the stock market to be overvalued. The survey results for the institutional investors are available semi-annually between 1989Q4 and 2001Q2 and quarterly thereafter. The survey results for the individual investors are

[32] https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/inflation-forecasts.

[33] The specific article can be found at: https://www.bloomberg.com/news/articles/2020-06-16/wall-street-is-far-from-dangerously-bullish-bofa-poll-shows

[34] For example, per our digitized data a gross 61.3% and 77.2% of survey investors perceive the market to be overvalued on March 2015 and June 2020, respectively. This corresponds closely to the 61.3% and 78.0% numbers reported in the public media for March 2015 and June 2020, respectively. See https://insurancenewsnet.com/oarticle/BofA-Merrill-Lynch-Fund-Manager-Survey-Finds-Concerns-of-Overvaluation-in-Both-E-a-612750 for the March 2015 numbers.

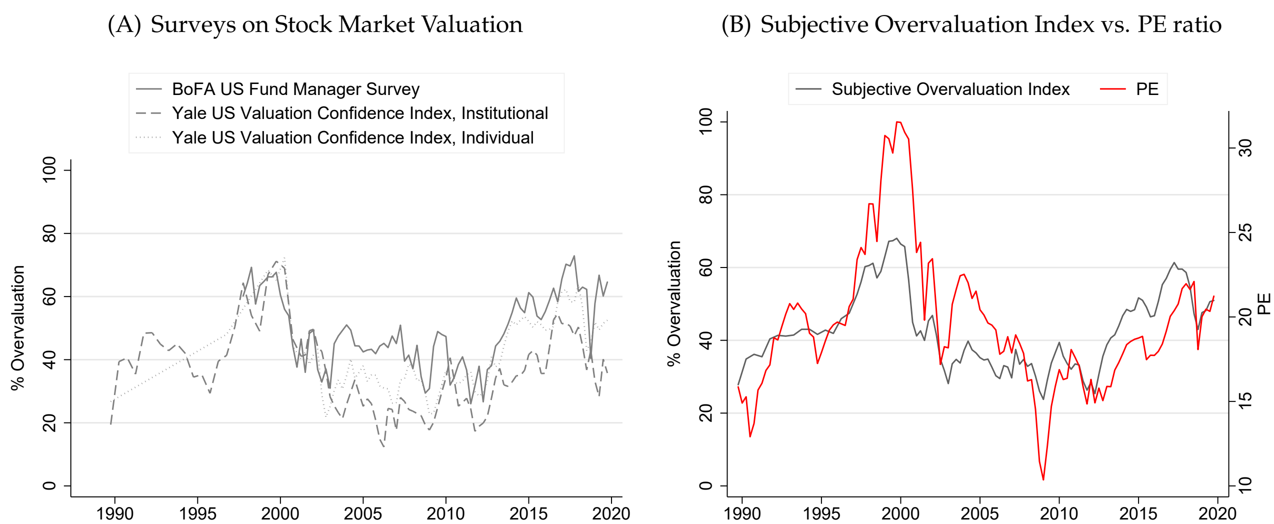
[35] https://www.bloombergquint.com/markets/most-bullish-fund-manager-survey-of-2020-has-this-message

[36] https://som.yale.edu/faculty-research-centers/centers-initiatives/international-center-for-finance/data/stock-market-confidence-indices/united-states-stock-market-confidence-indices

available in 1989Q4 and 1996Q4, semi-annually between 1999Q2 and 2001Q2 and quarterly thereafter. To obtain a quarterly series for both institutional and individuals investors from 1989Q4 onwards, we impute missing values by linear interpolation.

Panel A of Figure 3 plots the fraction of surveyed investors that perceive the stock market as overvalued for all three individual surveys. All three survey measures are significantly positively correlated with similar means and variances. During the dot-com boom, more than 60% of survey respondents reported that the market was overvalued in all three surveys. Conversely, during the depths of the financial crises, the fraction of surveyed investors that perceived the market to be overvalued dropped to below 40% in all surveys, i.e. more investors thought that the stock market was undervalued.

Figure 3: Subjective Stock Market Valuation



Note: The figure show the surveys on the valuation of the stock market. The surveys underlying Panel A are the Bank of America Fund Manager Survey, the Yale Valuation Confidence Index for institutional investors, and the Yale Valuation Confidence Index for wealth individual investors. These surveys ask participants whether the US stock market is valued too high (or too low) relative to a fair value and Panel A shows the fraction of survey respondents that perceive the market as being overvalued. Panel B compares the Subjective Overvaluation Index – we use the average of the fraction of survey respondents that perceive the market to be overvalued across the three surveys – to an objective measure of the stock market valuation – namely, the price to earnings ratio.

Subjective Overvaluation Index. To construct a measure of subjective stock market mispricing, we take the mean of all three surveys, which results in a quarterly measure of overvaluation from 1997Q4 onwards – the starting quarter of the BoFA Fund Manager Survey. We then extend the measure back to 1989Q4 by regressing it on the mean of the Yale U.S. Valuation Confidence Index for institutional investors and the Yale U.S. Valuation Confidence Index for individual investors and replacing any missing values with these fitted values.

Panel B of Figure 3 plots our Subjective Overvaluation Index against the stock market valuation –

as measured by the PE ratio. We observe a tight link between the two variables: Whenever the PE ratio is high (low), more (less) surveyed investors perceive the market to be overvalued. This correlation indicates that the majority of the surveyed investors behave fairly rational. This notion is also later confirmed when we relate the Subjective Overvaluation Index with the survey expectations of the investors that hold extrapolative expectations.

Table 1: Summary Statistics

	mean	sd	min	max	count
Subjective Overvaluation Index	0.43	0.11	0.24	0.68	121
BoFA US Fund Manager Survey	49.30	11.59	26.04	72.89	89
Yale US Valuation Confidence Index, Institutional	37.65	12.79	12.33	71.21	121
Yale US Valuation Confidence Index, Individual	42.14	11.82	21.77	72.32	121
PE	19.38	4.12	10.36	31.54	130
log(PE)	2.94	0.20	2.34	3.45	130
Cashflow extrapolators' expected 1yr growth – $\mathbb{E}_t^c[\Delta e_{t+1}]$	21.74	9.82	-16.48	40.06	130
Rational investors' expected 1yr growth – $\mathbb{E}_t^r[\Delta e_{t+1}]$	12.87	10.26	-27.01	32.15	130
Realized growth next year	12.87	17.35	-45.79	46.19	130
Cashflow extrapolators' expected LTG – $\mathbb{E}_t^c[LTG]$	12.38	1.81	9.54	18.32	130
Rational investors' expected LTG – $\mathbb{E}_t^r[LTG]$	7.63	0.80	5.58	9.23	130
Realized growth next 7years	7.59	2.27	0.97	13.00	106
Price extrapolators' expected 1yr returns – $\mathbb{E}_t^p[r_{t+1}]$	5.87	1.53	1.67	9.29	130
Gallup Bullish-Bearish	0.17	0.20	-0.32	0.54	93
Investors Intelligence Bullish-Bearish	17.01	14.80	-19.30	47.44	130
American Association Bullish-Bearish	7.75	14.87	-35.33	41.10	130
Negative surplus consumption ratio (Risk Aversion)	-1.37	0.98	-3.59	1.75	130
Rationally required 1yr returns – $\mathbb{E}_t^r[r_{t+1}]$	7.73	2.43	3.76	13.83	130
Long-term real rate	1.96	1.62	-0.82	5.43	130

Note: This table provides summary statistics of the variables used in the paper. The unit of observation is a quarter. The time period is from 1987:Q3 to 2019:Q4.

4 Empirics – Heterogeneous Investors and Stock Prices

4.1 Empirical Framework

Our stylized model links stock prices to the risk aversion and the expectations of heterogeneous investors. With our empirical proxies for these objects at hand, we want to examine this link in the data. Accordingly, we estimate the following linear empirical model

$$\begin{aligned}\log(PE_t) = \beta_0 &+ \beta^c_{\Delta e_{t+1}} \mathbb{E}_t^c\left[\Delta e_{t+1}\right] + \beta^c_{LTG} \mathbb{E}_t^c\left[LTG\right] + \beta^p_{r_{t+1}} \mathbb{E}_t^p\left[r_{t+1}\right] \\ &+ \beta^r_{\Delta e_{t+1}} \mathbb{E}_t^r\left[\Delta e_{t+1}\right] + \beta^r_{LTG} \mathbb{E}_t^r\left[LTG\right] + \beta_\gamma \text{RiskAversion}_t + \epsilon_t.\end{aligned} \tag{9}$$

where the dependent variable $log(PE_t)$ is the logarithm of the price-to-earnings ratio, $\mathbb{E}_t^c\left[\Delta e_{t+1}\right]$ are

Table 2: The Determinants of Stock Market Fluctuations

	Levels		2-year Changes	
	(1) pe	(2) pe	(3) Δ pe	(4) Δ pe
$\mathbb{E}_t^c[\Delta e_{t+1}]$	0.34*	0.06	0.47***	0.05
	(0.17)	(0.17)	(0.14)	(0.11)
$\mathbb{E}_t^c[LTG]$	6.45***	5.06***	3.84***	4.09***
	(1.43)	(1.08)	(1.43)	(0.86)
$\mathbb{E}_t^p[r_{t+1}]$	6.07***	4.50***	4.90***	3.82***
	(0.89)	(0.85)	(0.82)	(0.62)
$\mathbb{E}_t^r[\Delta e_{t+1}]$		−0.11		0.01
		(0.14)		(0.10)
$\mathbb{E}_t^r[LTG]$		1.53		2.24
		(1.70)		(1.40)
Risk Aversion (negative surplus consumption ratio)		−8.09***		−7.93***
		(1.72)		(1.56)
10y Real Rate		−0.05		−1.54
		(1.08)		(2.26)
R^2	0.77	0.84	0.68	0.81
N	130	130	122	122

Note: This table shows the results from estimating regression specification 9. The time period is from 1987:Q3 to 2019:Q4. The unit of observation is a quarter. Columns 1–2 are estimated in levels, while columns 3–4 are estimated in changes: we use the 2-year change for every variable. The dependent variable is the logarithm of the PE ratio. Further details about the construction of the variables are provided in Section 3. Newey-West standard errors are shown in parentheses. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01).

the short-term earnings growth expectations of cash flow extrapolating investors, $\mathbb{E}_t^c[LTG]$ are the long-term earnings growth expectations of the cash flow extrapolating investors, $\mathbb{E}_t^p[r_{t+1}]$ are the short-term return expectations of price extrapolating investors, $\mathbb{E}_t^r[\Delta e_{t+1}]$ and $\mathbb{E}_t^r[LTG]$ are the rational short-term and long-term earnings growth expectations, and RiskAversion_t is a measure of investors' risk aversion – for which our preferred measure is (the negative value of) the surplus consumption ratio. The expectations of cash flow and price extrapolating investors are backed out from survey data as further described in Section 3.

The results of estimating equation (9) via OLS are reported in Table 2. Column 1 includes only the extrapolative expectation measures. On their own, these variables are able to explain 77% of the variation in stock prices. In column 2, we follow the full specification and incorporate investors' risk aversion and rational earnings growth expectations. We find that cash flow extrapolators short-term earnings growth is insignificantly related with the level of the stock market. By contrast, cash flow extrapolators long-term earnings growth and price extrapolators expected returns are significantly positively to stock market levels. Concretely, a 1 percentage point increase in cash flow extrapolators long-term earnings growth and price extrapolators expected returns result in 5.1% and 4.5% increases in the PE ratio which approximately corresponds to an increase of the PE ratio by 1.01 and 0.89 units, respectively.[37] Both of these coefficients are economically meaningful, since the standard deviation of the PE ratio is 4.12 units. In accordance with most of the prior literature (e.g. Cochrane (2011)), we find that rational earnings growth cannot explain movements in the price to earnings ratio. Lastly, variation in consumption-based risk aversion – as proxied for by the habit factor – has also large explanatory power for the PE ratio: a one standard deviation increase in the habit factor lowers the price to earnings ratio by 1.51 units.[38] Together, all explanatory variables explain 84% of the time-series variation in the PE ratio.

Because the logarithm of the PE ratio – used as the dependent variable – is quite persistent,[39] one might be concerned about the amount of independent data points used in our estimation. To eliminate this concern we also estimate equation (9) in changes (columns 3 to 4). This yields qualitatively and quantitatively similar conclusions.

Importantly, there is no one-to-one mapping between the beta coefficients in our empirical model and the sensitivities implied by our stylized model. However, our regression coefficients, of course, have a meaning and we interpret their magnitudes in Section 5 when we also relate our results to the prior literature. For now, we treat them as reduced-form sensitivities of stock prices to the various factors.

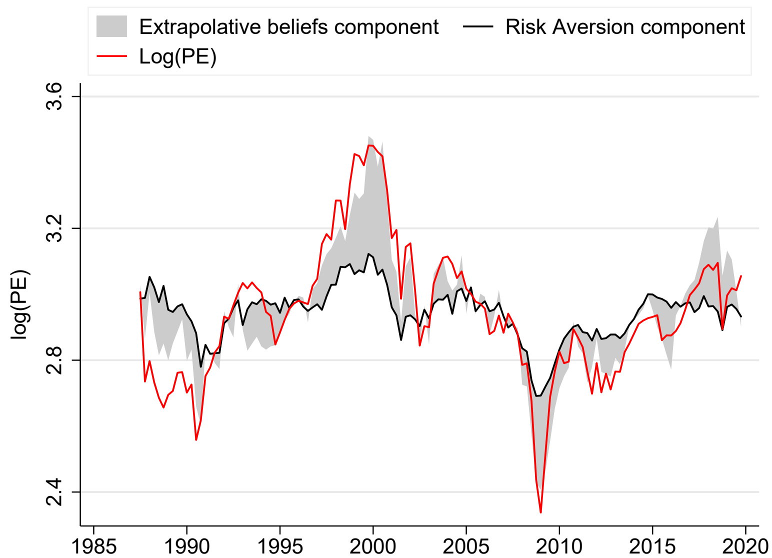
To illustrate how the various factors explain the time-series of stock prices, we plot the predictive values of the right-hand side objects of equation (9) in Figure 4. We focus thereby on movements in extrapolative beliefs and risk aversion, as rational cash flow explanations are insignificantly correlated with the price-to-earnings ratio. We obtain the stock price movement that is explained by risk aversion through $\hat{\beta}_\gamma \text{RiskAversion}_t$. Similarly, we obtain the component of stock prices that is explained by extrapolative beliefs according to $\hat{\beta}^c_{\Delta e_{t+1}} \mathbb{E}^c_t [\Delta e_{t+1}] + \hat{\beta}^c_{LTG} \mathbb{E}^c_t [LTG] + \hat{\beta}^p_{r_{t+1}} \mathbb{E}^p_t [r_{t+1}]$. The figure illustrates the high explanatory power of our variables for the stock market valuation. It also highlights how important extrapolative beliefs are for explaining stock market movements. In particular, the run-up in stock prices during the dot-com boom and the collapse in prices during the financial crisis are well explained by our proxies for investors' beliefs.

[37] Formally, exp(0.0506) = 1.052 and exp(0.053) = 1.046. Assuming the increase is relative to the mean PE ratio of 19.38 this corresponds to an increase of the PE ratio by 1.01 units and 0.89 units, respectively.

[38] Exp(-0.809) = 0.9222 which corresponds to a 7.77% decrease. Assuming the decrease is relative to the mean PE ratio of 19.38, this results in a decrease in the PE ratio of 1.51 units.

[39] The autocorrelation coefficient of log(PE) is 0.93 on the quarterly level.

Figure 4: The Determinants of Stock Market Fluctuations



Note: This figure compares the log PE ratio to the predicted values from the right-hand side objects in equation (9). The Risk Aversion predicted value corresponds to $\hat{\beta}_{\gamma}\text{RiskAversion}_t$ and is added to $\hat{\beta}_0$ to form the solid black line in the graph (i.e., the Risk Aversion component). The Extrapolative beliefs predicted value corresponds to $\hat{\beta}^c_{\Delta e_{t+1}}\mathbb{E}^c_t[\Delta e_{t+1}] + \hat{\beta}^c_{LTG}\mathbb{E}^c_t[LTG] + \hat{\beta}^p_{r_{t+1}}\mathbb{E}^p_t[r_{t+1}]$ and is added to the Risk Aversion component to form the gray shaded area in the graph (i.e., the Extrapolative beliefs component). The time period is from 1987:Q3 to 2019:Q4.

4.1.1 A Variance Decomposition of Stock Prices

Next, we want to quantify how much of stock price fluctuations can be ascribed to the different factors. We do this via a simple variance decomposition.

We compute the covariance of the right and left-hand side of equation (9) with $\log(PE_t)$ and divide both sides by the variance of $\log(PE_t)$. This results in

$$1 = \beta^c_{\Delta e_{t+1}} b^c_{\Delta e_{t+1},pe} + \beta^c_{LTG} b^c_{LTG,pe} + \beta^p_{r_{t+1}} b^p_{r_{t+1},pe} + \beta^r_{\Delta e_{t+1}} b^r_{\Delta e_{t+1},pe} + \beta^r_{LTG} b^r_{LTG,pe} + \beta_{\gamma} b_{\gamma,pe} + b_{\epsilon,pe}$$
$$\text{where } b^c_{\Delta e_{t+1},pe} = \frac{Cov\left(\mathbb{E}^c_t[\Delta e_{t+1}], \log(PE_t)\right)}{\mathbb{V}(\log(PE_t))} \quad \text{and so forth.} \tag{10}$$

Implementing this empirically is rather straightforward. First, we use the estimates $\hat{\beta}$ that we obtained by estimation (9) – as reported Table 2. Second, we separately regress each explanatory variable on the log PE ratio. Repeating this univariate regression for each explanatory variable yields the estimates $\hat{b}$. For example, to obtain an estimate of $b^c_{\Delta e_{t+1},pe}$, we regress $\mathbb{E}^c_t[\Delta e_{t+1}]$ on $\log(PE_t)$. Multiplying the estimate $\hat{b}^c_{\Delta e_{t+1},pe}$ with $\hat{\beta}^c_{\Delta e_{t+1}}$ then yields the fraction of stock price fluctuations that is explained by the short-term earnings growth expectations of cash flow extrapolating investors. Repeating this procedure for all variables yields the variance decomposition as outlined in equation (10).

Table 3: Variance Decomposition of Stock Prices

	$b = \frac{Cov(\,.\,,log(PE))}{\sigma^2(log(PE))}$	Variation explained (%) (1)	(2)
Price Decomposition			
$\mathbb{E}_t^c[\Delta e_{t+1}]$	0.10	0.58	
$\mathbb{E}_t^c[LTG]$	0.07	34.00	
$\mathbb{E}_t^p[r_{t+1}]$	0.05	22.52	
$\mathbb{E}_t^r[\Delta e_{t+1}]$	0.12	-1.29	
$\mathbb{E}_t^r[LTG]$	-0.00	-0.55	
Risk Aversion (negative surplus consumption ratio)	-0.04	29.26	
Real Rate	0.01	-0.06	
Residual		15.59	
Mispricing			
$(\mathbb{E}_t^c - \mathbb{E}_t^r)[\Delta e_{t+1}]$	-0.02		-0.11
$(\mathbb{E}_t^c - \mathbb{E}_t^r)[LTG]$	0.07		35.77
$(\mathbb{E}_t^p - \mathbb{E}_t^r)[r_{t+1}]$	0.08		33.67
$log(PE_t^r)$ + Residual			30.56

Note: This table provides a decomposition of fluctuations of the stock market as outlined in equation 10. Column b denotes the univariate regression coefficients, i.e. $\frac{Cov(\,.\,,log(PE))}{\sigma^2(log(PE))}$. The time period is from 1987:Q3 to 2019:Q4.

The results of the variance decomposition are shown in column 1 of Table 3. It shows that 34% of the variation in the stock market valuation can be explained by time-variation in cash flow extrapolators' long-term earnings growth. The return expectations of price extrapolating investors explain 23% of the variation. Thus, in total these two variables explain 57% of the fluctuations in the stock market valuation. Variation in risk aversion, as measured by the surplus consumption ratio or habit, explains another 29%. By contrast, rational earnings growth expectations contribute little to explaining stock price movements (Shiller, 1981; Cochrane, 2011).

Alternative Risk Aversion Measures. One concern might be that the surplus consumption ratio is not the appropriate risk aversion measure. To address this concern, we estimate our empirical framework using various other risk aversion measures that have been proposed by the literature including the VIX, the cay-factor (Lettau and Ludvigson (2001)), the AEM and HKM intermediary factors from (Adrian, Etula, and Muir (2014); He, Kelly, and Manela (2017)), and the Excess Bond Premium (Gilchrist and Zakrajšek (2012)). The estimation results and the variance decomposition are shown in Appendix Tables C.11 & C.12. We find that these alternative measures explain very little of actual stock price movements. In contrast, the variation explained by our extrapolative belief measures is almost unchanged when we use different risk aversion measures.

4.2 Stock Market Mispricing

Our results show that extrapolating investors are an important force in the stock market. We therefore want to understand how much of the stock price variation can be ascribed as mispricing.

Theoretically, our model defines mispricing as the deviation of the actual stock price from the price that would prevail if there were only rational investors in the stock market. Moreover, mispricing is a linear function of the cash flow wedges and the return wedge. Following this intuition, we define mispricing in our empirical framework at time t as

$$\text{Mispricing}_t = \beta^c_{\Delta e_{t+1}} \underbrace{\left(\mathbb{E}^c_t\left[\Delta e_{t+1}\right] - \mathbb{E}^r_t\left[\Delta e_{t+1}\right]\right)}_{\text{Short-term cash flow wedge}} + \beta^c_{LTG} \underbrace{\left(\mathbb{E}^c_t\left[LTG\right] - \mathbb{E}^r_t\left[LTG\right]\right)}_{\text{Long-term cash flow wedge}} + \beta^p_{r_{t+1}} \underbrace{\left(\mathbb{E}^p_t\left[r_{t+1}\right] - \mathbb{E}^r_t\left[r_{t+1}\right]\right)}_{\text{Return wedge}}, \tag{11}$$

where $\mathbb{E}^r_t\left[r_{t+1}\right]$ is the rationally required return that is obtained by regressing future excess returns on the surplus consumption ratio. The rest of the variables is defined as above. Thus, mispricing is determined by the expectation wedges multiplied with their stock price sensitivities.

Ultimately, we are interested in the fraction of stock market movements that are arise from mispricings. Using the variance decomposition approach, we can obtain this fraction by taking the covariance of the mispricing term with $\log(PE_t)$ and divide by the variance of $\log(PE_t)$, this yields the fraction of the total stock price variance that is explained by mispricing, %Mispricing,

$$\%\text{Mispricing} = \beta^c_{\Delta e_{t+1}}\left(b^c_{\Delta e_{t+1},pe} - b^r_{\Delta e_{t+1},pe}\right) + \beta^c_{LTG}\left(b^c_{LTG,pe} - b^r_{\Delta LTG,pe}\right) + \beta^p_{r_{t+1}}\left(b^p_{r_{t+1},pe} - b^r_{r_{t+1},pe}\right). \tag{12}$$

Column 2 of Table 3 reports estimates for this decomposition. Based on our estimates, stock market mispricing explains around 70% of all stock market fluctuations. Thus, if we abstract from price movements that are caused by these irrational beliefs, then stock prices would vary by roughly 70% less. Why is this number higher than the fraction of the stock prices variance that is explained by extrapolative beliefs, i.e. 57%? Here, we can again use the intuition from the model. As our estimates imply a non-zero share of price extrapolating investors in the economy, stock prices exhibit a higher sensitivity to movements in risk aversion movements than in a rational world.

4.2.1 Subjective Mispricing as Observed in Survey Data

Lastly, we make use of our novel survey data that asks investors about their perception of the current stock market valuation. As we have shown in Section 3, the survey reponses, i.e. the fraction of surveyed investors that perceive the stock market to be overvalued, are highly positively related with the price-to-earnings ratio, as one would expect from a rational investors. We therefore conjecture that there is a large fraction of rational people among the survey respondents.

This allows us to test another prediction of our stylized model. Theoretically, the model establishes a link between the percentage of rational traders who think the market is overvalued and the belief extrapolation in form of the cash flow and the return wedge. In the data, we test with the following specification

$$\text{Subjective Overvaluation Index}_t = \alpha_0 + \alpha^c_{\Delta e_{t+1}} \underbrace{\left(\mathbb{E}^c_t\left[\Delta e_{t+1}\right] - \mathbb{E}^r_t\left[\Delta e_{t+1}\right]\right)}_{\text{Short-term cash flow wedge}} + \alpha^c_{LTG} \underbrace{\left(\mathbb{E}^c_t\left[LTG\right] - \mathbb{E}^r_t\left[LTG\right]\right)}_{\text{Long-term cash flow wedge}} + \alpha^p_{r_{t+1}} \underbrace{\left(\mathbb{E}^p_t\left[r_{t+1}\right] - \mathbb{E}^r_t\left[r_{t+1}\right]\right)}_{\text{Return wedge}} + \omega_t, \tag{13}$$

where the dependent variable Subjective Overvaluation Index$_t$ corresponds to the percentage of investors who believe the market is overvalued averaged across the three valuation surveys as further explained in Section 3.

Table 4: Subjective Mispricing and Extrapolative Beliefs

	Dependent variable: Subjective Overvaluation Index							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{E}^c_t[\Delta e_{t+1}]$	0.04			−0.08				
	(0.18)			(0.16)				
$\mathbb{E}^c_t[LTG]$		0.58***		0.47***				
		(0.13)		(0.14)				
$\mathbb{E}^p_t[r_{t+1}]$			0.50***	0.33**				
			(0.18)	(0.14)				
$(\mathbb{E}^c_t - \mathbb{E}^r_t)[\Delta e_{t+1}]$					−0.19**			−0.13
					(0.08)			(0.08)
$(\mathbb{E}^c_t - \mathbb{E}^r_t)[LTG]$						0.42***		0.38**
						(0.15)		(0.16)
$(\mathbb{E}^p_t - \mathbb{E}^r_t)[r_{t+1}]$							0.39**	0.25
							(0.15)	(0.16)
R^2	0.00	0.35	0.25	0.44	0.03	0.18	0.13	0.27
N	121	121	121	121	121	121	121	121

Note: This table shows the results from estimating specification (13). The time period is from 1989:Q4 to 2019:Q4. The unit of observation is a quarter. All variables (including the dependent variable) are standardized. The construction of the variables is explained in Section 3. Newey-West standard errors are shown in parentheses. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01).

The results of the estimation are shown in Table 4. All variables are standardized to facilitate the interpretation of the coefficients. We start out in columns 1 to 4 by using only the extrapolative belief measures (without computing the wedges with respect to rational investors' expectations). In univariate regressions, we find that the extrapolators' expectations about long-term earnings growth and returns are significantly positive related to the Subjective Overvaluation Index. The regression coefficients imply that a one standard deviation increase in cash flow extrapolators' expected long-term earnings growth and price extrapolators' expected returns over the next one year coincide with 0.47 and 0.33 standard deviation increases in the fraction of the survey respondents who believe the stock market is overvalued, respectively. Running a multivariate regression using all the extrapolators' expectations (column 4) leads to similar conclusions. Together, extrapolators' expectations of earnings growth and returns are able to explain around 44% the variation in the dependent variable. In columns 5 to 8, we use the wedges between the extrapolators' and rational investors' expectations as the explanatory variables. The main conclusions are similar.

To conclude, in the surveys that ask certain investors about their perception of the current stock market valuation, more investors perceive the market to be overvalued exactly when other investors have overly optimistic cash flow growth and excessively high return expectations. This is exactly as predicted in the theory, and provides strong support for the fact that there are heteregenous investors in the stock market.

5 Discussion & Relationship to Literature

Our framework enables us to perform a price-earnings ratio decomposition when different types of investors are active in the stock market. This is a novel approach, as the prior empirical literature has, by and large, assumed that there is only one representative investors. We therefore link our estimates to the prior literature in this section. At the same time, we also discuss the magnitude of the beta coefficients that we obtained from estimating our empirical model.

5.1 Rational Cash Flow Forecasts

First, we link our results to the rational expectations literature that uses mostly realized earnings or dividend growth as proxy for rational forecasts. Most prominently, Cochrane (2011) performs a variant of our price-earnings decomposition by running the regression[40]

$$\sum_{j=1}^{K} \rho^{j-1} \Delta d_{t+j} = a + b_d pd_t + \epsilon_{t+K} \tag{14}$$

[40] In the presence of a constant dividend payout ratio, a price-dividend and price-earnings ratio decomposition are equivalent. Technically, Cochrane (2011) performs a decomposition of the dividend-price ratio instead of a price-dividend ratio, but in log-terms they are the inverse of each other.

for horizon $K = 15$. In this framework the estimate $\hat{b}_d$ corresponds to the fraction of variation in the price-dividend ratio that can be explained by future realized dividend growth (the equivalent in our framework is given by the estimate $\hat{\beta}^r_{\Delta e_{t+1}} \hat{b}^r_{\Delta e_{t+1},pe} + \hat{\beta}^r_{LTG} \hat{b}^r_{LTG,pe}$). Estimating equation (14), Cochrane (2011) finds that $\hat{b}_d$ is small and statistically insignificant. This is consistent with our results, as we find the estimate $\hat{\beta}^r_{\Delta e_{t+1}} \hat{b}^r_{\Delta e_{t+1},pe} + \hat{\beta}^r_{LTG} \hat{b}^r_{LTG,pe}$ to be small and statistically insignificant.

5.2 Survey Cash Flow Forecasts

In our price-earnings decomposition, we find that earnings forecasts from equity analysts, which we name "cash flow extrapolators" as their forecasts exhibit overreaction, have large explanatory power for stock price movements. Specifically, we find that 34% of stock price fluctuations can be explained by their long-term earnings growth expectations, while close to none can be explained by their short-run growth forecasts.

How do these results relate to the recent literature that also uses I/B/E/S earnings forecasts? Contrary to our result, De La O and Myers (2021) find that one year earnings growth forecasts can explain 41.7% of the variation in the S&P500 price-earnings ratio. Relatedly, Bordalo, Gennaioli, Porta, and Shleifer (2020) use both short-term and long-term I/B/E/S earnings growth forecasts and show that they can explain 61.7% of the variation in the price-earnings ratio (roughly 1.75 times our estimate). We now turn to explaining why our results differ from these prior studies.

5.2.1 Short-Term Earnings Growth

De La O and Myers (2021) rely on the Campbell and Shiller (1988) decomposition for earnings[41]

$$pe_t = const. + \sum_{j=1}^{\infty} \rho^{j-1} \Delta e_{t+j} - \sum_{j=1}^{\infty} \rho^{j-1} r_{t+j} \tag{15}$$

where pe_t is the log price-to-earnings ratio, and Δe_{t+j} and r_{t+j} are the earnings growth and the return in year $t + j$, respectively. Motivated by this decomposition, De La O and Myers (2021) deduce the variation in the price-earnings ratio due to one-year earnings growth expectations to be

$$\frac{Cov\left(\mathbb{E}^c_t\left[\Delta e^{DM}_{t+1}\right], pe^{DM}_t\right)}{Var(pe^{DM}_t)} = 41.7\%. \tag{16}$$

where pe^{DM}_t and $\mathbb{E}^c_t\left[\Delta e^{DM}_{t+j}\right]$ denote the price-earnings ratio and forecasted earnings growth measures used in their analysis. Nevertheless, when we perform the exact same calculation with our measures (Table 3), we find that $\hat{b}_{1,pe} = \frac{Cov(E^c_t[\Delta e_{t+1}], pe_t)}{Var(pe_t)} = 10\%$.

[41] The derivation of the Campbell and Shiller (1988) decomposition is shown in Appendix Section F. Here, the dividend payout ratio, de_{t+j}, is assumed to be constant for simplicity.

The reason for the substantial difference lies in the actual earnings measures that are used in the denominators of pe_t and Δe_{t+1}. In particular, De La O and Myers (2021) use GAAP net earnings excluding extraordinary items as reported by Compustat ("GAAP earnings"), whereas we use the average of the past three years of "Street" earnings (as reported by I/B/E/S).[42] Differences in actual earnings arise because GAAP earnings contain a range of one-time transitory items (beyond extraordinary items) that are not included by equity analysts when they make their forecasts (see Appendix Section D.5 for more information on the differences between GAAP earnings and I/B/E/S actuals).

We prefer using I/B/E/S actuals over GAAP earnings for two reasons. First, to obtain analysts short-term earnings growth estimates it makes more sense to compare "Street" estimates to "Street" actuals, i.e. to compare apples to apples. Second and more importantly, GAAP earnings have become increasingly volatile in recent decades (See Appendix Figure B.4).[43] Thus, scaling by past year's GAAP earnings in order to obtain earnings growth measures or price-to-earnings ratios is problematic. Specifically, most of the variation in the price to GAAP earnings ratio is driven by the GAAP earnings (the denominator) and not the price (the numerator). This leads to counterintuitive results, as we illustrate in Figure B.5. For example, the price to GAAP earnings ratio increases dramatically during the financial crisis (as a result of a decline in earnings), while our price-to-earnings ratio and the CAPE ratio show a decline in the stock market valuation consistent with the observed fall in stock prices.

This also explains why De La O and Myers (2021) find that 41.7% in the price-to-earnings ratio can be explained by short-term earnings growth expectations: prices and earnings forecasts are scaled by volatile GAAP earnings and this mechanically induces a highly positive covariance between $\mathbb{E}_t^c\left[\Delta e_{t+1}^{DM}\right]$ and pe_t^{DM} in equation (16). We further illustrate this point in Appendix Table C.13 where we replicate the results of De La O and Myers (2021) and show how their results change when we (a) our pe ratio and/or (b) our measure of expected one-year earnings growth (i.e., $E_t^c[\Delta e_{t+1}]$).

5.2.2 Short-Term and Long-Term Earnings Growth

Bordalo, Gennaioli, Porta, and Shleifer (2020) perform a similar analysis as De La O and Myers (2021), but rely on I/B/E/S earnings growth forecasts for both the short-term and long-term. In order to obtain a proxy for the entire term structure of expected earnings growth, $\sum_{j=1}^{\infty} \rho^{j-1} E_t^c\left[\Delta e_{t+j}^{BDGS}\right]$, the authors use the earnings growth forecasts at the one and two year horizons, and assume that earnings thereafter grow at the long-term earnings growth estimate (LTG) for years $t+3$ to $t+10$, before assuming that

[42] We use approximately identical measures of one-year ahead earnings forecasts $E_t^c[e_{t+1}]$ from I/B/E/S analysts, so the numerator cannot explain the difference in results.

[43] For example, the variance of GAAP earnings reported by Shiller has increased over tenfold between 1970-1994 to 1995-2019 from 2.7% to 30.5%.

earnings growth reverts to its time-series average. Therefore, the authors construct

$$\sum_{j=1}^{\infty} \rho^{j-1} E_t^c \left[\Delta e_{t+j}^{BDGS}\right] = \mathbb{E}_t^c \left[e_{t+1}^{BDGS}\right] + \rho \mathbb{E}_t^c \left[\Delta e_{t+2}^{BDGS}\right] + 7.06 \times \mathbb{E}_t^c \left[\Delta LTG\right] + const., \tag{17}$$

which uses $\sum_{j=3}^{10} \rho^{j-1} \mathbb{E}_t^c \left[\Delta e_{t+j}^{BDGS}\right] = 7.06 \times \mathbb{E}_t^c \left[\Delta LTG\right]$ for $\rho = 0.97$.[44] While Bordalo, Gennaioli, Porta, and Shleifer (2020) make explicit assumptions about the persistence of the earnings growth forecasts, we estimate the coefficient on $\mathbb{E}_t^c \left[\Delta LTG\right]$ directly using regression (9) and find it to be $\hat{\beta}_{LTG}^c$ = 5.06.

As such, our estimation approach implies a lower loading on the I/B/E/S long-term earnings growth estimate. This may be because (i) cash flow extrapolators' expectations do not extend as far as ten-years ahead, (ii) the fraction of cash flow extrapolators in the market is less than one – the underlying assumption in Bordalo, Gennaioli, Porta, and Shleifer (2020) is that all market participants hold the same expectations as the equity analysts in the I/B/E/S data) or (iii) some combination of the aforementioned reasons.

Finally, Bordalo, Gennaioli, Porta, and Shleifer (2020) also conduct a price-earnings decomposition and find that $\frac{Cov\left(\sum_{j=1}^{10} \rho^{j-1} E_t^c \left[\Delta e_{t+j}^{BDGS}\right], pe_t^{BDGS}\right)}{Var\left(pe_t^{BDGS}\right)} = 61.7\%$. Like in De La O and Myers (2021), this result is (almost entirely) driven by the use of GAAP earnings in the denominators for pe_t^{BDGS} and $E_t^c[\Delta e_{t+1}^{BDGS}]$.[45]

5.3 The Price Impact of Price Extrapolators

Our estimates imply that price extrapolators have a strong impact on aggregate stock prices. In particular, a 1% percentage point increase in the return expectations of the price extrapolators, as measured by survey data, translates into an increase in the price-to-earnings ratio by 0.89 units.

Next, we analyse this sensivity estimate in light of the recent literature that has estimated micro elasticities on the percentage point change in equity share to a one percentage change in expected return (Giglio, Maggiori, Stroebel, and Utkus, 2021) and macro elasticities (Gabaix and Koijen, 2021; Ben-David, Li, Rossi, and Song, 2021; Li, 2021) on the dollar change in the stock market wealth to one dollar in active equity flows.

To do so, note that our elasticity for the price extrapolators from equation (9), β_3, can be re-stated as (we drop time subscripts for covenience)

$$\beta_3 = \frac{\Delta \log(PE)}{\Delta \mathbb{E}^p[\text{1yr return}]} \tag{18}$$

[44]Chen, Da, and Zhao (2013) use very similar assumptions.

[45]This can be seen more readily in Panel A of Table 2 in Bordalo, Gennaioli, Porta, and Shleifer (2020) where they show that future realized earnings can explain 53.2% of the variation in their price-to-earnings ratio measure. This is at odds with the prior literature on rational expectations as explained above.

which can be re-stated in approximate form as[46]

$$\beta^{p}_{r_{t+1}} = \frac{\%\Delta P}{\Delta \mathbb{E}^{p}[\text{1yr return}]} \tag{19}$$

To link this to the literature, first assume w.l.o.g. there is only one share outstanding in the equity market, and note that according to demand-based theories (see Gabaix and Koijen (2021)) the dollar change in the stock market price ΔP to a dollar in active equity flows of the price extrapolators is given by

$$\Delta P = \delta \times \text{Equity Flows of Price Extrapolators} \tag{20}$$

where δ is the macro elasticity. Noting that by definition P is equal to total stock market wealth, we can re-state this as

$$\%\Delta P = \frac{\Delta P}{P} = \delta \times \frac{\text{Equity Flows of Price Extrapolators}}{\text{Stock Market Wealth}} \tag{21}$$

In accordance with our model in section 2, define $\mu_r \equiv \frac{\text{Stock Market Wealth of Price Extrapolators}}{\text{Stock Market Wealth}}$ to be the fraction of stock market wealth controlled by price extrapolators then we have[47,48]

$$\%\Delta P = \delta \times \mu_r \times \frac{\Delta\text{Equity Share}_p}{\text{Equity Share}_p} \tag{22}$$

where Equity Share_p corresponds to the proportion of the price extrapolators' wealth invested in the stock market (i.e., their equity share). As such, inserting equation (22) into equation (19) we have

$$\beta^{p}_{r_{t+1}} = \delta \times \mu_r \times \frac{1}{\text{Equity Share}_p} \times \frac{\Delta\text{Equity Share}_p}{\Delta \mathbb{E}^{p}[\text{1yr ret}]} \tag{23}$$

Define $\beta^{ME}_p \equiv \frac{\Delta\text{Equity Share}_p}{\Delta \mathbb{E}^{p}[\text{1yr ret}]}$ to be the elasticity of the price extrapolators change in equity share percentage to a one percentage point change in their expected returns. Then we can re-state equation (23) in terms of μ_r as

$$\mu_r = \frac{\beta^{p}_{r_{t+1}} \times \text{Equity Share}_p}{\delta \times \beta^{ME}_p} \tag{24}$$

We can rely on the literature to obtain empirical estimates for $\{\delta, \text{Equity Share}_p, \beta^{ME}_p\}$ and use these estimates along with our estimate of $\hat{\beta}^{p}_{r_{t+1}} = 4.5$ to obtain the implied value of μ_r, i.e. the fraction of

[46] Assuming earnings are constant which can be ensured by using moving-average of earnings when scaling stock prices.

[47] We use $\frac{\text{Equity Flows of Price Extrapolators}}{\text{Stock Market Wealth}} = \frac{\text{Stock Market Walth of Price extrapolators}}{\text{Stock Market Wealth}} \times \frac{\text{Wealth of Price extrapolators}}{\text{Stock Market Wealth of Price extrapolators}} \times \frac{\text{Equity Flows of Price Extrapolators}}{\text{Wealth of Price Estrapolators}} = \mu_r \times \frac{1}{EqShare_p} \times \Delta EqShare_p$.

[48] We slightly abuse notation here, as μ_r is the total share of wealth owned by price extrapolators in the model, while here it is the total share of equity wealth owned by price extrapolators.

investors who extrapolate prices. This also allows us to assess the degree of reasonableness for our estimate $\hat{\beta}^{p}_{r_{t+1}}$.

First, recent papers have estimated the macroelasticity in the U.S. stock market, $\hat{\delta}$, to be about 5 (Gabaix and Koijen, 2021; Ben-David, Li, Rossi, and Song, 2021; Li, 2021). Second, for an estimate of price extrapolators' equity share in their portfolio we draw on one of the survey sources used to form their return herein, namely the American Association of Individual Investors (AAII), who have surveyed investors on their equity share allocation from November 1987 onwards.[49] In this data, the average equity share was 61.2%.[50]

Third, Giglio, Maggiori, Stroebel, and Utkus (2021) obtain detailed information on the beliefs and the portfolio composition of a large subset of Vanguard retail investors and provide estimates for the elasticity of the equity share to the expected return $\hat{\beta}^{ME}_{p}$. In particular, in their main analysis they find this coefficient to be 1.164.[51] We deem this estimate to be a lower bound, as Vanguard investors are likely to be more passive than the average retail investor, particularly investors who trade actively on their extrapolative beliefs. Indeed, the average investors' turnover in Giglio, Maggiori, Stroebel, and Utkus (2021) is about one third of the turnover reported in Barber and Odean (2000) who obtain portfolio-level data on a large number of retail brokerage accounts. Thus, if we scale the baseline estimate by 3 we obtain an estimated elasticity very close to the elasticity of 3.6 that Giglio, Maggiori, Stroebel, and Utkus (2021) report for their subset of active investors. Accordingly, we use $\hat{\beta}^{ME}_{p} = 3.6$.

Using the aformentioned estimates for $\{\beta_3, \delta, \text{Equity Share}_p, \beta^{ME}_p\}$ we find that the implied fraction of price extrapolators in the market is $\mu_r = 15.3\%$. We conclude by noting that the price impact coefficient, β_3, seems to be not unreasonable.

5.4 Measurement Error and Bias in Survey Data

Our empirical approach relies on the usage of survey data as proxies for investors' expectations. This raises the typical concerns regarding the use of survey data. First, some survey respondents might not understand the question or answer it without care. This would then create measurement error. Second, the survey expectations might not reflect the expectations of actual stock market investors and would therefore be not relevant. Third, survey data may be biased. In turn, we argue why these concerns are unlikely to be issues in our setting.

On the first point, we use earnings forecasts from equity analysts whose job it is to conduct these

[49] Data: https://www.aaii.com/assetallocationsurvey

[50] This 61.2% allocation is also consistent with Giglio, Maggiori, Stroebel, and Utkus (2021) who obtain detailed information on the portfolio composition of a large subset of Vanguard retail investors and find their average equity share to be 67.5%.

[51] We use the estimate in column 3 of Table 3 which removes extreme outliers.

forecasts, often after extensive talk with company management and other industry experts. Additionally, there are also monetary incentives for these equity analysts to produce the most accurate forecasts. Measurement is therefore more likely to affect the three individual survey that ask investors about their bullishness about the short-term evolution of stock prices, although it seems unlikely that the surveyed investors do not understand the question. To address this issue, we take the first principal component of these surveys which removes any idiosyncratic noise or measurement error contained in the individual surveys. Moreover, we note that the three survey answers are consistent with each other, as they are highly positive correlated (see Appendix Table C.2) making it possible to extract a meaningful principal component.

Additionally, our results show that the economic outcome, i.e. the aggregate stock price, is highly correlated with the survey expectations as predicted by our theory. Thus, if there is still significant measurement error in our survey expectations, then this would only lower the estimate of the stock price sensitivity to these expectations (i.e. classical measurement error problem). This also addresses the concern that the survey expectations do not appropriately capture the expectations of stock market investors. If our expectation measures were not relevant, then we would not observe such a strong relationship and explanatory power for aggregate stock prices.

Finally, there is also the concern that the surveys are biased. For example, Kothari, So, and Verdi (2016) argue that equity analysts have a strong incentive to produce accurate forecasts, but also that they are incentivised to make slightly optimistic forecasts. However, for any bias to affect our results, this bias needs to co-move with stock prices. This seems rather unlikely. Regarding surveys on return expectations, Cochrane (2011) argues that these surveys reflect risk-adjusted expectations. The explanatory power of these surveys in our setting might then not come from the actual expectations about future returns themselves, but from the risk adjustment. However, we directly include controls for risk aversion in our empirical framework which invalidates this argument.

6 Conclusion

In this paper, we develop – based on a heterogeneous agent model – a quantitative framework that allows us to understand the sources of stock market fluctuations using observable quantities. We find that survey expectations of long-term cash flow growth and short-term returns explain a large fraction of stock market fluctuations highlighting the role of mispricings in the stock market. Our paper contributes to a growing literature that highlights the usefulness of survey data in explaining asset prices. More research efforts should be devoted to how the expectations of investors as observed in surveys are formed – something we are silent about.

References

Adrian, Tobias, Erkko Etula, and Tyler Muir, 2014. Financial intermediaries and the cross-section of asset returns. The Journal of Finance 69, 2557–2596.

Bansal, Ravi, and Amir Yaron, 2004. Risks for the long run: A potential resolution of asset pricing puzzles. The journal of Finance 59, 1481–1509.

Barber, Brad M, and Terrance Odean, 2000. Trading is hazardous to your wealth: The common stock investment performance of individual investors. The journal of Finance 55, 773–806.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer, 2015. X-capm: An extrapolative capital asset pricing model. Journal of financial economics 115, 1–24.

———, 2018. Extrapolation and bubbles. Journal of Financial Economics 129, 203–227.

Barberis, Nicholas, Ming Huang, and Tano Santos, 2001. Prospect theory and asset prices. The quarterly journal of economics 116, 1–53.

Barberis, Nicholas, Andrei Shleifer, and Robert W Vishny, 2005. A model of investor sentiment Princeton University Press.

Barro, Robert J, 2006. Rare disasters and asset markets in the twentieth century. The Quarterly Journal of Economics 121, 823–866.

Ben-David, Itzhak, Jiacui Li, Andrea Rossi, and Yang Song, 2021. Ratings-driven demand and systematic price fluctuations. Discussion paper, .

Bhandari, Anmol, Jaroslav Borovička, and Paul Ho, 2019. Survey data and subjective beliefs in business cycle models. Available at SSRN 2763942.

Bianchi, Francesco, Sydney C Ludvigson, and Sai Ma, 2020. Belief distortions and macroeconomic fluctuations. Discussion paper, National Bureau of Economic Research.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer, 2020. Expectations of fundamentals and stock market puzzles. Discussion paper, National Bureau of Economic Research.

Boudoukh, Jacob, Roni Michaely, Matthew Richardson, and Michael R Roberts, 2007. On the importance of measuring payout yield: Implications for empirical asset pricing. The Journal of Finance 62, 877–915.

Bradshaw, Mark T, Theodore E Christensen, Kurt H Gee, and Benjamin C Whipple, 2018. Analysts' gaap earnings forecasts and their implications for accounting research. Journal of Accounting and Economics 66, 46–66.

Bradshaw, Mark T, and Richard G Sloan, 2002. Gaap versus the street: An empirical assessment of two alternative definitions of earnings. Journal of Accounting Research 40, 41–66.

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 Robert J Shiller, 1988. The dividend-price ratio and expectations of future dividends and discount factors. The Review of Financial Studies 1, 195–228.

Cassella, Stefano, and Huseyin Gulen, 2018. Extrapolation bias and the predictability of stock returns by price-scaled variables. The Review of Financial Studies 31, 4345–4397.

Chen, Long, Zhi Da, and Xinlei Zhao, 2013. What drives stock price movements?. The Review of Financial Studies 26, 841–876.

Cochrane, John H, 2011. Presidential address: Discount rates. The Journal of finance 66, 1047–1108.

——— , 2017. Macro-finance. Review of Finance 21, 945–985.

Coibion, Olivier, and Yuriy Gorodnichenko, 2015. Information rigidity and the expectations formation process: A simple framework and new facts. American Economic Review 105, 2644–78.

Cutler, David M, James M Poterba, and Lawrence H Summers, 1990. Speculative dynamics and the role of feedback traders. .

De Bondt, Werner FM, and Richard Thaler, 1985. Does the stock market overreact?. The Journal of finance 40, 793–805.

De La O, Ricardo, and Sean Myers, 2021. Subjective cash flow and discount rate expectations. The Journal of Finance 76, 1339–1387.

De Long, J Bradford, Andrei Shleifer, Lawrence H Summers, and Robert J Waldmann, 1990. Noise trader risk in financial markets. Journal of political Economy 98, 703–738.

de Silva, Tim, and David Thesmar, 2021. Noise in expectations: Evidence from analyst forecasts. Discussion paper, National Bureau of Economic Research.

Gabaix, Xavier, and Ralph SJ Koijen, 2021. In search of the origins of financial fluctuations: The inelastic markets hypothesis. Discussion paper, National Bureau of Economic Research.

Gennaioli, Nicola, Yueran Ma, and Andrei Shleifer, 2016. Expectations and investment. NBER Macroeconomics Annual 30, 379–431.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus, 2021. Five facts about beliefs and portfolios. Discussion paper, National Bureau of Economic Research.

Gilchrist, Simon, and Egon Zakrajšek, 2012. Credit spreads and business cycle fluctuations. American economic review 102, 1692–1720.

Greenwald, Daniel L., Martin Lettau, and Sydney C. Ludvigson, 2014. Origins of stock market fluctuations. Discussion paper, .

Greenwald, Daniel L, Martin Lettau, and Sydney C Ludvigson, 2019. How the wealth was won: Factors shares as market fundamentals. Discussion paper, National Bureau of Economic Research.

Greenwood, Robin, and Andrei Shleifer, 2014. Expectations of returns and expected returns. The Review of Financial Studies 27, 714–746.

Haddad, Valentin, and Tyler Muir, 2021. Do intermediaries matter for aggregate asset prices?. Discussion paper, National Bureau of Economic Research.

He, Zhiguo, Bryan Kelly, and Asaf Manela, 2017. Intermediary asset pricing: New evidence from many asset classes. Journal of Financial Economics 126, 1–35.

Hirshleifer, David, Jun Li, and Jianfeng Yu, 2015. Asset pricing in production economies with extrapolative expectations. Journal of Monetary Economics 76, 87–106.

Kothari, Sagar P, Eric So, and Rodrigo Verdi, 2016. Analysts' forecasts and asset pricing: A survey. Annual Review of Financial Economics 8, 197–219.

La Porta, Rafael, 1996. Expectations and the cross-section of stock returns. The Journal of Finance 51, 1715–1742.

Lettau, Martin, and Sydney Ludvigson, 2001. Consumption, aggregate wealth, and expected stock returns. the Journal of Finance 56, 815–849.

Li, Jiacui, 2021. What drives the size and value factors?. Working Paper.

Malmendier, Ulrike, and Stefan Nagel, 2011. Depression babies: do macroeconomic experiences affect risk taking?. The quarterly journal of economics 126, 373–416.

Nagel, Stefan, and Zhengyang Xu, 2019. Asset pricing with fading memory. Discussion paper, National Bureau of Economic Research.

Shiller, Robert J, 1981. Do stock prices move too much to be justified by subsequent changes in dividends?. .

———, 2015. Irrational exuberance Princeton university press.

van Binsbergen, Jules H, Xiao Han, and Alejandro Lopez-Lira, 2020. Man vs. machine learning: The term structure of earnings expectations and conditional biases. Discussion paper, National Bureau of Economic Research.

Appendix for "Heterogeneous Investors and Stock Market Fluctuations"

A Theory – Derivations

A.1 Derivation of the Infinite-Horizon Model

We assume that investors live for only one period and then pass on their wealth to the next generation upon death. Moreover, we assume that the newborn investors are distributed with the same shares $(\mu_p, \mu_c, 1-\mu_p-\mu_c)$, inherit homogeneous wealth and have a total mass of 1 in every period.[1] Under these assumptions, investors still maximize their one-period utility as in the static model. We additionally introduce a risk-free rate such that the risk-free asset's gross return is R. For convenience, we further assume that the total payoff volatility σ is the same in every period.

Any fundamental investor i born in period t will choose the stock market shares λ_t^i to maximize

$$\max_{\lambda_t^i} (w_t - \lambda_t^i p_t)R + \lambda_t^i \mathbb{E}_t^i[p_{t+1} + d_{t+1}] - \frac{1}{2}\gamma_t(\lambda_t^i)^2\sigma^2,$$

where $w_t - \lambda_t^i p_t$ is their wealth invested into the risk-free asset. This yields a stock market demand for any fundamental investor i of

$$\lambda_t^i = \frac{\mathbb{E}_t^i[p_{t+1} + d_{t+1}] - Rp_t}{\gamma_t\sigma^2}.$$

Furthermore, we assume that the stock market demand of price extrapolators is unchanged by the introduction of the risk-free rate (and therefore given by (2)). We restate equation (2) here for convenience

$$\lambda_t^i = \frac{\mathbb{E}_t^p[r_{t+1}]}{\gamma_t\sigma^2}. \tag{A.1}$$

The equilibrium price of the stock market is then similar to the static model's equilibrium price (Equation (3)), but includes a term for the risk-free rate:

$$\begin{aligned} p_t = R^{-1}\Bigg(&\frac{1-\mu_c-\mu_p}{1-\mu_p}\left(\mathbb{E}_t^r[p_{t+1}] + \mathbb{E}_t^r[d_{t+1}]\right) + \frac{\mu_c}{1-\mu_p}\left(\mathbb{E}_t^c[p_{t+1}] + \mathbb{E}_t^c[d_{t+1}]\right) \\ &- \frac{1}{1-\mu_p}\gamma_t\sigma^2 + \frac{\mu_p}{1-\mu_p}\mathbb{E}_t^p[r_{t+1}]\Bigg). \end{aligned} \tag{A.2}$$

However, in the dynamic model the price in period $t+$ is an endogenous object that is given by iterating one step forward on equation (A.2). Consistent with this, any fundamental investor i's expectation about

[1] These assumptions prevent the rational investors from "taking over" the economy. Furthermore, it ensures that investors from becoming more rational (possibly by learning from their errors).

p_{t+1} is given by

$$\mathbb{E}_t^i[p_{t+1}] = R^{-1}\Bigg(\frac{1-\mu_c-\mu_p}{1-\mu_p}\left(\mathbb{E}_t^i[\mathbb{E}_{t+1}^r[p_{t+2}]+\mathbb{E}_{t+1}^r[d_{t+2}]]\right)+\frac{\mu_c}{1-\mu_p}\left(\mathbb{E}_t^i[\mathbb{E}_{t+1}^c[p_{t+2}]+\mathbb{E}_{t+1}^c[d_{t+2}]]\right)$$
$$-\frac{1}{1-\mu_p}\mathbb{E}_t^i[\gamma_{t+1}]\sigma^2+\frac{\mu_p}{1-\mu_p}\mathbb{E}_t^i[\mathbb{E}_{t+1}^p[r_{t+2}]]\Bigg). \tag{A.3}$$

Again, this illustrates the difference between fundamental investors and price extrapolators. While price extrapolators have short-sighted return expectations that are potentially independent of future cash flows, fundamental investors think carefully about the forces that drive stock prices in the future (the right-hand side objects of (A.3)), such as cash flows.

We make two assumptions that allow us to go further. First, we assume that risk aversion γ and the return expectations of the price extrapolator follow first-order autoregressive processes with parameter ρ. Thus,

$$\gamma_{t+1} = \bar{\gamma}+\rho(\gamma_t-\bar{\gamma})+\epsilon_{\gamma,t+1}, \tag{A.4}$$

where $\bar{\gamma}$ is the mean risk aversion of all investors in the economy. Similarly,

$$\mathbb{E}_{t+1}^p[r_{t+2}] = \bar{r}+\rho(\mathbb{E}_t^p[r_{t+1}]-\bar{r})+\epsilon_{p,t+1}, \tag{A.5}$$

where $\bar{r}$ is the mean return expectation of the price extrapolator. Consistent with these dynamics, the expectations of any fundamental investors i follow $\mathbb{E}_t^i[\mathbb{E}_{t+1}^p[r_{t+2}]] = \bar{r}+\rho(\mathbb{E}_t^p[r_{t+1}]-\bar{r})$ and $\mathbb{E}_t^i[\gamma_{t+1}] = \bar{\gamma}+\rho(\gamma_t-\bar{\gamma})$. Second, we assume that any fundamental investor thinks that she has correct beliefs and that any newborn fundamental trader in period $t+1$ will adopt her beliefs. Essentially, we assume $\mathbb{E}_t^i[\mathbb{E}_{t+k}^j[x_{t+k+1}]] = \mathbb{E}_t^i[x_{t+k}]$ for any fundamental trader j born in period $t+k$, where x_{t+k+1} can be the price or the dividend in period $t+k+1$.

Using these assumptions in (A.3) we obtain for fundamental investor i's expectation about p_{t+1}

$$\mathbb{E}_t^i[p_{t+1}] = R^{-1}\left(\mathbb{E}_t^i[p_{t+2}]+\mathbb{E}_t^i[d_{t+2}]-\frac{1}{1-\mu_p}\left[\bar{\gamma}+\rho(\gamma_t-\bar{\gamma})\right]\sigma^2+\frac{\mu_p}{1-\mu_p}\left[\bar{r}+\rho(\mathbb{E}_t^p[r_{t+1}]-\bar{r})\right]\right). \tag{A.6}$$

Using this (for $\mathbb{E}_t^r[p_{t+1}]$ and $\mathbb{E}_t^c[p_{t+1}]$) in the stock market pricing equation (A.2) yields

$$\begin{aligned} p_t =& \frac{1-\mu_c-\mu_p}{1-\mu_p}\left(\mathbb{E}_t^r[d_{t+1}]+R^{-1}\mathbb{E}_t^r[d_{t+2}]+R^{-1}\mathbb{E}_t^r[p_{t+2}]\right)\\ &+\frac{\mu_c}{1-\mu_p}\left(\mathbb{E}_t^c[d_{t+1}]+R^{-1}\mathbb{E}_t^c[d_{t+2}]+R^{-1}\mathbb{E}_t^c[p_{t+2}]\right)\\ &-\frac{1}{1-\mu_p}\left(\gamma_t+R^{-1}(1-\rho)\bar{\gamma}+R^{-1}\rho\gamma_t\right)\sigma^2+\frac{\mu_p}{1-\mu_p}\left(\mathbb{E}_t^p[r_{t+1}]+R^{-1}(1-\rho)\bar{r}+R^{-1}\mathbb{E}_t^p[r_{t+1}]\right). \end{aligned} \tag{A.7}$$

We can further iterate on the expectations of p_{t+2} and then continue iterating forward on the price until infinity to obtain

$$p_t = \kappa_0 + \frac{1-\mu_c-\mu_p}{1-\mu_p}\left(\sum_{j=1}^{\infty} R^{-j}\mathbb{E}_t^r[d_{t+j}] + \lim_{T\to\infty} R^{-T}\mathbb{E}_t^r[p_{t+T}]\right) + \frac{\mu_c}{1-\mu_p}\left(\sum_{j=1}^{\infty} R^{-j}\mathbb{E}_t^c[d_{t+j}] + \lim_{T\to\infty} R^{-T}\mathbb{E}_t^c[p_{t+T}]\right) - \frac{1}{1-\mu_p}\,\frac{1}{1-R^{-1}\rho}\gamma_t\sigma^2 + \frac{\mu_p}{1-\mu_p}\frac{1}{1-R^{-1}\rho}\mathbb{E}_t^p[r_{t+1}], \tag{A.8}$$

where $\kappa_0 = \kappa\frac{1}{1-\mu_p}\bar{\gamma} + \kappa\frac{\mu_p}{1-\mu_p}\bar{r}$ and $\kappa = \frac{R^{-1}(1-\rho)}{(1-R^{-1})(1-R^{-1}\rho)}$.

Equilibrium Price. Imposing the no-bubble condition, i.e. $\lim_{T\to\infty} R^{-T}\mathbb{E}_t^i[p_{t+T}] = 0\,\forall i$, we arrive at

$$p_t = \kappa_0 + \frac{1-\mu_c-\mu_p}{1-\mu_p}\sum_{j=1}^{\infty} R^{-j}\mathbb{E}_t^r[d_{t+j}] + \frac{\mu_c}{1-\mu_p}\sum_{j=1}^{\infty} R^{-j}\mathbb{E}_t^c[d_{t+j}] - \frac{1}{1-\mu_p}\,\frac{1}{1-R^{-1}\rho}\gamma_t\sigma^2 + \frac{\mu_p}{1-\mu_p}\frac{1}{1-R^{-1}\rho}\mathbb{E}_t^p[r_{t+1}]. \tag{A.9}$$

Stock Market Mispricing. As in the static model, we define the cash flow expectation wedge for any period j to be $\tilde{\mathbb{E}}_t^c[d_{t+j}] = \mathbb{E}_t^c[d_{t+j}] - \mathbb{E}_t^d[d_{t+j}]$ and the return expectation wedge in period t as $\tilde{\mathbb{E}}_t^p[r_{t+1}] = \mathbb{E}_t^p[r_{t+1}] - \gamma_t\sigma^2$. The "rational price" in the dynamic model is the price that would prevail in an economy with only rational investors. It is given by $p_t^r = \kappa_1 + \sum_{j=1}^{\infty} R^{-j}\mathbb{E}_t^r[d_{t+j}] - \frac{1}{1-R^{-1}\rho}\gamma_t\sigma^2$ where $\kappa_1 = \kappa\bar{\gamma}$.

Analogous to the static model, stock market mispricing is defined as the difference between the actual stock market price and the price that would prevail in an economy with only rational investors. Stock market mispricing is thus given by

$$p_t - p_t^r = \kappa_2 + \frac{\mu_c}{1-\mu_p}\sum_{j=1}^{\infty} R^{-j}\tilde{\mathbb{E}}_t^c[d_{t+j}] + \frac{\mu_p}{1-\mu_p}\frac{1}{1-R^{-1}\rho}\tilde{\mathbb{E}}_t^p[r_{t+1}], \tag{A.10}$$

where $\kappa_2 = \kappa\frac{\mu_p}{1-\mu_p}\bar{\gamma} + \kappa\frac{\mu_p}{1-\mu_p}\bar{r}$.

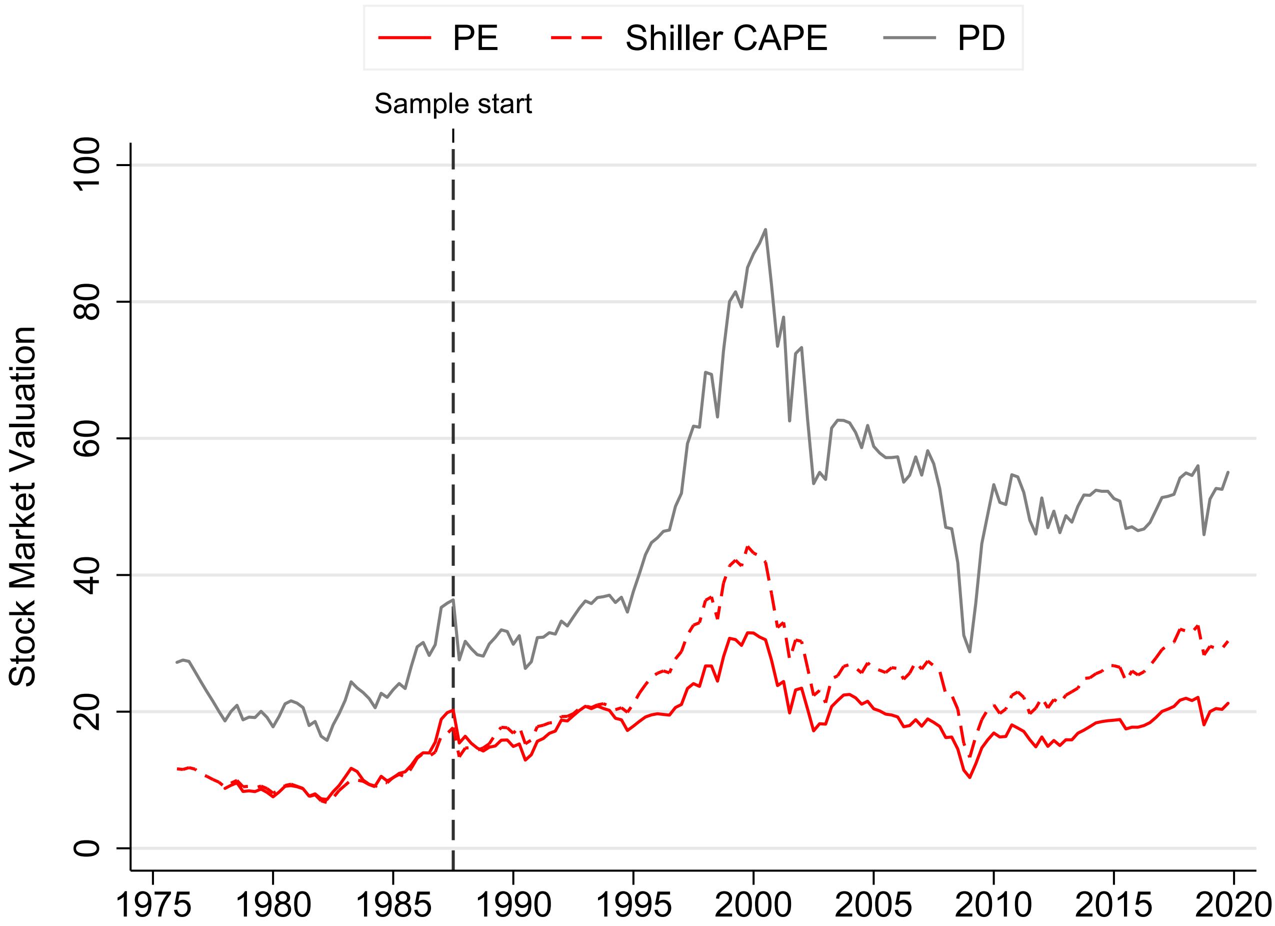
Stock Market Mispricing from the Perspective of the Rational Investors. To obtain the subjective mispricing from the perspective of the rational investors, we again expand the model by introducing signal dispersion within the group of rational investors. We assume that rational investor i's expectation about future dividends follow $\sum_{j=1}^{\infty} R^{-j}\mathbb{E}_t^r[d_{t+j}] + \epsilon_i$, where $\epsilon_i \sim \mathcal{N}(0,\psi)$.

The fraction of rational investors that perceive the market to be overvalued at time t – the subjective mispricing – is then given by

$$\%OV_t = \Phi\left(\frac{p_t - p_t^r}{\psi}\right) = \Phi\left(\frac{1}{\psi}\left[\kappa_2 + \frac{\mu_c}{1-\mu_p}\sum_{j=1}^{\infty} R^{-j}\tilde{\mathbb{E}}_t^c[d_{t+j}] + \frac{\mu_p}{1-\mu_p}\frac{1}{1-R^{-1}\rho}\tilde{\mathbb{E}}_t^p[r_{t+1}]\right]\right). \tag{A.11}$$

B Appendix Figures

Figure B.1: Comparison: Objective Valuation Measures (S&P500)

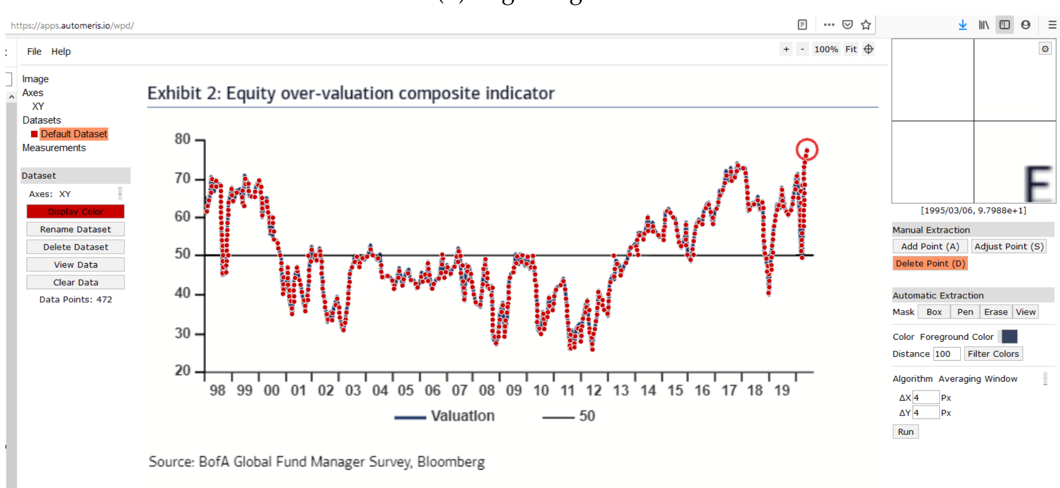


Note: Sample is S&P 500. PE corresponds to the ratio of the S&P500 index to average I/B/E/S reported actuals at the S&P500 level over the past three years. Shiller CAPE corresponds to the cyclically-adjusted price earnings ratio taken directly from Shiller's website where uses an average of the past 10 years earnings in the denominator. PD corresponds to the ratio of the S&P500 index to dividends reported by Shiller on his website, respecitvely. Further information on the comparison between these three valuation measures is given in section D.1. The Shiller data sources are available at: `http://www.econ.yale.edu/~shiller/data.htm`

Figure B.2: Digitizing BoFA Fund Manager Survey

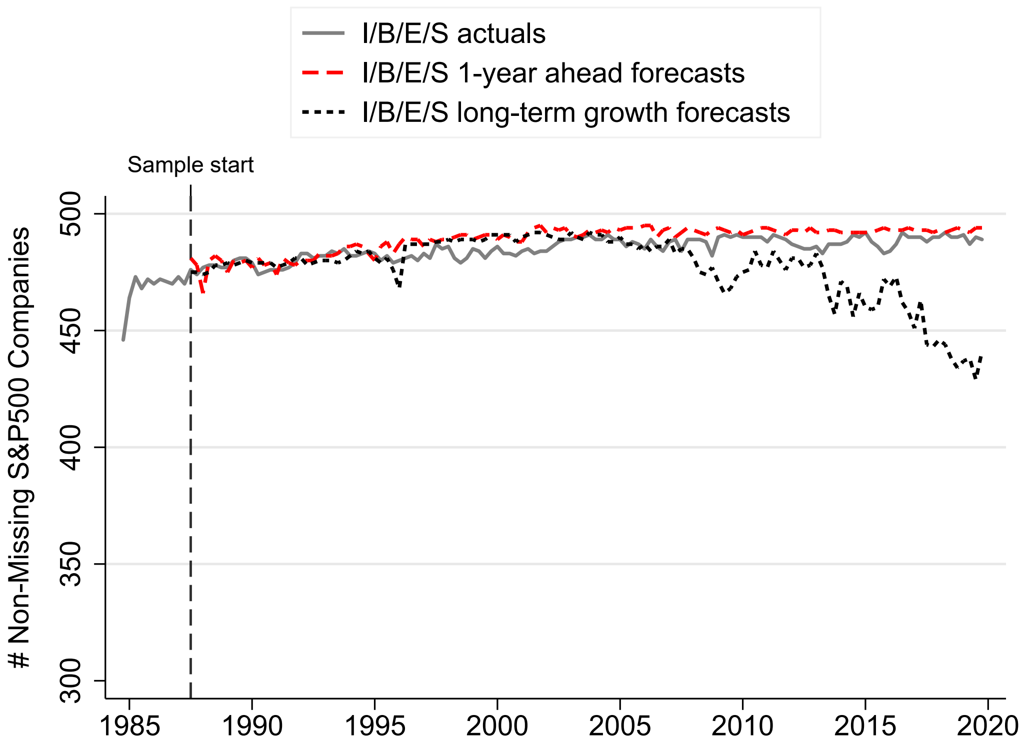
(A) BoFA Overvaluation Survey

(B) Digitizing Tool



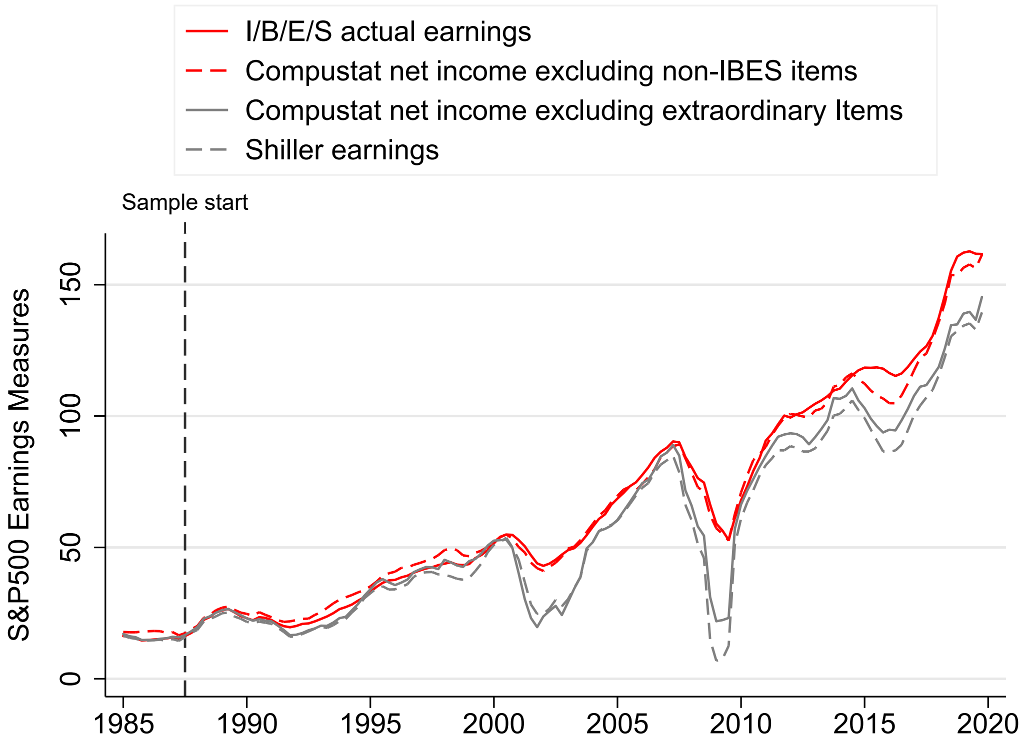
Note: This figure plots the two main graphs underlying our BoFA Fund Manager Survey data. Panel A shows the original graphical results of the Bank of America Fund Manager survey on the % of Fund Managers who believe the market is overvalued. We digitize the graphical results using a web based plot digitizer and Figure B.2 Panel B shows the result of this digitization process, where each "red point" corresponds to an estimate of the underlying data point. We download these digitized data points as a csv file and keep the last data point estimate within each quarter interval over the period 1997:Q4 to 2019:Q4. More information on this digitzation process is given in section D.2.

Figure B.3: Number of S&P500 Companies in our Sample



Note: This graph plots the number of S&P500 companies for which we have non-missing I/B/E/S actuals and I/B/E/S twelve month forecasts over the periods 1984:Q4 to 2019:Q4 and 1987:Q3 to 2019:Q4, respectively. Further information on the matching process which yields the sample of S&P500 companies is given in section D.3. Our period of analysis is 1987:Q3 to 2019:Q4.

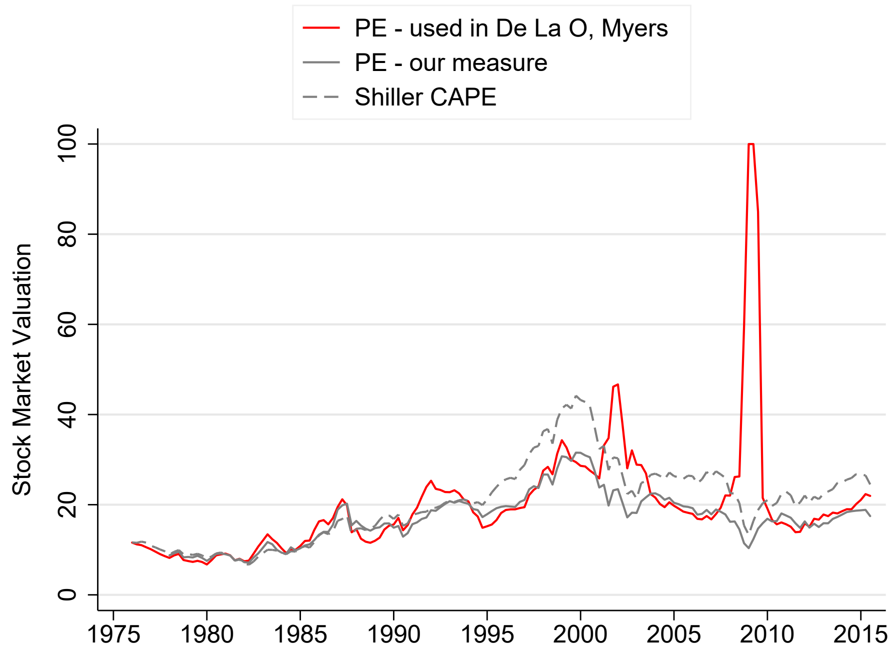
Figure B.4: Comparison: S&P500 Earnings Measures



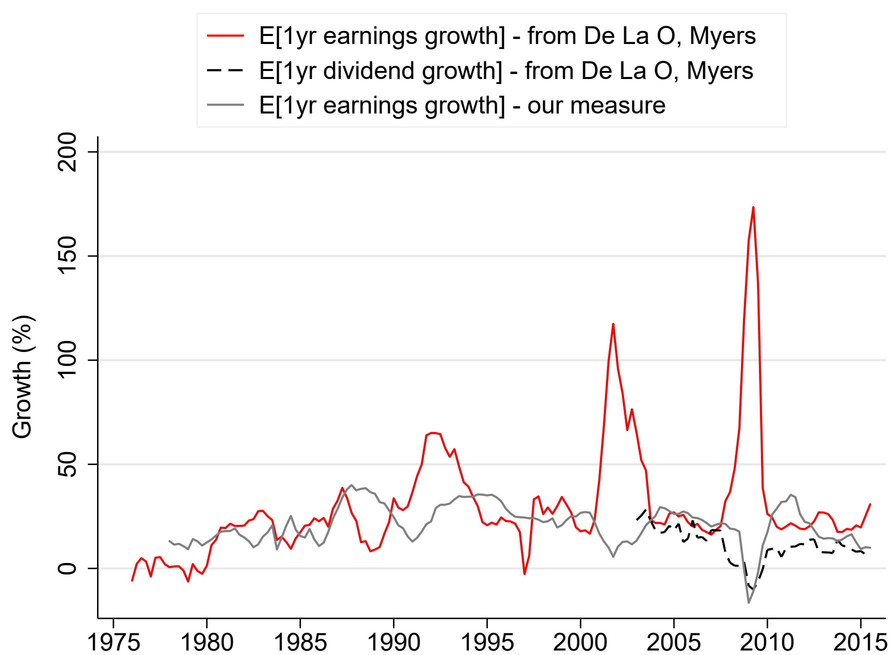
Note: This graph plots the normalized S&P500 earnings for (1) I/B/E/S actuals, (2) compustat net income excluding non-IBES items, (3) compustat net income excluding extraordinary items and (4) earnings as reported on Shiller's website. To construct these series we first only keep companies which have non-missing IBES actuals. We then aggregate (1) to (3) to the S&P500 level following the aggregation process outlines in Appendix D.4. More information on the differences between (1), (2) and (3) is given in section D.5.

Figure B.5: Comparison to De La O, Myers

(A) Valuation ratios



(B) Forecasts of short-term earnings growth



Note: This figure displays the differences underlying our results vis-a-vis De La O and Myers (2021). Panel A plot our valuation measure ("PE - our measure") against the price to earnings measured used in De La O and Myers (2021) ("PE - used in De La O, Myers") and the ten year cyclically adjusted price to earnings ratio posted on Shiller's website ("Shiller CAPE"). Panel B plots our measure of I/B/E/S analysts expected one year earnings growth ("E[1yr earnings growth] - our measure") against the I/B/E/S analysts expected one year earnings growth ("E[1yr earnings growth] - from De La 0, Myers") and dividend growth ('E[1yr dividend growth] - from De La 0, Myers") measures used in De La O and Myers (2021). More information on this comparison is given in section 5.2.

C Appendix Tables

Table C.1: Correlations between Valuation Ratios

Variables	PE	Shiller CAPE	PD
PE	1.00		
Shiller CAPE	0.91	1.00	
PD	0.83	0.91	1.00

Note: Correlations between the three valuation measures defined in section D.1. The period of analysis is is 1987:Q3 to 2019:Q4.

Table C.2: Correlations: Price Extrapolators' Expected Returns

Variables	(1)	(2)	(3)	(4)
(1) PE				
(2) $E_t^p[r_{t+1}]$	0.60			
(3) Gallup bullish-bearish	0.86	0.63		
(4) American Association bullish-bearish	0.40	0.84	0.35	
(5) Investors Intelligence bullish-bearish	0.31	0.80	0.22	0.46

Note: This table shows the pearson correlation coefficients between the PE ratio, price extrapolators expected one-year returns (i.e., $E_t^p[r_{t+1}]$) and the two components that comprise the extrapolators' expected one-year returns (i.e., Gallup, American Association and Investors' Intelligence bullish-bearish spreads). More detail on these variables is given in section 3. P-values are shown in parantheses.

Table C.3: Correlations: Subjective Overvaluation Index

Variables	(1)	(2)	(3)	(4)
(1) PE				
(2) Subjective Overvaluation Index	0.75			
(3) BoFA US Fund Manager Survey	0.59	0.88		
(4) US Valuation Confidence Index, Institutional	0.71	0.88	0.66	
(5) US Valuation Confidence Index, Individual	0.71	0.95	0.82	0.75

Note: This table shows the pearson correlation coefficients between the PE ratio, Subjective Mispricing Index and the three survey components that comprise the Subjective Mispricing Index (i.e., BoFA Fund Manager Survey, US Valuation Confidence Index, Institutional and US Valuation Confidence Index, Individual). More detail on these variables is given in section 3. P-values are shown in parantheses.

Table C.4: Gallup bull-bear spread

	Gallup bull-bear spread
	(1)
Wells Fargo / Gallup Investor Optimism Index	0.36***
	(0.04)
R^2	0.70
N	45

Note: Regression results used to extend the Gallup bull-bear spread. Refer to section 3.4 for more detail. The regression period is 1996:Q4 to 2007:Q3. Robust standard errors are shown in parentheses. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01)

Table C.5: $E_t^r[\Delta e_{t+1}]$

	Next year's earnings growth
	(1)
1yr Consumption growth	−4.36***
	(1.66)
1yr GDP growth	4.77***
	(1.49)
Industrial production growth	5.77***
	(0.88)
Constant	0.09***
	(0.03)
R^2	0.35
N	130

Note: Regression results used to construct rational investors' expected one year earnings growth (i.e., $E_t^r[\Delta e_{t+1}]$). Refer to section for 3.3 for more detail. The period of analysis is 1987:Q3 to 2019:Q4. Newey-West standard errors are shown in parentheses. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01)

Table C.6: $E_t^r[LTG]$

	Next 7yr earnings growth
	(1)
7yr Consumption growth	-2.61^{***}
	(0.84)
7yr GDP growth	2.16^{**}
	(1.05)
7yr Industrial production growth	0.54
	(0.45)
Constant	0.09^{***}
	(0.01)
R^2	0.14
N	106

Note: Regression results used to construct rational investors' expected long-term earnings growth (i.e., $E_t^r[LTG]$). Refer to section for 3.3 for more detail. The period of analysis is 1987:Q3 to 2015:Q4. Newey-West standard errors are shown in parentheses. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01)

Table C.7: $E_t^r[r_{t+1}]$

	Future 5yr ann. excess return
	(1)
Habit	-1.16
	(0.76)
$\mathbb{E}_t^c[LTG]$	-2.13^{***}
	(0.37)
$\mathbb{E}_t^c[\Delta e_{t+1}]$	0.07
	(0.06)
$\mathbb{E}_t^p[r_{t+1}]$	-0.46
	(0.40)
Constant	0.36^{***}
	(0.06)
R^2	0.40
N	118

Note: Regression results used to construct rational investors' one-year required risk-premia. To construct rational investors' one year required returns we then add back the one-year Treasury bill rate. Refer to section 3.5 for more detail. The period of analysis is 1987:Q3 to 2019:Q4. Newey-West standard errors are shown in parentheses. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01)

Table C.8: Extrapolative Earnings Expectations - CG Regressions

(A) Using I/B/E/S actuals ("Street earnings")

	Realized − Forecasted Earnings Growth			
	(1)	(2)	(3)	(4)
	1yr	3yr	4yr	5yr
$(E_t^c - E_{t-1}^c)[\Delta e_{t+1}]$	−1.15***			
	(0.17)			
$(E_t^c - E_{t-1}^c)[LTG]$		−3.21***	−2.77***	−2.07***
		(0.77)	(0.43)	(0.45)
R^2	0.46	0.19	0.23	0.22
N	122	114	110	106

(B) Using Compustat actuals (GAAP earnings)

	Realized − Forecasted Earnings Growth			
	(1)	(2)	(3)	(4)
	1yr	3yr	4yr	5yr
$(E_t^c - E_{t-1}^c)[\Delta e_{t+1}]$	−0.35			
	(0.24)			
$(E_t^c - E_{t-1}^c)[LTG]$		−9.46***	−7.45***	−5.85***
		(1.79)	(0.93)	(1.13)
R^2	0.06	0.35	0.37	0.37
N	122	114	110	106

Note: This table shows the coefficients from the Coibion and Gorodnichenko (2015) regressions for short-term (i.e., one-year) and longer-term earnings. In panel A we use I/B/E/S actuals for as our realized earnings measure whereas in panel we use Compustat net income excluding extraordinary items. More information on these regressions is given in section E. Sample period is 1987:Q3 to 2019:Q4. Newey-West standard errors are shown in parentheses. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01).

Table C.9: Extrapolative Return Expectations

	Price extrapolators' expected 1yr returns		Rational investors' required 1yr returns	
	(1)	(2)	(3)	(4)
log(Past 1yr Excess Return)	0.05***		-0.02*	
	(0.01)		(0.01)	
log(PE)		5.01***		-2.48**
		(0.45)		(0.96)
R^2	0.33	0.44	0.02	0.04
N	130	130	130	130

Note: This table shows the coefficients from regressions price extrapolators' expected one-year returns (i.e., $E_t^p[r_{t+1}]$) and rational investors' required one year returns (i.e., $E_t^r[r_{t+1}]$) on past one-year returns and the log of the PE ratio (PE_t). More information on the construction of $E_t^p[r_{t+1}]$, $E_t^r[r_{t+1}]$ and PE_t are given in sections 3.4, 3.5 and 3.1, respectively. Sample period 1987Q3 to 2019Q4. Newey-West standard errors are shown in parentheses. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01).

Table C.10: Robustness: Realized Earnings Growth for E_t^r[Earnings growth]

	Levels		2-year Changes	
	(1)	(2)	(3)	(4)
	pe	pe	Δ pe	Δ pe
$\mathbb{E}_t^c[\Delta e_{t+1}]$	0.45***	-0.06	0.53***	-0.00
	(0.11)	(0.18)	(0.15)	(0.13)
$\mathbb{E}_t^c[LTG]$	8.47***	7.97***	5.62***	4.59***
	(0.65)	(1.31)	(1.05)	(1.21)
$\mathbb{E}_t^p[r_{t+1}]$	5.66***	4.33***	4.91***	3.77***
	(0.62)	(0.87)	(0.89)	(0.90)
Future realized 1yr earnings growth		0.16		-0.01
		(0.12)		(0.07)
Future realized 7yr earnings growth		0.60		1.64***
		(0.70)		(0.44)
Risk Aversion (negative surplus consumption ratio)		-4.04		-6.10***
		(2.75)		(2.11)
R^2	0.86	0.89	0.77	0.86
N	106	106	98	98

Note: This table shows the results from estimating regression equation (9) using future realized one-year and five-year earnings growth as controls for rational investors' expected one-year and long-term earnings growth, respectively. The time period for columns 1–2 and 3–4 are 1987:Q3 to 2015:Q4 and XX, respectively. The unit of observation is a quarter. Columns 1–2 are estimated in levels, while columns 3–4 are estimated in changes: we use the absolute 2-year change for every variable. The dependent variable is the logarithm of the PE ratio as defined in section 3. Further details about the construction of the independent variables $E_t^c[\Delta e_{t+1}]$, $E_t^c[LTG]$ and $E_t^P[r_{t+1}]$ are also provided in Section 3. Newey-West standard errors are shown in parentheses. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01).

Table C.11: Additional Risk Aversion Proxies

	pe					
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{E}_t^c[\Delta e_{t+1}]$	0.34*	0.40**	0.28	0.33*	0.23	0.30*
	(0.17)	(0.20)	(0.18)	(0.18)	(0.20)	(0.16)
$\mathbb{E}_t^c[LTG]$	6.45***	7.97***	7.44***	6.46***	6.48***	6.94***
	(1.43)	(0.73)	(0.97)	(1.57)	(1.42)	(1.16)
$\mathbb{E}_t^p[r_{t+1}]$	6.07***	5.98***	5.12***	6.35***	6.57***	5.34***
	(0.89)	(0.67)	(0.84)	(1.18)	(0.90)	(1.06)
$\mathbb{E}_t^r[\Delta e_{t+1}]$		0.04	0.03	−0.01	0.04	−0.09
		(0.13)	(0.14)	(0.16)	(0.14)	(0.15)
$\mathbb{E}_t^r[LTG]$		1.73	1.24	2.08	2.03	1.76
		(1.87)	(1.98)	(2.22)	(1.93)	(1.39)
AEM Factor		−0.03*				
		(0.02)				
HKM Factor			0.02**			
			(0.01)			
VIX				−0.00		
				(0.02)		
CAY Factor					0.01	
					(0.02)	
EBP Spread						−0.03*
						(0.01)
R^2	0.77	0.84	0.81	0.78	0.78	0.78
N	130	121	125	130	130	130

Note: This table shows the results from estimating regression equation (9). The time period is from 1987:Q3 to 2019:Q4. The unit of observation is a quarter. The dependent variable is the logarithm of the PE ratio. The independent variables $E_t^c[\Delta e_{t+1}]$ $-E_t^r[LTG]$ are as defined in section 3. The risk premia proxies are: the VIX, the CAY-factor from Lettau and Ludvigson (2001), the AEM and HKM intermediary factors from Adrian, Etula, and Muir (2014) and He, Kelly, and Manela (2017), and the Excess Bond Premium from Gilchrist and Zakrajšek (2012). The risk premia proxies are standardized, while all other variables are not. The HKM and AEM factors are only available until 2018Q3 and 2017Q3, respectively. Newey-West standard errors are shown in parentheses. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01).

Table C.12: Variance Decomposition with Alternative Risk Aversion Proxies

	$b = \frac{Cov(\,.\,, log(PE))}{\sigma^2(log(PE))}$	Variation explained (%)				
		(1)	(2)	(3)	(4)	(5)
$\mathbb{E}_t^c[\Delta e_{t+1}]$	0.10	2.69	3.95	3.22	2.27	2.89
$\mathbb{E}_t^c[LTG]$	0.07	49.94	53.46	43.36	43.46	46.58
$\mathbb{E}_t^p[r_{t+1}]$	0.05	25.64	29.96	31.79	32.89	26.77
$\mathbb{E}_t^r[\Delta e_{t+1}]$	0.12	0.35	0.49	-0.65	0.43	-1.00
$\mathbb{E}_t^r[LTG]$	0.00	-0.45	-0.62	-0.75	-0.73	-0.63
HKM Factor	1.27	2.37				
AEM Factor	0.30		-0.88			
VIX	-0.78			0.04		
CAY Factor	-0.35				-0.45	
EBP	-1.58					3.71
Residual		19.93	16.72	22.40	22.13	21.69

Note: This table shows the stock price decomposition for additional risk premium measures. The time period is from 1987:Q3 to 2019:Q4. The unit of observation is a quarter. The risk premia proxies are: the VIX, the CAY-factor from Lettau and Ludvigson (2001), the AEM and HKM intermediary factors from Adrian, Etula, and Muir (2014) and He, Kelly, and Manela (2017), and the Excess Bond Premium from Gilchrist and Zakrajšek (2012). The risk premia proxies are standardized, while all other variables are not. The HKM and AEM factors are only available until 2018Q3 and 2017Q3, respectively.

Table C.13: Comparison to De La O and Myers (2021)

	$\mathbb{E}_t^{DM}[\Delta e_{t+1}]$ – De La O and Myers (2021) measure			$\mathbb{E}_t^{c}[\Delta e_{t+1}]$ – Our measure	
	(1)	(2)	(3)	(4)	(5)
pe – De La O and Myers (2021) measure	0.45***	0.44***		−0.04	
	(0.05)	(0.05)		(0.03)	
pe – our measure			0.14**		0.11***
			(0.06)		(0.01)
R^2	0.65	0.64	0.03	0.04	0.16
N	159	151	151	151	151

Note: Row 1 we use the log of the price to earinngs ratio in De La O and Myers (2021) (i.e., pe_t^{DM}), row 2 we use our log of the price to earnings ratio (i.e., pe_t), columns 1 – 3 we use the expected one-year earnings growth in De La O and Myers (2021) and columns 4–5 uses our measure of I/B/E/S analysts' expected one-year earnings growth. Column 1 uses the sample period from De La O and Myers (2021) (i.e., 1976:Q1 to 2015:Q3). The sample used in columns 2 to 5 runs from 1978:Q1 to 2015:Q3. Newey-West standard errors are shown in parentheses. Significance levels: *(p<0.10), **(p<0.05), ***(p<0.01).

D Data – Further Information

D.1 Comparison: Valuation Ratios

In Figure B.1, we provide a comparison of different valuation ratios for the S&P 500. PE corresponds to the price to the earnings ratio we use (refer to section 3.1 for definition). PD and Shiller CAPE corresponds to the price dividend ratio and ten year cyclically adjusted price earnings to ratio, respectively, as sourced from Robert Shiller's website.[2] We prefer using our PE ratio for the two reasons. First, share repurchases have become an increasingly large fraction of corporate payouts mitigating the usefulness of the price-dividend ratio as a stable indicator of valuation attractiveness overtime (e.g., see Boudoukh, Michaely, Richardson, and Roberts (2007)). Second, our PE ratio – in using the past three years I/B/E/S actuals in the denominator – gives a more up-to-date proxy for normalized earnings relative to the Shiller CAPE ratio (see section D.5 for more detail on I/B/E/S actuals and specifically their exclusion of many transitory items that are included in GAAP earnings). Looking at Figure B.1 we can see there is a structural upward shift in the pd ratio between pre-2000 and post-2000 (reflecting the shift away from dividends toward buybacks) whilst the Shiller CAPE closely tracks our measure. Although we prefer using our PE ratio measure, our results are robust to using any all three valuation measures as evidence by the high correlations between the three measures per table C.1.

D.2 Bank of America Fund Manager Survey

Figure B.2 Panel A shows the original results of the Bank of America Fund Manager survey on the % of Fund Managers who believe the market is overvaluaed, as reported in a public media Bloomberg article on June 16th 2020.[3]

To digitize these graphical results we use a web based plot digitizer.[4] We follow four steps to digitize the original graphical results. Firstly, we upload the png.image in Panel A to the web based plot digitizer as a 2-D (X-Y) plot. Secondly, we align the x-axis by selecting (12/31/1998, 12/31/2009) as our (x1,x2) co-ordinates, and (20,80) as our (y1 y2) co-ordinates. Thirdly, we use their pen tool to shade in the blue line plot, namely the area we want the digitizer to give us estimates of the underlying co-ordinates. Thirdly, we select the averaging window to be small enough so as that we obtain at least one data point estimate for each quarter interval.[5] Figure B.2 Panel B shows the result of this digitization process, where each "red point" corresponds to an estimate of the underlying data point. Finally, we download these digitized data points as a csv file and keep the last data point estimate within each quarter interval over the period 1997Q4 to 2019Q4.

To test the accuracy of the digitization process we compare our underlying digitized data points to survey points mentioned in two distinct public media articles in 2015 and 2020, respectively. In particular, per our digitized data a gross 61.3% and 77.2% of investors perceive the market to be overvalued on March 2015 and June 2020, respectively. This corresponds closely to the 61.3% and 78.0% numbers

[2] http://www.econ.yale.edu/~shiller/data.htm

[3] Article link: https://www.bloomberg.com/news/articles/2020-06-16/wall-street-is-far-from-dangerously-bullis

[4] Link to the digitizer we use: https://apps.automeris.io/wpd/

[5] Indeed, the averaging window is small enough such that in most cases we obtain at least one data point estimate for each month interval.

reported in the public media for March 2015 and June 2020, respectively.[6,7]

D.3 Matching of I/B/E/S to Compustat and CRSP.

First, we merge all unique Compustat identifiers to CRSP using the Compustat-to-CRSP linktable in the WRDS Database only keeping securities that have at one stage been in the S&P500 index.[8] Second, we merge the resulting data set to I/B/E/S on the permno identifier using the iclink file and drop any observations with missing identifiers (i.e., missing at least one of gvkey, permno and ticker). This yields a dataset of 1,581 unique permno, gvkey, ticker, link-date observations. Third, we separate any observation which has a duplicate identifier (i.e., duplicate gvkey, duplicate permno or duplicate ticker) from observations which have unique identifiers. For the set of duplicate observations we manually check each link (correcting any mistaken links) and ensure each duplicate observation will be unique in any quarterly period.[9] This yields a final set of 1,545 unique permno identifiers for any quarterly period. Per Figure B.3 – which shows the number of non-missing companies for which we have I/B/E/S data on actual earnings and earnings forecasts – we see that we obtain most of the S&P500 firms throughout our period of analysis.

D.4 Aggregating Earnings to the S&P500 Level

We follow De La O and Myers (2021) to aggregate I/B/E/S earnings to the S&P500 level. Define:

x_t : the set of companies in the S&P500 at time t

x_t^o : the set of companies in the S&P500 for which we have I/B/E/S actuals time t

x_t^1 : the set of companies in the S&P500 for which we have one year I/B/E/S forecasts at time t

$e_{i,t}^{IBES}$: I/B/E/S earnings per share of company i at time t

$P_{i,t}$: The price of company i at time t

$S_{i,t}$: The shares outstanding for company i at time t

Aggregation of actual earnings Note that the S&P500 divisor at each period t ("$Divisor_t$") satisfies the identity:

$$S\&P500_t^{Index} = \frac{\sum_{i \in x_t} P_{i,t} S_{i,t}}{Divisor_t} = \frac{MCAP_t^{SP500}}{Divisor_t}$$

At each point t we compute the total market cap and IBES actuals earnings of companies in the S&P500 for which we have non-missing IBES actuals earnings as:

[6] The 2015 article link is attached at the end of this footnote. Per the article, a net "23 percent in March" found equity markets to be overvalued. In gross percentage terms, this corresponds to 61.3 $(= \frac{100-23}{2} + 23)$ percent of investors who perceive the market to be overvalued. Link: https://insurancenewsnet.com/oarticle/BofA-Merrill-Lynch-Fund-Manager-Survey-Finds-Concerns-of-Overvaluation-in-Both-E-a-612750

[7] Refer to section 3.7 for the June 2020 article link which states "BofA Survey Finds 78% of Investors See Market as Overvalued"

[8] The link table is entitled ccmxpf-linktable.

[9] For example, the data set might contain two observations for ticker ABC – one between link-dates 1/1/1980:12/31/1980 and a second between link-dates 1/1/1990:12/31/1990 – but these ticker identifiers are unique in any given quarterly period.

$$MCAP_t^{IBES} = \sum_{i \in x_t^0} P_{i,t} S_{i,t} \qquad e_t^{IBES} = \sum_{i \in x_t^0} IBES_{i,t}^{\text{Act}} S_{i,t}$$

We then scale e_t^{IBES} earnings to the S&P500 level via the following normalization:

$$e_t^{S\&P500} = \frac{e_t^{IBES}}{MCAP_t^{IBES}} * S\&P500_t^{Index} \equiv \frac{e_t^{IBES}}{Divisor_t} * \frac{MCAP_t^{SP500}}{MCAP_t^{IBES}}$$

This equation makes clear the representative assumption we invoke - namely, that the ratio of total S&P500 earnings to S&P500 earnings for which we have only have non-missing IBES actuals is equal to the ratio of their market capitalizations. To show this is likely to hold up to a small approximation error, Figure B.3 plots the number of S&P500 companies for which we have non-missing IBES actuals over the period 1984 Q4 to 2019 Q4. As can be seen, the minimum and average number of companies is 446 and 484, respectively.

To validate the aggregation assumption we also aggregate (1) Compustat net income excluding extraordinary items to the S&P500 level only for those companies for which we have non-missing I/B/E/S actuals and plot them against (2) the S&P500 earnings reported by Shiller. As can be seen in Figure B.4, Compustat net income excluding extraordinary items closely tracks the Shiller earnings series validating the aggregation process.

At the same time, I/B/E/S actuals and Compustat net income excluding non-IBES items deviate – sometimes substantially – from these two aggregate earnings series because they exclude certain special items as described in Section D.5.

Aggregation of short-term earnings forecasts We aggregate twelve month earnings forecasts to the S&P500 level by following a similar process. To do so we assume that analysts expect that any changes in S&P500 constituents or shares outstanding that affect total earnings will be offset by changes in the divisor. This amounts to assuming that analysts expect that changes in constituents or shares outstanding do not affect the price-earnings ratio of the S&P500. A stronger assumption that would also be consistent with this aggregation methodology is to assume analysts do not expect any chage in constituents or shares outstanding over the following year. Formally, this assumption means that:

$$E_t[e_{t+1}^{SP500}] = E_t[\sum_{i \in x_{t+1}} e_{i,t+1} \frac{S_{i,t+1}}{Divisor_{t+1}}] = E_t[\sum_{i \in x_t} e_{i,t+1} \frac{S_{i,t}}{Divisor_t}]$$

Under both this assumption and that our set of companies for which we have twelve month ahead forecasts is representative of the S&P500, we can aggregate forecasted earnings per share to the S&P500 level via:

$$E_t[e_{t+1}^{SP500}] = \frac{E_t[e_{t+1}^{IBES}]}{MCAP_t^{IBES}} * S\&P500_t^{Index} \equiv \frac{E_t[e_{t+1}^{IBES}]}{Divisor_t} * \frac{MCAP_t^{SP500}}{MCAP_t^{IBES}}$$

where $E_t[e_{t+1}^{IBES}] = \sum_{i \in x_t^1} E_t[e_{i,t+1} S_{i,t}]$ corresponds to total forecasted twelve-month ahead earnings for the set of S&P500 companies for which we have non-missing forecasts at time t. Per Figure B.3 the minumum and average number of S&P500 companies for which we have twelve month ahead forecasts over our analysis period is 466 and 489, respectively, and so our aggregation assumption holds quite well.

D.5 I/B/E/S Actuals vs. GAAP Earnings

Equity analysts report forecasts of earnings based on "street earnings" instead of GAAP earnings. When comparing earnings forecasts to actual earnings, we use the actual Street earnings as reported by IBES. These earnings can deviate sometimes significantly from GAAP earnings (as reported in Compustat) because of various transitory items that are removed from the street earnings. Regarding this, the IBES guide states that *"...I/B/E/S receives an analyst's forecast after discontinued operations, extraordinary charges, and other non-operating items have been backed out. While this is far and away the best method for valuing a company, it often causes a discrepancy when a company reports earnings. I/B/E/S adjusts reported earnings to match analysts' forecasts on both an annual and quarterly basis. This is why I/B/E/S actuals may not agree with other published actuals; i.e., Compustat."* A large accounting literature has therefore argued that in order to obtain consistent forecast error, it is thus important to compare forecasts and actual earnings on the same basis (e.g, Bradshaw and Sloan, 2002; Bradshaw, Christensen, Gee, and Whipple, 2018).

Per Bradshaw and Sloan (2002) a large number of the charges that are excluded from street earnings include items that correspond to discontinued operations, extraordinary and non-recurring items and a line item referred to as "special items" in Compustat.[10] To show that this assertion holds quite well we aggregate (1) I/B/E/S actuals and (2) Compustat net income excluding discontinued, special and extraordinary items and non-recurring income taxes ("compustat net income excluding non-IBES items") to the S&P500 level only for those companies for which we have non-missing I/B/E/S actuals.[11,12] Figure B.4 plots I/B/E/S actuals (red line) against compustat net income excluding non-IBES items (dashed red line) where we see that both series track each other very closely with a correlation of 0.99.

D.6 Construction: Rational Investors' Expectation

More details underlying the construction of rational investors' expected one-year earnings (i.e., $E_t^r[\Delta e_{t+1}]$), expected long-term earnings growth (i.e., $E_t^r[LTG]$) and required returns (i.e., $E_t^r[r_{t+1}]$) are given in sections 3.3 , 3.3 and 3.5, respectively. The regressions underlying the construction of these rational investor expectations are given in Tables C.5 - C.7.

[10]Special items include restructuring charges, asset impairments, merger and acquisition charges and other significant non-recurring items.

[11]More details on the aggregation process is given in Appendix D.4.

[12]More specifically, the line items niq, xidoq, spiq and nrtxtq in compustat fundq correspond to net income, extraordinary itmes and discontinued operations, special items and non-recurring income taxes, respectively. We then compute it as "compustat net income excluding non-IBES items" = niq - xidoq - spiq - nrtxtq where we sum up the quarterly earnings over the past four quarters to obtain compustat net income excluding non-IBES items on a yearly basis.

E Additional Results

E.1 Are Survey Expectations Extrapolative?

In Table C.8, we run the Coibion and Gorodnichenko (2015) regressions for I/B/E/S analysts' expected one year earnings growth ($E_t^c[\Delta e_{t+1}]$) and long-term earnings growth ($E_t^c[LTG]$). In panel A, we use I/B/E/S actuals (i.e., street earnings) for realized earnings. In panel B, we use Compustat net income excluding extraordinary items (i.e., GAAP earnings). See section D.5 for more information on the difference between these two realized earnings measures. As can be seen, Panel B very closely matches the results reported in Table 6 of Bordalo, Gennaioli, Porta, and Shleifer (2020) where we see we have neither significant overreaction or underreaction for one-year earnings growth but large overreaction for longer-term earnings growth. When we more closely compare apples-to-apples (i.e., I/B/E/S actuals to I/B/E/S forecasts) and run the same regressions using I/B/E/S actuals we instead find significant overreaction for both one-year and longer-term earnings growth (albeit the degree of overreaction slightly smaller for longer-term earnings growth). We conclude that I/B/E/S analysts are cashflow extrapolators in both the short-term and longer term in that they overreact to recent good (bad) news such that realized earnings growth tend to be worse (better) than forecast over both short and long horizons.

In Table C.9 we repeat the Greenwood and Shleifer (2014) analysis to show that price extrapolators' expected are extrapolative. Per column (1) price extrapolators expected one-year returns (i.e., $E_t^p[r_{t+1}]$) are positively related to past one-year returns (r_t) such that when past returns have been good (bad) price extrapolators revise up (down) their expected returns (i.e., they extrapolate past returns). As a result, price extrapolators' expected one-year returns are significantly positively correlated with the log of the PE ratio. By contrast rational investors' expected returns decrease (increase) when past returns have been good (bad) and are significantly negatively correlated with the log of the PE ratio.

F Price Decompositions

F.1 Derivation of Campbell-Shiller Decomposition

Start with the definition of returns:

$$r_{t+1} = log(1 + R_{t+1}) = log\left(\frac{P_{t+1} + D_{t+1}}{P_t}\right)$$

We can re-arrange the RHS:

$$log\left(\frac{P_{t+1} + D_{t+1}}{P_t}\right) = log\left(\frac{\frac{P_{t+1}+D_{t+1}}{D_{t+1}}\frac{D_{t+1}}{D_t}}{\frac{P_t}{D_t}}\right) = log\left(\frac{\left(1 + \frac{P_{t+1}}{D_{t+1}}\right)\frac{D_{t+1}}{D_t}}{\frac{P_t}{D_t}}\right)$$

We can rewrite in term of the log-notation, i.e. $log(\frac{D_{t+1}}{D_t}) = \Delta d_{t+1}$ and $log(\frac{P_t}{D_t}) = pd_t$:

$$r_{t+1} = log\left(1 + \frac{P_{t+1}}{D_{t+1}}\right) + \Delta d_{t+1} - pd_t = log\left(1 + exp(pd_{t+1})\right) + \Delta d_{t+1} - pd_t$$

We can do a first-order Taylor approximation of the function $f(.) = log\left(1 + exp(.)\right)$ around the (time-series) mean of the pd-ratio $\overline{pd}$. The first derivative is given by $f'(.) = \frac{exp(.)}{1+exp(.)}$.

Therefore:

$$log\left(1 + exp(pd_{t+1})\right) \approx log\left(1 + exp(\overline{pd})\right) + \frac{exp(\overline{pd})}{1 + exp(\overline{pd})}(pd_{t+1} - \overline{pd})$$

Inserting in the above equation we arrive at

$$r_{t+1} \approx log\left(1 + exp(\overline{pd})\right) + \frac{exp(\overline{pd})}{1 + exp(\overline{pd})}(pd_{t+1} - \overline{pd}) + \Delta d_{t+1} - pd_t$$

Note that this equation holds up to the approximation error of the Taylor-series approximation. Now we can define the constants $k = log\left(1 + exp(\overline{pd})\right) + \frac{exp(\overline{pd})}{1+exp(\overline{pd})}\overline{pd}$ and $\rho = \frac{exp(\overline{pd})}{1+exp(\overline{pd})}$. If we assume a value of $\overline{PD} = 25$ then $\rho = \frac{exp(log(25))}{1+exp(log(25))} \approx 0.96$.

The equation now reads

$$r_{t+1} = k + \rho pd_{t+1} + \Delta d_{t+1} - pd_t + \epsilon_{t+1}$$

where ϵ_{t+1} is the approximation error.

Let's ignore the approximation error and re-arrange such that the log pd-ratio is on the LHS:

$$pd_t = k + \Delta d_{t+1} - r_{t+1} + \rho pd_{t+1} \tag{F.1}$$

Iterating forward (plugging in the RHS of equation F.1 for pd_{t+1}) and assuming that the terminal condition $\lim_{j\to\infty} \rho^j pd_{t+j} = 0$ holds, we arrive

$$pd_t = \frac{k}{1-\rho} + \sum_{j=1}^{\infty} \rho^{j-1} \Delta d_{t+j} - \sum_{j=1}^{\infty} \rho^{j-1} \Delta r_{t+j} \tag{F.2}$$

Note that this equation is an ex-post identity: It cannot be used to explain the current price-dividend ratio, but tells us that the current price-dividend ratio can explain either future dividend growth or future returns.

We can take expectations at time t:

$$pd_t = \frac{k}{1-\rho} + \sum_{j=1}^{\infty} \rho^{j-1} E_t \Delta d_{t+j} - \sum_{j=1}^{\infty} \rho^{j-1} E_t \Delta r_{t+j} \tag{F.3}$$

Interpretation: The stock market valuation (here: price-dividend ratio) depends on expectations of future dividend growth and future returns. This has to hold for any investors trading the stock market, otherwise he would buy and sell until the equation holds.