# Do Survey Expectations of Stock Returns Reflect Risk-Adjustments?*

Klaus Adam
University of Oxford, Nuffield College and CEPR

Dmitry Matveev
Bank of Canada

Stefan Nagel
University of Chicago, NBER, CEPR, and CESIfo

September 25, 2018

**Abstract**

Motivated by the observation that survey expectations of stock returns are inconsistent with rational return expectations under real-world probabilities, we investigate whether alternative expectations hypotheses entertained in the asset pricing literature are consistent with the survey evidence. We empirically test (1) the notion that survey forecasts constitute rational but risk-neutral forecasts of future returns, and (2) the notion that survey forecasts are ambiguity averse/robust forecasts of future returns. We find that these alternative hypotheses are also strongly rejected by the data, albeit for different reasons. Hypothesis (1) is rejected because survey return forecasts are not in line with risk-free interest rates and because survey expected excess returns are predictable. Hypothesis (2) is rejected because agents are not always pessimistic about future returns, instead often display overly optimistic return expectations. We speculate as to what kind of expectations theories might be consistent with the available survey evidence.

*We would like to thank Jarda Borovička, James Choi, Charles Manski and seminar participants at the CESifo Venice Summer Institute Workshop on "Expectation Formation" for helpful comments and suggestions. The views expressed in this paper do not necessarily represent those of the Bank of Canada. Nagel thanks the Fama-Miller Center at the University of Chicago for research support.

# 1 Introduction

Expectations play an important role in macroeconomics and asset pricing. The predominant approach in these fields is to impose the assumption of rational expectations, which equates the subjective probability distribution perceived by the agents within the model with the objective probability distribution perceived by an outside observer equipped with a large sample of data generated by the model. The rational expectations approach is elegant, internally consistent, and it eliminates the need to empirically study the formation of subjective expectations—but the assumption about expectations underlying it could be false. Recognizing this, Manski (2004) calls on researchers to collect survey data on expectations. Measurement of expectations allows researchers to consider alternatives to the rational expectations assumption in an empirically disciplined way.

A growing body of research in asset pricing follows this approach by examining survey data on investor stock market return expectations. This literature finds that the time-series dynamics of investor return expectations in surveys are in conflict with the predictions of influential rational expectations asset-pricing theories. Models like Campbell and Cochrane (1999) generate volatile asset prices and predictable returns by making risk premia counter-cyclical. By virtue of the rational expectations assumption, the subjective beliefs of investors in these models agree with the objective distribution, and hence these theories predict that the representative investor perceives counter-cyclical expected returns. In contrast, the survey evidence in Vissing-Jorgensen (2003), Bacchetta et al. (2009), Adam, Marcet and Beutel (2017) and Greenwood and Shleifer (2014) suggests that investor return expectations are pro-cyclical: Subjective expected returns are higher following high realized stock market returns and in times of high price-dividend ratios.

In these studies of return expectations, and more generally in much of the literature using survey measures of expectations, researchers interpret the elicited survey expectations as a representation of subjective beliefs that are distinct from respondents' preferences. More specifically, the typical interpretation assumes that people do not confound the probability of a state of the world with the desirability of this state when they answer survey questions about subjective beliefs. Recently, Cochrane (2011) suggested that a particular form of confounding of beliefs and preferences could reconcile rational expectations models with the survey evidence (see also Cochrane (2017)). Cochrane asserts that

> *"If people report the risk-neutral expectation, then many surveys make sence* [sic]*."* (p. 1068)

Under risk-neutral expectations, outcomes are weighted by their probabilities multiplied with the marginal utility associated with the respective outcome (and rescaled so that the weights sum to one). In other words, risk-neutral expectations give more weight than the actual (physical) expectation to outcomes in high marginal utility states.

While the traditional interpretation of non-confounding of beliefs and preferences may be plausible—after all, without evidence to the contrary, it seems natural start with the assumption that people report what they are asked to report—its validity is ultimately an empirical question. In this paper we therefore examine the empirical validity of Cochrane's assertion. Return expectations are an excellent setting for studying this issue because asset pricing theory provides a sharp prediction: risk-neutral expectations of returns on traded assets should equal the risk-free rate. For instance, if a survey respondent is asked to state the expected rate of return on a diversified portfolio of stocks over the next 12 months, Cochrane's risk-neutral expectations hypothesis predicts that the respondent should respond with the 12-month risk free rate as their expected rate of return.

We use several different surveys of individual investors, professional investors, and Chief Financial Officers (CFO) covering various sample periods from the 1980s until recently. The implication of the risk-neutral expectations hypothesis that expected returns equal the risk-free rate is strongly rejected. Unconditionally, survey expectations of stock market returns exceed the risk-free rate by 1 to 5 percentage points depending on the survey and sample period and the difference is highly statistically significant. Conditionally, the risk-neutral expectations hypothesis predicts that deviations of survey expected returns from the risk-free rate should be random measurement errors that are unrelated to cyclical variables. However, we find that for almost all surveys and forecast horizons, the stock market's price-dividend (P/D) ratio predicts the direction of the deviation. Specifically, for individual investors and CFOs, the deviation is pro-cyclical: expected stock returns exceed the risk-free rate when the P/D ratio is high.

Overall, the empirical evidence strongly rejects the hypothesis that survey respondents report a risk-neutral expectation. As we show, this is a rejection in a very general sense. We can allow for differences in opinion, heterogeneous preferences, and biased subjective beliefs and we still obtain that risk-neutral expected returns equal the risk-free rate. In the absence of trading frictions, heterogeneous individuals should adjust the risk-profile of their portfolio and their borrowing and lending such that their future time-discounted expected marginal utilities align with the current risk-free rate. As a consequence, their risk-neutral expected rates of return are all equal to the risk-free rate. The rejection in our tests thus implies that there exists no internally consistent probability measure that can reconcile the observed survey data with the risk-neutral expectations hypothesis in a frictionless setting.

These results also hold up if we replace Treasury yields with variable mortgage rates as alternative risk-free rate proxy. To the extent that there is a friction-induced wedge between borrowing and lending rates, a collateralized borrowing rate may be a better proxy for the risk-neutral expected return than a Treasury rate. However, using this alternative risk-free rate proxy does not materially change the results.

While the data clearly reject the risk-neutral expectations hypothesis for survey expected returns, this still leaves open the possibility that preferences and beliefs are confounded in other ways. For example, Manski (2017) reviews several examples from the literature in which individuals appear to overestimate the probability of extremely bad outcomes such as death or being a crime victim. In asset pricing, a distorted probability measure that overweights bad outcomes can be used to represent ambiguity aversion or concerns about model misspecification (Hansen and Sargent (2001)). Bhandari, Borovička, and Ho (2016) use survey expectations of macroeconomic variables to estimate such distorted probabilities. However, the fact that decision-making of ambiguity averse or robustness-seeking individuals can be modeled using a distorted probability measure does not imply that survey responses would necessarily reflect such distorted probabilities. For this reason, we investigate this issue empirically.

If ambiguity aversion or robustness-seeking is reflected in the beliefs elicited in surveys, we should find that survey expected returns should be pessimistically biased relative to rational expectations of returns. To test this prediction, we compare survey expected returns to realized rates of return. Unconditionally, we find that the expected stock market returns reported in surveys are approximately unbiased as forecasts of realized returns. Conditionally, taking into account predictable variation in the wedge between survey expected returns and realized returns, we find that investors are roughly as many times optimistic as they are pessimistic. These findings are inconsistent with the idea that survey expectations have a pessimistic bias due to ambiguity aversion or robustness concerns.

In summary, our findings show that the expected returns elicited in surveys are not risk-neutral expectations. More generally, we do not find evidence of marginal utility weighting or a distortion towards overweighting of bad outcomes in survey expectations. The fact that the (pro-cyclical) empirical time-series dynamics of expected returns reported in surveys differ starkly from the (counter-cyclical) predictions of leading rational expectations models therefore cannot be explained away by positing that individuals report risk-neutral expectations. It does not appear possible to simultaneously match asset price dynamics and the survey evidence without entertaining some departure from rational expectations, such as extrapolative expectations (Barberis et al. (2015)), learning about underlying trends in price growth (Adam et al. (2016, 2017)), or learning from life-time experience (Collin-Dufresne, Johannes, and Lochstoer (2016), building on Malmendier and Nagel (2011, 2016)).

In our analysis, we work with survey data that provides respondents' point expectations of stock market returns or stock price changes. A potential alternative would be to work with surveys that elicit quantiles of respondents' subjective distributions. Manksi (2017) highlights advantages of probabilistic expectations data and he surveys research based on such data. For the specific purpose in our paper, however, point expectations are preferable. The risk-neutral expectations hypothesis makes sharp predictions about point expectations of asset returns, but we would not be able to derive testable predictions about quantiles of subjective asset return distributions without additional auxiliary assumptions about the functional form of marginal utility. Nevertheless, our findings are still relevant for the literature on probabilistic survey expectations. Our evidence that point expectations do not appear to be distorted by risk-adjustments also provides support for interpreting the responses to probabilistic survey questions as an expression of respondents' subjective distribution without confounding of beliefs and preferences.

The remainder of the paper is structured as follows. In Section 2, we derive testable implications of the risk-neutral expectations and pessimistic expectations hypotheses. After describing the data sources in section 3, we present the empirical results in 4. Section 5 concludes.

# 2 Hypotheses about Survey Expectations

Let $R_{t+1}$ denote a one-period return on a stock market index realized over the period $t$ to $t+1$ and let $R_t^f$ denote the one-period return offered by a risk-free asset over the same periods. As a benchmark, we consider an investor $i$ who can freely trade in both instruments. The effects of trading constraints will be discussed below. The investor's first order conditions imply that the returns must satisfy

$$1 = E_t^{\mathcal{P}^i}[M_{t+1}^i R_{t+1}] \tag{1}$$

$$1 = E_t^{\mathcal{P}^i}[M_{t+1}^i] R_t^f, \tag{2}$$

where $M_{t+1}^i$ is the agents' one-period stochastic discount factor (SDF) from $t$ to $t+1$ and $E_t^{\mathcal{P}^i}$ is an expectations operator that is based on some (potentially subjective) probability measure $\mathcal{P}^i$. Specific asset pricing theories give rise to specific forms of the $SDF^i$ or make specific assumptions about $\mathcal{P}^i$, but we shall not be concerned with this here: the testable implications derived below will rely exclusively on equations (1) and (2) being satisfied for some $M^i$ and some probability measure $\mathcal{P}^i$. In fact, we can even allow $M^i$ and $\mathcal{P}^i$ to differ across investors.

Equations (1) and (2) assume that agents can (at the margin) freely trade in the stock and in the bond market. This is in line with the assumptions made in a wide range of representative agent asset pricing models, e.g., the ones in the tradition of Campbell and Cochrane (1999) or Bansal and Yaron (2004). Limited participation models, e.g., the one considered by Guvenen (2009), postulate that some agents only trade in the bond market but have no access to the stock market, while others have access to both markets. For agents without access to the stock market, equation (1) will not necessarily be satisfied. For agents with access to both markets, equations (1) and (2) both hold. To avoid that our empirical results are tainted by limited stock market access, we shall consider below mainly survey sources for which we know that survey respondents do have stock market access.

## 2.1 Risk-Neutral Expectations

Let $\mathcal{E}_t^i[.]$ denote the expectations of individual $i$ measured in a survey.

The hypothesis put forward by Cochrane (2011, 2017) is that survey returns are expectations of risk-neutral stock returns, i.e.,

$$\mathcal{E}_t^i[R_{t+1}] = E_t^{\mathcal{P}^i}\left[\frac{M_{t+1}^i}{E_t^{\mathcal{P}^i}\left[M_{t+1}^i\right]}R_{t+1}\right] + \varepsilon_t^i, \tag{3}$$

where the term pre-multiplying $R_{t,t+1}$ inside the expectations operator is a Radon-Nikodym term that transforms the 'physical' probability of future states, which enter the computation of the expectation $E_t^{\mathcal{P}^i}[\cdot]$, into a 'risk-neutral' or 'marginal-utility weighted' probability.[1] The measurement error $\varepsilon_t^i$ captures the fact that we measure

[1] Let $s_t$ denote the state in $t$ and $p^i(s_{t+1}|s_t)$ the physical probability (implied by $\mathcal{P}^i$) of transitioning from $s_t$ to state $s_{t+1}$ in $t+1$. The risk neutral probability $n^i(s_{t+1}|s_t)$ of reaching state $s_{t+1}$ given state $s_t$ is then

$$n^i(s_{t+1}|s_t) \equiv p^i(s_{t+1}|s_t)\frac{M^i(s_{t+1}|s_t)}{E_t^{\mathcal{P}^i}\left[M^i(s_{t+1}|s_t)\right]},$$

so that under the risk-neutral expectations hypothesis, we get

$$\mathcal{E}_t^i[R_{t+1}] = E_t^{n^i}[R_{t+1}] + \varepsilon_t,$$

where $E_t^{n^i}[\cdot]$ is the expectations operator that integrates over states using the probabilities $n^i(s_{t+1}|s_t)$.

the true expectations only with noise.

Note that Cochrane's (2011, 2017) hypothesis is stronger than what is stated in equation (3): it additionally postulates that $\mathcal{P}^i$ is a 'rational' or 'objective' probability measure. This additional constraint, however, turns out not to be relevant for the arguments that follow, prompting us to proceed with the more general case in which agents are allowed to have objective or subjective beliefs.

Equation (3) implies that future returns that materialize in states in which marginal utility and thus the SDF is high (low) are treated by agents as if they are more (less) likely than under the objective measure and thus lead to an upward (downward) 'distortion' of the expected returns relative to plain return expectations. Under the risk-neutral return hypothesis, survey expectations can thus look 'distorted', when wrongly interpreting them as plain return expectations. Therefore, the risk-neutral expectations hypothesis could potentially help reconcile rational expectations asset pricing theories with the survey evidence.

Equation (3) together with equations (1) and (2) implies

$$\mathcal{E}_t^i[R_{t+1}] = R_t^f + \varepsilon_t^i, \tag{4}$$

which shows that under the risk-neutral expectations hypothesis survey expectations of future stock returns must equal - up to a measurement error - the risk-free interest rate. This is intuitive, as under the risk-neutral measure all subjectively expected returns are identical and equal to the risk-free interest rate. This implication of the risk-neutral return hypothesis can be tested empirically. Specifically, since it is possible to interpret returns as nominal returns and the stochastic discount factor as a nominal discount factor, one can test equation (4) directly using nominal return expectations from surveys and nominal risk-free interest rates. A failure of equation (4) to hold will thereby imply that there exists no (objective or subjective) probability measure $\mathcal{P}^i$ that is consistent with the risk-neutral return hypothesis.

Let $\mathcal{E}_t[R_{t+1}]$ denote the mean (or median) of a cross section of survey return forecasts $\{\mathcal{E}_t^i[R_{t+1}]\}$. Since equation (4) holds for every investor, it also holds for the mean (median) of these survey return expectations, i.e.,

$$\mathcal{E}_t[R_{t+1}] - R_t^f = \varepsilon_t, \tag{5}$$

where $\varepsilon_t$ is the cross-sectional mean (median) of the individual measurement errors $\varepsilon_t^i$. We assume $E[\varepsilon_t] = 0$, but we allow $\varepsilon_t$ to be autocorrelated over time.

*Unconditional test.* We can then examine the risk-neutral expectations hypothesis by estimating $a$ in

$$\mathcal{E}_t[R_{t+1}] - R_t^f = a + \varepsilon_t. \qquad \text{(RN-U)} \tag{6}$$

and testing the null hypothesis $H_0 : a = 0$.

*Conditional test.* In addition to this prediction about unconditional means, the risk-neutral expectations hypothesis also implies a strong prediction about the time-series dynamics: Since the wedge $\mathcal{E}_t[R_{t+1}] - R_t^f$ should be zero except for measurement error, it should not be correlated with any covariates that are uncorrelated with the measurement error. For example, macroeconomic variables or standard market return predictors in asset pricing should not correlate with the wedge. On the other hand, this wedge would be strongly correlated with such variables under interesting alternative hypotheses. For example, if respondents report rational expectations without risk-adjustment, the wedge in reported expectations should be equal to the conditional market risk premium (absent measurement error). Any covariates $x_t$ that are correlated with wedge $\mathcal{E}_t[R_{t+1}] - R_t^f$ should predict $R_{t+1} - R_t^f$ with the same sign. We therefore consider the specification

$$\mathcal{E}_t[R_{t+1}] - R_t^f = a_0 + a_1' x_t + \varepsilon_t. \qquad \text{(RN-C)} \tag{7}$$

where the vector $x_t$ includes predictor variables that could, for example, capture variation in $E_t[R_{t+1}] - R_t^f$. The risk-neutral expectations hypothesis implies $H_0 : a_0 = 0 \wedge a_1 = 0$.

## 2.2 Pessimistic Expectations

A less extreme hypothesis than the risk-neutral expectations hypothesis is one where investor expectations are pessimistically biased, but not all the way down to the risk-free rate. For example, investors that are averse to ambiguity or that are seeking robustness make decisions *as if* they hold pessimistically biased expectations about asset returns compared with rational expectations. The SDF in these models can be represented as $M_{t+1}^i = A_{t+1}^i Q_{t+1}^i$, where $A_{t+1}^i$ is a conventional marginal utility-based SDF and $Q_{t+1}^i$, with $E_t[Q_{t+1}^i] = 1$, can be viewed as a belief distortion that overweights bad states of the world.

Whether expectations reported in surveys reflect these belief distortions is an open question. Bhandari, Borovička, and Ho (2016)), for example, assume so. However, this need not be the case. Even if investor choices and asset prices can be accurately characterized by viewing $Q_{t+1}^i$ as a belief distortion, this does not imply that when investors are asked to state their expectations in a survey, they report the expectations distorted by $Q_{t+1}^i$. Whether they do so is an empirical question that we investigate here. Resolving this issue is important for interpretation of survey measures of expectations and also for the empirical measurement of ambiguity aversion.

The hypothesis that the belief distortion affects survey expectations implies

$$\mathcal{E}_t^i[R_{t+1}] = E_t[Q_{t+1}^i R_{t+1}] + \varepsilon_t^i. \tag{8}$$

Since $R$ is the return on a risky asset that pays off systematically more in good states, its payoff is negatively correlated with $Q$. Therefore,

$$\mathcal{E}_t^i[R_{t+1}] = E_t[R_{t+1}] + cov(Q_{t+1}^i, R_{t+1}) + \varepsilon_t^i < E_t[R_{t+1}] + \varepsilon_t^i \tag{9}$$

i.e., the belief distortion leads to a pessimistic bias in expectations.[2]

Since (9) holds for every individual, it also holds for the mean (or median)

$$\mathcal{E}_t[R_{t+1}] < E_t[R_{t+1}] + \varepsilon_t \tag{10}$$

where $\varepsilon_t$ is the cross-sectional mean (median) of the individual measurement errors $\varepsilon_t^i$.

*Unconditional test.* Based on (10) we can then examine the pessimistic expectations hypothesis by estimating the average pessimism bias $b \equiv E[\mathcal{E}_t(R_{t+1}) - E_t(R_{t+1})]$ in

$$\mathcal{E}_t[R_{t+1}] - E_t[R_{t+1}] = b + e_t, \tag{11}$$

where $e_t = \mathcal{E}_t(R_{t+1}) - E_t(R_{t+1}) - b$ is a composite residual with $E[e_t] = 0$ that contains the measurement error $\varepsilon_t$ as well as the time-varying part of the beliefs wedge that is not due to $\varepsilon_t$. We allow $e_t$ to be serially correlated. The inequality (10) implies $H_0 : b < 0$.

However, $E_t[R_{t+1}]$ is unobservable. One approach is to substitute in $R_{t+1} = E_t[R_{t+1}] + \eta_{t+1}$, which yields

$$\mathcal{E}_t[R_{t+1}] - R_{t+1} = b + e_t - \eta_{t+1} \qquad \text{(PE-U1)} \tag{12}$$

A potential problem with this approach is that average realized returns can be substantially different from rational conditional expected returns over extended periods of time. Since some of our survey series are quite short, this could be a serious problem. For example, the 1990s were a period in which return predictions from forecasting regressions based on the dividend yield (which indicated low expected returns) differed substantially from (high) average realized returns during this period. If survey expectations are on average below realized returns during this period, this

[2] To the extent that ambiguity aversion is responsible for much of the equity premium, then $M_{t+1}^i \approx \frac{1}{R_{F,t}} Q_{t+1}^i$. If so, we are back to the risk-neutral expectations hypothesis, where $\mathcal{E}_t^i[R_{t+1}] \approx R_t^f$.

may not be an indication of pessimistic beliefs but instead reflect sampling error. Put differently, since even somewhat precise estimation of expected returns requires very long sample periods, just replacing $E_t[R_{t+1}]$ with realized returns could be very inefficient.

An alternative and likely more efficient approach is to substitute the expected returns $E_t[R_{t+1}]$ in (11) by the fitted value, $\hat{E}_t[R_{t+1}]$, from the predictive regression

$$R_{t+1} = k_0 + k_1' z_t + u_t, \tag{13}$$

where $z_t$ includes commonly used return predictors, e.g., the dividend yield, and $u_t$ is a potentially serially correlated residual. The fitted value is also a noisy estimate of the conditional expected returns, but it should be substantially more precise than the average realized return over a relatively short time period. In particular, to yield more precise estimates, the first stage regression to generate $\hat{E}_t[R_{t+1}]$ could be run on a sample that is much longer than the time series of survey expectations. Such a longer time series helps reducing small-sample biases in the predictive regression.

For this approach to be valid in terms of consistency, $z_t$ does not necessarily have to be in the information set of survey respondents, as this would not bias the estimate of the unconditional mean wedge $b$.[3] Using this approach, we estimate the coefficient $b$ in

$$\mathcal{E}_t[R_{t+1}] - \hat{E}_t[R_{t+1}] = b + e_t + \omega_t, \qquad \text{(PE-U2)} \tag{14}$$

where $\omega_t \equiv E_t[R_{t+1}] - \hat{E}_t[R_{t+1}]$, and test $H_0 : b < 0$. Since $\hat{E}_t[R_{t+1}]$ is a generated variable, we need to adjust the standard errors accordingly. Appendix A provides the asymptotic distribution.

*Conditional test.* Pessimism due to ambiguity aversion would not only imply pessimism relative to rational expectations on average, but also conditionally, period by period. As already indicated in (11), the expectations wedge may be time-varying. Since the inequality (10) holds in every state, it also holds if we condition on a vector of covariates $x_t$, with $E[\varepsilon_t|x_t] = 0$, so that

$$E\{\mathcal{E}_t[R_{t+1}] - E_t[R_{t+1}]|x_t\} < 0 \tag{15}$$

Assuming that this conditional expectation is approximately linear in $x_t$, we have

$$\mathcal{E}_t[R_{t+1}] - E_t[R_{t+1}] = b_0 + b_1' x_t + e_t, \tag{16}$$

[3]However, if $z_t$ is not in the information set, then statistical power is lost, as any deviation of $\hat{E}_t[R_{t+1}]$ from $E_t[R_{t+1}]$ adds noise.

where $e_t = \varepsilon_t + \mathcal{E}_t[R_{t+1}] - E_t[R_{t+1}] - E\{\mathcal{E}_t[R_{t+1}] - E_t[R_{t+1}]|x_t\}$ and so $E[e_t|x_t] = 0$.[4] The pessimism hypothesis implies that $b_0 + b_1' x_t < 0$. For this linear model to be consistent with this inequality, $x_t$ needs to be suitably bounded. The rational expectations alternative implies $b_0 = 0 \wedge b_1 = 0$.

We can again follow two approaches to deal with the unobservability of $E_t[R_{t+1}]$ in (16). The first approach is to substitute in $R_{t+1} = E_t[R_{t+1}] + \eta_{t+1}$, which yields

$$\mathcal{E}_t[R_{t+1}] - R_{t+1} = b_0 + b_1' x_t + e_t - \eta_{t+1} \qquad \text{(PE-C1)} \qquad (17)$$

For $E[\eta_{t+1}|x_t] = 0$ to hold, we require that $x_t$ is in the information set of survey respondents. As before, this approach suffers from the fact that $R_{t+1}$ is an extremely noisy proxy for $E_t[R_{t+1}]$.

The second approach substitutes the fitted value, $\hat{E}_t[R_{t+1}]$, from the predictive regression (13), which yields

$$\mathcal{E}_t[R_{t+1}] - \hat{E}_t[R_{t+1}] = b_0 + b_1' x_t + e_t + \omega_t \qquad \text{(PE-C2)}, \qquad (18)$$

where $\omega_t = E_t[R_{t+1}] - \hat{E}_t[R_{t+1}]$, as before in the unconditional case above. To have $E[\omega_t|x_t] = 0$, we need that $x_t$ is in the information set of survey respondents and is included in $z_t$. We show in Appendix A how to obtain the asymptotic distribution of the estimator in this case.

# 3 Data

We use data on stock market return expectations from several different data sets, which, to the best of our knowledge, cover all available quantitative data on U.S. stock market return expectations of individuals who are not professional forecasters. Table 1 provides summary statistics for the different survey data sets.

## 3.1 Survey data sources

The first data set is the Duke CFO Global Business Outlook, a quarterly survey conducted by Duke University's Fuqua School of Business and CFO magazine. As Table 1 shows, the sample contains about 400 observations per quarter. Respondents

[4] At this point, we do not need to assume that $x_t$ is in the time-$t$ information set of survey respondents. If it's not, all it means is that we cannot replace $E\{E_t[R_{t+1}]|x_t\}$ by $E[R_{t+1}|x_t]$. However, for implementation of the estimation in terms of observables below, we will need the assumption that $x_t$ is in the survey respondents information set.

in the survey provide the rate of return they expect on the S&P500 index over the next year. We obtain the median and mean responses from this survey.

The second data set is the UBS/Gallup survey.[5] The survey is based on a nationally representative sample. But to participate in the survey, respondents need to hold stocks, bonds, or mutual funds of a combined value of at least $10,000. We use data from February 1999 onwards when the survey was conducted on a regular monthly basis until 2007 with about 700 observations per month. We also observe whether the respondent household holds more than $100,000 in stocks, bonds or mutual funds. As Table 1 shows, this subsample of wealthy households accounts for somewhat less than half of the sample. We use data from two survey questions about expected returns. The first questions asks about the return that the respondent expects from an investment in the stock market during the next twelve months. This question is available until April 2003. The second expectations question asks for the return that the respondents expect on their own portfolio. This question was in the survey until October 2007.

The third data set is a series constructed by Nagel and Xu (2018) that uses additional surveys to extend the UBS/Gallup survey forward and backward in time. We use the data from 1987m6 onwards when observations are available monthly without gaps. In this series, the missing market return expectation in the UBS/Gallup survey from 2003 to 2007 is imputed from the own portfolio return expectation as the fitted value from a regression of expected market returns on own portfolio expectations in the part of the sample where both are available. This series further includes mean one-year return expectations from Ameriks et al. (2016) (one survey in 2014). The series is then extended using data from surveys that do not have percentage return expectations, but coarser measures of investor beliefs. This is done by regressing the available return expectations on the average reported probability of a rise in the stock market in the Michigan Survey of Consumers (available from 2002 to 2016) and the fitted value is used to extend the percentage expectations series. This extended series is then regressed on a measure of the proportion of respondents expecting a rise in the stock market in surveys conducted by the Conference Board (1987-2016, monthly) and the fitted value used to extend the expected return series.

The final two data sets are surveys are from Robert Shiller and the Investor Behavior Project at Yale University.[6] The surveys are based on two samples: wealthy individual investors and institutional investors. Each individual response includes the

[5] The archive is available at `http://ropercenter.cornell.edu/ubs-index-investor-optimism/`

[6] The surveys are available at `http://som.yale.edu/faculty-research/our-centers/international-center-finance/data`

Table 1: Summary Statistics of Survey Data

The table shows summary statistics for the survey data sets we use in this study. We aggregate the individual survey responses in terms of means or medians within monthly or quarterly time periods as shown in the second column. The last two columns show the time series mean and standard deviation of the aggregated mean or median percentage expected return series.

| Survey source | Aggreg. | Periods | Forecast Horizon | Sample | Avg. # obs. per period | Aggreg. | $\mathcal{E}[R]$ Mean | $\mathcal{E}[R]$ S.D. |
|---|---|---|---|---|---|---|---|---|
| CFO | quarter | 2000q3 - 2016q1 | 1yr | | 390 | mean | 5.72 | 1.56 |
| | | | | | | median | 5.14 | 1.25 |
| UBS own | month | 1999m2 - 2007m10 | 1yr | all | 702 | mean | 10.31 | 2.68 |
| | | | | | | median | 8.28 | 2.31 |
| | | | 1yr | >100k | 310 | mean | 10.16 | 2.71 |
| | | | | | | median | 8.55 | 2.29 |
| UBS market | month | 1999m2 - 2003m4 | 1yr | all | 706 | mean | 10.76 | 3.16 |
| | | | | | | median | 8.73 | 2.76 |
| | | | | >100k | 311 | mean | 10.47 | 3.27 |
| | | | | | | median | 8.85 | 2.69 |
| UBS extended | month | 1972m8 - 2016m2 (w/ gaps) | 1yr | | n/a | mean | 9.46 | 2.23 |
| Shiller individual | quarter | 1999m1 - 2015m8 | 3m | | 75 | mean | 0.89 | 1.30 |
| | | | | | | median | 1.34 | 1.23 |
| | | | 6m | | 77 | mean | 2.13 | 1.60 |
| | | | | | | median | 2.70 | 1.28 |
| | | | 1yr | | 81 | mean | 5.09 | 2.98 |
| | | | | | | median | 5.67 | 1.97 |
| | | | 10yr | | 76 | mean | 37.05 | 24.19 |
| | | | | | | median | 22.55 | 27.07 |
| Shiller professional | quarter | 1999m1 - 2015m8 | 3m | | 60 | mean | 0.57 | 1.41 |
| | | | | | | median | 1.37 | 1.25 |
| | | | 6m | | 63 | mean | 2.00 | 1.92 |
| | | | | | | median | 3.71 | 1.34 |
| | | | 1yr | | 69 | mean | 5.12 | 3.08 |
| | | | | | | median | 7.36 | 2.87 |
| | | | 10yr | | 65 | mean | 70.74 | 25.84 |
| | | | | | | median | 56.85 | 20.18 |

day on which the survey was completed. The data set starts in January 1999 and we use data until August 2015. The average number of responses per quarter is 75 (60) for the individual (institutional) investor data set. Survey respondents are asked to forecast the percentage change in the Dow Jones Industrial Index over various horizons (3 months, 6 months, 1 year and 10 years).

## 3.2 Matching with returns data

In the tests of the risk-neutral expectations hypothesis, we compare survey expectations to risk-free rates over a matched maturity. For maturities from 3 to 6 months, our baseline tests use daily U.S. Treasury Bill yields, obtained from the FRED database at the Federal Reserve Bank of St. Louis. For maturities from one to ten years, we use daily zero-coupon yields from Gürkaynak et al. (2007).[7] We also explore alternative specifications in which we use the 1-year adjustable mortgage rate from the FRED database as a proxy for individuals' collateralized borrowing rates. Since this series has been discontinued at the end of 2015, the empirical results using mortgage interest rate have 2015 as their sample end. We convert all yields into effective yields over the relevant maturity.

For each survey source and survey forecast horizon we use the risk-free interest of a corresponding maturity prevailing at the reported survey date. For the CFO survey, we know a reference date, which is a day very close to when the survey was sent (by fax). The survey administrators request a response within a few days from this reference date. We match the survey responses with the risk-free rate on the reference date. For the UBS/Gallup survey, we know the 2-3 week period in which the survey took place and we use the average daily yield during this period. The extended UBS/Gallup series is monthly and we match survey expectations with the yield at the end of the month preceding the survey month. For the Shiller data sets, for which we observe the response date of each individual survey response, we match each survey response with the interest rate prevailing on the day of the survey. Since the Shiller surveys ask about price growth on the DJIA, but the test of the risk-neutral expectations hypothesis requires a total return, we adjust the price growth series by adding the DJIA dividend yield at the end of the month preceding the survey date, adjusted for the relevant forecasting horizon assuming the dividend yield stays constant.

In tests of the pessimism hypothesis, we compare survey expectations with returns or price growth of stock market indices. In case of the CFO survey, we match the

[7]The periodically updated data is available at `http://www.federalreserve.gov/pubs/feds/2006`

survey expectations with the total return on the S&P500 index over the one-year period starting from the reference date. For the UBS/Gallup survey, we use the return of the same series over the one-year period starting from first day of interview period. For the monthly extended UBS/Gallup series, we use a one-year total return on the S&P500 index from the end of the month prior to the survey month. For the Shiller surveys, we use price growth on the DJIA realized over the relevant horizon starting from the date of individual response.

Having computed the wedges between survey expectations and returns, we then aggregate the data to time series by computing means or medians within months or quarters as shown in the second and seventh column of Table 1.

### 3.3 Matching with conditioning variables

In several tests, we use the price-dividend (P/D) ratio as a conditioning variable (RN-C, PE-C1, PE-C2). We use the S&P500 P/D ratio for the CFO and UBS surveys, the CRSP value-weighted index P/D ratio for the UBS extended series and the PE-C2 test with the Shiller survey, and the DJIA P/D ratio for the PE-C1 test with the Shiller surveys. In terms of timing, we use the P/D ratio measured at the end of the last month preceding survey reference date for the CFO survey, at the end of the last month preceding first day of interview period for the UBS survey, and at the end of the last month preceding the date of the individual response for the Shiller surveys.

Furthermore, some tests use an estimate of objective conditional expected returns $\hat{E}_t[R_{t+1}]$ that we construct from a predictive regression with the price dividend ratio (PE-U2, PE-C2). We use monthly returns or price growth and the S&P500 P/D ratio for the CFO survey and the CRSP value-weighted index P/D ratio for all the other tests. When we construct the fitted value $\hat{E}_t[R_{t+1}]$, we do so using a P/D ratio that is timed relative to the survey date in the same way as explained above, with the exception of the Shiller survey where we use the P/D ratio at the end of the quarter preceding the interview quarter. We construct multi-period return forecasts by using multi-period realized returns on the left-hand side of equation (13).

## 4 Empirical Results

### 4.1 Risk-neutral expectations hypothesis

Table 2 reports results for the RN-U test based on equation (6) using Treasury rates as risk-free rates. This test looks at the most basic implication of the risk-neutral expectations hypothesis: Are subjective expected returns on average equal to the risk-free rate over the forecast horizon? As the table shows, the answer is a clear no. For surveys from all sources and horizons except the Shiller individual investor survey using medians over a 10-year horizon, the subjective expected returns elicited in the surveys exceed risk-free rates by several percentage points. As the $t$-statistics and $p$-values show, we can reject the risk-neutral expectations hypothesis at extremely high levels of significance. There is also a remarkable degree of consistency across different types of survey respondents. Subjective expected returns exceed risk-free rates, contradicting the risk-neutral expectations hypothesis, for business practitioners (CFOs), professional investors, wealthy individuals, and individual investors.

The risk-neutral expectations hypothesis not only implies that the unconditional average subjective expected excess return is equal to the risk free rate, but also that this equality holds conditionally, state by state. The RN-C tests based on (7) reported in Table 3 shed light on this conditional version of the risk-neutral expectations hypothesis. We use Treasury rates as risk-free rates and the P/D ratio as the regressor $x_t$ that could drive time-variation in subjective expected excess returns under the alternative hypothesis. We obtain small-sample bias-adjusted coefficient estimates and simulate $F$-statistics under the null hypothesis of risk-neutral expectations, as described in Appendix B.

The main test of the risk-neutral expectations hypothesis is the test of the joint hypothesis $a_0 = 0$ and $a_1 = 0$. As the results in Table 3 show, this hypothesis is rejected at the 5% level for all but 4 of the 27 survey series. Like the unconditional tests in Table 2 the conditional tests here indicate that there is a substantial wedge between the subjective expectations of returns and risk-free rates.

To what extent does this wedge vary with the P/D ratio? The bias-adjusted point estimates of $a_1$ point to an interesting difference between individual and professional investors. For almost all survey series with individual investor respondents and CFOs, the estimates of $a_1$ either indicate a statistically significant positive relationship of subjective expected excess returns to the P/D ratio, or the estimates are not significantly different from zero. In contrast, the subjective expected excess returns of professional investors in the Shiller survey are, with one exception (mean series at 10-year horizon), negatively related to the P/D ratio at the 5% significance level. For both individuals, CFOs, and professionals, however, the joint hypothesis of $a_0 = 0$ and $a_1 = 0$ is overwhelmingly rejected.

The statistically weak relationship between subjective expectations and the P/D ratio in Table 3 contrasts with the much stronger relationship documented in Greenwood and Shleifer (2014). The key difference is that here we examine subjective expected *excess* returns while Greenwood and Shleifer use expectations of total re-

Table 2: Unconditional Test of the Risk-Neutral Expectations Hypothesis

This table presents tests of the RN-U hypothesis. The column labeled $a$ reports the mean of the subjective expected return (in terms of percent) in excess of the risk-free rate (based on Treasury securities) over the relevant horizon. The $t$-statistics and $p$-values are based on a Newey-West estimator with 4 lags for quarterly data and 12 lags for monthly data.

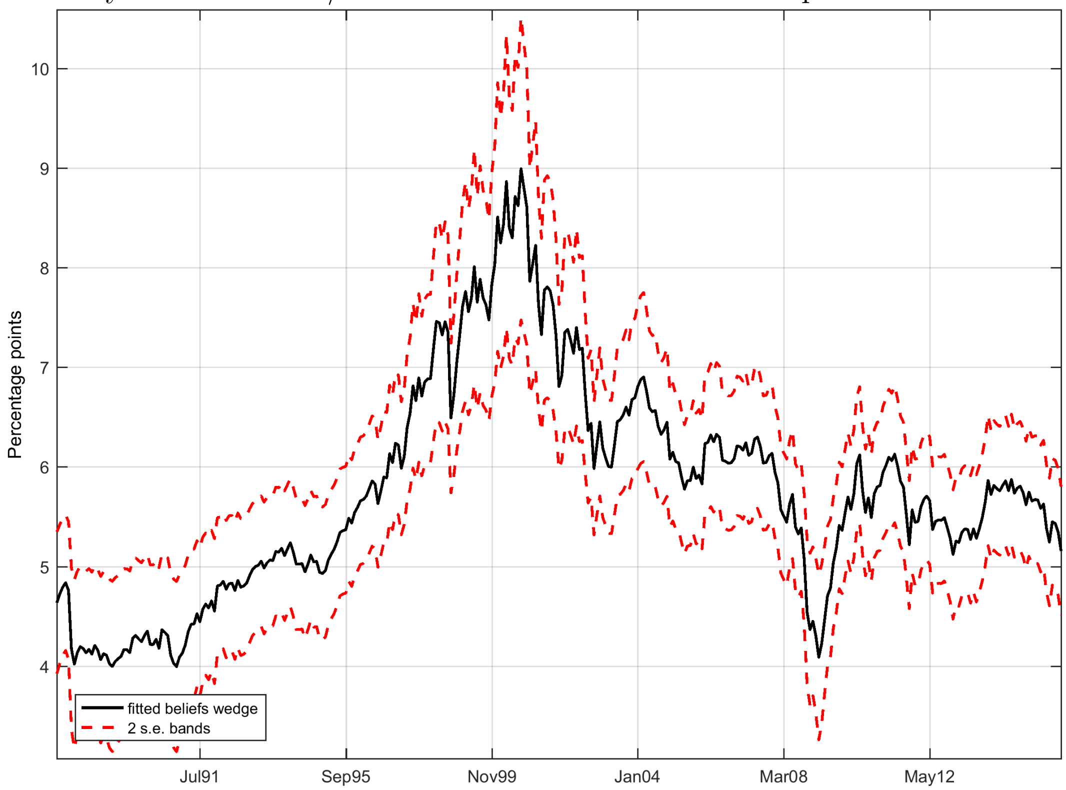
| Survey Source | | | $a$ | $t$-statistic | $p$-value for $H_0 : a = 0$ |
|---|---|---|---|---|---|
| CFO | | mean | 3.89 | 9.47 | 0.0000 |
| | | median | 3.55 | 8.43 | 0.0000 |
| UBS own | all | mean | 6.55 | 12.53 | 0.0000 |
| | | median | 4.52 | 9.99 | 0.0000 |
| | >100k | mean | 6.40 | 12.36 | 0.0000 |
| | | median | 4.79 | 10.82 | 0.0000 |
| UBS market | all | mean | 6.64 | 13.31 | 0.0000 |
| | | median | 4.61 | 14.13 | 0.0000 |
| | >100k | mean | 6.36 | 12.29 | 0.0000 |
| | | median | 4.74 | 15.80 | 0.0000 |
| UBS extended | | | 5.80 | 20.10 | 0.0000 |
| Shiller individual | 3m | mean | 1.00 | 4.71 | 0.0000 |
| | | median | 1.45 | 6.73 | 0.0000 |
| | 6m | mean | 2.29 | 7.98 | 0.0000 |
| | | median | 2.86 | 11.22 | 0.0000 |
| | 1yr | mean | 5.02 | 9.26 | 0.0000 |
| | | median | 5.81 | 13.14 | 0.0000 |
| | 10yr | mean | 8.90 | 2.34 | 0.0194 |
| | | median | -5.33 | -1.19 | 0.2341 |
| Shiller professional | 3m | mean | 0.68 | 2.28 | 0.0223 |
| | | median | 1.48 | 5.19 | 0.0000 |
| | 6m | mean | 2.16 | 3.82 | 0.0001 |
| | | median | 3.86 | 8.94 | 0.0000 |
| | 1yr | mean | 5.24 | 5.23 | 0.0000 |
| | | median | 7.43 | 8.41 | 0.0000 |
| | 10yr | mean | 42.47 | 10.79 | 0.0000 |
| | | median | 28.88 | 7.88 | 0.0000 |

Table 3: Conditional Test of the Risk-Neutral Expectations Hypothesis

This table presents tests of the RN-C hypothesis where we regress subjective expected returns in excess of Treasury rates on the lagged price-dividend ratio. The columns labeled $a_0$ and $a_1$ report the intercept and slope coefficients, respectively, from these regressions. We report the $a_1$ estimates multiplied by a factor of 1000. They are bias-adjusted for small samples, as described in Appendix B. The last two columns report Monte Carlo $p$-values obtained by simulating $F$-statistics under the null hypothesis, as described in Appendix B.

| Survey Source | | | $a_0$ | $a_1$ $\cdot 10^3$ | $p$-value for $H_0 : a_0 = a_1 = 0$ | $p$-value for $H_0 : a_1 = 0$ |
|---|---|---|---|---|---|---|
| CFO | | mean | -1.39 | 8.66 | 0.0044 | 0.2307 |
| | | median | 0.08 | 5.48 | 0.3937 | 0.2628 |
| UBS own | all | mean | 1.09 | 6.57 | 0.0000 | 0.0589 |
| | | median | -2.65 | 8.92 | 0.1860 | 0.0057 |
| | >100k | mean | 0.02 | 7.86 | 0.0000 | 0.0302 |
| | | median | -2.03 | 8.48 | 0.0106 | 0.0160 |
| UBS market | all | mean | -0.52 | 7.96 | 0.0000 | 0.0059 |
| | | median | -0.91 | 6.17 | 0.0000 | 0.0221 |
| | >100k | mean | -1.72 | 9.00 | 0.0000 | 0.0148 |
| | | median | -0.36 | 5.66 | 0.0000 | 0.2243 |
| UBS extended | | | 2.24 | 5.79 | 0.0000 | 0.0002 |
| Shiller individual | 3m | mean | 0.03 | 1.95 | 0.0029 | 0.4759 |
| | | median | 0.04 | 2.66 | 0.0000 | 0.0852 |
| | 6m | mean | 2.61 | -0.47 | 0.0000 | 0.8115 |
| | | median | 3.66 | -1.43 | 0.0000 | 0.5086 |
| | 1yr | mean | 10.46 | -9.77 | 0.0000 | 0.0141 |
| | | median | 9.18 | -6.02 | 0.0001 | 0.5184 |
| | 10yr | mean | 34.51 | -50.86 | 0.9759 | 0.4601 |
| | | median | -1.58 | -11.40 | 0.8236 | 0.8359 |
| Shiller professional | 3-months | mean | 4.26 | -6.45 | 0.0005 | 0.0068 |
| | | median | 4.99 | -6.34 | 0.0000 | 0.0015 |
| | 6m | mean | 9.48 | -13.20 | 0.0000 | 0.0001 |
| | | median | 9.58 | -10.19 | 0.0000 | 0.0005 |
| | 1yr | mean | 19.51 | -25.89 | 0.0000 | 0.0052 |
| | | median | 20.39 | -23.54 | 0.0000 | 0.0016 |
| | 10yr | mean | 83.88 | -76.17 | 0.0000 | 0.0844 |
| | | median | 76.47 | -87.97 | 0.0000 | 0.0200 |

Figure 1: Fitted values from regression (7) of subjective expected returns in excess of Treasury rates on the P/D ratio in the UBS extended sample




turns as dependent variable. During our sample period, the P/D ratio and Treasury rates are positively correlated and hence subtracting the risk-free rate from the subjective expected return weakens the positive relationship with the P/D ratio, see also the analysis in Adam, Marcet and Beutel (2017), which presents regression estimates for both excess returns and plain returns.

However, for the UBS extended series, where we have the longest time series, and hence more statistical power than for the shorter series, we can still reject the hypothesis $a_1 = 0$ with a high level of statistical confidence. Figure 1 shows the fitted values of subjective expected returns in excess of Treasury rates based on the bias-adjusted point estimates of $a_0$ and $a_1$ along with two-standard-error bands from the RN-C test regression (7) for the UBS extended sample. The figure shows that the estimated subjective conditional expected excess return is, in conflict with the predictions of the risk-neutral expectations hypothesis, far above zero throughout the whole sample. The lower boundary of the two-standard-error bands never include zero anywhere. Since the survey expectations in a given period likely contain substantial measurement error, the projection in Figure 1 might provide a better description of the time-series dynamics of the true subjective expectations than the raw measures of subjective expected excess returns. But even the raw series of $\mathcal{E}_t[R_{t+1}] - R_t^f$ never dips below zero.

The results we have presented so far could potentially be rationalized under the risk-neutral return hypothesis if individuals are borrowing constrained and face very high shadow interest rates, with the consequence that their risk-neutral expectations of stock returns could be substantially higher than Treasury rates. However, the fact that the RN-U and RN-C tests reject the risk-neutral expectations hypothesis for the subsample of wealthy investors in the UBS survey, the Shiller individual investor survey (which is also based on a sample of wealthy individuals), CFOs, and professional investors in the Shiller survey cast doubt on this alternative explanation. First, for these samples of wealthy investors and professionals, the borrowing constraints story does not appear plausible—especially if it requires shadow rates that are four or more percentage points above Treasury rates. Second, the point estimates of $a$ for the CFO, UBS wealthy, and Shiller individual survey at the 1-year horizon in Table 2 are very similar to the $a$ estimated from the full UBS sample, which suggests that the on average less wealthy individuals in the full UBS sample are not systematically different from wealthy investors and professionals in terms of their subjective expected return.

A more subtle friction-based explanation could be that there is a wedge between borrowing and lending rates. If the borrowing margin is relevant for many households, the Treasury rates that we have used so far may not be the relevant interest rates in households' Euler equation (2). Instead, borrowing rates may be more relevant. For this reason, we re-run the RN-U and RN-C tests with subjective expected excess returns calculated relative to one-year adjustable mortgage rates. We use a collateralized borrowing rate rather than an unsecured borrowing rate to avoid contamination by a substantial credit spread.

Table 4 presents the results. The basic message from this table is that there isn't much difference to the earlier tests with excess returns relative to Treasury rates. For example, in the unconditional test, we still reject the risk-neutral expectations hypothesis for all (one-year horizon) series, just as we did in Table 2. The point estimates for $a$ are slightly smaller, but all of them are still substantially greater than zero by about one to five percentage points. Similarly, for the conditional test of the hypothesis $a_0 = a_1 = 0$, we reject the null at a 5% level for all but one of the survey expectations series. Overall, taking into account the potential effects of differences in borrowing and lending rates doesn't help much to rescue the risk-

Table 4: Unconditional and Conditional Tests of the Risk-Neutral Expectations Hypothesis: Adjustable Mortgage Rates as Risk-free Rate

This table presents tests of the RN-U and RN-C hypotheses using one-year adjustable mortgage rates as risk-free rate to compute excess returns. We use only survey series where the forecasting horizon is one year. For the unconditional test RN-U, the column labeled $a$ reports the mean of the subjective expected excess return. The corresponding $p$-values are based on a Newey-West estimator with 4 lags for quarterly data and 12 lags for monthly data. For the conditional test RN-C, the columns labeled $a_0$ and $a_1$ report the intercept and slope coefficients, respectively, from regressions of subjective expected excess returns on the lagged price-dividend ratio. We report the $a_1$ estimates multiplied by a factor of 1000. They are bias-adjusted for small samples, as described in Appendix B. The last two columns report Monte Carlo $p$-values obtained by simulating $F$-statistics under the null hypothesis, as described in Appendix B.

| Survey Source | | | Unconditional | | Conditional | | | |
|---|---|---|---|---|---|---|---|---|
| | | | $a$ | $p$-val. $H_0:$ $a=0$ | $a_0$ | $a_1$ $\cdot 10^3$ | $p$-val. $H_0:$ $a_0=a_1=0$ | $p$-val. $H_0:$ $a_1=0$ |
| CFO | | mean | 1.65 | 0.0000 | -4.63 | 9.87 | 0.0086 | 0.0220 |
| | | median | 1.24 | 0.0008 | -4.05 | 8.37 | 0.0087 | 0.0287 |
| UBS own | all | mean | 5.13 | 0.0000 | -0.98 | 7.29 | 0.0012 | 0.0747 |
| | | median | 3.10 | 0.0000 | -4.92 | 9.95 | 0.0394 | 0.0049 |
| | >100k | mean | 4.98 | 0.0000 | -2.18 | 8.76 | 0.0001 | 0.1043 |
| | | median | 3.37 | 0.0000 | -4.29 | 9.52 | 0.0557 | 0.0033 |
| UBS market | all | mean | 5.03 | 0.0000 | -4.19 | 10.00 | 0.0002 | 0.0045 |
| | | median | 3.00 | 0.0001 | -7.20 | 11.35 | 0.0019 | 0.1163 |
| | >100k | mean | 4.74 | 0.0000 | -6.82 | 12.77 | 0.0000 | 0.0007 |
| | | median | 3.13 | 0.0000 | -7.13 | 11.39 | 0.0010 | 0.0313 |
| UBS extended | | | 4.28 | 0.0000 | 0.68 | 5.67 | 0.0000 | 0.0133 |
| Shiller individual | 1yr | mean | 3.09 | 0.0000 | 1.79 | 2.50 | 0.0000 | 0.6854 |
| | | median | 3.66 | 0.0000 | 3.28 | 0.79 | 0.0000 | 0.7146 |
| Shiller professional | 1yr | mean | 3.11 | 0.0001 | 14.02 | -19.76 | 0.0028 | 0.0693 |
| | | median | 5.33 | 0.0000 | 14.79 | -17.19 | 0.0000 | 0.0256 |

neutral expectations hypothesis.

## 4.2 Pessimism hypothesis

The results so far suggest that subjective expectations of returns exceed risk free rates by substantial amounts, which is inconsistent with the risk-neutral expectations hypothesis. However, this still leaves the possibility that their expectations are pessimistically biased relative to objective expectations of returns under the real-world probability measure. The alternative is that survey respondents simply do what the survey asks them to do: provide the expected return under their perceived real-world probability measure. In this section, we report results from tests of this pessimistic expectations hypothesis.

Table 5 reports the results from unconditional tests. The set of columns labeled PE-U1 presents results from estimating equation (12). A positive (negative) estimate for the coefficient $b$ indicates that the subjective return expectation exceeds on average (falls on average short of) the realized returns. The table also reports $p$-values for a one-sided test of the weak pessimism hypothesis $b \leq 0$. It shows that one cannot reject weak pessimism for more than two thirds of the subjective expectations series. However, the weak pessimism hypothesis includes $b = 0$ and only one third of the $t$-statistics turn out to be negative. Moreover, none of the $t$-statistics would allow to reject the alternative null hypothesis of unconditional optimism. If anything, there is thus a tendency towards unconditional optimism rather than pessimism. Overall, there is little evidence of deviations from unconditional unbiasedness of return expectations. In particular, the mean bias is also not significantly different from zero for the UBS extended series for which we have the longest time series and (among the 1-year horizon series) the smallest standard error and highest statistical power. Subjective expected returns thus appear to be, on average, close to unbiased, which is inconsistent with a substantial pessimism bias.

One concern with the PE-U1 tests could be that realized returns are just too noisy to provide much statistical power when we compare subjective expected returns with realized returns, especially given the relatively short sample for which survey expectations are available. This issue is addressed in the set of columns labeled PE-U2 in table 5, which compares the subjective return expectations with fitted values from a regression of realized returns on the lagged dividend yield estimated over the period 1926-2017. The reported estimates in table 5 show that almost all point estimates of $b$ move closer to zero. The absolute value of the $t$-statistics, however, are often bigger than with PE-U1 test because the standard errors of the estimation

Table 5: Unconditional Tests of the Pessimism Hypothesis

This table presents unconditional tests of the pessimism (PE-U) hypothesis. The columns labeled with PE-U1 report the mean of the subjective expected return in excess of the realized return over the forecast horizon along with the associated $t$-statistic and $p$-value based on a Newey-West estimator with 4 lags for quarterly data and 12 lags for monthly data. The columns labeled with PE-U2 report the mean of the subjective expected return in excess of the fitted value from a regression of the relevant return or price growth for each survey on the lagged P/D ratio, with regression parameters estimated with data from 1926-2017. The $t$-statistics and $p$-values in this case are computed based on the asymptotic approximation outlined in Appendix A, including a Newey-West estimator of the covariance matrix using 4 lags for quarterly data and 12 lags for monthly data.

| | | | PE-U1 | | | PE-U2 | | |
|---|---|---|---|---|---|---|---|---|
| Survey Source | | | $b$ | $t$-stat. | $p$-val. $H_0: b \leq 0$ | $b$ | $t$-stat. | $p$-val. $H_0: b \leq 0$ |
| CFO | | mean | -1.61 | -0.43 | 0.6663 | -0.60 | -0.37 | 0.6425 |
| | | median | -3.69 | -1.05 | 0.8526 | -1.70 | -1.11 | 0.8663 |
| UBS own | all | mean | 7.59 | 1.62 | 0.0526 | 7.09 | 2.79 | 0.0026 |
| | | median | 5.56 | 1.21 | 0.1141 | 5.06 | 2.06 | 0.0199 |
| | >100k | mean | 13.69 | 2.01 | 0.0222 | 6.94 | 2.73 | 0.0031 |
| | | median | 12.07 | 1.81 | 0.0351 | 5.33 | 2.17 | 0.0148 |
| UBS market | all | mean | 13.97 | 2.07 | 0.0193 | 9.84 | 3.17 | 0.0008 |
| | | median | 11.94 | 1.79 | 0.0366 | 7.81 | 2.64 | 0.0042 |
| | >100k | mean | 13.69 | 2.01 | 0.0222 | 9.55 | 3.06 | 0.0011 |
| | | median | 12.07 | 1.81 | 0.0351 | 7.94 | 2.70 | 0.0035 |
| UBS extended | | | -1.86 | -0.72 | 0.7636 | 2.10 | 1.06 | 0.1454 |
| Shiller individual | 3-months | mean | -0.55 | -0.68 | 0.7513 | -0.51 | -1.01 | 0.8428 |
| | | median | -0.31 | -0.38 | 0.6497 | -0.07 | -0.14 | 0.5563 |
| | 6m | mean | -0.01 | -0.01 | 0.5030 | -0.61 | -0.69 | 0.7548 |
| | | median | 0.41 | 0.25 | 0.4030 | -0.04 | -0.05 | 0.5192 |
| | 1yr | mean | 0.52 | 0.17 | 0.4323 | 0.28 | 0.18 | 0.4290 |
| | | median | 1.11 | 0.36 | 0.3578 | 0.87 | 0.57 | 0.2841 |
| | 10yr | mean | 11.53 | 0.50 | 0.3088 | 8.26 | 0.48 | 0.3158 |
| | | median | -2.01 | -0.08 | 0.5326 | -6.73 | -0.36 | 0.6389 |
| Shiller professional | 3-months | mean | -0.59 | -0.62 | 0.7334 | -0.83 | -1.69 | 0.9542 |
| | | median | 0.05 | 0.05 | 0.4794 | -0.04 | -0.08 | 0.5301 |
| | 6m | mean | -0.08 | -0.05 | 0.5181 | -0.74 | -0.86 | 0.8044 |
| | | median | 1.12 | 0.65 | 0.2587 | 0.96 | 1.17 | 0.1216 |
| | 1yr | mean | 0.58 | 0.20 | 0.4218 | 0.31 | 0.22 | 0.4118 |
| | | median | 2.60 | 0.89 | 0.1864 | 2.56 | 1.86 | 0.0318 |
| | 10yr | mean | 43.18 | 2.27 | 0.0115 | 40.35 | 2.77 | 0.0028 |
| | | median | 25.81 | 1.43 | 0.0761 | 25.61 | 1.93 | 0.0268 |

become smaller, consistent with an increase in statistical power.[8] Overall, the picture remains mixed. Some estimates of $b$ are larger than zero, while some are smaller, but again with the tendency towards optimism being slightly more prevalent: we can reject the weak pessimism hypothesis at the 5% level for 11 of the 27 series, while the weak optimism hypothesis is rejected in only one instance. By and large, this test again suggests that subjective expectations are in unconditional terms not far from being unbiased.

Table 6 presents conditional tests of the pessimism hypothesis. Using the estimated coefficients from equations (17) and (18), it tests for all observed price-dividend and expectations pairs, whether one can reject at the 5% level conditional pessimism (thus implying optimistic expectations) or reject conditional optimism (thus implying conditional pessimism). It then reports the share of observations for which such rejections can be achieved. For the remaining share of observations, i.e., one minus the reported share of rejections of optimism and pessimism, no definite conclusion can be reached at the considered significance level.

The PE-C1 version of the test in Table 6, which is based on equation (17), looks at the predicted wedge between subjective expected returns and subsequently realized returns; the PE-C2 version of the test, which is based on equation (18), calculates the wedge as the difference between the subjective expected return and the fitted value from a predictive regression of realized returns, where the predictive regression uses the lagged price-dividend ratio and is estimated over the period 1926-2017.

The first notable result from Table 6 is that many observations can neither be classified as optimistic nor as pessimistic. From a purely statistical point of view, subjective expected returns thus are many times in the vicinity of objective expected returns. Nevertheless, for a substantial share of observations we can reject return pessimism. This is especially true for the UBS surveys. For other surveys, e.g., the Shiller individual surveys, the share of rejections of optimism and pessimism are roughly balanced (PE-C1) or tilted in favor of rejecting optimism (PE-C2). Overall, whether tests reject more often optimism or more often pessimism appears to depend on the survey source. Since surveys cover different sample periods, this suggests that the direction of rejections depends on the sample years. This conjecture is supported by Figure 2, which uses the UBS extended sample and reports the fitted values of subjective expected returns in excess of the estimated objective expected returns on the P/D ratio, along with two-standard-error bands from the PE-C2 test regression (18). The deviations of subjective expected returns from estimated objected expected returns is pro-cyclical. In boom times, like the late 1990s, investors

[8]This increase in statistical power occurs despite the fact that the standard errors are adjusted for the estimation uncertainty coming from the first-stage predictive regression.

Table 6: Conditional Tests of the Pessimism Hypothesis

This table presents tests of the PE-C hypothesis where we regress on the lagged P/D ratio the subjective expected returns in excess of realized returns (PE-C1) or the subjective expected return in excess of the fitted value from a regression of realized returns on the lagged P/D ratio (PE-C2), where the last regression is estimated over the period 1926-2017. Based on these regression estimates, we determine, at every point in time, whether we can reject the pessimism hypothesis (weakly negative predicted subjective excess return) or the optimism hypothesis (weakly positive predicted subjective excess return) at a 5% level. In the PE-C1 case, the regressions are bias-adjusted as described in Appendix B. In the PE-C2 case, the $t$-statistics for this test are computed based on the asymptotic approximation outlined in Appendix A, including a Newey-West estimator of the covariance matrix using 4 lags for quarterly data and 12 lags for monthly data.

| | | | PE-C1 | | PE-C2 | |
|---|---|---|---|---|---|---|
| | | | Reject pessimism | Reject optimism | Reject pessimism | Reject optimism |
| | | | (share of observations) | | (share of observations) | |
| CFO | | mean | 0.1212 | 0.1667 | 0.0968 | 0.2258 |
| | | median | 0.0645 | 0.1774 | 0.0172 | 0.3621 |
| UBS own | all | mean | 0.3504 | 0.0256 | 0.3714 | 0.0000 |
| | | median | 0.3419 | 0.0256 | 0.2190 | 0.0000 |
| | >100k | mean | 0.3504 | 0.0256 | 0.3429 | 0.0000 |
| | | median | 0.3504 | 0.0256 | 0.2286 | 0.0000 |
| UBS market | all | mean | 0.8889 | 0.0000 | 0.5686 | 0.0000 |
| | | median | 0.5714 | 0.0000 | 0.2745 | 0.0000 |
| | >100k | mean | 0.4444 | 0.0000 | 0.5294 | 0.0000 |
| | | median | 0.4762 | 0.031 | 0.2941 | 0.0000 |
| UBS extended | | | 0.1563 | 0.3125 | 0.1335 | 0.1278 |
| Shiller indiv. | 3-months | mean | 0.1692 | 0.5538 | 0.0000 | 0.2769 |
| | | median | 0.1846 | 0.4154 | 0.0000 | 0.1231 |
| | 6m | mean | 0.2154 | 0.2154 | 0.0000 | 0.2615 |
| | | median | 0.2308 | 0.1692 | 0.0000 | 0.1231 |
| | 1yr | mean | 0.1846 | 0.2154 | 0.0462 | 0.1538 |
| | | median | 0.0154 | 0.7692 | 0.1077 | 0.0923 |
| | 10yr | mean | 0.1077 | 0.1077 | 0.2286 | 0.4857 |
| | | median | 0.7077 | 0.0000 | 0.0857 | 0.4286 |
| Shiller prof. | 3-months | mean | 0.1231 | 0.4000 | 0.0000 | 0.0769 |
| | | median | 0.1538 | 0.0000 | 0.0000 | 0.0000 |
| | 6m | mean | 0.1846 | 0.2308 | 0.0000 | 0.0000 |
| | | median | 0.1385 | 0.0923 | 0.0000 | 0.0000 |
| | 1yr | mean | 0.1846 | 0.0308 | 0.0000 | 0.0000 |
| | | median | 0.1846 | 0.0615 | 0.0000 | 0.0000 |
| | 10yr | mean | 0.2615 | 0.3077 | 0.4571 | 0.0000 |
| | | median | 0.1846 | 0.6615 | 0.3143 | 0.0000 |

Figure 2: Fitted values from regression (18) of subjective expected returns in excess of the estimated objective expected returns on the P/D ratio in the UBS extended sample. The objective expected returns are estimated from regression (13) of realized returns on P/D ratio over a longer sample 1926-2017.


are too optimistic. Following crashes, like in early 2009, investors are too pessimistic. The UBS/Gallup survey includes observations from the late 90s, but no observations from the financial crisis and its aftermath. In contrast, the Shiller surveys include observations from the financial crisis and subsequent years. This partly explains why these other surveys on average have a lower share of optimistic observations than the UBS/Gallup survey.

Overall, the results in Tables 5 and 6 show that while subjective expectations are in unconditional terms close to unbiased, there is a substantial time-varying conditional bias. This conditional bias flips sign and is often an optimism bias, inconsistent with the pessimistic expectations hypothesis.

# 5 Conclusion

Our empirical findings show that subjective stock return expectations from a number of different surveys are not consistent with the idea that survey respondents report expectations under a risk-neutral probability measure. We show that both the unconditional and conditional properties of subjective return expectations deviate substantially from the prediction of the risk-neutral expectations hypothesis which predicts that subjective expected returns are equal to a maturity-matched risk-free rate. Allowing for differences in borrowing and lending rates or restricting the sample to individuals who are unlikely to be borrowing-constrained does not change this basic conclusion.

More generally, we don't find evidence that individuals report "risk-adjusted" expectations that are pessimistically distorted relative to the empirical distribution of stock returns. Unconditionally, average subjective expected returns are close to average realized returns without a significant bias. Conditionally, there are substantial deviations of subjective expected returns from the objective expected returns generated by empirical predictability regressions, but these deviations are optimistic in some periods and pessimistic in others, and they cancel out on average.

Our results therefore suggest that the predictable time-variation in the subjective expectations error around this unconditional mean of approximately zero is the most interesting property of these aggregated return expectations series that research should further investigate. For example, learning from experience (Malmendier and Nagel (2011, 2016)), return extrapolation (Barberis et al. (2015)), or learning from price growth (Adam et al. (2016, 2017)) could contribute to these time-varying subjective expectations errors.

# References


Adam, K., A. Marcet, and J. Beutel (2017): "Stock Price Booms and Expected Capital Gains," *Amercian Economic Review*, Vol. 107 No. 8, 2352–2408.

Adam, K., A. Marcet, and J. P. Nicolini (2016): "Stock Market Volatility and Learning," *Journal of Finance*, 71(1), 33–82.

Ameriks, J., G. Kézdi, M. Lee, and M. D. Shapiro (2016): "Heterogeneity in Expectations, Risk Tolerance, and Household Stock Shares," Discussion paper, University of Michigan.

Bacchetta, P., E. Mertens, and E. V. Wincoop (2009): "Predictability in Financial Markets: What Do Survey Expectations Tell Us?," *Journal of International Money and Finance*, 28, 406–426.

Bansal, R., and A. Yaron (2004): "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 59, 1481–1509.

Barberis, N., R. Greenwood, L. Jin, and A. Shleifer (2015): "X-CAPM: An Extrapolative Capital Asset Pricing Model," *Journal of Financial Economics*, 115(1), 1–24.

Bhandari, A., J. Borovička, and P. Ho (2016): "Identifying Ambiguity Shocks in Business Cycle Models Using Survey Data," Discussion paper, National Bureau of Economic Research.

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

Cochrane, J. H. (2011): "Presidential Address: Discount Rates," *Journal of Finance*, LXVI, 1047–1108.

——— (2017): "Macro-Finance," *Review of Finance*, 21, 945–985.

Collin-Dufresne, P., M. Johannes, and L. A. Lochstoer (2016): "Parameter Learning in General Equilibrium: The Asset Pricing Implications," *American Economic Review*, 106, 664–698.

Greenwood, R., and A. Shleifer (2014): "Expectations of returns and expected returns," *Review of Financial Studies*, 27(3), 714–746.

Gürkaynak, R. S., B. Sack, and J. H. Wright (2007): "The U.S. Treasury yield curve: 1961 to the present," *Journal of Monetary Economics*, 54(8), 2291 – 2304.

Guvenen, F. (2009): "A Parsimonious Macroeconomic Model for Asset Pricing," *Econometrica*, Vol. 77.

Hansen, L. P., and T. J. Sargent (2001): "Acknowledging Misspecification in Macroeconomic Theory," *Review of Economic Dynamics*, 4(3), 519–535.

Malmendier, U., and S. Nagel (2011): "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking," *Quarterly Journal of Economics*, 126, 373–416.

——— (2016): "Learning from Inflation Experiences," *Quarterly Journal of Economics*, 131(1), 53–87.

Manski, C. F. (2004): "Measuring Expectations," *Econometrica*, 72(5), 1329–1376.

——— (2017): "Survey Measurement of Probabilistic Macroeconomic Expectations: Progress and Promise," in *NBER Macroeconomics Annual 2017*, ed. by M. S. Eichenbaum, and J. Parker, vol. 32. University of Chicago Press.

Nagel, S., and Z. Xu (2018): "Asset Pricing with Fading Memory," Discussion paper, University of Chicago.

Stambaugh, R. F. (1999): "Predictive Regressions," *Journal of Financial Economics*, 54, 375–421.

Vissing-Jorgensen, A. (2003): "Perspectives on Behavioral Finance: Does "Irrationality" Disappear with Wealth? Evidence from Expectations and Actions," in *2003 Macroeconomics Annual*, Boston. NBER.

Wooldridge, J. M. (2001): *The Econometrics of Cross-Section and Panel Data*. MIT Press, Cambridge MA.

# Appendix

## A Asymptotic distribution of test statistics

This appendix describes asymptotic distributions of test statistics that take into account uncertainty due to the generated regressor in the tests PE-U2 and PE-C2.

### A.1 Test PE-U2

Using $\hat{E}_t$ as short form for $\hat{E}_t[R_{t+1}]$ and $\mathcal{E}_t$ for $\mathcal{E}_t[R_{t+1}]$, we can rewrite (14) as

$$\mathcal{E}_t = b + \hat{E}_t + e_t - (\hat{E}_t - E_t) \tag{A.1}$$

and we can treat this as a generated regressors problem with the coefficient on $\hat{E}_t$ known and equal to one.

With the first stage regression written as

$$R_{t+1} = Z_t'k + u_{t+1} \tag{A.2}$$

where $Z_t \equiv (1, z_t')'$, $k = (k_0, k_1')'$ and $\hat{E}_t = Z_t'\hat{k}$.

The derivation follows Wooldridge (2001), Appendix 6A.

Let $T_1$ be the sample size of the survey data and $T_2$ the size of the sample used to estimate $\hat{E}_t[R_{t+1}]$, with $T_2 > T_1$. When we take limits, $T_1 \to \infty$, we keep $\phi = T_1/T_2$ fixed, so that $T_2 \to \infty$ at the same rate.

Usual OLS calculations yield

$$\sqrt{T_1}(\hat{b} - b) = T_1^{-1/2} \sum_{t=T_2-T_1+1}^{T_2} \left[e_t - (\hat{E}_t - E_t)\right] \tag{A.3}$$

where

$$T_1^{-1/2} \sum_{t=T_2-T_1+1}^{T_2} (\hat{E}_t - E_t) = \frac{1}{T_1} \frac{\sqrt{T_1}}{\sqrt{T_2}} \left( \sum_{t=T_2-T_1+1}^{T_2} Z_t' \right) \sqrt{T_2}(\hat{k} - k) \tag{A.4}$$

$$= \sqrt{\phi} G \sqrt{T_2}(\hat{k} - k) + o_p(1) \tag{A.5}$$

and where $G = E[Z']$ is the probability limit of $T_1^{-1} \sum_{t=T_2-T_1+1}^{T_2} Z_t'$, and $o_p(1)$ is a sequence that converges in probability to 0 when $T_1 \to \infty$.

We have

$$\sqrt{T_2}(\hat{k}-k) = \hat{C}^{-1}T_2^{-1/2}\sum_{t=1}^{T_2} Z_t u_{t+1} \tag{A.6}$$

$$= \hat{C}^{-1}T_2^{-1/2}\left(\sum_{t=T_2-T_1+1}^{T_2} Z_t u_{t+1} + \sum_{t=1}^{T_2-T_1} Z_t u_{t+1}\right) \tag{A.7}$$

with

$$\hat{C} = \frac{1}{T_2}\sum_{t=1}^{T_2} Z_t Z_t', \qquad \text{where} \qquad \hat{C} \xrightarrow{p} E[Z_t Z_t'] \tag{A.8}$$

Substituting back into (A.3) and denoting $C \equiv E[Z_t Z_t']$, we get

$$\sqrt{T_1}(\hat{b}-b) = T_1^{-1/2}\left\{\sum_{t=T_2-T_1+1}^{T_1}\left[e_t - \phi G C^{-1} Z_t u_{t+1}\right]\right\}$$
$$-(T_2-T_1)^{-1/2}\sqrt{\phi(1-\phi)}GC^{-1}\left(\sum_{t=T_2-T_1+1}^{T_2} Z_t u_{t+1}\right) + o_p(1) \tag{A.9}$$

By the central limit theorem, as $T_1 \to \infty$, with $\phi$ fixed, $\sqrt{T_1}(\hat{b}-b)$ is asymptotically normal.

In the simplest case with $Z_t u_{t+1}$ and $e_t$ uncorrelated at all leads and lags, and with $e_t$ serially uncorrelated, we get an asymptotic variance of $\sqrt{T_1}(\hat{b}-b)$ of

$$\text{Var}[e_t] + \phi G C^{-1} E[Z_t Z_t' u_{t+1}^2] C^{-1} G'. \tag{A.10}$$

With correlation between $e_t$ and $Z_t u_{t+1}$, the asymptotic variance becomes

$$\text{Var}[e_t] + \phi G C^{-1} E[Z_t Z_t' u_{t+1}^2] C^{-1} G' - 2\sqrt{\phi} G C^{-1}\text{Cov}(e_t, Z_t u_{t+1}). \tag{A.11}$$

We can estimate the first term in (A.10) and (A.11) as

$$\frac{1}{T_1}\sum_{t=T_2-T_1+1}^{T_2} \hat{e}_t^2, \tag{A.12}$$

the expectation in the second term in (A.10) and (A.11) as

$$\frac{1}{T_2}\sum_{t=1}^{T_2} Z_t Z_t' \hat{u}_{t+1}^2, \tag{A.13}$$

and the covariance term in (A.11) as

$$\frac{1}{T_1}\sum_{t=T_2-T_1+1}^{T_2} \hat{e}_t Z_t \hat{u}_{t+1}, \tag{A.14}$$

with

$$\hat{e}_t = \mathcal{E}_t - \hat{E}_t - \hat{b}, \qquad \hat{u}_{t+1} = R_{t+1} - R_t^f - Z_t'\hat{k}, \tag{A.15}$$

and $\phi = T_1/T_2$.

The empirical implementation is based on further generalization of the above cases. To account for serial correlation of $\{e_t\}$, the first term in (A.10) and (A.11) is replaced by

$$\Omega = \lim_{T_1\to\infty} \Omega_T, \quad \text{where} \quad \Omega_T \equiv \frac{1}{T_1} E\left[\sum_{t=T_2-T_1+1}^{T_2} e_t \sum_{t=T_2-T_1+1}^{T_2} e_t\right] \tag{A.16}$$

To account for serial correlation if $\{u_{t+1}\}$, the expectation in the second term in (A.10) and (A.11) is replaced by

$$W = \lim_{T_2\to\infty} W_T, \quad \text{where} \quad W_T \equiv \frac{1}{T_2} E\left[\sum_{t=1}^{T_2} Z_t u_{t+1} \sum_{t=1}^{T_2} Z_t' u_{t+1}\right] \tag{A.17}$$

Finally, to account for serial correlation between $e_t$ and $Z_t u_{t+1}$, the covariance in the third term in (A.11) is replaced by

$$V = \lim_{T_1\to\infty} V_T, \quad \text{where} \quad V_T \equiv \frac{\sqrt{\phi}}{T_1} E\left[\sum_{t=T_2-T_1+1}^{T_2} e_t \sum_{t=1}^{T_2} Z_t u_{t+1}\right] \tag{A.18}$$

The variance terms introduced by (A.16)–(A.18) should be estimated using HAC-estimators. The empirical implementation employs the Newey-West estimator.

## A.2 Test PE-C2

The derivation follows the same lines as above in the PE-U2 case. Bring $E_t[R_{t+1}]$ to the RHS of (16) and treat like a generated regressor with coefficient constrained to one. The only difference to earlier derivation to the PE-U2 case is that here we estimate slopes as well as a constant. As a result, we get

$$\sqrt{T_1}(\hat{b} - b) = D^{-1}\frac{X'e}{\sqrt{T_1}} - \sqrt{\phi}D^{-1}\tilde{G}C^{-1}\frac{Z'u}{\sqrt{T_2}} + o_p(1), \tag{A.19}$$

where $b \equiv (b_0, b_1')'$, $Z \equiv [Z_1 \ldots Z_{T_2}]'$, $X_t \equiv (1, x_t')'$, $X \equiv [X_{T_2-T_1+1} \ldots X_{T_2}]'$, $D \equiv E[X_t X_t']$, $\tilde{G} \equiv E[X_t Z_t']$, $u \equiv [u_2 \ldots u_{T_2+1}]'$, and $e \equiv [e_{T_2-T_1+1} \ldots e_{T_2}]'$. Following the empirical analysis in the paper, in what follows we assume that the set of regressors in $z_t$ and $x_t$ is the same so that $D = \tilde{G} = C$ and (A.19) turns into

$$\sqrt{T_1}(\hat{b} - b) = C^{-1}\left[\frac{X'e}{\sqrt{T_1}} - \sqrt{\phi}\frac{Z'u}{\sqrt{T_2}}\right] + o_p(1), \tag{A.20}$$

The asymptotic variance of $\sqrt{T_1}(\hat{b} - b)$ is as follows

$$C^{-1}\left[\tilde{\Omega} + \phi W - 2\sqrt{\phi}\tilde{V}\right]C^{-1}, \tag{A.21}$$

where the variance term $\tilde{\Omega}$ is defined by

$$\tilde{\Omega} = \lim_{T_1 \to \infty} \tilde{\Omega}_T, \quad \text{where} \quad \tilde{\Omega}_T \equiv (1/T_1)E[X'ee'X], \tag{A.22}$$

the variance term $W$ is defined by

$$W = \lim_{T_1 \to \infty} W_T, \quad \text{where} \quad W_T \equiv (1/T_2)E[Z'uu'Z], \tag{A.23}$$

and the variance term $\tilde{V}$ is defined by

$$\tilde{V} = \lim_{T_1 \to \infty} \tilde{V}_T, \quad \text{where} \quad \tilde{V}_T \equiv \left(\sqrt{\phi}/T_1\right)E[X'eu'Z]. \tag{A.24}$$

One should pick consistent estimators of $\tilde{\Omega}_T$, $W_T$, and $\tilde{V}_T$ depending on the properties of sequences $\{e_t\}$ and $\{u_{t+1}\}$. As in the PE-U2 test, the numerical implementation uses the Newey-West estimator.

# B Small-sample bias adjustments of test estimates

This appendix describes the small-sample bias adjustments performed in the tests RN-C and PE-C1.

Consider the system

$$\mathcal{E}_t[R_{t+1}] - R_t^f = a_0 + a_1 x_{t-1} + u_t, \tag{A.25}$$
$$x_t = \rho_0 + \rho_1 x_{t-1} + \xi_t, \tag{A.26}$$
$$u_t = \chi u_{t-1} + \eta_t + \lambda \xi_t, \tag{A.27}$$

where $\eta_t$ and $\xi_t$ are independent i.i.d. normal random disturbances. Equation (A.25) is the RN-C test regression equation (7) from the main text, but here for a scalar regressor $x_{t-1}$ which is the P/D ratio in our empirical implementation. In the case of the PE-C1 test, we replace $R_t^f$ and $u_t$ in equation (A.25) by $R_{t+1}$ and $u_{t+1}$ so that (A.25) becomes the PE-C1 test regression equation (17) from the main text. Equation (A.26) captures the fact that $x_t$ is persistent. Equation (A.27) allows for non-zero covariance in the residuals of equations (A.25) and (A.26), and for persistence in the residuals of equation (A.25). In the special case without persistence ($\chi = 0$), equation system (A.25)-(A.27) reduces to one similar as that considered in Section 2 in Stambaugh (1999), although there with realized returns as dependent variable. Since the empirical evidence suggests that $\chi > 0$, we consider the more general case.

For the empirically plausible case with $\chi \neq 0$, the regressor $x_{t=1}$ and the residual $u_t$ in equation (A.25) depend both on lagged values of $\xi_t$, whenever $\lambda \neq 0$. OLS estimates of the coefficients $(a_0, a_1)$ in equation (A.25) are then asymptotically biased. Furthermore, even if $\chi = 0$, OLS estimates suffer from a small-sample bias for $\lambda \neq 0$ (Stambaugh (1999)). To address both issues, we proceed as follows:

1. Estimate equation (A.26) using OLS and perform Monte-Carlo simulations to correct for the small-sample bias of the OLS estimates of $\rho_0$ and $\rho_1$.

2. Lag equation (A.25) by one period, multiply by $\chi$, and subtract the result from equation (A.25). This delivers

$$\begin{aligned} \mathcal{E}_t[R_{t+1}] - R_t^f &= a_0(1-\chi) + \chi(\mathcal{E}_{t-1}[R_t] - R_{t-1}^f) \\ &\quad + a_1 x_{t-1} - \chi a_1 x_{t-2} + \lambda \xi_t + \eta_t, \end{aligned} \tag{A.28}$$

   which can be estimated using non-linear least squares (NLLS), given the observed explanatory variables and the estimates of $\xi$ from step 1. This delivers consistent estimates for $\chi, \lambda, \sigma_\eta^2, a_0$ and $a_1$.

In the RN-C test the remaining steps are as follows.

3. Compute the value of the $F$-statistic from our hypothesis tests $H_0 : a = b = 0$ and $H_0 : b = 0$.

4. Derive the small sample distribution of the $F$-statistic. This is done by simulating (A.25)-(A.27), using our parameter estimates (with some of the parameters changed to the value they assume under the considered null). Estimate (A.28) on the simulated data and compute the $F$-statistic for the simulated data.

5. Compute the small-sample bias corrected estimates of $a_0$ and $a_1$ by using the NLLS estimates from step 4 to perform a bias correction of the NLLS estimates from step 2. These bias corrected estimates of $a_0$ and $a_1$ are reported in the tables of the main text.

In the PE-C1 test the remaining steps are as follows.

3. Simulate (A.25)-(A.27), using parameter estimates from Step 2. Estimate (A.28) on the simulated data.

4. Compute the small-sample bias corrected estimates of $a_0$ and $a_1$ by using the NLLS estimates from step 3 to perform a bias correction of the NLLS estimates from step 2. Use these bias corrected estimates of $a_0$ and $a_1$ and the covariance matrix from step 2 to test for pessimism or optimism of observations at every point in time.