Subjective Beliefs and Asset Prices

Renxuan Wang

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2021

Abstract

Subjective Beliefs and Asset Prices

Renxuan Wang

Asset prices are forward looking. Therefore, expectations play a central role in shaping asset prices. In this dissertation, I challenge the rational expectation assumption that has been influential in the field of asset pricing over the past few decades. Different from previous approaches, which typically build on behavioral theories originated from psychology literature, my approach takes data on subjective beliefs seriously and proposes empirically grounded models of subjective beliefs to evaluate the merits of the rational expectation assumption. Specifically, this dissertation research: 1). collects and analyzes data on investors' actual subjective return expectations; 2). builds models of subjective expectation formation; 3). derives and tests the models' implications for asset prices. I document the results of the research in two chapters.

In summary, the dissertation shows that investors do not hold full-information rational expectations. On the other hand, their subjective expectations are not necessarily *irrational*. Rather, they are bounded by the information environment investors face and reflect investors' personal experiences and preferences. The deviation from fully-rational expectations can explain asset pricing anomalies such as cross-sectional anomalies in the U.S. stock market.

In the first chapter, I provide a framework to rationalize the evidence of extrapolative return expectations, which is often interpreted as investors being irrational. I first

document that subjective return expectations of Wall Street (sell-side, buy-side) analysts are contrarian and counter-cyclical. I then highlight the identification problem investors face when they form return expectations using imperfect predictors through Kalman Filters. Investors differ in how they impose subjective priors, the same way rational agents differ in different macro-finance models. Estimating the priors using surveys, I find Wall Street and Main Street (CFOs, pension funds) both believe persistent cash flows drive asset prices but disagree on how fundamental news relates to future returns. These results support models featuring heterogeneous agents with persistent subjective growth expectations.

In the second chapter, I propose and test a unifying hypothesis to explain both cross-sectional return anomalies and subjective return expectation errors: some investors falsely ignore the dynamics of discount rates when forming return expectations. Consistent with the hypothesis: 1) stocks' expected cash flow growth and idiosyncratic volatility explain significant cross-sectional variation of analysts' return forecast errors; 2). a measure of mispricing at the firm level strongly predicts stock returns, even among stocks in the S&P500 and at long horizon; 3). a tradable mispricing factor explains the CAPM alphas of 12 leading anomalies including investment, profitability, beta, idiosyncratic volatility and cash flow duration.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I wish to express my deepest gratitude to all of the members of the dissertation committee: Paul Tetlock, Robert Hodrick, Stijn Van Nieuwerburgh, Andrew Ang and Zhongjin Lu. Without you, the dissertation would not have been possible.

Paul, I learned from you how to conduct research at the highest level. You taught me to think about big picture problems, to be objective and thorough about facts and to always have common sense when solving abstract models. I am fortunate to have you as my adviser.

Bob, thank you for your kindness and invaluable advice about academic research and life in general. You supported me during the hardest times of my Ph.D. and kept on encouraging me to continue with my research.

Stijn, your advice and encouragement during the last part of my Ph.D. and the job market were the reason that kept me going. I wish I had met you earlier. Your comments have contributed enormously to these chapters and my thinking towards research in general.

Andrew, thank you for all your support and conversations throughout the years. You have been an inspiration to me long before I came to Columbia. Looking back, I may have been better off, had I followed your advice more strictly.

Zhongjin, I truly appreciate the discussions and advice regarding research and academic life. I admire your passion and attitude towards research into financial markets.

I am thankful for the encouragement and guidance from other faculty members at the Finance Division at Columbia. In particular, I would like to thank Wei Jiang for her invaluable advice, mentoring and inspiring discussions about culture and life; Michael Johannes for admitting me into Columbia and providing me the opportunity to teach EMBA courses; Gur Huberman for the stimulating discussions during a joint research project; Lars Lochstoer for teaching me how to think deeply about new facts and problems; Suresh Sundaresan for the encouragement and pulling me back to the Ph.D. program. I also acknowledge the financial support from the Finance Division of Columbia Business School, the Dutch Government and the Prince Bernhard Culture Fund.

I would also like to thank my colleagues and friends, who supported me throughout the Ph.D. journey. In particular, I would like to thank Adrien Alvero for the time we spent together, not only during those research projects but also the good times we had on the Upper West Side, Mannheim and South America; Tuomas Tomunen for working on several projects together and the squash games; Milan Vidojevic for the interesting discussions about our joint research project and the knowledge about factors; Kevin Guo for the intense discussions and debates about financial markets and the trip to Las Vegas; Berend Coster for the moral support and road trips throughout America and Europe; Paul Bekker for the econometric training before the Ph.D.; Vinzenz Ziesemer for motivating me to pursue the Ph.D. and the discussions about career and life; Ward Bortz for the support during difficult times and friendship. In addition, the thesis is partially inspired by the substantial time spent on 9 West 57th Street, which revealed to me the real New York finance. I am grateful for the friendship of Preet Patel and Richard Penman.

Last but not least, I am indebted to my family. My parents, Shuxiang and Zhiwu, have given me the unconditional support throughout my life, even after I have made unconventional decisions that they did not understand. Their support has provided me the strength to think independently and to embark on a unique journey of my own. This journey would have been lonely and impossible without my wife Yan, who has always been supportive and encouraging from Europe to the U.S. and now onto Asia. Finally, I am grateful for Celine, who joined us recently and gave me extra motivation to carry on.

Dedication

To Celine, Yan and my parents.

Chapter 1: Subjective Return Expectations

This chapter studies how investors form subjective return expectations. Return expectation is a key input in investors' investment decisions, which ultimately drive asset prices. Furthermore, the parameters governing how investors form return expectations reflect their perceptions of risk and uncertainty, which are central in financial economics. Yet, despite its importance, our understanding towards how return expectations are formed is surprisingly limited.[1]

Instead of investors' actual (subjective) return expectations, the literature has devoted much effort in studying objective return expectations. Objective return expectations are statistical measures of what would have been a good proxy for future returns ex-post, from the point of view of an econometrician who analyzes historical data.[2] The key finding is that future returns are predictable in-sample by fundamental-price ratios, such as dividend-price or earnings-price ratios. The finding motivates macro-finance models featuring fully rational agents (Campbell and Cochrane (1999) and Bansal and Yaron (2004)).[3] In these models, agents rationally hold counter-cyclical return expectations as they require higher (lower) risk premium to hold stocks during recessions (expansions).

However, objective return expectations can differ from investors' actual (subjective) return expectations. Indeed, direct survey evidence shows that return expectations among a number of investors appear exclusively extrapolative and even pro-cyclical.[4] The evidence further motivates a new class of models featuring *irrational* extrapolative agents (Barberis

[1] Noticing the lack of studies in finance and macroeconomics on subjective expectations, Brunnermeier et al. (2021), and Manski (2018) call for more research in this area.

[2] The literature is often referred to as the return predictability literature, which is large. See Cochrane (2011a) and Koijen and Van Nieuwerburgh (2011) for reviews.

[3] More specifically, the rational agent assumption means that investors know the true parameter values and the objective probability distributions that govern the data generating process of the environment.

[4] See for example Vissing-Jorgensen (2003), Greenwood and Shleifer (2014), Andonov and Rauh (2020).

et al. (2015a), Adam, Marcet, and Beutel (2017), Hirshleifer, Li, and Yu (2015), and Choi and Mertens (2019)),[5] These models show promise in capturing key empirical moments of both asset prices and subjective beliefs.

These new models also motivate two open questions. First, in reality*, who* hold contrarian beliefs to accommodate demand from extrapolative agents, if any one? Second, and more importantly, *why* would investors hold extrapolative beliefs in light of the evidence of return predictability?

This paper addresses these questions. At a high level, the results in this paper demonstrate that subjective return expectations can appear extrapolative or contrarian, and neither expectation needs to be *irrational*. Instead, different return expectations could all be a result of investors optimally learning from past information. They are different because investors are bounded by the information environment and their own personal preference and/or experiences.

I start by documenting new facts about subjective return expectations. Wall Street analysts, which include sell-side and buy-side analysts, hold contrarian and counter-cyclical return expectations. Together with existing evidence of extrapolative return expectations, the new evidence suggests Wall Street and Main Street (retail investors, CFOs, pension funds) persistently disagreeing with each other. These new results provide a micro-foundation for contrarian investors in models like Barberis et al. (2015a) and paint a more complete picture of who hold what expectations in real world.

Subsequently, I propose a return expectation formation framework to understand these facts together with evidence of return predictability. The framework has two key elements. First, Bayesian investors face a predictive system with imperfect predictors, such as commonly used dividend-price or earning-price ratios. Second, different investors form and update their return expectations based on Kalman Filter. Such return expectation formation process means investors are minimizing their perceived forecast errors using past informa-

[5]More specifically, extrapolation means that agents' return expectations are positive related to past realized returns and/or past cash flow news.

tion. This is a realistic setting consistent with the task facing econometricians, who have information for past data.[6]

I demonstrate through simulations that the return expectation formation framework can generate the observed heterogeneous expectation dynamics, including those of extrapolative, contrarian, pro- and counter- cyclical patterns. The main reason that optimizing investors agree to disagree is that they encounter a parameter identification issue known to researchers in the return predictability literature.[7] Intuitively, the issue arises because one observable change in return predictors such as dividend-price ratio can be theoretically driven by two latent economic forces, namely expected return and/or expected future cash flow growth. Consequently, a Bayesian investor needs to impose subjective priors on certain parameters that govern these latent processes in order to form a unique return expectation. I find that analytically, investors need to impose prior beliefs on a). relative importance of cash flow news in driving valuations compared to discount rate news b). correlation between cash flow news and expected return news. Moderate differences in the two priors beliefs can result in different return expectations.[8]

Similar to the way different investors have their own subjective prior beliefs, leading macro-finance models also make different assumptions regarding these two unidentifiable parameters. For example, in Campbell and Cochrane (1999), the rational representative agent believes cash flow process is i.i.d. while their expected return process is persistent due to their habit utility function. Such an assumption implies discount rate shocks and cash flow shocks are perfectly negatively correlated. On the other hand, in models featuring

[6]The setting with imperfect predictor is a realistic one and has been considered in studies of objective return expectations. Pástor and Stambaugh (2009) studies exactly this setting of predictive system with imperfect predictors. More generally, the imperfect predictor assumption is rooted in the weak out-of-sample predictive power of most return predictors, such as dividend-price ratio and the small R-squared found in most of the return predictive regressions that employ different predictors. For related discussions, see Welch and Goyal (2008) and Ang and Bekaert (2007).

[7]Papers that discuss this identification issue include Cochrane (2008), Koijen and Van Nieuwerburgh (2011), Pástor and Stambaugh (2009), and Rytchkov (2012).

[8]Interestingly, academic researchers are still debating about the objective values of these parameters. The debate about the persistence parameter centers around the predictability of cash flows, see Bansal, Kiku, and Yaron (2012) and Beeler (2012). For a discussion on the correlation between discount rate and cash flow shocks, see Lochstoer and Tetlock (2020).

persistent long-run risks, agents hold objective (Bansal and Yaron (2004)) or subjective beliefs (Collin-Dufresne, Johannes, and Lochstoer (2016)) that cash flow process contains a persistent component. In these models, the correlation between expected cash flow news and expected return news depends on the relative magnitude of the representative agent's inter-temporal elasticity of substitution (IES) and relative risk aversion.[9]

Building on the expectation formation framework, I use survey data to identify these unobservable prior beliefs to distinguish asset pricing models. My main estimation results are as follows. First, both Wall Street (sell-side analysts) and Main Street (CFOs, pension funds) believe cash flow process as more persistent than discount rate process. Moreover, expected cash flow process is the main economic force driving asset price variations. Second, Wall-Street believes positive cash flow news leads to lower future returns while Main-Street believes the opposite. These results support models featuring heterogeneous agents with persistent expected fundamental process.

Why do different investors have different sets of prior beliefs in the first place? The literature suggests personal experiences as a key variable that drives subjective beliefs about economic variables (Malmendier and Nagel (2016)) and risk appetite (Malmendier and Nagel (2011)). These two channels echo the two prior beliefs highlighted in the expectation formation framework proposed in this paper. Taking advantage of the breadth of sell-side analysts survey, I find that more experienced analysts are indeed more contrarian. This result further supports models featuring agents learning from experiences, such as Collin-Dufresne, Johannes, and Lochstoer (2017) and Nagel and Xu (2019a).[10]

The rest of the paper is organized as follows. I document new facts about subjective return expectations in Section 1.1; I present and demonstrate the expectation formation framework

[9] More specifically, when an agent has a high IES (e.g. IES >1) relative to risk aversion, she prefers to consume less in light of higher expected growth, leading to a lower expected return. On the other hand, when her IES is small relative to risk aversion (IES <1), she prefers to consume more now facing positive growth news, leading to her to believe risk premium is higher and positively related to expected growth news.

[10] In Collin-Dufresne, Johannes, and Lochstoer (2016), subjective expected consumption growth is dependent on the number of years an agent learns about the endowment process. In Nagel and Xu (2019a), the subjective expected consumption growth is related to an agent's experienced payouts.

in Section 1.2; I estimate prior beliefs governing the expectation formation process in Section 1.3 before conclude in 1.4.

Related Literature

This paper relates to three strands of literature. First, this paper is related to the literature that uses survey data to study empirical properties of subjective return expectations. Vissing-Jorgensen (2003), Greenwood and Shleifer (2014) study surveys conducted for CFOs, retail investors and consumers, and recently, Andonov and Rauh (2020) studies return expectations of pension funds. This paper contributes to the literature by studying surveys on sell-side analysts. Sell-side analysts play an important role in financial markets, and they conduct research as their profession, including publishing price targets. The frequency as well as the number of responses in analyst return expectation data also exceed other surveys studied in the literature.

Second, the theoretical return expectation framework proposed in this paper contributes to a growing literature that models subjective expectation formation. More specifically, it is most related to studies that model agents as econometrician who learns from past information to form their expectations. In Branch and Evans (2010), agents form their future price expectation by linearly projecting past information to future prices.[11] In their framework, the agents are not using the full information available, especially the information in distant past. Furthermore, the agents form expectations by minimizing squared forecast errors in an overly parsimonious linear model. In contrast, the agents in the framework developed in this paper use all information and form expectation using the Kalman Filter. Another closely related paper is Adam, Marcet, and Beutel (2017), which develop an expectation formation framework in which agents learn from past price movements to form expectations about future capital gains. Agents in their setting form extrapolative return expectations, which is consistent with survey evidence. The framework proposed in the current paper can generates

[11] Their framework is a form of constant-gain adaptive learning, which comes from a large macroeconomics literature about learning, summarized in Evans and Honkapohja (2012)

heterogeneous return expectation dynamics, not only extrapolative return expectations, and allows agents to observe the same information. Additionally, in the current framework agent can use different predictors, such as past dividend price ratios, as opposed to only past returns.

Within the literature on expectation formation, this paper also relates to studies that emphasize personal experiences and cognitive biases in shaping investors' beliefs. In particular, Nagel and Xu (2019a) propose a model where agents have fading memories when forming expectations about unconditional mean consumption growth and show it can match important asset pricing moments and lead to return expectations that closes to a constant. Bordalo, Gennaioli, and Shleifer (2018) develop a diagnostic expectation formation framework, in which agents suffer from Kahneman and Tversky's representativeness heuristic. The current paper finds experiences of analysts matter for expectation formation, thus confirms the findings that personal experiences matter for expectation formation. On the other hand, the agents considered in this paper are forming expectation through Kalman Filter, using all available information in the past. The main deviation from full rational expectation is that agents do not conduct Bayesian analysis to find optimal priors in sample, as in Pástor and Stambaugh (2009).

Finally, the paper also contributes to the literature that tries to explain asset pricing puzzles by allowing agents to deviate from rational expectation. Specifically, this paper uses subjective return expectations to differentiate asset pricing models featuring alternative investor beliefs. O and Myers (2020) also uses sell-side analysts' earnings and CFO's return expectations to test asset pricing models. Their approach uses the decomposition of price to earnings/dividend ratios. The tests conducted in this paper makes use of the expectation formation framework to uncover a richer set of moments, such as correlation between (subjective) discount rate and cash flow news, in addition to variance decomposition of price earnings ratio. This allows me to make finer distinction between asset pricing models.

1.1 Heterogeneous Return Expectations and the Contrarian Wall Street Analysts

In this section, I document new facts to extend our understanding about subjective return expectations. First, I assemble a comprehensive set of return expectations and demonstrate that the market structure of return expectations form clusters within Wall Street (sell-side, buy-side) and Main Street (CFOs, consumers). Expectations are negatively correlated between the two clusters and Wall Street expectations are more inline with measures of objective return expectations. Second, I zoom in on sell-side analysts' and establish that they have contrarian return expectations on the market, firm and analyst level.

1.1.1 Data Sources and Measuring Wall-Street Analysts Return Expectations

Table 1.1 summarizes the data sources. While other subjective return expectations have been studied in the literature[12], the current paper document new facts about buy-side and sell-side analyst return expectations, which I briefly describe in this subsection. More details about data sources are provided in Appendix A.2.

The buy-side return expectation is from the asset management firm Grantham, Mayo& van Otterloo (GMO), which publishes a 7-year asset class forecasts each quarter on their website.[13] The reason to use GMO's return expectations are two-folds. First, seldom do any of the buy-side firms publish their return expectations and GMO is the only one that has a long-term historical account of return forecasts available since the second quarter of 2000. Second, GMO runs a large asset allocation fund and return expectations are important for these funds. Tower (2010) document that GMO's return expectations actually predict the returns in Vanguard mutual funds which have different asset classes and styles. Their expected returns for equities are available on their website since 2017. For pre-2017 data, I hand-collected the data from the internet.

[12] Examples include Greenwood and Shleifer (2014), Ben-David, Graham, and Harvey (2013), Vissing-Jorgensen (2003), and Adam, Matveev, and Nagel (2021a).

[13] https://www.gmo.com/americas/research-library/

Table 1.1: Data Sources for Subjective and Objective Return Expectations

Measures of Subjective Return Expectation	
Who	**Source**
Sell-Side Analysts	I/B/E/S detailed unadjusted price targets
Buy-Side Analysts	Grantham, Mayo & Van Oterlook (GMO) 7-year Asset Class Forecasts
Institutional Investors (Pension)	Shiller Survey/Yale University
Survey of Professional Forecasters	the Federal Reserve
CFOs	Duke University CFO Global Business Outlook
Retail Investors	Shiller Survey/Yale University
Consumers	University of Michigan Consumer Surveys
Proxies for Objective Return Expectations	
Proxy	**Source**
S&P500 Price Dividend Ratios	GlobalX and CRSP
PE ratio	Prof. Robert Shiller's Website
Consumption-Wealth Ratio	Prof. Martin Lettau's Website

Notes: The sample period is 2002-01-01 to 2018-12-31, which all data are available. Appendix A.2 provides more details about the data sources. Appendix A.3 describes how index level price dividend ratios are constructed.

Aggregate (S&P500-level and firm-level) sell-side analysts return expectations are constructed using individual analyst price targets. The expected returns are computed by dividing individual analyst's price targets by the daily closing price on the day the estimates was issued and subtracted by 1,[14] or

$$\mu_{i,f,d}^{A} = \frac{P_{i,f,d}^{A,12}}{P_{f,d}} - 1$$

where $P_{i,f,d}^{A,12}$ is the price target of analyst i for firm f, issued at day d and $P_{f,d}$ is the closing price of the firm f. The superscript "A" denotes the 12-month ahead estimates.[15]

Firm-level return expectations are simple averages of analyst-level return expectation and market-level return expectations are market-cap weighted firm-level return expectations. Details of how these expectations are constructed are in Appendix B.5.

The data set of sell-side analysts price targets has comprehensive coverage, which include forecasts of about 2700 analysts from 236 brokerage firms at a point in time on average. I detail the coverage and the summary statistics in Appendix A.1.2.

1.1.2 Heterogeneous Return Expectations: Wall-Street vs Main-Street

Table 1.2a shows correlations among surveys of different parties. The correlation matrix clearly displays a two-cluster structure: Wall Street vs. Main-Street. Sell-Side analysts, buy-side analysts, professional forecasters at the Fed form a Wall-Street cluster, whose return expectations are positively correlated with each other. On the other hand, retail investors, CFOs and consumers form another cluster, which are conventionally thought of as Main Street investors. The latter cluster were consistent with results in Greenwood and Shleifer (2014). On the other hand, correlations of return expectations between these two clusters are negative, with consumers and buy-side analysts have a -69% at one extreme.

[14] The same formula is used in Brav and Lehavy (2003a) and Da and Schaumburg (2011)

[15] Notice this methodology ensures there is no mechanical relation between mean estimated expected returns and the level of prices. On each issuing date the analyst has the freedom to pick her own price target since she observes the prices. 12-month ahead estimates are the most commonly issued horizon. The other horizons, although available, have much poorer coverage.

The last row of Table 1.2a reports how different subjective return expectations react to past 6 month returns. Wall Street appears to have a contrarian view and Main Street extrapolates. While the extrapolative expectations have been documented before sell-side analysts' contrarian expectations have not been documented in the literature, to the best of my knowledge. To show this results are robust, I provide detailed analysis in Section 1.1.3.

Panel 1.2b shows correlations between subjective return expectations and proxies for objective return expectations considered in the literature. Wall Street analysts' expectations are negatively (positively) correlated with price-fundamental ratios (CAY), which means they are counter-cyclical. In particular, for a commonly used objective return expectation (Cochrane (2011a)), which is the fitted value of future returns regressed on CAY and log(P/D) ratios, the Wall Street and Main Street return expectations are positively and negatively correlated with high magnitude.

Table 1.2: Heterogeneous Subjective Return Expectations

(a) Correlations Between Different Subjective Expectations

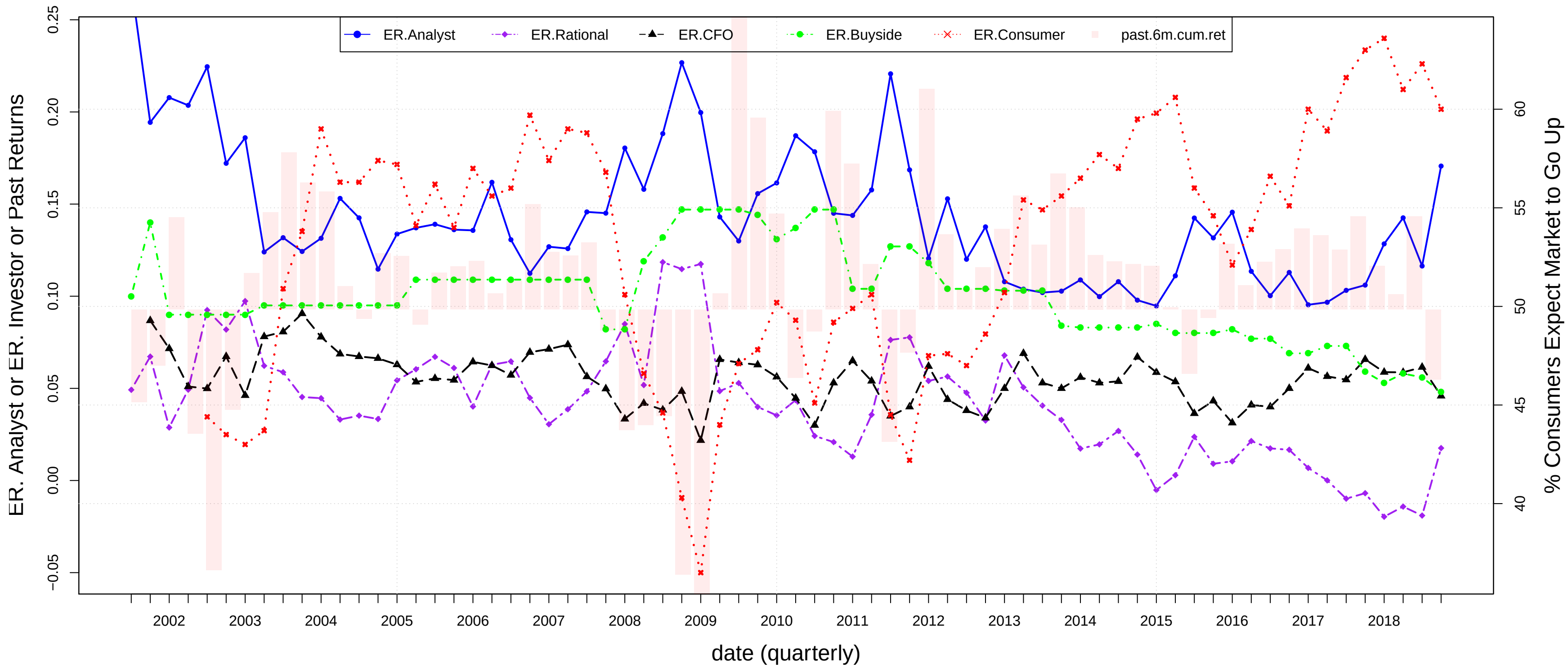
	ER.analyst	ER.buy.side	Prof.Forecaster	Shiller.institutional	Shiller.retail	ER.CFO	ER.consumer(pct)
ER.analyst	1.00						
ER.buy.side	0.41***	1.00					
Prof.Forecaster	0.45***	0.53***	1.00				
Shiller.institutional	0.21*	0.26**	0.50***	1.00			
Shiller.retail	0.00	-0.20	0.30**	0.33**	1.00		
ER.CFO	-0.27**	-0.06	0.35***	0.20*	0.68***	1.00	
ER.consumer(pct)	-0.68***	-0.69***	-0.31**	-0.37***	0.04	0.38***	1.00
past.6m.cum.ret	-0.65***	-0.09	-0.11	0.00	0.08	0.48***	0.37***

(b) Correlation Between Analyst Expectations and Proxies of Objective Returns Expectations

	ER.Analyst	ER.Buy.Side	ER.CFO	ER.Consumer
ER.Analyst	1.00			
ER.Buy.Side	0.43***	1.00		
ER.CFO	-0.34***	-0.13	1.00	
ER.Consumer	-0.68***	-0.69***	0.38***	1.00
ER.Rational	0.61***	0.58***	-0.20	-0.73***
CAY	0.54***	0.45***	0.08	-0.55***
log(P/D)	-0.23*	-0.36***	0.63***	0.52***
CAPE	-0.51***	-0.81***	0.42***	0.90***
GS10.pct	0.28**	0.19	0.51***	0.08

Notes: “ER.Rational” is fitted values of regressing future 12-month returns on CAY and log(P/D) from 1970-2019. “*”, “**” and “***” represent significant level at 10%, 5% and 1%, respectively. Data are based on quarterly series. “ER.consumer(pct)” represents return expectations from Michigan Survey for households asking "percent chance” their investment would increase next year. More details about these surveys are in Appendix A.2.

Figure 1.1: Market Structure for Return Expectations



Notes: "ER.Rational" is fitted values of regressing future 12-month returns on CAY and log(P/D) from 1970-2019. "past.6m.cum.ret" are the realized cumulative returns from the past 6 months, plotted as bars. "ER.consumer" are return expectations from Michigan Survey for households asking "percent chance" their investment would increase next year, which is measured on the right axis. More details about these surveys are in Appendix A.2.

Different return expectations within each cluster are also loading differently on past returns and fundamental/price ratios as well as risk-free rate. For example, buy-side analysts are more related to objective return expectations and are less driven by past returns, while sell-side analysts are more influenced by past returns and risk-free rate. Since fundamental-price ratios are much more persistent than past realized returns, sell-side analysts' return expectations should be more volatile than buy-side analysts'. Similar patterns can be found between CFOs and consumers.

Figure 1.1 visualizes these rich expectation dynamics. Indeed, sell-side analysts' return expectations are more volatile than that of buy-side analysts the CFOs. Furthermore, the persistent disagreement between consumers and sell-side analysts also stands out.

1.1.3 Contrarian Return Expectations of Sell-side Analysts

I demonstrate the contrarian feature of sell-side analysts' return expectations is a robust finding. I run the time-series regression of aggregate analyst return expectations, $\mu^A_{m,t}$, on two-month lagged cumulative $k-month$ past market returns, $R_{m,t-2,t-k}$, and other (lagged) control variables, X_{t-2} in the regression: [16]

$$\mu^A_{m,t} = a + bR_{m,t-2,t-k} + cX_{t-2} + e_t \tag{1.1}$$

A negative coefficient b would mean analysts expect the market to have a negative expected return when the market has yielded a positive return in the past, i.e. contrarian expectation. Since the dependent variables are persistent, I use Newey-West standard errors with 12-month lag to correct for auto-correlations. For control variables, I include the 10-year U.S. treasury yields, price dividend ratios as well as analyst aggregate long-term growth

[16] The independent variables are lagged by two months when entering into the regression to prevent the estimates of b from being contaminated by stale analysts forecasts. When constructing individual analyst return expectations, the analyst price targets are at most 2 months old by construction. Therefore, lagging 2 months when running the predictive regression makes sure all future return expectations are out-of-sample. As an example, when using the past 6-month cumulative returns at the end of June 2005 to predict analyst's aggregate return expectations end of August 2005, the oldest analyst return expectation is constructed using price targets and stock price in early July 2005.

measure to proxy for expected future earnings growth.

Panel (b) in Table 1.3 shows the estimation results for regression 1.1. Panel (a) shows the empirical distribution of the key variables in the regression to help interpret the magnitude of the coefficients. The coefficients on past returns are negative across all of the specifications and are significant both statistically and economically. In Column 1, for one standard deviation (percent) increase in the past 6 month cumulative returns, the next month analyst return expectations decrease by 1.7% (0.16%), with a t-stat of 3.8. Since the monthly volatility of analyst return expectation is only 3.5% the estimate means the economic magnitude of the contrarian effects is also large.

The contrarian effect does not only apply to short-term past realized returns. Column 2 and 3 in the table shows that the past 36 month returns have almost the same predictive power as the 6 month cumulative returns. In fact, one standard deviation increase in past 36 month returns decreases the future analyst return expectation by 1.8%, on top of the past 6 month returns. Furthermore, past 6-month cumulative returns and 36-month cumulative returns together explain up to 41% of the time-variation in monthly analyst return expectations. This high R-squared further demonstrate the economic magnitude of the contrarian effect. Results from Column 2 and 3 also raise the question of which horizon of past returns matter most to analyst future return expectation. I investigate this question in Appendix A.4.1.

The contrarian results are not much affected when including other control variables, as shown in Column (4). The 10-year treasury yield is the only variable that has a (marginal) significance in predicting analyst's return expectations. To understand the magnitude of the coefficient on treasury yield, consider as a benchmark, analysts believe risk-free rate is a constant and risk-free rate is a part of the expected future return. In this case, the coefficient should be 1. Therefore, a estimated coefficient of 0.843 means analysts expect a persistence in risk-free rate process.

The contrarian effect are not a result of aggregation, nor staleness of the analyst forecasts.

Table 1.3: Aggregate Analyst Return Expectations and Past Returns

(a) Summary Statistics: Monthly Aggregate Expectation Data (SP500 firms)

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
ER.analyst	208	0.142	0.037	0.087	0.115	0.160	0.270
tot.ret.1m	208	0.006	0.041	−0.168	−0.015	0.031	0.109
past.6m.cum.ret	208	0.031	0.111	−0.427	−0.010	0.089	0.388
past.36m.cum.ret	208	0.156	0.302	−0.434	−0.108	0.371	0.858
GS10	208	0.033	0.011	0.015	0.023	0.042	0.053
log(P/D)	208	3.946	0.152	3.289	3.875	4.031	4.305
LTG	208	0.115	0.014	0.081	0.105	0.122	0.153
avg.nr.firms.ER.analyst	208	490.178	7.076	459	486	495	500

(b) Monthly Regression: Aggregate Analyst Return Expectations on Past Returns

	Dependent variable:			
	Aggregate Analyst Return Expectations			
	(1)	(2)	(3)	(4)
past 6m cum.ret	−0.163***		−0.125***	−0.115***
	(0.043)		(0.033)	(0.038)
past 36m cum.ret		−0.062***	−0.049***	−0.038**
		(0.022)	(0.015)	(0.018)
GS10				0.867*
				(0.497)
Log(P/D)				−0.015
				(0.042)
Analyst LTG Estimate				0.382
				(0.370)
Constant	0.146***	0.151***	0.152***	0.135
	(0.006)	(0.009)	(0.006)	(0.129)
Observations	206	206	206	206
R^2	0.266	0.284	0.428	0.516
Adjusted R^2	0.263	0.281	0.422	0.504
Residual Std. Error	0.030 (df = 204)	0.030 (df = 204)	0.027 (df = 203)	0.025 (df = 200)
F Statistic	74.001*** (df = 1; 204)	81.031*** (df = 1; 204)	75.862*** (df = 2; 203)	42.706*** (df = 5; 200)

Note: *p<0.1; **p<0.05; ***p<0.01
SEs are Newey-West with 12 month lag

Notes: In Panel (b) Time-series regression of aggregate analyst return expectations, $\mu^A_{m,t}$, on past returns. I include (2-month lagged) 6-month past cumulative returns, $R_{m,t-2,t-6}$ and (2-month lagged) 36-month past cumulative returns, $R_{m,t-2,t-36}$ and other 2-month lagged control variables. "lag.2m" denotes the variables are lagged by 2 months before entering the regressions. Sample period: 2002-03-01 to 2018-12-31, a total of 202 months. ER.analyst: value-weighted analyst return expectation for the SP500 index; tot.ret.1m: one-month total return on SP500 index; past.6m.cum.ret: 6-month cumulative returns on the SP500 index; past.36m.cum.ret: 36-month cumulative returns on the SP500 index; GS10: yield on 10-year constant maturity treasury; Log(P/D): log price dividend ratio of SP500 index; LTG: value-weighted analyst long-term growth expectation for the SP500 index; avg.nr.firms.ER.analyst: average monthly number of firms that have analyst return expectations in the SP500 index.

Appendix A.4.1 and A.4.3 show that the contrarian results hold at firm level and at analyst level, respectively. Remarkably, the magnitude of the contrarian effects are similar at each level. Furthermore, the results on the analyst level are based on analysts whoever issues their forecasts for the first-time ever, thus eliminating concerns that the contrarian effect is simply a result of stale analyst forecasts. Appendix A.1.3 provides detailed analysis on the timing and frequencies of analysts' price target forecasts.

1.2 A Framework for Subjective Return Expectation Formation

I propose an expectation formation framework to understand the subjective return expectation dynamics observed in the data. I start by describing the environment faced by investors in this framework. I present the subjective return expectation dynamics in this environment for an investor who minimizes his/her perspective forecast errors through Kalman Filter, and show through simulation the framework can generated the rich patterns observed. Subsequently, I use a simplified system (where investors only consider dividend yield and returns) to demonstrate the parameter identification problem faced by investors. Finally, I identify the prior beliefs investors need to impose for arriving at unique expectations.

1.2.1 The Environment

There are three types of shocks ϵ_d, ϵ_g and ϵ_μ, which represent news about current dividend, expected future cash flow growth and discount rates, respectively. They follow multivariate normal distribution

$$\begin{pmatrix} \epsilon_{d,t+1} \\ \epsilon_{\mu,t+1} \\ \epsilon_{g,t+1} \end{pmatrix} \sim N \left[0, \begin{pmatrix} \sigma_d^2 & \sigma_{\mu d} & \sigma_{gd} \\ \sigma_{\mu d} & \sigma_\mu^2 & \sigma_{\mu g} \\ \sigma_{gd} & \sigma_{\mu g} & \sigma_g^2 \end{pmatrix} \right] \tag{1.2}$$

Dividend growth in the next quarter, Δd_t contains a potentially persistent component g_t:

$$\Delta d_{t+1} = g_t + \epsilon_{d,t+1} \tag{1.3}$$

$$g_{t+1} = E_g(1-\phi) + \phi g_t + \epsilon_{g,t+1} \tag{1.4}$$

Furthermore, let $\mu_t = E_t(r_{t+1})$ be the discount rate process and it follows an autoregressive process

$$\mu_{t+1} = (1-\beta)E_r + \beta\mu_t + \epsilon_{\mu,t+1} \tag{1.5}$$

Besides, there exist a vector of return predictors such as price earnings ratios, whose values are correlated with the three shocks. Denote their values as x_t, they follow

$$x_{t+1} = (I-A)E_x + Ax_t + \epsilon_{x,t+1} \tag{1.6}$$

x_t can be used to predict future returns because of its correlations with $\epsilon_{\mu,t+1}$ via

$$\begin{pmatrix} \epsilon_{d,t+1} \\ \epsilon_{\mu,t+1} \\ \epsilon_{g,t+1} \\ \epsilon_{x,t+1} \end{pmatrix} \sim N \left[0, \begin{pmatrix} \sigma_d^2 & \sigma_{\mu d} & \sigma_{gd} & \sigma'_{dx} \\ \sigma_{\mu d} & \sigma_\mu^2 & \sigma_{\mu g} & \sigma'_{\mu x} \\ \sigma_{gd} & \sigma_{\mu g} & \sigma_g^2 & \sigma'_{gx} \\ \sigma_{dx} & \sigma_{\mu x} & \sigma_{gx} & \sigma_x^2 \end{pmatrix} \right] \tag{1.7}$$

This environment is consistent with the literature of return predictability, where researchers aim to construct proxies for objective expected returns (Van Binsbergen and Koijen (2010), Pástor and Stambaugh (2009)). Although simple, the setup encapsulates asset pricing models featuring either or both persistent cash flow or discount rate process.

1.2.2 Investors' Subjective Return Expectation Formation Process

Investors do not observe the news directly and neither do they know the exact values of the parameters that govern the data generating process described in (1.5), (1.6) and (1.7). Instead, they observe changes in predictors such as fundamental-price ratios and past returns and estimate the news and parameter values from available information. Based on their own estimates, they update their subjective return expectations through Kalman Filter, or:

$$\widetilde{ER}_{t|t} = \widetilde{ER}_{t|t-1} + \widetilde{m}_t * (r_t - \widetilde{ER}_{t|t-1}) + \widetilde{n}_t * \epsilon_{x,t} \tag{1.8}$$

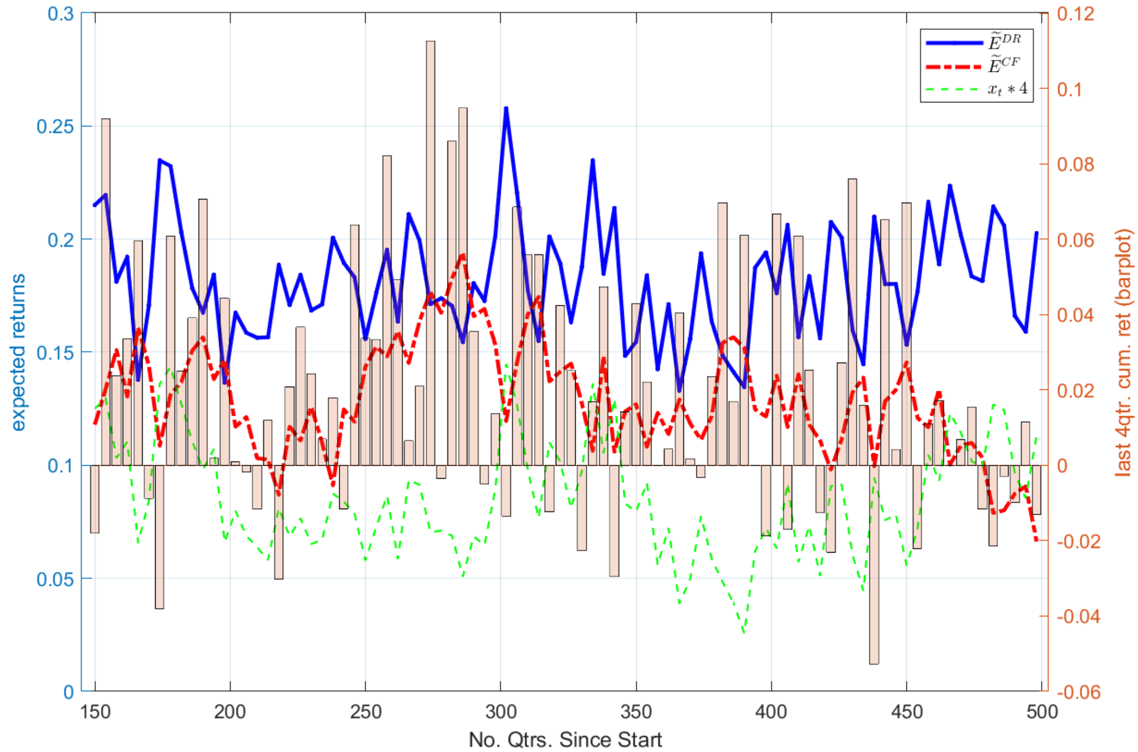
where $\widetilde{ER}_{t|t} = \tilde{E}(r_t|\mathcal{F}_t)$ is the subjective return expectation and $\mathcal{F}_t$ denotes the agent's information set, which contains values of all past predictors and past realized returns up to and including time t; "$\tilde{\cdot}$" means that the expectation depends on agent's own subjective beliefs; $\epsilon_{x,t}$ is the innovation in predictors defined (1.6). $\widetilde{m}_t$ and $\widetilde{n}_t$ are functions of both parameters in the system (1.3) to (1.7) and values of realized returns and $\epsilon_{x,t}$. The time subscript in $\widetilde{m}_t$ and $\widetilde{n}_t$ capture the fact that agents can learn over time and adjust their expectation formation process over time. See Appendix A.7 for detailed derivation to arrive at Equation (1.8) and expressions for m_t and n_t as well as their steady-state values.

The return expectation in (1.8) connects investors' subjective returns expectations to past realized returns and return predictors through agents' optimization process. Thus, this framework rationalizes why the observed pattern that return expectations from surveys are related to past returns and fundamental-price ratios: from the perspective of investors, they are simply trying to project the most accurate return expectations based on news in $\epsilon_t = (\epsilon_{d,t}, \epsilon_{\mu,t}, \epsilon_{g,t}, \epsilon_{x,t})'$.[17]

The simple expectation formation framework can generate the heterogeneous return expectations dynamics similar to the one we observed empirically in Figure 1.1. I confirm

[17] In Appendix (A.6), I demonstrate in Figure (A.7) that it is advantageous for investors to use realized returns to forecast future returns, even though past return as a standalone predictor does not forecast future returns in linear predictive regressions. In fact, in studies such as Van Binsbergen and Koijen (2010), past returns are used together with dividend yields to predict future returns.

Figure 1.2: Different Return Expectations vs. Predictors and Past Returns (Simulated Data)



Note: Simulated return expectations are based on Equation (1.8). $\widetilde{ER}_t^{DR}$ denotes the (annualized) return expectation of an investor who believes that asset prices are 97% driven by discount rate variations; $\widetilde{ER}_t^{CF}$ denotes the (annualized) return expectation of an investor who believes asset prices are 99% driven by expected cash flow growth variations; $x_t * 4$ are the dividend yields times 4; orange bars represent the cumulative past 6 months returns. Dividend yields and past returns are simulated based on moments calibrated to historical data. More details of the simulations are in Appendix A.6.

the framework's ability through simulation in Figure 1.2. In this figure, two investors form different return expectations following (1.8), even though they observe the same information -past realized returns and dividend yields, and both investors try to minimize their forecast errors. I explain how and why this framework can generate such return expectations next in Section 1.2.3.

1.2.3 Understanding Return Expectation Formation Process: A Simple One-Predictor System

Why Do Subjective Return Expectations Differ?

The reason behind persistent disagreement in investors' subjective return expectations is that investors face a parameter identification problem when predictors for future returns are not perfect. Intuitively, when the magnitude of return predictability is small, as found in the literature,[18] different interpretations of the data may persist. The expectation formation framework provides an analytical base to understand why and how these differences persist. In this framework, imperfect predictors mean that agents believe none of shocks in predictors x_t , $\epsilon_{x,t}$ has a correlation with $\epsilon_{\mu,t}$ of absolute value of 1 and the corresponding persistent parameter in A equals β. This leads to a parameter identification problem. Below, I demonstrate the parameter identification problem using a simple example of investors only considering dividend yield as a predictor.

In this case, investors observe both the dividend yields and past returns to extract shocks to expected returns ϵ_μ in order to update their return expectations. Assuming investors understand the present value relationship, their perceived system becomes

[18] Typically the R2 in regressions of future returns on predictors are small. Typically, when the future returns are of short horizon, such as one year, the R2 of these regressions are smaller than 10%. Furthermore, as discussed in Welch and Goyal (2008) the out-of-sample predictive powers of these predictors are also poor.

$$r_{t+1} = \mu_t + \epsilon_{d,t+1} - \rho\kappa_\mu\epsilon_{\mu,t+1} + \rho\kappa_g\epsilon_{g,t+1} \tag{1.9}$$

$$dp_{t+1} = (1-\phi)B_{dp} + \phi dp_t + \kappa_\mu(\beta-\phi)\mu_t + \kappa_\mu\epsilon_{\mu,t+1} - \kappa_g\epsilon_{g,t+1} \tag{1.10}$$

$$\begin{pmatrix} \epsilon_{d,t+1} \\ \epsilon_{\mu,t+1} \\ \epsilon_{g,t+1} \end{pmatrix} \sim N\left[0, \begin{pmatrix} \sigma_d^2 & \sigma_{\mu d} & \sigma_{gd} \\ \sigma_{\mu d} & \sigma_\mu^2 & \sigma_{\mu g} \\ \sigma_{gd} & \sigma_{\mu g} & \sigma_g^2 \end{pmatrix}\right] \tag{1.11}$$

where $\kappa_\mu = \frac{1}{1-\rho\beta}$ and $\kappa_g = \frac{1}{1-\rho\phi}$ and ρ=0.94 and B_{dp} is a constant.

The system in Equation (1.9) and (1.10) present a parameter identification problem because investors need to separate three shocks, $\epsilon_{d,t}$, $\epsilon_{\mu,t}$ and $\epsilon_{g,t}$ from two observables in the data, namely innovations in realized returns (u_{t+1}) and dividend to price ratios (v_{t+1}), or

$$u_{t+1} = \epsilon_{d,t+1} - \rho\kappa_\mu\epsilon_{\mu,t+1} + \rho\kappa_g\epsilon_{g,t+1} \tag{1.12}$$

$$v_{t+1} = \kappa_\mu\epsilon_{\mu,t+1} - \kappa_g\epsilon_{g,t+1} \tag{1.13}$$

As a result of this parameter identification issue, different investors can persistently disagree on what their subjective value of expected return shocks are and how the expected return process evolves. Since these are not uniquely pinned down by the data, prior beliefs about parameter values in the system is necessary to form a unique return expectation. I discuss in the next subsection *what* prior beliefs need to be set subjectively by investors.

The parameter identification problem outlined here has been discussed in the literature, because researchers who try to forecast future returns using Kalman Filters also face the same issue.[19] To avoid/solve this problem, the literature either simply impose a value to the unidentified parameters or conduct Bayesian analysis to examine what values of prior beliefs are the most accurate in terms of fitting the historical data.[20] Essentially, these exercises

[19] See for example: Cochrane (2008), Pástor and Stambaugh (2009), Koijen and Van Nieuwerburgh (2011), and Rytchkov (2012).

[20] As an example, Van Binsbergen and Koijen (2010) assume the parameter σ_{gd} to be zero and Pástor and

are researchers trying to impose priors to uniquely pin down their return expectations.

Adding to the system more predictors will not completely resolve the problem, because other predictors are also imperfect and introduce more noises. As an example, adding CAY theoretically introduces shocks to payout ratio and leverage process. Even though aggregate consumption is correlated to aggregate dividend, the portion of consumption that paid out to shareholders as dividend can vary over time. This new shock means the parameter identification problem persists.

What Drives Differences in Return Expectations?

I show that the two prior beliefs investors hold drive their return expectations: 1). how important is expected cash flows news for driving asset valuation when compared to discount rate news; 2). whether positive cash flow news means negative or positive for future returns. Quantitatively, when investors believe asset prices are mostly driven by fundamentals, their return expectations are much easier to appear as extrapolative than contrarian, and vice versa. I explain these results below and more analysis can be found in Appendix A.8.

To understand intuitively why these two particular priors can lead to pro- and counter-cyclical return expectations, consider extracting the latent expected return process from the dividend-price ratio alone, through

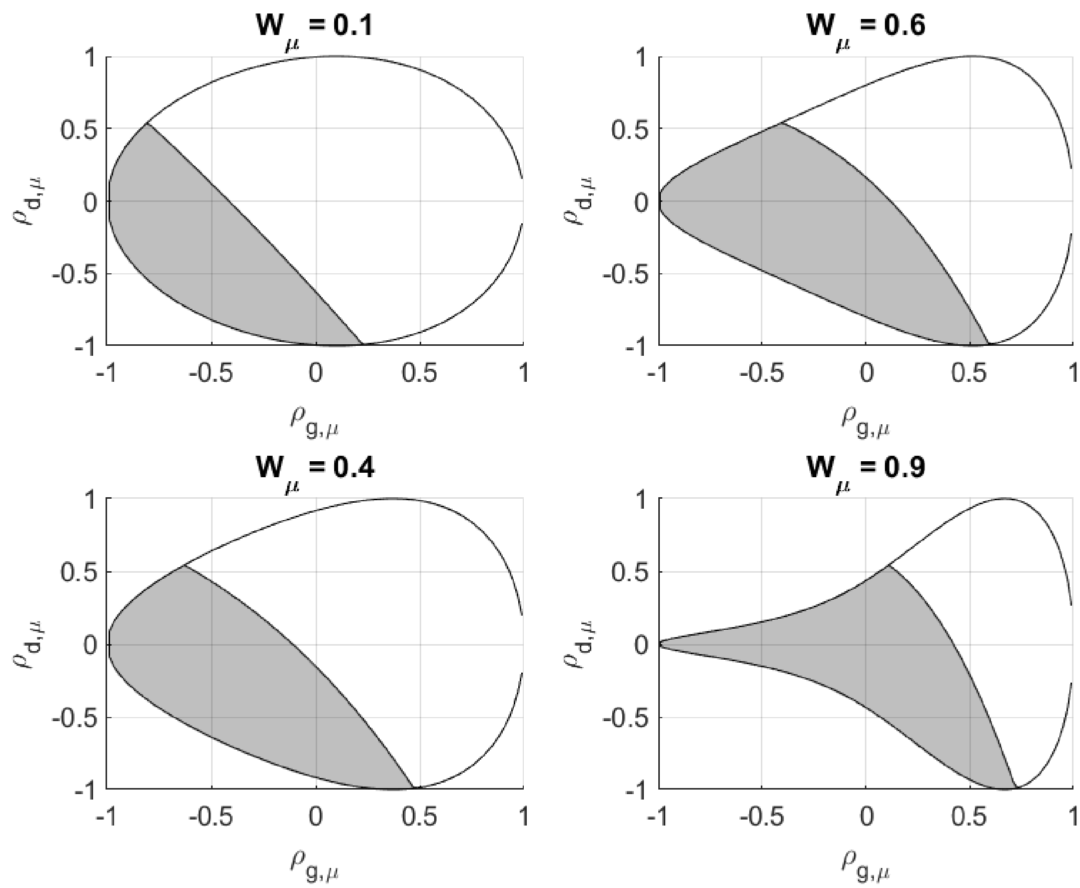
$$dp_t = B_{dp} + \frac{1}{1-\rho\beta}\mu_t - \frac{1}{1-\rho\phi}g_t \tag{1.14}$$

From the point of view of an econometrician, to identify the process μ_t from g_t, she needs to specify 1). how persistent the expected return process is compared to the cash flow process and 2). how the shocks to these two processes are correlated. In the case that one believes most of price-dividend moves are due to a shocks to expected cash flow growth, because the cash flow process is much more persistent ($\phi > \beta$), she would more likely believe a positive

Stambaugh (2009) finds that when econometricians impose the correlations between u_{t+1} and v_{t+1} to be strongly negative, dividend yields have be better performance when forecasting future returns.

change in dividend price ratio to be a result of lowered expected future cash flow growth. On top of that, if she also believes negative cash flows shocks are typically associated with negative future returns, she would lower her return expectations. In this case, her return expectations are negatively related to positive changes in dividend price ratio, therefore appearing to be pro-cyclical.

Figure 1.3: Prior Beliefs, Parameter Space and Return Expectation



Note: Grey shaded area represents the parameter space for $(\rho_{d,\mu}, \rho_{g,\mu})$ in which an investor appear to be contrarian, or $\widetilde{m} < 0$; white area within closed space in each graph are parameter space in which an investor would appear extrapolative. Each subplot in has a fixed level of (relative) discount rate volatility, defined as $W_\mu = \frac{\kappa_\mu \sigma_\mu}{\sigma_{v,t}}$. Area within closed loops are feasible parameter space for $(\rho_{d,\mu}, \rho_{g,\mu}, W_\mu)$ that satisfies the condition that correlation matrix of $(\rho_{d,\mu}, \rho_{v,d}, \rho_{v,\mu})$ needs to be semi-positive definite. This constraint puts bounds on the value of $\rho_{d,\mu}$, through $\rho_{d\mu} \in [-\sqrt{1-\rho_{v\mu}}, \sqrt{1-\rho_{v\mu}}]$.

Figure 1.3 demonstrates how different priors would impact investors' subjective return expectations. More specifically, the figure plots the possible value pairs of priors on parameters $(\rho_{d,\mu}, \rho_{g,\mu}, W_\mu)$ in order for a forecaster to appear contrarian/extrapolative.[21] $W_\mu := \frac{\kappa_\mu \sigma_\mu}{\sigma_v}$ denotes the volatility of discount rate shocks (numerator) as proportion to the shocks to dividend yields. The area within the closed lines are the feasible parameter space for the parameters $(\rho_{d,\mu}, \rho_{g,\mu}, W_\mu)$ and the shaded area within each of the feasible area is the parameter space in which an investor will appear to be contrarian,[22]while the white area in each closed loop is the parameter space in which an investor will appear extrapolative. Each subplot in Figure 1.3 has a fixed level of (relative) discount rate volatility (W_μ).

The figure provides the following insights. First, investors will mostly likely appear extrapolative when they interpret (expected) cash flow news as positively related to future returns. Within each subplot, the upper right corner, where $\rho_{g\mu}$ and $\rho_{d\mu}$ take on higher values are regions in which investors expectations would appear extrapolative. This is intuitive, as investors are essentially extrapolating from current cash flow news for future returns, if cash flow news is important to them.

Second, the more an investor considers expected cash flow to be important for asset prices, the less likely the investor will appear to be a contrarian. Quantitatively, when expected future cash flows is the dominant force, for example as in the the top-left panel ($W_\mu = 0.1$), as long as $\rho_{g\mu} > 0.15$, investors would appear to be extrapolative, no matter how negative the value of $\rho_{d\mu}$ they believe in. On the other hand, when the forecaster believe discount rate is more important (bottom-right plot), all investors who believe that $\rho_{g\mu} < 0.1$ will appear contrarian (gray area).

Finally, the figure also shows how the expectations framework can accommodate rich

[21]Notice W_μ is not the same as the discount rate variation as percentage of total dividend yield variance. However, they are positively related to each other, up to scaling by the persistent parameters. Appendix A.8 provides more detailed discussion on this subject.

[22]Following the condition that

$$\rho_{d\mu}^2 - 2\rho_{v\mu}\rho_{vd}\rho_{d\mu} + (1 - \rho_{vd}^2)(1 - \rho_{v\mu}^2) \leq 0$$

and

$$\rho_{v,d} \approx 0$$

dynamics in return expectations. Even in the case where all investors believe asset prices are driven by fundamentals (top-left panel of Figure 1.3), some could appear contrarian while the others extrapolative, because of their different beliefs in $\rho_{g,\mu}$, for example. For more technical discussion about how these parameters related to return expectations, see Appendix A.8. Next, I use survey data to back out prior beliefs of different investors, as these parameters are important for differentiating asset pricing models.

1.3 Identify Prior Subjective Beliefs From Survey Data

Using subjective return expectation data, the framework allows for the identification of investors' prior beliefs, which are crucial assumptions in asset pricing models. First, I describe how to estimate the prior beliefs in this framework. Subsequently, I apply the estimation methodology to selected survey data and discuss the implications of the estimates to asset pricing theories. Finally, I discuss what makes the prior beliefs different by providing evidence that personal experiences of sell-side analysts impact their return expectations.

1.3.1 Estimation Framework

Thanks to the observable surveys, or Equation (1.17) below, all parameters governing investors' return expectation process are identifiable, including those in the variance-covariance matrix in (1.2). We have the following systems of equations:

$$\hat{r}_{t+1} = \hat{\mu}_t + \epsilon_{\Delta d,t+1} - \rho\kappa_\mu(\beta)\epsilon_{\mu,t+1} + \rho\kappa_g(\phi_g)\epsilon_{g,t+1} \tag{1.15}$$

$$\hat{dp}_{t+1} = \phi_g \hat{dp}_t + \kappa_\mu(\beta - \phi_g)\tilde{\mu}_t + \kappa_\mu\epsilon_{\mu,t+1} - \kappa_g\epsilon_{g,t+1} \tag{1.16}$$

$$\hat{\mu}^A_{t+1} = \beta\hat{\mu}^A_t + L(\beta)\epsilon_{\mu,t+1} \tag{1.17}$$

$$\hat{x}_{t+1} = A\hat{x}_t + \epsilon_{x,t+1} \tag{1.18}$$

where the "ˆ" denotes that variables are demeaned and the shocks follow multivariate normal as in 1.19:

$$\begin{pmatrix} \epsilon_{d,t+1} \\ \epsilon_{\mu,t+1} \\ \epsilon_{g,t+1} \\ \epsilon_{x,t+1} \end{pmatrix} \sim N \left[0, \begin{pmatrix} \sigma_d^2 & \sigma_{\mu d} & \sigma_{gd} & \sigma'_{dx} \\ \sigma_{\mu d} & \sigma_\mu^2 & \sigma_{\mu g} & \sigma'_{\mu x} \\ \sigma_{gd} & \sigma_{\mu g} & \sigma_g^2 & \sigma'_{gx} \\ \sigma_{dx} & \sigma_{\mu x} & \sigma_{gx} & \sigma_x^2 \end{pmatrix} \right] \tag{1.19}$$

which allows me to estimate the system based on maximum-llikelihood, where $\kappa_\mu(\beta) = \frac{1}{1-\rho\beta}$, $\kappa_g(\phi_g) = \frac{1}{1-\rho\phi_g}$ and $L(\beta) = \sum_{k=0}^{3} \beta^k$. $\hat{\mu}_{t+1}^A$ is the observed (demeaned) 12-month return expectations (and the superscript "A" denotes it's annual), which is the quarterly return expectations $\hat{\mu}_t$ rolling forward, following the dynamics of $\hat{\mu}_t$. The return expectations follow Equation (1.17), because the demeaned quarterly return expectations follow

$$\hat{\mu}_{t+1} = \beta\hat{\mu}_t + \epsilon_{\mu,t+1}$$

Since the main interests are in the covariance matrix and the persistent parameters, I use demeaned returns and expectation data, $\hat{r}_{t+1}$ and $\hat{\mu}_{t+1}^A$. More details about the estimate procedure is documented in A.9.

1.3.2 Estimation Results

Parameter Estimates

Table 1.4 presents the parameter estimates based on return expectations of sell-side analysts, pension funds ("Shiller Institutional") and CFOs.[23] I chose these three series to apply the estimation framework for the following reasons. First, these series show relatively distinct correlations among each other as shown in Table 1.2a. Second, these three series are based on surveys with similar questions, so there are less scaling issues.[24]

[23] See Appendix A.2 for more details about how these data are constructed.

[24] The buy-side analyst surveys are based on a different forecasting horizon (7 years) while the data on households with the question of "What do you think is the percent chance that this one thousand dollar investment will increase in value in the year ahead, so that it is worth more than one thousand dollars one

Table 1.4: Parameter Estimates Return Expectation Dynamics

Investor	**Sell-Side Analyst**	**Shiller Institutional**	**CFO**
Panel A: Estimates of Structural Parameters			
ϕ	0.929	0.920	0.918
	(0.03)	(0.018)	(0.024)
β	0.478	0.674	0.604
	(0.058)	(0.09)	(0.166)
σ_d	0.022	0.021	0.021
	(0.002)	(0.002)	(0.002)
σ_μ	0.014	0.010	0.005
	(0.002)	(0.002)	(0.001)
σ_g	0.008	0.012	0.012
	(0.002)	(0.002)	(0.002)
$\rho_{d,\mu}$	-0.474	-0.0152	0.156
	(0.201)	(0.059)	(0.354)
$\rho_{d,g}$	-0.00222	0.121	0.127
	(0.308)	(0.059)	(0.42)
$\rho_{g,\mu}$	-0.552	0.349	0.675
	(0.134)	(0.07)	(0.165)
Panel B: Implied Parameters			
W_μ	0.304	0.326	0.152
$\widetilde{m}$	-0.103	0.173	0.059
$\widetilde{n}$	0.029	0.167	0.019
κ_μ	1.816	2.731	2.316
κ_g	7.883	7.377	7.303
$\rho_{v,\mu}$	0.756	-0.028	-0.575
$\sqrt{Q}$	0.010	0.012	0.005

As analyzed in the previous section, the key parameters are first, the persistent parameters (ϕ and β), which impact the relative importance of discount rate news compared to cash flow news (W_μ); second, the correlations between cash flows and expected returns ($\rho_{g,\mu}$, $\rho_{d,\mu}$). Differences in these values are crucial for how return expectations would appear.

First, all three types of investors believe cash flow process is persistent, in fact, more persistent than the discount rate process. In particular, sell-side analysts believe the cash flow process has a persistent parameter of 0.93, the highest among the three, although the differences are small among them. The persistence of the return expectations, however, differ significantly across participants.[25] Sell-side analysts return expectations (0.478) are much less persistent and more volatile than those of the CFOs (0.604). In fact, judging from the standard errors of the parameter estimate σ_μ, we can't reject the null hypothesis that CFOs' return expectations are different from a constant, consistent with the results of O and Myers (2020). This leads to beliefs that discount rate news are less important in driving valuations in valuation ratios than cash flow news, i.e. $W_\mu < 0.5$, which means that all of the three types of investors would be placed on the two left panels in Figure 1.3. Among different investors, CFOs' estimates for discount rates importance is the least among the three, making them the most prone to extrapolation.

Second, the three investor types have very different estimates on how cash flow news impact future returns, especially about what news about expected cash flow growth means for future returns ($\rho_{g,\mu}$). Sell-side analysts, in particular, would revise down their return expectations in light of positive news about cash flows ($\rho_{g,\mu} = -0.552$ and $\rho_{d,\mu} = 0 - 0.474$) while CFOs hold opposing beliefs ($\rho_{g,\mu} = 0.675$, $\rho_{d,\mu} = 0.156$). Pension funds' views are in between the two. The parameter estimates of $\rho_{v,\mu}$ are consistent with the correlations in Table 1.2. Sell-side analysts return expectations are counter-cyclical ($\rho_{v\mu} > 0$) while those of CFO's are pro-cyclical ($\rho_{v\mu} < 0$); pension funds' return expectations are little correlated

year from now?" Instead, data on pensions, sell-side analysts and CFO's are all based on question of the percent increase in prices of stocks.

[25]Notice that this parameter estimate is close to auto-correlation parameter estimates from a simple OLS regression.

with dividend yields. Relating the finding to Figure 1.3: CFOs' beliefs would place them into the upper right corner in the top-left panel while sell-side analysts would be on the lower-left grey area in the bottom-left panel.

In addition, the volatility parameters of the unexpected cash flow shocks, σ_d are large and similar for different participants at around 2% per quarter. This is reasonable because the main difference between the return series and the price dividend ratio series is due to the unexpected cash flow shocks. As a result, investors should be able to almost identify the value of the unexpected cash flow shock volatility from the two series.

Subjective Variance Decomposition of Returns and Dividend-Price Ratios

The estimates also shed light on investors' beliefs about why returns and prices move, which are of great interests for researchers. Therefore, I compute variance decomposition of price-dividend ratios and returns and present the results for the three types of investors in Table 1.5 and 1.6, respectively.

Table 1.5: Variance Decomposition for Dividend to Price Ratios

	μ_t	g_t	$-2Cov(\mu_t, g_t)$	Var(dp)
Sell-side Analyst				
Variance	0.37%	2.01%	-0.34%	2.03%
Portion of Returns	18.21%	98.73%	-16.93%	100.00%
Institutional				
Variance	0.73%	3.15%	-1.84%	2.03%
Portion of Returns	35.85%	154.80%	-90.65%	100.00%
CFO				
Variance	0.10%	2.69%	-0.75%	2.03%
Portion of Returns	4.71%	132.23%	-36.94%	100.00%

Table 1.6: Variance Decomposition for Quarterly Unexpected Returns

	ϵ_d	ϵ_μ	ϵ_g	$-2Cov(\epsilon_\mu, \epsilon_g)$	$-2Cov(\epsilon_d, \epsilon_\mu)$	$2Cov(\epsilon_d, \epsilon_g)$	$r_{t+1} - \mu_t$
Sell-side Analyst							
Variance	0.04%	0.05%	0.36%	0.16%	0.05%	0.01%	0.60%
Portion of Returns	7.33%	9.00%	59.15%	26.03%	7.96%	1.82%	100.00%
Shiller Institutional							
Variance	0.04%	0.06%	0.62%	-0.13%	0.00%	0.05%	0.60%
Portion of Returns	6.58%	10.10%	102.62%	-21.91%	0.17%	8.16%	100.00%
CFO							
Variance	0.04%	0.01%	0.64%	-0.12%	-0.01%	0.05%	0.60%
Portion of Returns	6.60%	2.10%	106.30%	-19.89%	-1.26%	8.52%	100.00%

Consistent with the magnitude of W_μ, market participants believe variations in future cash flows are the dominant force in driving returns and asset prices, instead of discount rates, as shown in Table 1.5.[26] This is mainly due to the higher persistence of the cash flow expectation processes, as opposed to the volatility of the shocks, as evident by the fact that shocks to cash flows take up a smaller portion of variance of returns (Table 1.6). I show in Section A.10 that such decomposition is robust when using analysts' own cash flow expectations directly, instead of the implied cash flow expectation in the estimation.

In fact, as shown in Table 1.5, all of the market participants think that the level of expected returns and cash flow expectations are positively correlated, i.e. people believe a higher expected future fundamental growth is accompanied by higher expected return. This positive correlation holds also for the short-term shocks to expectations, or $Cov(\epsilon_\mu, \epsilon_g)$, with the exception of sell-side analysts, who believe that these two shocks are negatively correlated at the quarterly frequency.

[26]In fact, this view is consistent with the argument put forward by Bordalo et al. (2020), which shows that a lot of stock market puzzles are driven by biased expectations about market fundamentals.

Which Predictors Are Important For Return Expectations?

I find that Shiller's CAPE ratio seem to be the most influential predictors from the perspective of investors, although investors interpret it differently.

Table 1.7: Correlations Between Innovations in Predictors and Different News

Investor	Sell-Side Analyst	Shiller Institutional	CFO
Panel A: Correlation between Innovations in Predictors and Expected Returns			
$\rho_{CAY,\mu}$	0.395	0.018	-0.334
	(0.132)	(0.089)	(0.112)
$\rho_{CAPE,\mu}$	-0.866	-0.076	0.614
	(0.031)	(0.109)	(0.091)
$\rho_{E_g,\mu}$	-0.045	-0.055	0.114
	(0.161)	(0.082)	(0.122)
Panel B: Correlation between Innovations in Predictors and Expected Cash Flow Growth			
$\rho_{CAY,g}$	-0.210	-0.257	-0.295
	(0.147)	(0.138)	(0.148)
$\rho_{CAPE,g}$	0.834	0.820	0.900
	(0.046)	(0.048)	(0.024)
$\rho_{E_g,g}$	0.445	0.296	0.299
	(0.145)	(0.142)	(0.142)

Table 1.7 shows how different market participants interpret the signals from different well-known predictors. Following the logic of the present value relation, investors could interpret positive shock to price to fundamental ratios as a sign of either a higher future expected cash flow or lower future returns or both.

Panel A shows the correlation between the shocks to different predictors and shocks to expected returns. Shiller's CAPE ratio is an important predictor considered by both sell-side analysts and CFOs, albeit of different sign. Next, for sell-side analysts, the consumption-wealth ratio or CAY, is a positive predictor for future returns, while for CFOs and consumers, this measure is a negative predictor. Relatively speaking, Analyst's earnings growth forecasts are less important in forming return expectations.[27]

Panel B shows the correlation between shocks to different predictors and expected future cash flows. Contrasting with results on expected returns, people seem to interpret predictors

[27]Notice that although the sign is negative, the correlation between earnings growth expectation shocks and discount rate shocks are not significantly different from zero for sell-side analysts, which might seem to contradict the result from the model estimation on $\rho_{g,\mu}$. I discuss further this point further in Section A.10.

similarly. In particular, higher CAPE ratio and lower CAY are interpreted as signs of higher future cash flows while higher expected future earnings growth from analysts are signs of higher future cash flow growth, as expected.

These results are a confirmation of the model assumption: that different people might interpret the same predictor as being a different signal for future returns. Due to the parameter identification problem, this persistent difference in attitude towards predictors ultimately result in the differences in return expectations dynamics.

1.3.3 Discussions

Implications for Equilibrium Asset Pricing Theories

The empirical estimates of investors' subjective beliefs provide new moments on investors' beliefs to distinguish asset pricing models. Below, I discuss how these estimates are related to asset pricing models.

First, all three types of investors believe expected future cash flow is more persistent and the dominant force driving asset prices and returns. This evidence is at odds with models where rational agents believe expected cash flows are i.i.d. (Campbell and Cochrane (1999),Barberis et al. (2015a)).[28] On the other hand, the results support models in which agents believe there exists a persistent component in cash flow process. This includes the long-run risk models (Bansal and Yaron (2004) and Pohl, Schmedders, and Wilms (2021)) and models featuring agents learning about fundamentals (Collin-Dufresne, Johannes, and Lochstoer (2016) and Nagel and Xu (2019a)). Additionally, the variance decomposition results reveal that the three types of investors all seem to over-estimate how much cash flow variation contribute to the variation in asset prices when compared to objective measures of return expectations, which show discount rate variation should contribute most to asset

[28] In Barberis et al. (2015a), the rational agents hold contrarian return expectations to accommodate the extrapolative demand of irrational traders. In a way, the finding that expectations of sell-side analysts are contrarian and counter-cyclical provides a micro foundation for who are the rational contrarian, who understand cash flows are i.i.d..

price variation.[29]

Second, different types of investors hold heterogeneous beliefs regarding how cash flow news impacts future returns. This finding provides support for models featuring heterogeneous agents in a long-run risk environment (Pohl, Schmedders, and Wilms (2021)) or parameter learning (Collin-Dufresne, Johannes, and Lochstoer (2017)). Of course, this results do not rule out the possibility that a subset of investors with a particular prior is driving the asset prices, which would support models such as Adam, Marcet, and Beutel (2017) or Nagel and Xu (2019a).

Why would different investors form different priors in the first place? Models featuring parameter learning , such as Collin-Dufresne, Johannes, and Lochstoer (2017), emphasizes the micro-founded channel of personal experiences. Differences in personal experiences could lead to different risk appetite (Malmendier and Nagel (2011)) or different subjective beliefs about future economic variables (Malmendier and Nagel (2016)). Based on the current expectation formation framework, these two channels would lead to different return expectations in terms of being more or less contrarian, for example. I test this hypothesis next, taking advantage of the breadth of analyst expectation data.

What Drives Differences in Prior Beliefs? The Role of Personal Experience

Using surveys on individual analysts, I find that more experienced analysts tend to be more contrarian. The results highlight individual experiences as one potential channel that drives differences in analysts' prior beliefs, and further supports models featuring parameter learning.

I run the following regression

$$\mu^a_{i,f,t} = \alpha_t + \alpha_f + bR_{f,t-6} + cX_{i,f,t} + \beta R_{f,t-6} * X_{i,f,t} + \epsilon_{i,f,t} \tag{1.20}$$

[29]In Cochrane (2011a) and Koijen and Van Nieuwerburgh (2011), discount rate variations contribute to more than 100% of dividend price ratio variation. See Table A.14b for a direction comparison in this sample between subjective and objective variance decomposition.

where $\mu^a_{i,f,t}$ is the analyst-level return expectations or their deviations from the consensus; $R_{f,t-6}$ is the (one-month-lagged) past 6 month return of firm f; α_t and α_f are time and firm fixed effects; $X_{i,f,t}$ are analyst individual-level variables such as personal experiences, at the time the analyst is issuing the expected return for firm f.

The parameter estimates of interests is β, which measures how much *more contrarian* an analyst is, when the analyst's personal experience variables $X_{i,f,t}$ increases by one unit. The contrarian magnitude is measured in terms of deviation from consensus return expectation, compared with other analysts issuing the return expectation for the same firm during the same month.

When selecting personal experience variables, I consider the literature has documented that people learn from personal experiences (Malmendier and Nagel (2011), Malmendier and Nagel (2016)) and information rigidity or sticky expectation (Mankiw and Reis (2002), Bouchaud et al. (2019)), I use the number of months an analyst experience recession, the number of years of experience as an analyst as well as the number of stocks an analyst covers. To that end, I construct a comprehensive analyst-level data set on return and earnings expectations, which I document in more details in Appendix A.5.

Table 1.8 shows the estimates for Equation (1.20), in which the interacting variable with past returns is number of years of experience an analyst has up to the time of issuance. [30] The left two columns consider the deviation from consensus as dependent variable while the right two columns use expected returns. Within each group of the same dependent variables, the right column regression excludes the recession months defined by NBER.

Across different specifications, the estimates on β are significantly negative. Since the standard deviation of the analyst number of years variable is about 7 years, the β estimate in Column (1) can be interpreted as an older analyst with 1 year of more experience in the industry would be lower than the consensus by 7.2% compared to the young analyst

[30] In unreported tables, I also consider interacting with the number of firms covered as well as the number of months in recession. These two variables do not show up statistically significant when interacting with past returns.

Table 1.8: Analyst No. of years in the industry and return expectations

	Dependent variable:			
	Deviation From Consensus		Expected Returns	
	(1)	(2)	(3)	(4)
No Yrs Experience	−0.042	−0.050	−0.009	−0.006
	(0.053)	(0.053)	(0.053)	(0.052)
No Firms Covered by Analyst	0.022**	0.022**	0.028***	0.029***
	(0.009)	(0.009)	(0.009)	(0.009)
No Months Recession	0.026	0.029*	0.021	0.022
	(0.017)	(0.017)	(0.017)	(0.017)
Past 6m log.ret	0.132	0.469*	−3.853***	−3.321***
	(0.250)	(0.256)	(0.522)	(0.524)
Past 6m log.ret x No Yrs Experience	−0.072***	−0.095***	−0.135***	−0.217***
	(0.022)	(0.024)	(0.036)	(0.034)
Including Recession Months	Yes	No	Yes	No
Firm-Fixed Effects?	No	No	Yes	Yes
Month-Fixed Effects?	Yes	Yes	Yes	Yes
Observations	1,280,473	1,054,913	1,332,314	1,096,782
R^2	0.024	0.017	0.353	0.392
Adjusted R^2	0.024	0.017	0.348	0.387
Residual Std. Error	27.697 (df = 1280264)	24.545 (df = 1054743)	29.265 (df = 1323792)	25.821 (df = 1088634)
Note:				*p<0.1; **p<0.05; ***p<0.01

Note: "Deviation from Consensus" is the difference between the analyst's own expected return, subtracted the firm-level mean consensus expected return from the week before the issuance of the analyst's own expected returns. "no.yrs.experience" measures at the time an analyst issues an expected return for stock f, the number of years since he/she first issues an forecasts (EPS/Price Target) when issuing for the same firm during the same fiscal quarter, given the same level of past 6 months cumulative returns. These results are consistent with the hypothesis that analysts learn from the markets and adapt their expectations over time.

1.4 Conclusion

Investors' subjective return expectations play a central role in their decision process and asset pricing. Recent evidence of extrapolative return expectations based on surveys seems to paint some investors as irrational. The results presented in this paper show that return expectations are not always extrapolative. Furthermore, through analyzing investors' expectation formation process, I find that even if they form extrpolative return expectations, their expectation formation process does not need to be irrational. The disagreement between return expectations of Wall Street and Main Street mirror discussions held within academia: the key parameters that lead to differences in return expectations are assumed to take on different values in prominent maco-finance models.

The paper aims to advance our understanding towards subjective return expectations on four fronts, all of which leave open more questions that worth investigation. First, the contrarian and counter-cyclical return expectations of both sell-side and buy-side analysts show that more work need to be done in collecting subjective return expectations. The expectations of buy-side analysts I collect is only one of many funds. Dahlquist and Ibert (2021) made progress on this front by collecting large set of return expectations from asset management companies.

Second, the theoretical expectation formation framework developed here does not show implications for equilibrium pricing. What would be the equilibrium asset pricing implications when agents are forming expectations based on the framework developed here, especially when taking seriously the parameter estimates shown in Section 1.3, is a natural next question.

Third, the parameter estimates based on the framework suggest there is at least a subset of investors who underestimate how volatile and dynamic discount rate process is, when compared to the objective variance decomposition results. Does that mean there is systematic misvaluation by investors in their investment process? As a start, Renxuan (2020a) explores the idea that investors underestimate dynamics of discount rates and find that this mispricing can explain many of asset pricing anomalies in the cross-section.

Fourth, the evidence of analysts' personal experience driving differences in subjective beliefs is consistent with the literature on expectation formation, and the framework proposed here further shows they could be linked to either beliefs about future fundamental or preferences. This calls for further studies based on micro-level evidence to understand the distinction between and importance of these two channels.

Chapter 2: Asset Prices When Investors Ignore Discount Rate Dynamics

Many studies in the asset pricing literature document stock return predictability in the cross section: several firm-level characteristics, such as profitability ratios, idiosyncratic return volatility, asset growth and cash flow duration, significantly predict future returns. Trading strategies taking advantage of the predictability evidence result in higher returns after adjusting for their market risk (alpha), a phenomenon known as cross-sectional anomalies. These anomalies are not only at odds with leading theories such as the Capital Asset Pricing Model (CAPM), but can also lead to significant inefficiencies in the real economy.[1] Yet, despite years of effort, we are still in search of a *unified* explanation for why these anomalies arise and persist.

Recent development in asset pricing calls for researchers to take data on subjective beliefs seriously when interpreting asset prices.[2] Such development helps to guide and discipline the search for explanations but at the same time further raises the bar for newly proposed ones: they should not only be able to explain empirical moments on prices or returns but also data on investors' subjective beliefs.[3]

This paper proposes and tests a hypothesis about how investors' form subjective return expectation. The key assumption of the hypothesis is that some investors falsely ignore the dynamics of discount rates when forming their return expectations. I term the hypothesis

[1]The point about the real economy was discussed extensively in Binsbergen and Opp (2019).

[2]See the recent review article of Brunnermeier et al. (2021) for a summary.

[3]Most of the actions along this line of inquiry has been on aggregate asset prices, instead of the cross-section. Partial list of models that proposes new belief formation process that deviating from rational expectation assumption to understand aggregate asset prices: Hirshleifer, Li, and Yu (2015), Barberis et al. (2015b), Adam, Marcet, and Beutel (2017), and Bordalo, Gennaioli, and Shleifer (2018) and Collin-Dufresne, Johannes, and Lochstoer (2016) Nagel and Xu (2019b). For the cross-section, existing study include Bouchaud et al. (2019) and Bordalo et al. (2019), which focus on subjective earnings expectations and explain the anomalies on profitability and analysts' long-term growth forecasts, respectively.

the "Constant Discount Rate" (CDR) hypothesis.

In sum, I find that a). data on analysts' return expectations and firm fundamentals are consistent with the prediction of the CDR hypothesis; b). a tradable factor constructed based on the hypothesis can explain the CAPM alphas of 12 leading cross-sectional anomalies. Thus, the CDR hypothesis can serve as a unifying explanation for cross-sectional anomalies that is also consistent with data on investors' subjective beliefs.

Evidence of how investors actually form return expectations in practice lends support to the CDR hypothesis. As an example, bond yields, which are a common proxy for expected returns, are solved using bond prices based on the constant discount rate assumption. Furthermore, when estimating the Implied Cost of Capital (ICC) of stocks, which are often used as a proxy for expected returns in practice and in academic literature (Pástor, Sinha, and Swaminathan (2008) and Gebhardt, Lee, and Swaminathan (2001)), investors apply the same heuristic and assume constant discount rates.[4] Renxuan (2020b) estimates subjective decomposition of returns and asset prices based on surveys of investor return expectations and finds sell-side analysts, CFOs and pension funds all underestimate the dynamics of discount rates in driving asset prices. These results make the CDR hypothesis seem more plausible.

At a deeper level, the CDR hypothesis is motivated by a large literature on heuristic, which refers to the mental processes of people looking for solutions that are practical or satisfactory, rather than optimal, when facing complex problems. The literature starts from the seminal work of Simon (1956) and suggests that people may trade off accuracy for efficiency and simplicity.[5] Naturally, one would expect investors to use heuristics when making investment decisions. After all, financial markets are extremely complex and dynamic, and not

[4]Damodaran (2012), Koller, Goedhart, Wessels, et al. (2010) are textbooks on stock valuation. Most of the treatment of the Discounted Cash Flow Models assume constant discount rates.

[5]See Gigerenzer and Gaissmaier (2011) for a review on the literature. Notice that such a "accuracy-effort trade-off" heuristic in psychology literature is distinct from the "representativeness" and "conservatism" heuristics proposed Tversky and Kahneman (1974), which can generate over- and under- reaction in financial markets, respectively (Barberis, Shleifer, and Vishny (1998)). Yet, despite its prominence in psychology literature, the finance literature has largely ignored its application in financial markets, compared to how the work of Tversky and Kahneman (1974) has been applied.

all investors have the time or ability to solve complicated mathematical problems. Indeed, ignoring the dynamics of discount rates greatly simplifies investors' valuation process.[6] However, applying such a heuristic also biases investors' return expectations because discount rates do vary over time (Cochrane (2011b)). The results found in the paper support the view that cross-sectional asset pricing anomalies are simply a result of approximation errors from investors' heuristic decision-making process, a form of bounded rationality.

How exactly the CDR assumption leads to cross-sectional asset pricing anomalies is not obvious, so I start by developing a framework to formalize the hypothesis and understand how the CDR assumption could lead to biases in expectation and mispricing. Intuitively, investors with the CDR assumption overestimate the impact of cash flow news has on stock prices, because they fail to understand a dynamic discount rate will *offset* part of the impact cash flow news has on stock prices.[7] As a realistic scenario, when Tesla, Inc (TSLA) announces that it is to develop a new battery business, it opens up a new revenue stream which might lead to higher cash flow growth in the future. As a result, TSLA's stock price should go up because of the positive cash flow news. However, with the new battery business, TSLA's balance sheet also becomes riskier and the market requires a higher premium to hold it, which lowers the stock price. On the other hand, investors with the CDR belief will interpret the lowered stock price as cheap, or high expected return, because they fail to take into account the interaction between dynamic discount rates and cash flows.

The CDR assumption leads to biases of different degrees across stocks because the biases are incurred at each payout period and stocks are long-term assets with different payout horizons. As a result, two firm-level fundamental characteristics, namely expected cash flow growth and volatility, which proxy for stocks' cash flow duration or convexity, respectively, would drive cross-sectional differences in return expectation biases. Furthermore, any firm characteristics that forecast firm future cash flow growth and/or volatility would forecast

[6] See Ang and Liu (2004) for a more comprehensive valuation model, which takes into account the time-variation of discount rates. As they show, allowing time-variation in discount rate leads to much more complicated valuation formula.

[7] Empirically, these two shocks are positively correlated on the stock level.

return expectation biases of investors with CDR beliefs.

In the presence of investors who hold CDR beliefs (the CDR investors), equilibrium asset prices could exhibit cross-sectional anomalies. Intuitively, investors holding CDR beliefs incur return expectation biases, which leads them to buy too much/little of certain stocks, causing over/under valuation compared to the CAPM benchmark. Since the most overvalued stocks (high cash flow growth and uncertainty) also exhibit high-degree of comovement in their asset payoffs (Ball, Sadka, and Sadka (2009), Herskovic et al. (2016)), rational arbitrageurs who are averse to taking systematic risk do not trade aggresively against the CDR investors.[8] As a result, mispricing persists in equilibrium.[9]

Subsequently, I test the implications of the CDR hypothesis using data on investors' subjective beliefs, firm fundamentals and asset prices and find supporting evidence. First, CDR investors' return expectation biases should be higher for stocks with higher expected cash flow growth and cash flow volatility. Using sell-side analysts return expectations, I find that analysts' long-term growth expectations and idiosyncratic volatility, which proxy for stocks' expected cash flow growth and cash flow volatility, respectively, strongly positively predict future analysts' return forecast errors. Furthermore, these two characteristics alone explain 34% of cross-sectional variations of average log forecast errors among all stocks. In addition, the biases in analysts return forecasts are mostly positive, consistent with the CDR hypothesis.

Second, a measure of mispricing on the firm-level based on the hypothesis strongly predicts future stock returns. Roughly speaking, the measure is the difference between the CDR implied expected return subtracted by a version of expected return implied by the conditional CAPM. For the former, I chose the ICC model developed by Pástor, Sinha, and

[8]This potentially could be due to their high exposure to aggregate discount rate shocks: stocks with higher cash flow duration are more exposed to aggregate discount rate shocks. More detailed discussion on this channel can be found in Santos and Veronesi (2010) and Lettau and Wachter (2007a).

[9]The mechanism is similar to the one disussed in Kozak, Nagel, and Santosh (2018), except here the cash flow growth and uncertainty are the key factors that drive the co-movement in asset fundamentals, while in their setting the characteristics are not explicitly specified.

Swaminathan (2008) (henceforth PSS) as a measure.[10] For the latter, I use a measure of dynamic beta times a constant.[11] Consistent with the CDR hypothesis, the measure negatively predicts future stock returns and the economic magnitude of the predictability is large: the stocks with the highest overvaluation significantly underperforms those with the least overvaluation even within the S&P 500 universe (FF-5 alpha of 6% with a t-stat of 3.53) and the underperformance persists after 5 years.

Third, a tradable factor mimicking portfolio based on the CDR hypothesis explains the CAPM-alphas of 12 prominent cross-sectional anomalies (9 out of 11 in Stambaugh and Yuan (2017)). The set of anomalies include the most robust persistent ones found in the literature, such as investment, profitability, beta, idiosyncratic volatility and cash flow duration. These characteristics all forecast future expected cash flow growth and/or idiosyncratic volatility with the signs consistent with the CDR hypothesis.

After the literature review, the rest of the paper is organized as follows. First, I develop a theoretical framework to formalize the CDR hypothesis and to guide empirical analysis in 2.1. This section provides an expression for return expectation bias under CDR (Section 2.1.3) and stocks' average CAPM-alpha (Section 2.1.4). Next, I empirically test the hypothesis in Section 2.2. This section starts with implications on subjective beliefs (Section 2.2.1), followed by implications on asset prices (2.2.2). I conclude in Section 2.3.

Related Literature

This paper contributes to literature that aims to explain asset pricing anomalies by relaxing the rational expectation assumptions of agents. This literature typically uses non Bayesian expectations grounded in psychology to explain the behavioral biases of irrational

[10]The main reason for using this model instead of other ICC models is that PSS is more applied in the finance literature, see for example, Chen, Da, and Zhao (2013). I do examine other models in the Appendix, which includes the model of Gebhardt, Lee, and Swaminathan (2001) and find similar results for the main tests.

[11]The beta is pre-estimated CAPM betas of Welch (2019). The constant is the average market return subtracted by an adjustment due to market-level bias caused by the CDR assumption. The resulting measure of mispricing uses analysts earnings forecasts, stock prices, payout ratios.

market participants. For example, investors overreact or underreact to news, which can lead to return reversal and momentum (Daniel, Hirshleifer, and Subrahmanyam (1998), Barberis, Shleifer, and Vishny (1998), Hong and Stein (1999)). Furthermore, Bordalo et al. (2019) and Bouchaud et al. (2019) propose "diagnostic expectation" and "sticky expectation" dynamics, respectively, and they use them to explain the abnormal returns of analysts' long-term-growth-estimates-sorted portfolios and profitability-sorted portfolios, respectively. Most recently, Barberis, Jin, and Wang (2020) try to use prospect theory to explain cross section anomalies. The current paper contributes to the literature by proposing a new type of investor belief that deviate from rational expectation, namely the CDR heuristic. Such a belief has never been studied in the literature. Furthermore, I show the CDR heuristic can be a ubiquitous channel through which many anomalies can occur.

This paper also contributes to the literature that tries to understand how investors form their subjective return expectations. One strand of the literature studies investor surveys, such as Greenwood and Shleifer (2014), Adam, Marcet, and Beutel (2017) and Adam, Matveev, and Nagel (2021b). The key findings are that investors' subjective return expectations deviate from the rational expectations typically assumed in workhorse asset pricing models. As an example, the results in Greenwood and Shleifer (2014) show that return expectations from CFOs and retail investors are strongly pro-cyclical, a result which is the opposite of the counter-cyclical "risk premium" in models such as Campbell and Cochrane (1995). The other strand of the literature uses fund flows to infer the asset pricing models investors use, for example, Berk and Van Binsbergen (2016) and Barber, Huang, and Odean (2016), although Jegadeesh and Mangipudi (2020) disputes the validity of their results. This paper contributes to the literature by using asset pricing moments to infer what kind of expected return models investors could use. The empirical results in the current paper support the hypothesis that the expected return models investors use fail to take into account the dynamic nature of future returns and therefore restrict the set of potential candidate (subjective) return expectation models.

Finally, the paper also contributes to the recent literature that aims to explain a wide set of anomalies using a small number of factors. Recent examples include Fama and French (2016), Hou, Xue, and Zhang (2015), Stambaugh and Yuan (2017), Daniel, Hirshleifer, and Sun (2019). The current paper contributes to the literature by showing that investors' constant discount rate assumption can be an important force in driving many asset pricing anomalies, both theoretically and empirically. Instead of extracting a small number of factors from a larger set of anomalies, as in Fama and French (2016) and Stambaugh and Yuan (2017), this paper starts from a single expectation dynamic and constructs the factor using expectations data.

2.1 The Constant Discount Rate (CDR) Hypothesis

In this section, I develop the CDR hypothesis and derive its implications for subjective beliefs and asset prices to guide the empirical analysis. I start with a stylized example to provide intuition on how the CDR assumption can lead to biases in return expectations. I then extend the simple example into a more realistic setting and derive an analytical expression of the biases, which is a function of firms' fundamental characteristics. Subsequently, I show how biases in return expectations can lead to mispricing in equilibrium and derive an expression of a stock's CAPM-alpha, which is a function of the expectation bias.

2.1.1 The CDR Investors' Investment Process

An investor who holds the CDR beliefs (the CDR investor) forms her expectations as follows. First, she looks at a stock's multiple, such as price-dividend or price-earnings ratio. Additionally, she also projects the firm's fundamentals, including expected future cash flows as well as the uncertainty of the cash flows. Finally, she forms her return expectations by combining the two pieces of information she gathered: she evaluates what the current expected return is given the multiple and projected cash flows based on a present value model. Crucially, the model she uses does not assume discount rate would vary over time, which ultimately leads to return expectation biases.

The setting is supported by how investors form return expectations in practice. As shown in Mukhlynina and Nyborg (2016), practitioners mostly use price multiples and projected future cash flows to infer future returns of stocks. The logic is that given a projected future cash flow growth, one stock should have a higher expected returns if its prices are low. The process can be justified by the famous Campbell-Shiller decomposition, which states that the (log) price dividend ratios of different stocks differ because of either different expected future cash flow growth, expected future returns or both. [12]

2.1.2 A Stylized Example: the Three-period Case

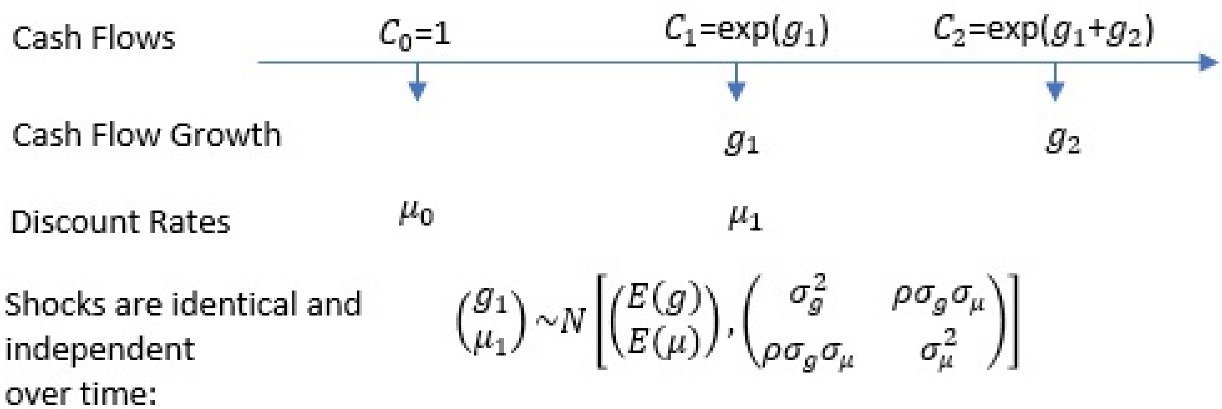
A stock is expected to pay risky dividends for two periods after paying \$1 at period 0. The (log) dividend is expected to grow at a stochastic rate of $g_t, t = 1, 2.$ and the discount rate used to price the stock for the two cash flows are μ_0 and μ_1, respectively. Figure 2.1 illustrates this example. While time 0 discount rate is known to investors, the discount rate at time 1, μ_1, is stochastic. This is because the risk of the stock and the market may change in period 1, and the discount rate should reflect the uncertainty. More importantly, discount rate shocks are correlated with cash flow shocks and follow a bi-variate normal distribution in this example. The fair price of the stock at time 0, after the dividend payout, should be $P_0 = P_0^{(1)} \left[1 + E_0(C_1) exp\left(-E(\mu) + \frac{1}{2}\sigma_\mu^2 - \rho\sigma_\mu\sigma_g\right)\right]$, where $P_0^{(1)}$ is the present value of the period 1 cash flow.[13] On the other hand, the CDR Investors, who ignore the discount rate volatility, would interpret the price through $P_0 = P_0^{(1)} \left[1 + E_0(C_1) exp\left(-\tilde{\mu}_1\right)\right]$, where $\tilde{\mu}_1$ is

[12] Although this decomposition provides an intuitive framework to understand asset prices, it is a heuristic based on a first-order Taylor approximation that ignores the discount rate volatility. In fact, the discount rate volatility does not impact prices in this framework.

[13] More specifically,

$$\begin{aligned}
P_0 &= e^{-\mu_0} E_0(e^{g_1}) + E_0(e^{-\mu_0 - \mu_1} C_2) \\
&= e^{-\mu_0} e^{E(g) + \frac{1}{2}\sigma_g^2} + e^{-\mu_0 + 2E(g) + \sigma_g^2} e^{-E(\mu) + \frac{1}{2}\sigma_\mu^2 - \rho\sigma_\mu\sigma_g} \\
&=: P_0^{(1)} \left[1 + E_0(C_1) exp\left(-E(\mu) + \frac{1}{2}\sigma_\mu^2 - \rho\sigma_\mu\sigma_g\right)\right]
\end{aligned}$$

Figure 2.1: A two-period example



their subjective belief on how the market is discounting the stock.[14] Consequently, the CDR heuristic leads to a bias of $b = \sigma_\mu(-\frac{1}{2}\sigma_\mu + \rho\sigma_g)$.[15]

Empirically, the biases in return expectations b are positive on average. According to Vuolteenaho (2002), discount rate shocks are positively correlated with cash flow shocks at the firm level, and the magnitude of cash flow shocks are much larger than discount rate shocks. In particular, based on his estimates,[16] the CDR investors should have a bias (per year) of $b = 0.14 \times (-\frac{1}{2} \times 0.14 + 0.47 \times 0.29) = 0.0092$ for a typical stock, which means that the CDR investors would on average consider the prices as too cheap and consequently buy more of the stock, creating overpricing.

2.1.3 Biases in the Infinite-period Case

A more realistic setting is to consider stocks paying out in infinite periods. In this case, firm-specific fundamental characteristics, namely cash flow growth and uncertainty,

[14] More specifically,

$$\begin{aligned} P_0 &= e^{-\mu_0}E_0(C_1) + e^{-\mu_0-\tilde{\mu}_1}E_0(C_2) \\ &= e^{-\mu_0}E_0(C_1) + e^{-\mu_0+2E(g)+\sigma_g^2}e^{-\tilde{\mu}_1} \\ &=: P_0^{(1)}\left[1 + E_0(C_1)exp\left(-\tilde{\mu}_1\right)\right] \end{aligned}$$

[15] In the infinite-period, dynamic case, the unconditional bias becomes $b^i = \left\{1 - exp\left[\sigma_\mu^i(\sigma_\mu^i - \rho^i\sigma_c^i)\right]\right\} exp(g^i + \frac{1}{2}(\sigma_c^i)^2)$ for stock i. I discuss this more in detail in 2.1.4.

[16] More specifically, see Table III and Panel B of Vuolteenaho (2002).

drive biases in return expectations in the cross section. Appendix B.1 derive the analytical expressions for the biases in this general setting.

Intuitively, this is because investors incur these biases each period as in the three-period case, and stocks differ in the timing (duration) and uncertainty (convexity) of their future cash flows, which means the single period biases are compounded to a different degree across stocks. Stocks with higher cash flow growth have higher cash flow duration, because most of their cash flows are further into the future and therefore have longer payout horizon. As for convexity, higher cash flow volatility means a stock's price will be a more convex function of the discount rate.[17] In the bond valuation context, the same level of bias in the discount rate will translate into a larger misvaluation for bonds with higher duration and/or higher convexity. Since stocks' cash flow growth and uncertainty exhibit large heterogeneity in the cross-section, the cross-sectional differences in misvaluations due to the CDR assumption are likely large.[18][19]

I confirm the intuition by considering a more general framework in Appendix B.1. The analysis shows that the biases in return expressions can be analytically linked to firm-level fundamental characteristics including expected cash flow growth and cash flow (idiosyncratic) volatility. More specifically, the unconditional bias b^i, where i denote the stock, is given by

$$b^i = \delta^i exp(g^i + \frac{1}{2}(\sigma_c^i)^2) \tag{2.1}$$

and g^i and σ_c^i are expected growth and volatility of the cash flow growth, respectively. δ^i is

[17]See Dechow, Sloan, and Soliman (2004a), Weber (2018) and Gormsen and Lazarus (2019) for related discussions about the duration channel. Notice here the convexity goes to the other direction compared to the conventional bond convexity because here the convexity is measuring stock price's relationship with its cash flows volatility. See Pástor and Veronesi (2006) for a discussion on how cash flow uncertainty impacts stock valuations.

[18]For example, TSLA's long-term cash flows are a magnitude faster and more uncertain than those of Coca-Cola: Sell-side analysts' long-term growth expectation for TSLA is at 74% as of May 2020, compared to 2.93% for Coca-Cola. Furthermore, it is reasonble to think that the cash flows of TSLA is much more uncertain than Coca-Cola.

[19]In this example, the shocks are i.i.d. over time. In a more realistic case when discount rate shocks are persistent, the biases are likely to be larger. Intuitively, when CDR investors ignore the discount rate dynamics, they are also ignoring the long and persistent effects that the volatility may imply.

defined as

$$\delta^i = 1 - exp\left[\sigma_\mu^i(\sigma_\mu^i - \rho^i\sigma_c^i)\right] \tag{2.2}$$

which also depends on volatility of stock's discount rates (σ_μ^i) and correlations between the discount rate and cash flows (ρ^i).

Based on previous estimates in the literature, including those of Vuolteenaho (2002), δ^i should be positive, which means CDR investors' return expectation biases are on average positive and should increase with both expected cash flow growth and cash flow volatility. I provide more discussions about the signs and the magnitude of the biases in Appendix 2.2.1. Moreover, I verify empirically the signs and magnitude of biases based on both analysts' return expectations and price implied measures of constant discount rate and confirm they are positive in Section 2.2.1 and 2.2.2, respectively.

2.1.4 Biased Return Expectations and Equilibrium Asset Prices

When some investors hold CDR beliefs and make their investment decisions based on their own return expectations, a stock's CAPM-alpha should depend on the bias in CDR investors' return expectation as well as the share of the CDR investor in the economy. Intuitively, a positive bias in return expectations should lead an investor to buy more of a stock, causing over-valuation and low CAPM-alpha. More CDR investors would exacerbate the misvaluation in equilibrium.

More formal analysis in Appendix B.2 confirms this intuition. There, I study a multi-asset economy in which some investors with biased return expectations (CDR investors) trade with risk-averse rational investors (arbitrageurs). The setting is similar to the one studied in Kozak, Nagel, and Santosh (2018).[20] More specifically, the unconditional CAPM-alpha of stock i in the model is given by

[20] In Kozak, Nagel, and Santosh (2018), they study a model where the covariance matrix of the payoffs of assets are driven by several principle components. In my model, these biases are linked to firm-level fundamental characteristics, which are the driving force behind comovement in fundamental payoffs.

$$\alpha^i = \theta(-b^i + \beta^i b^M) \tag{2.3}$$

where θ is the share of CDR investors and b^i is the return expectation bias that potentially equal to the one defined in (2.1).[21] The β^i are commonly defined CAPM beta and b^M is the aggregate bias CDR investors hold on the market level. So the expected return on the stock i is

$$E(R^i_{t+1}) - R_f = \alpha^i + \beta^i \left[E(R^M_{t+1}) - R_f\right] \tag{2.4}$$

2.1.5 Testable Implications

The theoretical analysis yields two sets of testable implications under the CDR hypothesis for investors' subjective beliefs and asset prices.

The expression in Equation (2.1) and (2.2) leads to testable implications on investor beliefs. First, if the CDR is true, CDR investors' unconditional return expectation biases across different firms should be largely explained by proxies of stocks' expected cash flow growth and cash flow volatility. Second, if the term δ^i is positive, which is also empirically testable, the biases should increase with expected cash flow growth and volatility.

Equation (2.3) and (2.4) suggest that if one can measure the bias b^i on the stock level based on CDR, she can test the CDR based directly using asset prices and returns. First, the measure of CDR-induced bias should also be positive and related to expected cash flow growth and cash flow volatility. Second, the measure for the bias or an ex-ante measure of CAPM-alpha according to (2.3) should predict stocks' realized CAPM-alpha. Finally, if the CDR is true, CAPM-alphas of all assets should be explained by a factor mimicking portfolio that is based on the ex-ante measure of CAPM-alpha. In fact, the loading on the portfolio should equal to 1.

I examine the extent to which these implications are true in Section 2.2.

[21] The model developed in Appendix B.2 is a general one, which could be due to any form of return expectation bias. When the bias is measured empirically based on the logic of the CDR, the asset pricing test will be a test of the CDR hypothesis, which is what I do in the next sections.

2.2 Evaluating The Constant Discount Rate Hypothesis

In this section, I test the CDR hypothesis empirically, guided by the theoretical framework developed in the previous section. I start by testing the implications on investors' subjective beliefs based on sell-side anaylsts' return expectation data. Next, I develop a measure of mispricing due to the CDR assumption and use it to test implications for asset pricing.

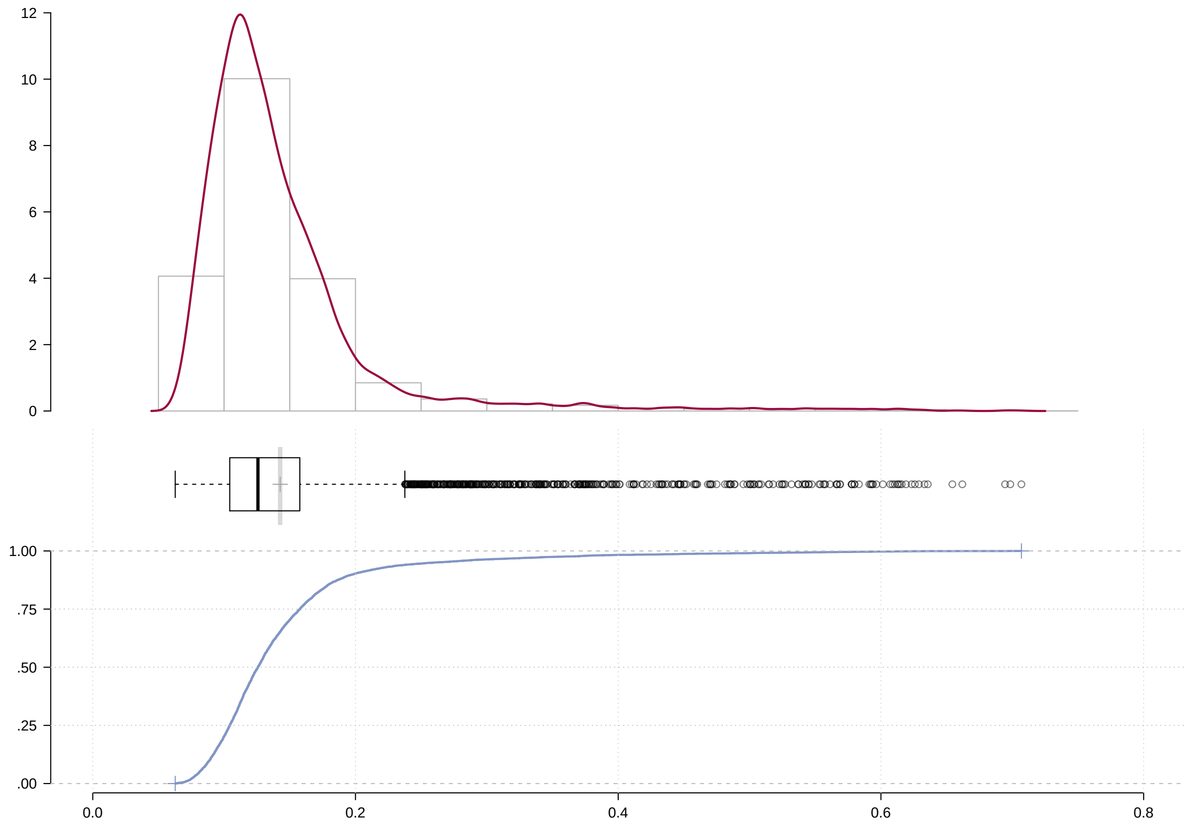
2.2.1 Testing Implications on Subjective Beliefs

Measuring Subjective Return Expectation Biases Using Sell-Side Analysts' Price Targets

The CDR hypothesis does not make prediction about who are the CDR investors. I test the implications of CDR hypothesis using sell-side analysts' return expectations data. Renxuan (2020b) find that sell-side analysts do underestimate the volatility of discount rates on the aggregate. Furthermore, survey evidence from Mukhlynina and Nyborg (2016) shows that sell-side analysts do consider the commonly used Discounted Dividend Model (DDM) as the main approach they use for valuation. Additionally, sell-side analysts' return expectations have comprehensive coverage on the firm-level, which is unique when compared to surveys on CFOs or households, for example.

The firm-level return expectation biases are defined as 12-month realized price returns on a particular stock, minus the sell-side analysts' ex-ante firm-level consensus return expectations at the end of each calendar quarter. The sell-side analysts' return expectations are defined as price targets divided by current prices subtracted by 1. Details about the data set as well as the construction are documented in Appendix B.5. Firm-level average return expectation biases are computed as the time-series average over the entire history of firm-level biases.

Figure 2.2: Distribution of average firm-level analyst forecast errors of 12-month ahead returns



Notes: The top and bottom panel plot the empirical probability distribution function (P.D.F.) and cumulative distribution function (C.D.F.) of average sell-side analysts' return forecast errors, respectively. Dark bar in the middle represents the median while the gray bar with cross represents the mean. x-axis is the value of the average biases while y-axis denotes probability in percentage point. The forecast errors are constructed based on sell-side analysts' 12 month price targets subtracted by realized average returns. More details about how the return expectations are computed are documented in Appendix B.5. Firm level forecast errors are averaged over time to arrive at average forecast error per firm. The sample period is from 1999-Q2 to 2018-Q4.

The Sign and Magnitude of Subjective Return Expectation Biases Analysis in the simple example of Section 2.1.2 makes clear that the sign and magnitude of the return expectation biases are crucial for the predictions of the CDR hypothesis. Therefore I verify the sign before proceeding to test other implications.

Figure 2.2 plots the empirical distribution of the average firm-level return expectation bias of sell-side analysts, together with the mean and median (bars in the middle). Subjective return expectations of sell-side analysts are on average positive at firm level and right skewed.

The empirical results are consistent with the findings in the literature, which have previously documented the positive biases of sell-side analysts.[22] The literature has mostly attributed the positive bias to analysts' own incentive, such as their own career concerns (Hong and Kubik (2003)). The CDR hypothesis provides an alternative interpretation to such a positive bias: analysts have a higher return forecasts because they ignore discount rate dynamics, which might be a honest mistake.

The Cross-Sectional Variation in Return Expectation Biases and Firm Characteristics

The CDR hypothesis predicts that the CDR investors' cross-sectional variations of unconditional return expectation biases should be driven by expected cash flow growth and volatility. Furthermore, given the positive signs of the return expectation biases, the biases should increase with these measures of these two characteristics.

I test the hypothesis by regressing average firm-level sell-side analysts' return forecast errors on analysts' long-term growth estimates and idiosyncratic volatility, which are proxies for the expected cash flow growth and volatility, respectively. Table 2.1 presents the regression results for both the entire stock universe with analyst return expectation coverage and the S&P 500 universe. I also contrast the results with regressions using four other firm-level characteristics know to be related to stock average returns and volatility.

[22] Papers which document large positive bias of analyst price targets include Brav and Lehavy (2003b) and Engelberg, McLean, and Pontiff (2019a).

The results support the CDR hypothesis. First, as Column (1) shows, both expected cash flow growth and volatility are strongly positively correlated with average return expectation errors. Analysts' biases are more positive if a stock has higher long-term growth expectation and/or higher idiosyncratic volatility. Furthermore, the two characteristics alone explain 34% of the cross-sectional variation, as indicated by the R2 of the regression. As a benchmark, when using four other characteristics to explain average return forecast errors (Column (2)), the R2s are much lower. The results are robust across different universes (Column (3) and (4)) and also hold for panel predictive regressions using quarterly data (see Internet Appendix).

2.2.2 Testing Implications on Asset Prices and Returns

Measuring CDR-induced Misvaluation

I propose an intuitive measure for the bias $\widehat{b}_t^i$ that an CDR investor would incur. The measure is the difference between the Implied-Cost-Capital developed (ICC) by Pástor, Sinha, and Swaminathan (2008) $\widehat{\Pi}_t^i$, and the product of dynamic beta $\widehat{\beta_t^i}$ developed by Welch (2019) and a constant that equal to average market excess return ($\widehat{E(R_t^m)}$), or

$$\widehat{b}_t^i = \widehat{\Pi}_t^i - \widehat{\beta_t^i}\widehat{E(R_t^m)} \tag{2.5}$$

The ICC captures the essence of the return expectation of a CDR investor: it is computed using price and projected cash flows based on a present value formula that ignores the volatility of the discount rate. The second term in Equation (2.5) is a proxy for the "true" dynamic expected return. What a true expected return is has still been hotly debated in the literature. Here I take a stand similar to that of Binsbergen and Opp (2019).[23]

Equipped with a measure of bias, I follow Equation (2.3) to construct a measure of misvaluation for individual stocks, $\widehat{\alpha}_t^i$. The misvaluation measure takes the following form:

[23] I also test the hypothesis using other proxies of true expected returns, but the results do not vary qualitatively.

Table 2.1: The cross-sectional determinants of average firm-level forecast errors of sell-side analysts

	Dependent variable:			
	average log forecast errors			
	(1)	(2)	(3)	(4)
CF growth	0.285***		0.347***	
	(0.010)		(0.031)	
I-Vol	0.520***		0.510***	
	(0.019)		(0.040)	
Investment		0.038***		0.139***
		(0.005)		(0.031)
Proftability		−0.002**		−0.001
		(0.001)		(0.001)
Beta		0.125***		0.284***
		(0.017)		(0.026)
B/M		−0.057***		0.027
		(0.013)		(0.024)
universe	all	all	SP500	SP500
Observations	4,691	3,945	814	1,005
R^2	0.336	0.045	0.323	0.152
Adjusted R^2	0.336	0.044	0.321	0.149
Residual Std. Error	0.291 (df = 4688)	0.347 (df = 3940)	0.217 (df = 811)	0.247 (df = 1000)
F Statistic	1,186.115*** (df = 2; 4688)	46.029*** (df = 4; 3940)	193.568*** (df = 2; 811)	44.940*** (df = 4; 1000)

Note: *p<0.1; **p<0.05; ***p<0.01

Notes: "Average log forecast errors" are the log of sell-side analysts' 12-month return forecast errors defined in Section 2.2.1. Independent variables are time-series average at the firm level based on quarerly data. "CF growth" is the average analyst long-term growth estimates; "I-Vol" are idiosyncratic return volatility measured using 60 days of daily returns and Fama-French 3 factor model; "Inestment" are change in total assets from the fiscal year ending in year t-2 to the fiscal year ending in t-1, divided by t-2 total assets at the end of each June using NYSE breakpoints; ""beta" are measured using the last 5 years of monthly returns; "B/M" are book-to-market ratio defined as in Fama and French (2015). The sample period is from 1999-Q2 to 2018-Q4.

$$\widehat{\alpha}_t^i = -\widehat{b}_t^i + \widehat{\beta}_t^i \widehat{b}_t^M \quad (2.6)$$

$$= -\left[\widehat{\Pi}_t^i - \widehat{\beta}_t^i \widehat{E(R_t^m)}\right] + \widehat{\beta}_t^i \left[\widehat{\Pi}_t^m - \widehat{E(R_t^m)}\right] \quad (2.7)$$

$$= -\widehat{\Pi}_t^i + \widehat{\beta}_t^i \left[\widehat{E(R_t^m)} - \widehat{b}_t^M\right] \quad (2.8)$$

Equation (2.8) can be estimated empirically. I follow closely the procedure developed by Pástor, Sinha, and Swaminathan (2008) to estimate Π_t^i.[24] Appendix B.6 details the procedure I use to estimate the ICC.[25]I estimate the dynamic β_t^i using the methodology proposed by Welch (2019).[26]I fix $\widehat{E(R_t^m)} = 0.064$, which is the average of market returns in the post-war sample and $\widehat{b_t^M}$=-2.3%, which is the calibration provided by Hughes, Liu, and Liu (2009).[27]

Summary Statistics of the Misvaluation Measure I use I/B/E/S summary file for analyst earnings and price targets forecasts, COMPUSTAT annual for balance sheet variables and CRSP for shares outstanding, share adjustment as well as price and return related variables. More detailed descriptions of the data sources are in B.4.

Compared to the mostly commonly used CRSP-COMPUSTAT universe, the universe used here cover only about 40% of the number of firms and has larger firms. This is because analysts typically only cover larger firms and the results presented are not concerning microcaps.

Estimating firm-level misvaluation requires 6 firm-level variables, 1 industry-level variable

[24] To examine that the results are robust, I consider alternative models developed in the literature, such as Gebhardt, Lee, and Swaminathan (2001) and find similar results.

[25] One concern with using these models are that the analyst estimates are not unbiased. However, as shown in Hou, Dijk, and Zhang (2012) and Wang (2015), compared to statistical models proposed in Hou, Dijk, and Zhang (2012), the analysts are not worse than statistical modes when predicting future cash flows in the same universe that have analyst coverage, especially for large cap stocks, where analysts are better in accuracy.

[26] Welch (2019) shows that empirically, his measure is superior than other estimates in some dimensions, including better performance in capturing the future realized beta.

[27] See Appendix B.1.3 for more detailed discussion on the sign and magnitude of market-level bias.

and 1 aggregate variable. The firm-level variables are: 3 analyst's consensus forecasts for a firm's earnings of current fiscal year (FY1), the next fiscal year (FY2) and the fiscal year thereafter (FY3); 1 analyst's consensus long-term forecast (LTG); 1 payout ratio, which is based on the firm's previous year total dividend to firm's net income and market β. The industry-level variable is the average LTG based on 48 Fama-French industry classification. The aggregate variable is the long-term average of gdp growth, which ranges from 7% to 6% over the 35 years in the sample. Based on these 5 inputs, I compute the implied cost of capital $\Pi_{i,t}$ and the entire term structure of a firm's payout ratio $PB_{i,t+s}$. More details on the estimation procedure is documented in Appendix (B.6).

One important point to note is that the estimation of firm-level misvaluation α^i does not include any anomaly variables that I will try to explain, except for β^i. However, β^i is mechanically positively related to α^i (since $\left[\widehat{E(R_t^m)} - (-\widehat{b^m})\right]$>0 in my estimation), while the misvaluation factor, is able to explain the "low-beta" anomaly. Therefore, the result that misvaluation factor being able to explain the anomaly returns of characteristics sorted portfolios can not be attributed to mechanically using these underlying characteristics when constructing the misvaluation measure.

Implied cost of capital Π^i are highly persistent, with an AR(1) coefficient of 0.92 based on quarterly data. I study the persistence of misvaluation α_t^i in more detail in Section 2.2.2. Table B.2 presents the summary statistics for firm-level quarterly estimates of Π_t^i, together with variables that are used to construct it. The statistics are inline with those presented in Chen, Da, and Zhao (2013), which also estimates the ICC based on Pástor, Sinha, and Swaminathan (2008).

Magnitude of Misvaluation Table 2.2a shows the empirical distribution of the 7246 average firm-level misvaluations, or $\sum_{t+1}^{T} \widehat{\alpha}_t^i / T$, which are the empirical estimate for $E(\hat{\alpha^i}_t)$.

First, all the misvaluation has a negative sign. This is reasonable because when ignoring the volatility of the discount rates, investors actually underestimates on average the discount rate and therefore over price stocks. In fact, this results are consistent with the calibration results from Ang and Liu (2004), in which they find the average misvaluation among value-growth portfolios to be about -15%.

The CDR implied misvaluation has a cross-sectional standard deviation of more than 7% per year. Notice that the statistics presented in Table (2.2) potentially underestimate the magnitude of the cross-sectional dispersion of misvaluation of dynamically sorted portfolios. As shown in the second row of Table 2.2a, the average time-series quarterly variation on the firm-level misvaluation is 2.3%, which translates to 4.6% per year. When constructing dynamically rebalanced portfolio annually, one would expect the spreads in ex-ante misvaluation spreads to go up substantively [28].

To provide an intuition for the time series as well as cross-sectional variations of misvaluations, Figure 2.3 plots the time series variation of $\hat{\alpha}_t^i$ for three firms. Besides the variations, the figure also shows the persistent nature of misvaluation.

Misvaluation and Firm Characteristics Under CDR, misvaluation α^i equals to the average CAPM-alpha of the stock. The CDR thus links a firm's CAPM-alpha directly to its characteristics through the relationship between misvaluation and characteristics.

To understand what drives the variation in misvaluation, note a stock's misvaluation or α^i , under CDR, can be decomposed as follows:

$$\alpha^i = \beta^i b^m - b^i$$

Table 2.2b shows that the cross-sectional variation in misvaluation is mainly driven the variation in biases b^i due to CDR. The standard errors of the bias is more than three times that of the expected return based on conditional CAPM.

The biases can be further decomposed into two separate components as

[28]for example 5.5% + 4.6%*2 = 14.7% when using the inter-quartile range and time-series standard errors as an indication.

$$
\begin{aligned}
b^i &= \delta^i \times \lambda^i \\
\delta^i &= 1 - exp\left[\sigma^i_\mu(\sigma^i_\mu - \rho^i \sigma^i_c)\right] \\
\lambda^i &= exp(g^i + \frac{1}{2}(\sigma^i_c)^2)
\end{aligned} \tag{2.9}
$$

As discussed before, the sign of the bias depends on δ^i, which is hard to estimate. However, λ^i involves only fundamental expectation variables that are commonly measured based on analysts' growth forecasts. I use analysts' long-term growth estimates to estimates g^i and its 36 month volatility to estimate σ^i_c. Since we observe estimates for b^i and λ^i, I back out the values of δ^i using (2.9).

Table 2.2c shows that the cross-sectional variation in biases b^i is mainly driven by λ^i, or characteristics related to expected fundamentals. Compared to to δ^i, λ^i has two times much the standard errors (9.5% vs. 4.7%).

These empirical results helps to understand cross-sectional relationship between firm characteristics and CAPM alphas. Under CDR, the CAPM-alpha comes entirely from misvaluation, which in turn is mostly driven by g^i and σ^i_c, via λ^i. As a result, through the CDR channel, certain characteristics can predict future returns or CAPM alphas only because these characteristics can predict future fundamental growth or fundamental volatility.

The sign of δ^i determines the relationship between fundamental characteristics and future CAPM alpha. Figure 2.4 shows the empirical distribution of δ^i. The figure confirms the previous conjecture that δ^i is mostly positive, as a result of a dominant cash flow news and positive correlation between cash flow and discount rate news on the firm level. This result also leads to a prediction of the CDR hypothesis with respect to the sign when using characteristics to forecast future CAPM-alphas: firm characteristics that positively predict future cash flow growth and/or volatility in the cross-section will predict negatively the CAPM-alphas.

Table 2.2: Empirical Distributions of Key Variables for Mis-valuation (α^i)

The table presents the empirical distributions of the measure of misvaluation defined in Equation (2.6) and its consisting variables. Data are quarterly firm-level data from 1986-01 to 2018-12. Empirical distributions are summarized based on average variable values over the entire time series for each firm. "ts.sd(α_t^i) is the standard deviation of the quarterly misvaluation measure for each firm over its entire history. "N" is the number of firms. $E(b_t^i)$ is calculated based on $\widehat{\Pi_t^i} - \widehat{\beta_t^i}0.064$ and $E(\lambda_t^i)$ is estimated based on (2.9) using analysts' long-term growth estimates for g^i and its 36 months (with minimum 12 months) volatility to estimate σ_c^i . $E(\delta_t^i)$ is calculated by dividing b_t^i by λ_t^i. The sample is winsorized at 0.5% and 0.95%. Rank correlations are Spearman rank correlations calculated using quarterly firm-level data based on the whole sample.

(a) Empirical Distribution of $E(\widehat{\alpha}_t^i)$

variable	mean	std	min	p25	median	p75	max	N
$E(\alpha_t^i)$	-0.165	0.072	-0.739	-0.181	-0.149	-0.126	-0.070	7246.000
ts.sd(α_t^i)	0.023	0.018	0.000	0.012	0.019	0.029	0.188	7246.000

(b) Empirical Distribution of $E(\widehat{\beta}_t^i\hat{\lambda})$ and $E(\hat{b}^i)$

variable	mean	std	min	p25	median	p75	max	N
$\beta^i E(R^m)$	0.063	0.023	-0.007	0.047	0.062	0.078	0.172	7246.000
$E(b_t^i)$	0.080	0.071	-0.059	0.040	0.064	0.098	0.633	7246.000

(c) Empirical Distribution of Variables in $E(\hat{b}^i)$

variable	mean	std	min	p25	median	p75	max	N
$E(\lambda_t^i)$	1.160	0.095	1.065	1.112	1.137	1.174	2.058	7246.000
$E(\delta_t^i)$	0.065	0.047	-0.055	0.036	0.056	0.084	0.328	7246.000
$E(g_t^i)$	0.143	0.071	0.063	0.104	0.126	0.158	0.707	7246.000
$E(\sigma_{c,t}^i)$	0.047	0.042	0.000	0.021	0.034	0.058	0.389	7246.000

(d) Rank Correlation Between $\widehat{\alpha}_t^i$ and Consisting Variables

$cor(\alpha_t^i, \mu_t^i)$	$cor(\alpha_t^i, b_t^i)$	$cor(\alpha_t^i, \lambda_t^i)$
-0.319	-0.601	-0.962

(e) Rank Correlation Between b_t^i and Consisting Variables

$cov(b_t^i, \lambda_t^i)$	$cov(b_t^i, g_t^i)$	$cov(b_t^i, \sigma_{c,t}^i)$
0.747	0.776	0.034

Figure 2.3: Evolution of α_t^i of specific firms

The figure plots the quarterly time series of misvalution measure $\widehat{\alpha}_t^i$ of three companies. "AA" : Alcoa Corporation; "AMZN": Amazon.com Inc; "BA": Boeing CO.

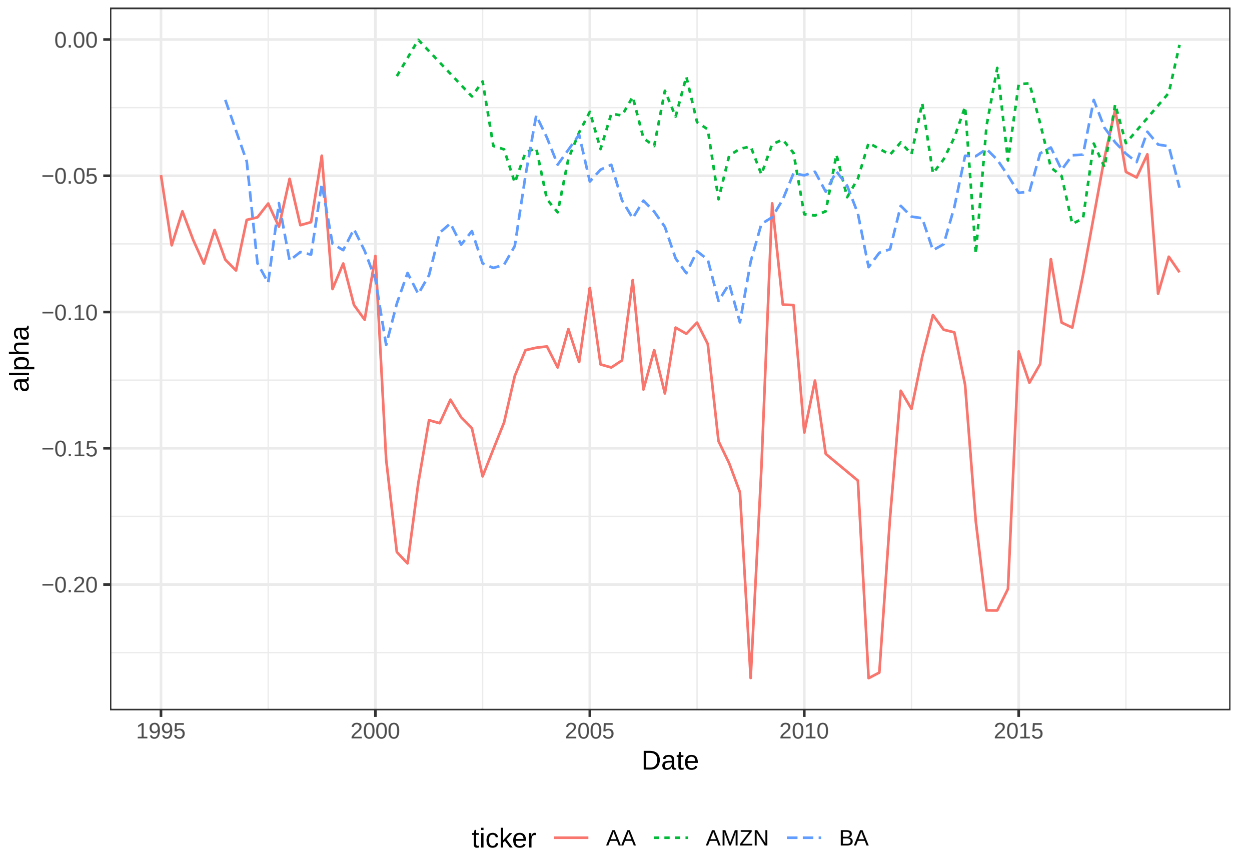
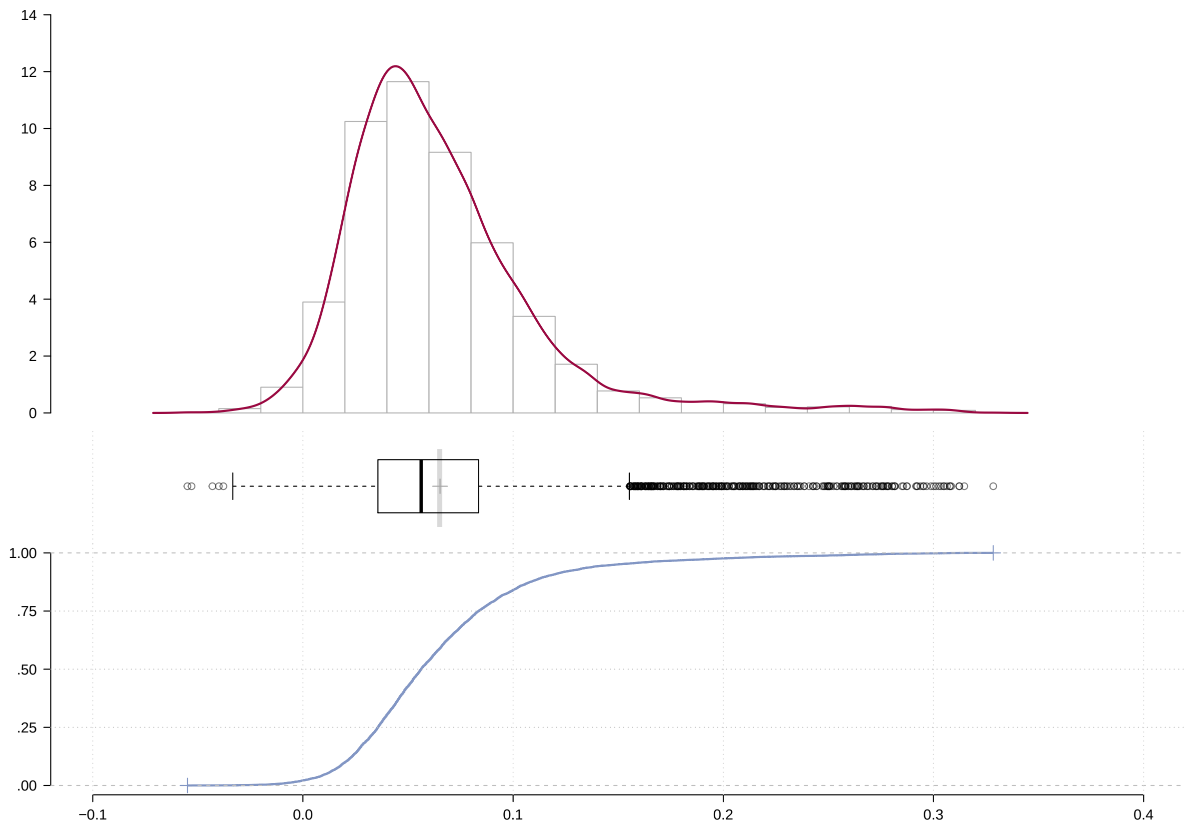


Figure 2.4: Probability Distribution Function (Top Panel) and Cumulative Distribution Function of $E(\delta_t^i)$

$E(\delta_t^i)$ is the firm-level time-series average of δ_t^i, which calculated by dividing b_t^i by λ_t^i. The sample is winsorized at 0.5% and 0.95%.



Misvaluation Sorted Portfolios

I test the first asset pricing implication of CDR, namely, the misvaluation measure $\widehat{\alpha}^i$ should positively predict a stock's CAPM-alpha. Furthermore, the average spreads in average realized CAPM-alphas should be close to the magnitude suggested by the spreads in ex-ante misvaluation measures.

To test this hypothesis, I follow the convention in the asset pricing literature (for example Fama and French (2015)) to sort stocks into quantile portfolios based on the misvaluation measure $\widehat{\alpha}^i$. I form portfolios at the end of June each year, using the available information up to that point[29]. I rebalance every month based on firms' market capitalization (value-weight) every month[30]. Effectively, the holding period of the trading strategy is 12 month. Table 2.3 presents the results.

Table 2.3 presents the detailed results with respect to the portfolios sorted based on msivaluation measure. These results support the hypothesis that investors do use constant discount rate in practice which leads to misvaluation.

First, the results in Panel A shows stocks which are most overvalued due to CDR have significant lower realized CAPM-alphas . The difference in realized CAPM-alphas between the least overvalued ("High") and most overvalued portfolios are 0.8% per month (9.6% per year). The spread in CAPM-alphas are statistically significant with a t-stat of 5. As shown in Panel B, since the spreads in the rational expected returns (μ^i) are moderate, most overvalued portfolios end up having a lower average returns than less overvalued stocks, which amounts to 0.8% per month (8.4% per year). In fact, the return spreads also have a Fama-French 5 factor alpha of 0.69% per month (8.28% per year), with a t-statistics over 5.

Second, the magnitude of the spreads between the misvaluations of high and low portfolios are very close to those of the realized CAPM-alpha spreads. This is consistent with the

[29]For the measure, the variables used are at latest available more than 2 weeks before being used to construct the ICC measures.

[30]I also presents the portfolio sorts using equal weights in B.7.1, which shows a larger spreads in CAPM-alpha.

prediction of CDR, that CAPM alpha should equal to misvaluation. As shown in Panel A, the firm-level average misvaluation within each portfolio ("Avg. $\widehat{\alpha}_t^i$") has a spreads between high and low of 0.77% compared to the realized CAPM-alpha of 0.80%. Furthermore, Panel C shows that what drives the misvaluation spread is consistent with what was analyzed in Section (2.2.2). The biases due to CDR, or b^i drive the difference in $\widehat{\alpha}^i$. And λ^i appear to have a bigger role in explaining b^i than δ^i, which are positive across all portfolios. Finally, among characteristics that consists of λ^i, growth expectation seems to be more important than growth volatility (σ_c^i) in driving the spreads of misvaluation across portfolios.

Additionally, notice the realized CAPM alphas of the portfolios are negative, except for the portfolio with the highest α^i. This might seem puzzling because the value-weighted CAPM-alphas should add up to zero , by construction. The reason for this result is two-folds. First, stocks with higher analyst coverage have on average lower returns, as shown in Hong, Lim, and Stein (2000) and Diether, Malloy, and Scherbina (2002). Stocks with a valid misvaluation measure need to have substantial analyst coverage, which is a dynamic universe smaller than the CRSP universe used to construct market excess returns. More specifically, to have a valid measure of misvaluation, I require the firm to have valid analyst forecasts for short-term (1 and 2 fiscal-year ahead) and long-term earnings. Second, stocks with a higher α^i are significantly larger than those with lower values, which further exacerbates the asymmetry among CAPM-alphas. More details about this point is presented in Appendix B.4, Table B.1.

The results in Table (2.3) also highlights two important questions to be addressed.

First, as shown in Panel A, the spread in the values of misvaluation across portfolios, does not seem to converge fast after the portfolio formation. From 12 months to 60 months after the formation of the portfolio, only narrows by 0.12% per month, or 1.44% per year. This means that the mispricing is highly persistent. A persistent effect means the misvaluation has a larger economic significance since it has implication for the long-run asset returns. I explore this further in Section (2.2.2) next.

Second, most overvalued portfolios tend to be smaller firms, as shown in Panel C. This result is intuitive as bigger companies receive more attention and attract more analysts to analyze. Prices should be more efficient and the probability of being misvalued should decrease. However, it does beg the question of whether and to what extent the misvaluation still present in the large caps. Cross-sectional phenomenon that only hold in small caps carry less economic significance for asset pricing theories, especially in recent years when large caps dominate the market. I therefore further investigate this issue in Section 2.2.2.

The Persistence of Misvaluation I demonstrate the economic significance of misvaluation due to CDR by showing that the misvaluation has persistent impact on asset prices.

First, $\widehat{\alpha}_t^i$ is a persistent variable. The pooled panel regression based on annual data shows $\widehat{\alpha}_t^i$ has an AR(1) coefficient of 0.948 (standard errors 0.006, clustered by firm and year). This means that the misvaluation measure has a half life of more than 13 years.

Consistent with the highly persistent measure $\widehat{\alpha}_t^i$, the turnover of the trading strategy constructed based on misvaluation has low turnover. Table 2.4a shows that the average portfolio turnover for the long and short side only amounts to 2% or less than 24% annually. Compared to the trading strategies analyzed in Novy-Marx and Velikov (2015), the 2% turnover would place the misvaluation trading strategy into the lowest turnover category, on par with profitability and only less than portfolios sorted based on size. This result means that transaction costs will unlikely render the CAPM alpha to zero.

Investors do not counteract the misvaluation effect quickly and stocks in the most over- (under-) valued portfolios under- (out-) perform even after 5 years after portfolio formation. Table 2.4b shows the returns of the trading strategy based on misvaluation for different holding periods. As shown in the table, the High-Low portfolio's CAPM-alpha still show up to be highly significant even for holding periods exceeding 60 months. The reduction in the return spreads as well as statistically significance amounts to 0.21% per month, from 12 months to 60 months holding period. This magnitude is also inline with the decay of the

Table 2.3: Pre-estimated Mis-valuation ($\widehat{\alpha}_t^i$) Sorted Portfolios and Realized Average Stock Returns (1986-06 to 2019-12)

The table presents the statistics related to portfolios sorted based on misvaluation measure created in (2.6). All numbers are expressed in percentages unless otherwise stated. Returns and alphas are based on monthly frequency.
Stocks are sorted into quantile portfolios based on the misvaluation measure $\widehat{\alpha}^i$ at the end of June each year, using the available information up to that point. Portfolios are rebalanced every month based on firms' market capitalization (value-weight). "Low" denote the portfolio with lowest $\widehat{\alpha}_t^i$. "High-Low" are excess returns of a portfolio that goes long on stocks with the highest $\widehat{\alpha}_t^i$ and short those with the lowest $\widehat{\alpha}_t^i$.
Panel A presents the average misvaluation after the portfolio formation for the next 12 months, 60 months as well as the average values for firms in the portfolio through the firm's lives.
Panel B presents statistics related to portfolio returns. "mean ex.ret" are monthly returns over 3 month treasury rates; "SE" are standard errors which are shown in brackets. "SR" are monthly Sharpe Ratios. "FF-5 alpha" denote Fama-French 5 factor alphas. "num_stocks" are average number of stocks included in the portfolio over time.
Panel C presents characteristics (value-weighted) associated each of the portfolio. g^i and σ_c^i are average portfolio analysts long-term growth expectation (LTG) as well as 36-month rolling volatility of the LTG.

	Low	**2**	**3**	**4**	**High**	**High - Low**
Panel A: Ex-ante Misvaluation vs. Realized Portfolio CAPM Alpha						
	Ex-ante Misvaluation					
Nxt. 12m $\widehat{\alpha}_t^i$	-1.98	-1.31	-1.18	-1.05	-1.01	0.98
Nxt. 60m $\widehat{\alpha}_t^i$	-1.91	-1.28	-1.17	-1.07	-1.05	0.86
Avg. $\widehat{\alpha}_t^i$	-1.80	-1.24	-1.14	-1.05	-1.03	0.77
Realized Portfolio	Realized Portfolio CAPM Alpha					
CAPM alpha	-0.80	-0.39	-0.26	-0.09	0.01	0.80
SE CAPM alpha	(0.14)	(0.10)	(0.09)	(0.06)	(0.06)	(0.16)
Panel B: Realized Portfolio Return Statistics						
mean ex.ret	-0.03	0.27	0.33	0.48	0.66	0.70
SE ex.ret	(6.12)	(4.93)	(4.56)	(4.24)	(4.85)	(3.29)
SR	-0.01	0.05	0.07	0.11	0.14	0.21
FF-5 alpha	-0.63	-0.34	-0.41	-0.23	0.06	0.69
SE FF-5 alpha	(0.11)	(0.09)	(0.08)	(0.06)	(0.06)	(0.13)
Panel C: Portfolio Characteristics						
Mkt.Cap (Million)	15379.69	33550.85	38340.65	47129.95	88655.92	73276.23
b^i	14.32	7.29	5.51	3.85	1.41	-12.91
μ^i	7.14	6.22	6.26	6.29	7.59	0.45
π^i	21.46	13.51	11.77	10.14	9.00	-12.46
λ^i	125.02	114.65	112.63	110.77	109.54	-15.48
δ^i	10.92	6.30	4.85	3.45	1.27	-9.65
σ_c^i	5.71	3.27	2.73	2.44	2.90	-2.80
g^i	21.46	13.51	11.77	10.14	9.00	-12.46

spreads in misvaluation measure it self, which amounts to 0.12% from 12 to 60 months.

For value-weighted portfolios, the persistence mainly comes from the continuing under-performace of stocks which are mostly over-valued due to CDR. While for the equally weighted portfolios, both long and short side continue to out- and under- perform. This result means there might exist an interaction between size and misvaluation measure. I investigate this in the next subsection.

Misvaluation In Different Size Segments of the Market I show that the misvaluation due to CDR also presents within the universe consists of the largest companies. Since a few large companies take up dominant share of the stock market, the finding that misvaluation presents in this part of the market means the channel of mispricing suggested by CDR is economically important.

Table 2.5 shows results of conducting an independent 3 by 3 double sort based on a stock's size and misvaluation. The CAPM alphas for the spread between most and least over-valued portfolio within the smallest companies is 1.08% per month (12.96% per year). However, even within the largest segment of the stock market, where the average market cap is more than 26 billion, the spread in CAPM-alpha is still 0.63% per month (7.56% per year), with a t-statistics of close to 5.

To further examine the economic significance of the CDR channel of misvaluation, I conduct the same portfolio sorting exercise within the SP500 universe, which contains the biggest U.S. companies and accounts for about 80% of all U.S. market capitalization available, as of Sep. 2020. Table 2.6 shows that even within this universe, the spread in CAPM alpha between the most and least overvalued stocks are 0.39% per month (4.68% per year). In fact, the FF-5 alpha is higher, at 0.53% per month, thanks to the fact that the returns load strongly negatively on the SMB factor.

Finally, portfolio characteristics in both Table 2.6 and 2.5 show that the spreads in misvaluations are in line with those of realized portfolio CAPM alphas, consistent with the

Table 2.4: The Persistence of Misvaluation

The table demonstrates the persistence of misvaluation and its persistent effect in asset prices. The pooled panel regression based on annual data shows $\widehat{\alpha}_t^i$ has an AR(1) coefficient of 0.948 (standard errors 0.006, clustered by firm and year).
Panel (a) calculates the portfolio's annualized turnover, or monthly turnover multiplied by 12. Panel (b) calculates the CAPM-alphas, value and equal weighted, of portfolios sorted based on $\widehat{\alpha}_t^i$ at the end of June, starting from 1986-06 and ending in 2018-12. The CAPM-alphas are calculated by regressing the excess returns of the portfolios on market returns based on the universe of stocks that have estimated $\widehat{\alpha}_t^i$. The reason for using this universe is to take into account the negative CAPM-alphas of stocks with higher analyst coverage.

(a) Portfolio Turnover: Misvaluation Sorted Portfolios

Portfolio	**short-side**	**2**	**3**	**4**	**long-side**	**avg.long.short**
ann.turnover	28.56%	36.44%	31.80%	27.43%	19.28%	23.92%

(b) Holding Period Returns of Misvaluation Sorted Portfolios

	portfolio holding periods (in month)					
	12	24	36	48	60	72
Panel A: CAPM-alpha's of value-weighted portfolios						
low α^i	-0.612	-0.524	-0.524	-0.593	-0.461	-0.568
[t-stat]	[-4.212]	[-3.356]	[-3.323]	[-3.663]	[-3.399]	[-3.529]
high α^i	0.147	0.092	0.082	0.096	0.071	0.071
[t-stat]	[2.775]	[1.793]	[1.782]	[2.262]	[1.603]	[1.841]
High - Low	0.760	0.616	0.606	0.689	0.531	0.638
[t-stat]	[4.646]	[3.573]	[3.543]	[3.964]	[3.56]	[3.755]
Panel B: CAPM-alpha's of equal-weighted portfolios						
low α^i	-0.626	-0.588	-0.642	-0.639	-0.627	-0.632
[t-stat]	[-2.877]	[-2.719]	[-2.964]	[-2.943]	[-2.931]	[-2.91]
high α^i	0.384	0.368	0.352	0.355	0.343	0.351
[t-stat]	[3.376]	[3.215]	[3.263]	[3.293]	[3.278]	[3.395]
High - Low	0.984	0.929	0.954	0.969	0.930	0.943
[t-stat]	[5.827]	[5.47]	[5.569]	[5.638]	[5.773]	[5.524]

prediction of the CDR. For example, for the SP500 universe, the model predicts the CAPM alpha would amounts to about 5% per year, while the realized CAPM alphas is at 4.68% per year.

Table 2.5: Mean Excess Returns of Size and Misvaluation Sorted Portfolio (Value Weighted, 1986-06-01 to 2018-12-31)

This table shows the returns and characteristics for 3 by 3 portfolios independently sorted based on misvaluation measure $\widehat{\alpha}_t^i$ in (2.6) and market capitalization from June previous year. All returns, alphas and their standard errors are monthly and expressed in percentages. "1_1" denote the portfolio with lowest market capitalization from June in the previous year and lowest α_t^i , respectively, while "3_1" denotes portfolios with the highest market capitalizationand lowest α_t^i. Portfolios are value weighted each month. "SE" are standard errors which are shown in brackets. "mean ex.ret" are monthly returns over 3 month treasury rates; "SR" are monthly Sharpe Ratios. "FF-5 alpha" denote Fama-French 5 factor alphas. "num_stocks" are average number of stocks included in the portfolio over time. Post portfolio formation average characteristics: "nxt.12m.alpha" the average misvaluation measure 12 month after portfolio formation, "pi" the implied cost of capital, "mu" is the average beta times 0.064, "LTG" analyst's long-term growth estimates, "sd(LTG)" 36 month rolling volatility of LTG.

stats	1_1	1_2	1_3	high-low.small	2_1	2_2	2_3	high-low.mid	3_1	3_2	3_3	high-low.large
mean ex.ret	0.33	0.92	1.41	1.08	0.08	0.58	1.07	0.99	0.05	0.36	0.6	0.56
SE ex.ret	(6.86)	(6.26)	(7)	(2.48)	(6.48)	(5.54)	(5.81)	(2.09)	(5.48)	(4.35)	(4.57)	(2.68)
SR	0.05	0.15	0.2	0.43	0.01	0.1	0.18	0.48	0.01	0.08	0.13	0.21
CAPM beta	1.26	1.14	1.25	-0.01	1.28	1.1	1.17	-0.12	1.14	0.94	1.02	-0.12
SE CAPM beta	(0.05)	(0.04)	(0.05)	(0.03)	(0.04)	(0.03)	(0.03)	(0.02)	(0.03)	(0.02)	(0.01)	(0.03)
CAPM alpha	-0.44	0.22	0.64	1.08	-0.73	-0.1	0.36	1.07	-0.65	-0.22	-0.02	0.63
SE CAPM alpha	(0.21)	(0.19)	(0.22)	(0.13)	(0.16)	(0.14)	(0.14)	(0.1)	(0.11)	(0.07)	(0.04)	(0.13)
FF-5 alpha	-0.39	0.15	0.56	0.95	-0.74	-0.23	0.25	0.97	-0.39	-0.39	0	0.39
SE FF-5 alpha	(0.1)	(0.09)	(0.13)	(0.13)	(0.08)	(0.06)	(0.07)	(0.1)	(0.1)	(0.07)	(0.04)	(0.12)
num_stocks	429.1	208.07	134.39		225.19	281.74	256.84		116.49	274.87	372.7	
ME (million)	219.77	245.87	264.56		912.54	972.94	1024.18		26798.68	46560.25	79038.81	
nxt.12m.alpha	-0.19	-0.14	-0.13		-0.21	-0.14	-0.13		-0.2	-0.14	-0.12	
pi	0.17	0.12	0.1		0.18	0.12	0.1		0.17	0.12	0.09	
mu	0.06	0.06	0.07		0.07	0.06	0.07		0.07	0.06	0.07	
LTG	0.22	0.17	0.16		0.22	0.16	0.15		0.18	0.14	0.13	
sd(LTG)	0.06	0.05	0.05		0.06	0.04	0.04		0.04	0.03	0.03	

Table 2.6: Misvaluation (α^i) Sorted Portfolios and Realized Average Stock Returns for SP500 Firms (1986-06 to 2018-12)

The table presents the statistics related to portfolios sorted based on misvaluation measure created in (2.6) for firms in SP500 universe. All numbers are expressed in percentages unless otherwise stated. Returns and alphas are based on monthly frequency. Stocks are sorted into quantile portfolios based on the misvaluation measure $\widehat{\alpha}^i$ at the end of June each year, using the available information up to that point. Portfolios are rebalanced every month based on firms' market capitalization (value-weight). "Low" denote the portfolio with lowest $\widehat{\alpha}_t^i$. "High-Low" are excess returns of a portfolio that goes long on stocks with the highest $\widehat{\alpha}_t^i$ and short those with the lowest $\widehat{\alpha}_t^i$. "fwd_12m_alpha" are the average misvaluation measure 12 months after portfolio formation. "CAPM alpha" are calculated by regressing portfolio excess returns to returns to the universe of S&P500 stocks that have the estimates of $\widehat{\alpha}_t^i$ available.

stats	Low	2	3	4	High	High - Low
mean ex.ret	0.2	0.39	0.47	0.56	0.65	0.45
SE ex.ret	(4.78)	(4.53)	(4.26)	(4.34)	(5.13)	(2.97)
SR	0.04	0.09	0.11	0.13	0.13	0.15
CAPM beta	0.99	0.96	0.91	0.94	1.11	0.11
SE CAPM beta	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
CAPM alpha	-0.29	-0.08	0.02	0.1	0.1	0.39
SE CAPM alpha	(0.1)	(0.08)	(0.07)	(0.07)	(0.08)	(0.15)
FF-5 alpha	-0.42	-0.4	-0.24	-0.2	0.11	0.53
SE FF-5 alpha	(0.1)	(0.09)	(0.08)	(0.07)	(0.09)	(0.15)
ME	43818.98	42726.38	54376.9	60586.93	110437.88	
fwd_12m_alpha	-0.17	-0.14	-0.12	-0.12	-0.12	

Explaining Cross-sectional Anomalies Using Factor-Mimicking Portfolio

Second, a factor-mimicking portfolio constructed based on misvaluation measure should be able to explain completely the CAPM-alphas of portfolios that sorted on characteristics that predict future misvaluation. Under CDR, the only reason for these characteristics based anomalies to generate CAPM-alphas is investors use constant discount rate, which cause them to overvalue stocks associated with these characteristics.

Choosing Anomalies When choosing cross-sectional anomalies, I consider portfolios sorted on profitability, asset growth, market beta, (idiosyncratic) volatility, cash flow duration. Furthermore, I also consider the anomalies that were included to construct the two mispricing factors in Stambaugh and Yuan (2017). I chose these anomalies because they have generated significant interests both in the academic literature and in practice. The strong interests could come from both their significance for academic theories (for example market beta, volatility) and persistent, robust empirical performance[31]. I provide some more background behind choosing these anomalies below.

First, the beta (for example Fama and French (1992)) and (idiosyncratic) volatility (for example Haugen and Heins (1975) and Ang et al. (2006)) anomalies generated much interests mainly because it directly speaks to the failure of CAPM and the breaks the commonly believed positive risk-reward relationship in financial market. An extensive and continuing effort has been proposed to explain the low risk anomalies (for example Black (1992), Frazzini and Pedersen (2014) Schneider, Wagner, and Zechner (2020)).

Furthermore, I include the anomalies based on profitability (Novy-Marx (2013), Fama and French (2015), Hou, Xue, and Zhang (2015)) and asset growth (Cooper, Gulen, and Schill (2008), Fama and French (2015), Hou, Xue, and Zhang (2015)), which predicts future returns with positive and negative sign, respectively, because they are are shown by recent literature to be able to summarize to a large degree the average returns of the cross-section, as shown in Fama and French (2016) and Hou, Xue, and Zhang (2015). Various theories have been proposed to explain these anomalies, both behavioral and rational (for example Bouchaud et al. (2019) for behavioral and Hou, Xue, and Zhang (2015) for rational).

Finally, I chose the cash flow duration factor (Dechow, Sloan, and Soliman (2004b), Weber (2018) and Gonçalves (2019)), which negatively predicts future stock returns, because of it is related directly to the future cash growth and has theoretical significance due to the theories related to the term structure of equity (see for example Lettau and Wachter

[31] I do not include the well-known "value" and "size" anomalies here because for the sample I consider (post 1986-06), it does not have a significant CAPM-alpha.

(2007b) and Croce, Lettau, and Ludvigson (2014)), which is important for linking the macro-finance theories to explain the time-series of aggregate stock returns to the cross-section (for exampleSantos and Veronesi (2010) Binsbergen and Koijen (2015)).

In order to show that the CDR hypothesis is indeed an important channel through which mispricing occur, I also consider the two mispricing factors constructed by Stambaugh and Yuan (2017). The two factors are constructed based on 11 anomalies and are shown in their paper to have strong power to explain the factors constructed in the literature. I examine the explanation power of the misvaluation factor on the two composite factors as well as the 11 consisting anomalies underlying these two factors.

Constructing the Misvaluation Factor To explain the anomaly portfolio returns, I first construct a factor-mimicking portfolio of misvaluation. I follow a similar procedure as employed by Fama and French (2015). First, I conduct 3 by 3 independent sort based on market capitalization and $\widehat{\alpha}_t^i$. Within each of the size tercile, which are of small, medium and large-cap stocks, I subtract the returns of stocks with the highest $\widehat{\alpha}_t^i$ by the stocks with the lowest $\widehat{\alpha}_t^i$ to obtain the returns of long-short portfolio. More specifically, the Constant Discount Rate (CDR) factor is

$$\begin{aligned} CDR_t = & \frac{1}{3}(R_t^{high,small} + R_t^{high,mid} + R_t^{high,big}) \\ & - \frac{1}{3}(R_t^{low,small} + R_t^{low,mid} + R_t^{low,big}) \end{aligned} \tag{2.10}$$

I show the return statistics of the factor together with its cumulative returns in Table 2.7 and 2.5, respectively. The factor has a volatility of 6.3% annually, with a mean realized return of 10.8%. The realized return mainly comes from the short leg, which contains stocks most overvalued.

The cumulative return graph shows that the strong performance of the CDR factor is not concentrated in a specific period over the past 33 years, confirming the results in the previous section that the misvaluation effect is persistent.

Figure 2.5: Cumulative Returns of CDR Factor (In Log Scale)

Notes: Sample period is 1986-07-01 to 2018-12-31. Stocks are sorted independently into 3 by 3 tarciles based on market capitalization and $\widehat{\alpha_t^i}$ at the end of each June. The portfolios are rebalanced each month based on market capitalization. the CDR factor is the "high minus low" and constructed by

$$CDR_t = \frac{1}{3}R_t^{high} - \frac{1}{3}R_t^{low}$$

where $R_t^{high} = \frac{1}{3}(R_t^{high,small} + R_t^{high,mid} + R_t^{high,big} - 3R_t^f)$ and $R_t^{low} = \frac{1}{3}(R_t^{low,small} + R_t^{low,mid} + R_t^{low,big} - 3R_t^f)$.

Table 2.7: Return Statistics CDR Factor

Notes: Sample period is 1986-07-01 to 2018-12-31. Stocks are sorted independently into 3 by 3 tarciles based on market capitalization previous June and $\widehat{\alpha_t^i}$ at the end of each June. The portfolios are rebalanced each month based on market capitalization. the CDR factor is constructed by

$$CDR_t = \frac{1}{3}R_t^{high} - \frac{1}{3}R_t^{low}$$

where $R_t^{high} = \frac{1}{3}(R_t^{high,small} + R_t^{high,mid} + R_t^{high,big} - 3R_t^f)$ and $R_t^{low} = \frac{1}{3}(R_t^{low,small} + R_t^{low,mid} + R_t^{low,big} - 3R_t^f)$.

	CDR	low $\hat{\alpha}$	high $\hat{\alpha}$
Annulized Return	0.108	-0.004	0.110
Annualized Std. Dev.	0.063	0.208	0.190
Annualized Sharpe	1.704	-0.021	0.578

Explaining Five Prominent Anomalies The Constant Discount Rate Hypothesis predicts that the CAPM alphas of individual assets should be completely consumed by the CDR factor. To test this hypothesis, I construct long short anomaly portfolios based on the five characteristics and regress the return of the portfolios on the market excess return and CDR factor defined in (2.10):

$$R_t^i = \alpha^i + CDR_t\beta_{CDR}^i + (R_t^m - R^f)\beta_m^i + \epsilon_t^i \tag{2.11}$$

The constant discount rate hypothesis predicts that all the alphas are jointly zero or

$$H_0^{CDR}: \quad \alpha^i = 0 \quad \forall i = 1, ..N$$

I test the hypothesis using the GRS tests. I also examine the alphas of single anomaly portfolios.

To eliminate the errors due to replication, I download the anomaly portfolios from official sources. More specifically, I download portfolios sorted based on beta, variance and residual variance sorted portfolios directly from Ken French's website[32]and the cash flow duration

[32]Beta are measured using the last 5 years of monthly returns; variances are historical variance based on

Table 2.8: Anomalies Portfolio Alpha/Beta Before/After Controlling for CDR Factor

Sample period is 1986-07-01 to 2018-12-31. In Panel A, the GRS test statistics are presented, which test the null hypothesis that all α^is in Equation 2.11 are jointly zero under the CAPM or the model where market factor together with CDR factors are included. Panel B presents the tests for individual assets in Equation 2.11. Panel B1: the long short anomaly portfolios are regressed on market excess returns over 3 month treasuries. Panel B2: long short anomaly portfolios are regressed on (value-weighted) market excess returns and CDR factor defined in Equation (2.10). "beta" are measured using the last 5 years of monthly returns; "prof" are operating profitability defined in Fama and French (2015); "res.var" are measured using 60 days of daily returns and Fama-French 3 factor model; "asset.growth" are the change in total assets from the fiscal year ending in year t-2 to the fiscal year ending in t-1, divided by t-2 total assets at the end of each June using NYSE breakpoints; "cf.dur" are cash flow duration measure defined in Weber (2018), a composite measure based on sales and book values. Except for the "cf.dur", all other portfolios are downloaded from Ken French's website and are long-short (value-weghted) portfolios constructed by subtracting portfolio with the lowest decile of beta, var, res.var, asset growth by the highest decile and subtracting the highest profitability portfolio by the lowest profitability portfolio. Decile portfolios of "cf.dur" are downloaded on Michael Weber's website and the portfolio ends on 2014-06-30 and are equally weighted.

Panel A: GRS. Test Ressults

Model	**CAPM**	**Mkt + CDR**
GRS-stat	5.422	1.003
P-value	0.000	0.416

Panel B: Tests on Single Anomaly Portfolios

	Predicting Volatility of Growth			**Predicting Future Growth**	
	beta	**res.var**	**prof**	**asset.growth**	**cf.dur**
Panel B1: CAPM alpha of anomaly portfolios					
CAPM Alpha (%)	0.565	1.246	0.721	0.488	1.261
t-statistics	[2.288]	[3.777]	[3.593]	[2.909]	[4.124]
CAPM Beta	-1.046	-0.971	-0.456	-0.177	-0.432
t-statistics	[-18.8]	[-13.063]	[-10.077]	[-4.688]	[-6.471]
Panel B2: CAPM alpha of anomaly portfolios after controlling for CDR factor					
CAPM Alpha (%)	-0.129	-0.114	0.174	0.085	0.296
t-statistics	[-0.484]	[-0.337]	[0.799]	[0.466]	[0.926]
CAPM Beta	-0.983	-0.849	-0.406	-0.141	-0.345
t-statistics	[-18.013]	[-12.243]	[-9.137]	[-3.759]	[-5.393]
Loading on CER	0.748	1.467	0.590	0.434	1.036
t-statistics	[5.706]	[8.807]	[5.522]	[4.815]	[6.804]

sorted portfolios from Michael Weber's website[33]

The GRS test results in Panel A of Table 2.8 show we can not reject the hypothesis: the CDR factor explains the CAPM-alphas of all five anomaly portfolios. More specifically, the GRS test statistics based on the CDR factor is just above 1, which has a p-value of 0.42, compared to the GRS test statistics of 5.4 under CAPM, which confirms these portfolios have high CAPM-alphas.

Examining the tests on each of the five single anomalies, Table 2.8 shows that all of the standalone portfolio's CAPM-alphas become statistically insignificant from zero after the inclusion of CDR factor. Furthermore, all of the anomalies load strongly on the CDR factor, with point estimates on loadings equal to 0.43 at the least (asset growth).

The magnitude of the reduction is large, especially for the idiosyncratic variance and cash flow duration factor, which amount to 1.36% and 0.96% per month based on the point estimates, as shown in Panel B of Table 2.8. This large reduction in CAPM-alphas is confirmed by these anomaly portfolio's large loadings on the CDR factor, which are 1.47 and 1.04 for the residual variance and cash flow duration factor, respectively. The strong explanation power of the CDR factor on these two particular anomalies makes intuitive sense since residual variances closely mimics the cash flow growth volatility (σ_c^i) in the model while cash flow duration aims to predict the future cash flow growth (g^i).

Misvaluation Factor and Mispricing Factors in Stambaugh and Yuan (2017) I show that the explanation power of the mispricing channel suggested by the CDR hypothesis goes beyond the five anomalies analyzed before. Stambaugh and Yuan (2017) constructs two mispricing factors based on 11 cross-sectional anomalies. They show that these 2 factors have superior performance compared to factor models constructed by Fama and French (2015) and Hou, Xue, and Zhang (2015) in summarizing the cross-section of average stock returns. I therefore examine to what extent the single misvaluation factor can explain the CAPM-

the past 60 days of daily returns and residual variances are measured using 60 days of daily returns and Fama-French 3 factor model.

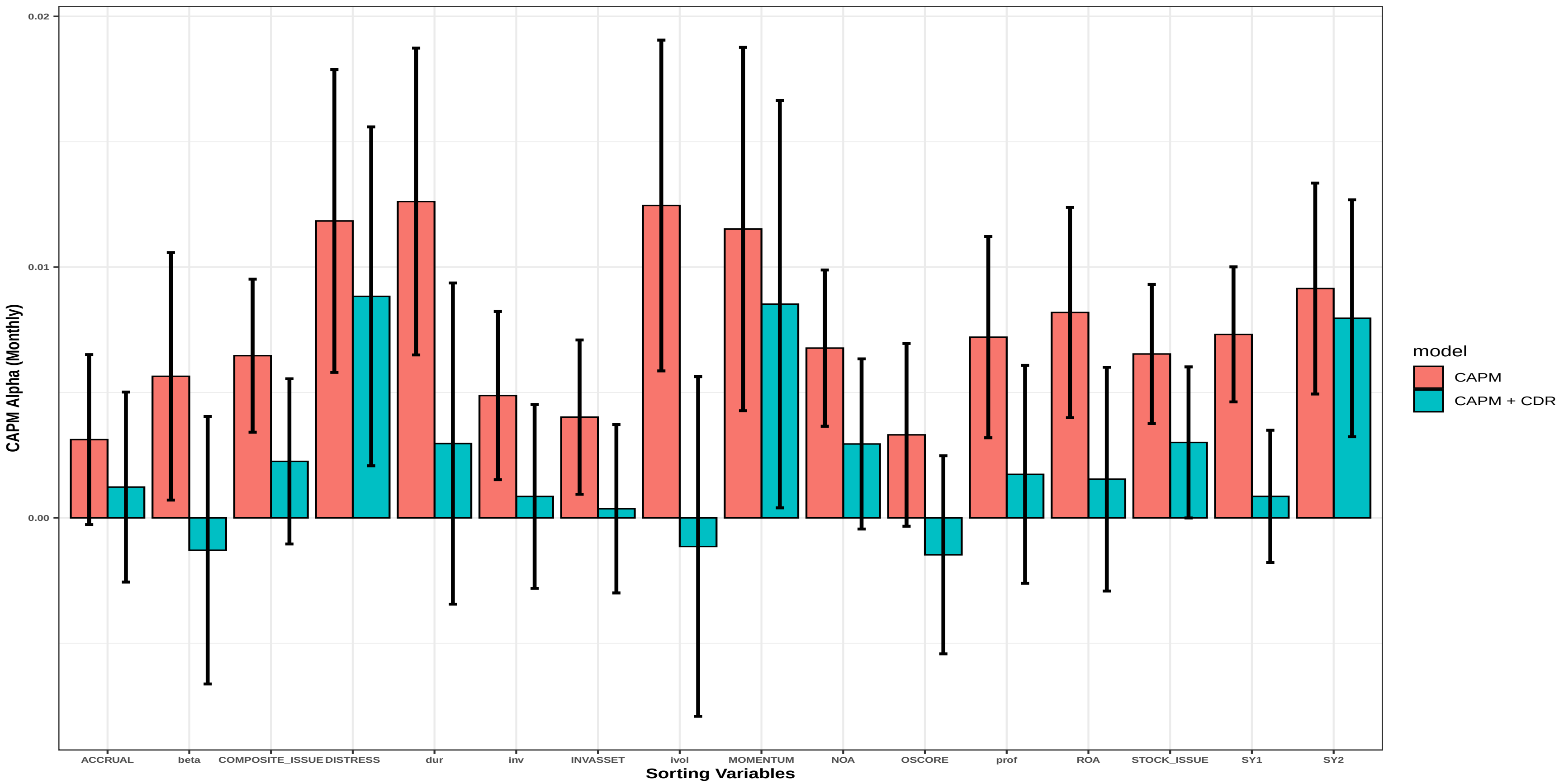
[33] The details of the measure is described in Weber (2018).

alphas of the misvaluation factor as well as the underlying 11 anomalies, 9 of which have not been included in the 5 anomalies examined earlier.

Figure 2.6 shows the CAPM-alpha together with its the standard errors of the estimates before and after including a CDR factor, for the total of 14 anomalies considered (11 of them form the basis for the 2 mispricing factors in Stambaugh and Yuan (2017)) together with the two mispricing factors. For all the 14 standalone anomalies as well as the two mispricing factors ("SY1" and "SY2"), the misvaluation factor constructed in this paper reduces their CAPM-alphas. In fact, for all but the momentum and distress factor, the CAPM-alphas become insignificant after regressing on the misvaluation factor. As a result, the CAPM-alphas of the first mispricing factor (SY1) is completely explained by the misvaluation factor. The second mispricing factor still remains unexplained by the CDR, mainly due to the momentum and distress factor. This makes sense as the misvaluation factor is a persistent, long-term factor, while the momentum and distress has been shown to have high turnover and mainly generate anomalous returns in the short term.

Figure 2.6: CAPM Alpha of Long-Short Anomaly Portfolios Before/After Controlling for CDR Factor

The figure plots the CAPM-alphas of the 2 mispricing factors constructed in Stambaugh and Yuan (2017) together 11 anomalies that are used to construct the factors, together with the duration, beta, and residual variance anomalies, before and after regressing on the CDR factor.
Sample period is 1986-07-01 to 2016-12-31. “CAPM” is the intercept when regressing long short anomaly portfolios on market excess returns over 3 month treasuries. “CAPM + CDR” is the intercept when regressing long short anomaly portfolios on (value-weighted) market excess returns and CDR factor defined in Equation (2.10). Two standard deviations above and below the estimates are indicated.
Long-short anomaly portfolio returns whose labels are in capital letters are downloaded from Robert Stambaugh’s webiste. ”beta”, “inv”, “ivol”, “prof” are downloaded from Ken French’s website and “dur” is downloaded from Michael Weber’s website. “ACCURAL” is the accrual anomaly of Sloan (1996); “beta” are measured using the last 5 years of monthly returns; “prof” are operating profitability defined in Fama and French (2015); “ivol” are measured using 60 days of daily returns and Fama-French 3 factor model; “inv” are the change in total assets (asset growth) as in Fama and French (2015) and Cooper, Gulen, and Schill (2008); “cf.dur” are cash flow duration measure defined in Weber (2018), a composite measure based on sales and book values. “COMPOSITE_ISSUE”: composite equity issuance of Daniel and Titman (2006); “STOCK_ISSUE”: the equity issuance measure of Loughran and Ritter (1995); “DISTRESS” : the distress risk measures of Campbell, Hilscher, and Szilagyi (2008); “OSCORE”: Ohlson’s O-score Ohlson (1980); “NOA” Net Operating Asset defined in Hirshleifer et al. (2004); “MOMENTUM”: momentum variable defined in Jegadeesh and Titman (1993); “INVASSET”: investment to assets defined in Titman, Wei, and Xie (2013); “SY1” the “MGMT” factor constructed in Stambaugh and Yuan (2017), which include Net stock issues, composite equity issues, accruals, net operating assets, asset growth, investment to assets; “SY2” the “PERF” factor i n Stambaugh and Yuan (2017), which includes distress, O-score, Momentum, profitability, return on assets.



Firm Characteristics and Misvaluation

Finally, the CDR also implies that characteristics that predict future anomalous returns (CAPM-alphas) should also predict future misvaluation (α_t^i), with the same sign. Since the cross-sectional variation in misvaluation is mostly driven by λ^i, these characteristics should forecast either expected future fundamental growth (g_t^i) or fundamental volatility ($\sigma_{c,t}^i$) or both, with an opposite sign as they predict future CAPM-alphas.

To test the first hypothesis, I run the following predictive panel regressions with date fixed effects:

$$y_t^i = a + B' X_{t-1}^i + f_t + \epsilon_{i,t}$$

where X_{t-1}^i are a vector of characteristics include market beta, volatility, idiosyncratic volatility, profitability, asset growth, cash flow duration and f_t are date fixed-effects. y_t^i are either the firm's future misvluation $\widehat{\alpha}_t^i$, analyst's long-term growth estimates g_t^i or the future volatility of analysts' long-term growth estimates, $\sigma_{c,t}^i$.

When predicting future misvaluation, the CDR hypothesis predicts that the predictive coefficients are all negative and significant, except for profitability, which should be positive and significant. This is because except for profitability, all the other characteristics positive predict firms' future CAPM-alphas.

Panel (a) and first column of Panel (b) of Table 2.9 confirm the prediction of the CDR hypothesis. All of the characteristics show up to have significant predictive power for future misvaluation, and the coefficients have signs correspond exactly to their CAPM-alphas. Quantitatively, the R-squared is high, at more than 21%. Comparing across different characteristics, beta and residual variances show up to have the strongest predictive power, followed by asset growth, profitability and cash flow duration.

I use expected growth and growth volatility as dependent variables in the regression and present the results in the Panel (b) of Table 2.9. These results show that the five characteristics generally can predict both the growth level and volatility, even though these

Table 2.9: Misvaluation and Firm Characteristics

Data are quarterly firm-level data from 1985-Q1 to 2018-Q4. Misvaluation captures the under- over- valuation following CDR, as defined in Equation (2.6). In Panel (a), Spearman rank correlations are calculated. In Panel (b), results from panel regression with date fixed effects

$$y_t^i = a + B' X_{t-1}^i + f_t + \epsilon_{i,t}$$

are presented, with standard errors clustered at firm-quarter level. Both dependent and independent variables are transformed into cross-sectional percentiles to avoid outliers and for the ease of interpretation. Expected growth and growth volatility are defined as analyst long-term growth expectation and 36-month rolling volatility of long-term growth expectation, respectively, both downloaded from IBES data base. "beta" are from Welch (2019) downloaded from Ivo Welch's website; "residual variances" are constructed using 60 days of daily returns (with minimum of 20 days) and Fama-French 3 factor model; "asset.growth" are the change in total assets from the fiscal year ending in year t-2 to the fiscal year ending in t-1, divided by t-2 total assets at the end of each June using NYSE breakpoints; "cf.dur" are cash flow duration measure defined in Gonçalves (2019), downloaded from Andrei Goncalves' webiste; "Op.Prof" are firms' operating profitability defined in Fama and French (2015). Both financials and utilities are excluded from the panel regressions and each firm needs to have a minimum of two years available in COMPUSTAT.

(a) Pair-wise Rank Correlation Misvaluation and Firm Characteristics

lag.asset.growth	lag.op.prof.	lag.res.var	lag.beta	lag.cf.dur
-0.155	0.116	-0.288	-0.401	-0.102

(b) Panel Regressions: Future misvaluation, expected growth, growth volatility on firm characteristics

	Dependent variable:		
	misvaluation	*expected.growth*	*growth.vol*
	(1)	(2)	(3)
lag.beta	-0.248^{***}	0.077^{***}	0.121^{***}
	(0.013)	(0.013)	(0.015)
lag.asset.growth	-0.071^{***}	0.079^{***}	0.062^{***}
	(0.007)	(0.008)	(0.009)
lag.op.prof.	0.042^{***}	-0.037^{***}	-0.160^{***}
	(0.009)	(0.010)	(0.012)
lag.res.var	-0.273^{***}	0.268^{***}	0.236^{***}
	(0.010)	(0.010)	(0.014)
lag.cf.dur	-0.046^{***}	0.040^{***}	0.051^{***}
	(0.010)	(0.010)	(0.011)
Observations	108,790	108,790	77,466
R^2	0.218	0.119	0.146
Adjusted R^2	0.216	0.117	0.145
Residual Std. Error	0.245 (df = 108615)	0.260 (df = 108615)	0.262 (df = 77312)
Note:			*p<0.1; **p<0.05; ***p<0.01

two variables are moderately correlated (24% correlation in the pooled sample). One notable characteristics is profitability, which predicts future misvaluation mainly due to its ability to forecast negatively a firm's future growth volatility. For a firm's operating profitability to increase by 1 percent in ranking in the cross-section, its future growth volatility decreases by 16% while expected growth only decrease by 3.7%. This result is intuitive: profitable firms typically have stable cash flows and are unlikely to incur high cash flow volatility in the future. Similar pattern holds for low beta firms.

2.3 Conclusion

This paper proposes and tests a unifying hypothesis to explain cross-sectional asset pricing anomalies: some investors falsely ignore the dynamics of discount rates when forming return expectations. The empirical findings in this paper show the potential impact of the mispricing due to the constant discount rate assumption is economically significant and many prominent asset pricing anomalies can be explained. Besides, data on analysts' return forecasts and firms' fundamentals are consistent with the predictions of the CDR hypothesis too. The results are also consistent with the aggregate time-series estimates provided by Renxuan (2020b), which shows that a large set of investors underestimate the importance of discount rate in driving the dynamic of asset prices on the market levell.

The results presented in this paper also have implications for the investment community. In particular, these results provide useful suggestions to those who employ conventional Discounted Cash Flow (DCF) models to value stocks: they could improve the accuracy of their expected returns by adjusting their estimates using the misvaluation measure developed in this paper.

The paper assumes that biases subjective return expectation would directly translate into over/under investments by the CDR investors through their own portfolio optimization model. This assumption is not warranted. A natural next step is to examine closely how subjective expectations are translated into changes in investors' investment decisions, extending the methods and data considered in Koijen and Yogo (2019).

Another potential venue of research is to connect the CDR expectation to other subjective expectation formation process proposed in the literature, such as the ones proposed in Bordalo et al. (2019) or Bouchaud et al. (2019), both of which concern expectations for fundamentals. Are these expectation formation processes consistent with each other or mutually exclusive?

Finally, further examination on the impact of misvaluation on real economy is also promising. Dessaint et al. (2021) find evidence supporting the idea that investors' using CAPM distort the prices in M&A markets. If the channel of misvaluation suggested in this paper is valid and long lasting, those firms who receive much higher valuation than it should due to CDR, should have a lower cost of equity capital, which could ultimately impact its real activities. A logical next step is to estimate the model proposed in Binsbergen and Opp (2019) to evaluate the loss of efficiency due to the misvaluation.

References

[1] Klaus Adam, Albert Marcet, and Johannes Beutel. "Stock Price Booms and Expected Capital Gains". In: *The American economic review* 107.8 (2017), pp. 2352–2408.

[2] Klaus Adam, Dmitry Matveev, and Stefan Nagel. "Do survey expectations of stock returns reflect risk adjustments?" In: *Journal of Monetary Economics* 117 (2021), pp. 723–740.

[3] Klaus Adam, Dmitry Matveev, and Stefan Nagel. "Do survey expectations of stock returns reflect risk adjustments?" In: *Journal of Monetary Economics* 117 (2021), pp. 723–740.

[4] Aleksandar Andonov and Joshua D Rauh. "The return expectations of institutional investors". In: (2020).

[5] Andrew Ang and Geert Bekaert. "Stock Return Predictability: Is it There?" In: *The review of financial studies* 20.3 (May 2007), pp. 651–707.

[6] Andrew Ang and Jun Liu. "How to discount cashflows with time-varying expected returns". In: *The Journal of Finance* 59.6 (2004), pp. 2745–2783.

[7] Andrew Ang et al. "The cross-section of volatility and expected returns". In: *The Journal of Finance* 61.1 (2006), pp. 259–299.

[8] Efstathios Avdis and Jessica A Wachter. "Maximum likelihood estimation of the equity premium". In: *Journal of financial economics* 125.3 (Sept. 2017), pp. 589–609.

[9] Ray Ball, Gil Sadka, and Ronnie Sadka. "Aggregate earnings and asset prices". In: *Journal of Accounting Research* 47.5 (2009), pp. 1097–1133.

[10] Ravi Bansal, Dana Kiku, and Amir Yaron. "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices". In: *Critical Finance Review* 1.1 (2012), pp. 183–221.

[11] Ravi Bansal and Amir Yaron. "Risks for the long run: A potential resolution of asset pricing puzzles". In: *The journal of Finance* 59.4 (2004), pp. 1481–1509.

[12] Brad M Barber, Xing Huang, and Terrance Odean. "Which Factors Matter to Investors? Evidence from Mutual Fund Flows". In: *The review of financial studies* 29.10 (Oct. 2016), pp. 2600–2642.

[13] Nicholas Barberis, Andrei Shleifer, and Robert Vishny. “A model of investor sentiment”. In: *Journal of financial economics* 49.3 (Sept. 1998), pp. 307–343.

[14] Nicholas Barberis et al. “X-CAPM: An extrapolative capital asset pricing model”. In: *Journal of financial economics* 115.1 (Jan. 2015), pp. 1–24.

[15] Nicholas Barberis et al. “X-CAPM: An extrapolative capital asset pricing model”. In: *Journal of financial economics* 115.1 (Jan. 2015), pp. 1–24.

[16] Nicholas C Barberis, Lawrence J Jin, and Baolian Wang. *Prospect theory and stock market anomalies.* Tech. rep. National Bureau of Economic Research, 2020.

[17] Jason Beeler. “The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment”. In: *Critical Finance Review* 1.1 (2012), pp. 141–182.

[18] Itzhak Ben-David, John R Graham, and Campbell R Harvey. “Managerial miscalibration”. In: *The Quarterly Journal of Economics* 128.4 (2013), pp. 1547–1584.

[19] Jonathan B Berk and Jules H Van Binsbergen. “Assessing asset pricing models using revealed preference”. In: *Journal of Financial Economics* 119.1 (2016), pp. 1–23.

[20] Jules H van Binsbergen and Christian C Opp. “Real Anomalies”. In: *The Journal of finance* 74.4 (Aug. 2019), pp. 1659–1706.

[21] Jules van Binsbergen and Ralph Koijen. “The term structure of returns: Facts and theory”. In: *Annual Review of Financial Economics forthcoming* (2015).

[22] Fischer Black. “Beta and return”. In: *Journal of portfolio management* 1 (1992).

[23] Pedro Bordalo, Nicola Gennaioli, and Andrei Shleifer. “Diagnostic Expectations and Credit Cycles”. In: *The Journal of finance* 73.1 (Feb. 2018), pp. 199–227.

[24] Pedro Bordalo et al. “Diagnostic Expectations and Stock Returns”. In: *The Journal of finance* 74.6 (Dec. 2019), pp. 2839–2874.

[25] Pedro Bordalo et al. *Expectations of fundamentals and stock market puzzles.* Tech. rep. National Bureau of Economic Research, 2020.

[26] Jean-philippe Bouchaud et al. “Sticky expectations and the profitability anomaly”. In: *The Journal of Finance* 74.2 (2019), pp. 639–674.

[27] William A Branch and George W Evans. “Asset Return Dynamics and Learning”. In: *The review of financial studies* 23.4 (Jan. 2010), pp. 1651–1680.

[28] Alon Brav and Reuven Lehavy. "An empirical analysis of analysts' target prices: Short-term informativeness and long-term dynamics". In: *The Journal of Finance* 58.5 (2003), pp. 1933–1967.

[29] Alon Brav and Reuven Lehavy. "An empirical analysis of analysts' target prices: Short-term informativeness and long-term dynamics". In: *The Journal of Finance* 58.5 (2003), pp. 1933–1967.

[30] Markus Brunnermeier et al. "Review Article: Perspectives on the Future of Asset Pricing". In: *The review of financial studies* 34.4 (Mar. 2021), pp. 2126–2160.

[31] John Y Campbell. *A variance decomposition for stock returns.* Tech. rep. National Bureau of Economic Research, 1990.

[32] John Y Campbell. "A Variance Decomposition for Stock Returns". In: *The Economic Journal* 101.405 (Mar. 1991), pp. 157–179.

[33] John Y Campbell and John H Cochrane. *By force of habit: A consumption-based explanation of aggregate stock market behavior.* Tech. rep. National Bureau of Economic Research, 1995.

[34] John Y Campbell and John H Cochrane. "By force of habit: A consumption-based explanation of aggregate stock market behavior". In: *Journal of political Economy* 107.2 (1999), pp. 205–251.

[35] John Y Campbell, Jens Hilscher, and Jan Szilagyi. "In Search of Distress Risk". In: *The Journal of finance* 63.6 (Dec. 2008), pp. 2899–2939.

[36] Long Chen, Zhi Da, and Xinlei Zhao. "What drives stock price movements?" In: *The Review of Financial Studies* 26.4 (2013), pp. 841–876.

[37] James J Choi and Thomas M Mertens. "Extrapolative expectations and the equity premium". In: *Available at SSRN 3462056* (2019).

[38] James Claus and Jacob Thomas. "Equity premia as low as three percent? Evidence from analysts' earnings forecasts for domestic and international stock markets". In: *The Journal of Finance* 56.5 (2001), pp. 1629–1666.

[39] John H Cochrane. "State-space vs. VAR models for stock returns". In: *Unpublished paper, Chicago GSB* (2008).

[40] John H Cochrane. "Presidential Address: Discount Rates". In: *The Journal of finance* 66.4 (Aug. 2011), pp. 1047–1108.

[41] John H Cochrane. “Presidential address: Discount rates”. In: *The Journal of Finance* 66.4 (2011), pp. 1047–1108.

[42] Randolph B Cohen, Christopher Polk, and Tuomo Vuolteenaho. “The price is (almost) right”. In: *The Journal of Finance* 64.6 (2009), pp. 2739–2782.

[43] Pierre Collin-Dufresne, Michael Johannes, and Lars A Lochstoer. “Parameter Learning in General Equilibrium: The Asset Pricing Implications”. In: *The American economic review* 106.3 (Mar. 2016), pp. 664–698.

[44] Pierre Collin-Dufresne, Michael Johannes, and Lars A Lochstoer. “Asset Pricing When ‘This Time Is Different’”. In: *The review of financial studies* 30.2 (Feb. 2017), pp. 505–535.

[45] Michael J Cooper, Huseyin Gulen, and Michael J Schill. “Asset growth and the cross-section of stock returns”. In: *The Journal of finance* 63.4 (2008), pp. 1609–1651.

[46] Mariano M Croce, Martin Lettau, and Sydney C Ludvigson. “Investor information, long-run risk, and the term structure of equity”. In: *Review of Financial Studies* (2014), hhu084.

[47] Zhi Da, Keejae P Hong, and Sangwoo Lee. “What Drives Target Price Forecasts and Their Investment Value?” In: *Journal of Business Finance and Accounting* 43.3-4 (Mar. 2016), pp. 487–510.

[48] Zhi Da and Ernst Schaumburg. “Relative valuation and analyst target price forecasts”. In: *Journal of Financial Markets* 14.1 (Feb. 2011), pp. 161–192.

[49] Magnus Dahlquist and Markus Ibert. “How Cyclical Are Stock Market Return Expectations? Evidence from Capital Market Assumptions”. Availalbe at SSRN: `http://dx.doi.org/10.2139/ssrn.3763796`. 2021.

[50] Aswath Damodaran. *Investment valuation: Tools and techniques for determining the value of any asset*. Vol. 666. John Wiley & Sons, 2012.

[51] Kent Daniel, David Hirshleifer, and Avanidhar Subrahmanyam. “Investor Psychology and Security Market Under- and Overreactions”. In: *The Journal of finance* 53.6 (Dec. 1998), pp. 1839–1885.

[52] Kent Daniel, David Hirshleifer, and Lin Sun. “Short- and Long-Horizon Behavioral Factors”. In: *The review of financial studies* (June 2019).

[53] Kent Daniel and Sheridan Titman. “Market reactions to tangible and intangible information”. In: *The Journal of Finance* 61.4 (2006), pp. 1605–1643.

[54] Patricia M Dechow, Richard G Sloan, and Mark T Soliman. “Implied equity duration: A new measure of equity risk”. In: *Review of Accounting Studies* 9.2-3 (2004), pp. 197–228.

[55] Patricia M Dechow, Richard G Sloan, and Mark T Soliman. “Implied Equity Duration: A New Measure of Equity Risk”. In: *Review of Accounting Studies* 9.2 (June 2004), pp. 197–228.

[56] Olivier Dessaint et al. “CAPM-based company (mis) valuations”. In: *The Review of Financial Studies* 34.1 (2021), pp. 1–66.

[57] Karl B Diether, Christopher J Malloy, and Anna Scherbina. “Differences of Opinion and the Cross Section of Stock Returns”. In: *The Journal of finance* 57.5 (2002), pp. 2113–2141.

[58] James Durbin and Siem Jan Koopman. *Time series analysis by state space methods*. Oxford university press, 2012.

[59] Joseph Engelberg, R David McLean, and Jeffrey Pontiff. “Analysts and Anomalies”. Aug. 2019.

[60] Joseph Engelberg, R David McLean, and Jeffrey Pontiff. “Analysts and anomalies”. In: *Journal of Accounting and Economics* (2019), p. 101249.

[61] George W Evans and Seppo Honkapohja. *Learning and expectations in macroeconomics*. Princeton University Press, 2012.

[62] Eugene F Fama and Kenneth R French. “The Cross-Section of Expected Stock Returns”. In: *The Journal of finance* 47.2 (June 1992), pp. 427–465.

[63] Eugene F Fama and Kenneth R French. “Profitability, investment and average returns”. In: *Journal of financial economics* 82.3 (Dec. 2006), pp. 491–518.

[64] Eugene F Fama and Kenneth R French. “A five-factor asset pricing model”. In: *Journal of financial economics* 116.1 (Apr. 2015), pp. 1–22.

[65] Eugene F Fama and Kenneth R French. “Dissecting anomalies with a five-factor model”. In: *The Review of Financial Studies* 29.1 (2016), pp. 69–103.

[66] Andrea Frazzini and Lasse Heje Pedersen. “Betting against beta”. In: *Journal of financial economics* 111.1 (Jan. 2014), pp. 1–25.

[67] William R Gebhardt, Charles M C Lee, and Bhaskaran Swaminathan. “Toward an Implied Cost of Capital”. In: *Journal of Accounting Research* 39.1 (June 2001), pp. 135–176.

[68] Gerd Gigerenzer and Wolfgang Gaissmaier. “Heuristic decision making”. In: *Annual review of psychology* 62 (2011), pp. 451–482.

[69] Dan Gode and Partha Mohanram. “Inferring the Cost of Capital Using the Ohlson–Juettner Model”. In: *Review of Accounting Studies* 8.4 (Dec. 2003), pp. 399–431.

[70] Andrei Gonçalves. “The short duration premium”. In: *Available at SSRN 3385579* (2019).

[71] Niels Joachim Gormsen and Eben Lazarus. “Duration-driven returns”. In: *Available at SSRN 3359027* (2019).

[72] R Greenwood and A Shleifer. “Expectations of returns and expected returns”. In: *The review of financial studies* (2014).

[73] Robert A Haugen and A James Heins. “Risk and the rate of return on financial assets: Some old wine in new bottles”. In: *Journal of Financial and Quantitative Analysis* (1975), pp. 775–784.

[74] Bernard Herskovic et al. “The common factor in idiosyncratic volatility: Quantitative asset pricing implications”. In: *Journal of financial economics* 119.2 (Feb. 2016), pp. 249–283.

[75] David Hirshleifer, Jun Li, and Jianfeng Yu. “Asset pricing in production economies with extrapolative expectations”. In: *Journal of monetary economics* 76 (Nov. 2015), pp. 87–106.

[76] David Hirshleifer et al. “Do investors overvalue firms with bloated balance sheets?” In: *Journal of Accounting and Economics* 38 (2004), pp. 297–331.

[77] Harrison Hong and Jeffrey D Kubik. “Analyzing the analysts: Career concerns and biased earnings forecasts”. In: *The Journal of Finance* 58.1 (2003), pp. 313–351.

[78] Harrison Hong, Terence Lim, and Jeremy C Stein. “Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies”. In: *The Journal of finance* 55.1 (Feb. 2000), pp. 265–295.

[79] Harrison Hong and Jeremy C Stein. “A unified theory of underreaction, momentum trading, and overreaction in asset markets”. In: *The Journal of finance* 54.6 (1999), pp. 2143–2184.

[80] Kewei Hou, Mathijs A van Dijk, and Yinglei Zhang. “The implied cost of capital: A new approach”. In: *Journal of Accounting and Economics* 53.3 (June 2012), pp. 504–526.

[81] Kewei Hou, Chen Xue, and Lu Zhang. "Digesting anomalies: An investment approach". In: *The Review of Financial Studies* 28.3 (2015), pp. 650–705.

[82] John Hughes, Jing Liu, and Jun Liu. "On the relation between expected returns and implied cost of capital". In: *Review of Accounting Studies* 14.2 (Sept. 2009), pp. 246–259.

[83] Ravi Jagannathan and Zhenyu Wang. "The conditional CAPM and the cross-section of expected returns". In: *The Journal of finance* 51.1 (1996), pp. 3–53.

[84] Narasimhan Jegadeesh and Chandra Sekhar Mangipudi. "What Do Fund Flows Reveal about Asset Pricing Models and Investor Sophistication?" In: *The review of financial studies* (Apr. 2020).

[85] Narasimhan Jegadeesh and Sheridan Titman. "Returns to buying winners and selling losers: Implications for stock market efficiency". In: *The Journal of finance* 48.1 (1993), pp. 65–91.

[86] Ralph SJ Koijen and Stijn Van Nieuwerburgh. "Predictability of returns and cash flows". In: *Annu. Rev. Financ. Econ.* 3.1 (2011), pp. 467–491.

[87] Ralph SJ Koijen and Motohiro Yogo. "A demand system approach to asset pricing". In: *Journal of Political Economy* 127.4 (2019), pp. 1475–1515.

[88] Tim Koller, Marc Goedhart, David Wessels, et al. *Valuation: measuring and managing the value of companies*. Vol. 499. john Wiley and sons, 2010.

[89] Serhiy Kozak, Stefan Nagel, and Shrihari Santosh. "Interpreting factor models". In: *The Journal of Finance* 73.3 (2018), pp. 1183–1223.

[90] Martin Lettau and Sydney C Ludvigson. "Understanding trend and cycle in asset values: Reevaluating the wealth effect on consumption". In: *american economic review* 94.1 (2004), pp. 276–299.

[91] Martin Lettau and Jessica A Wachter. "Why is long-horizon equity less risky? a duration-based explanation of the value premium". In: *The Journal of Finance* 62.1 (2007), pp. 55–92.

[92] Martin Lettau and Jessica A Wachter. "Why Is Long-Horizon Equity Less Risky? A Duration-Based Explanation of the Value Premium". In: *The Journal of finance* 62.1 (Feb. 2007), pp. 55–92.

[93] Lars A Lochstoer and Paul C Tetlock. "What drives anomaly returns?" In: *The Journal of Finance* 75.3 (2020), pp. 1417–1455.

[94] Tim Loughran and Jay R Ritter. "The new issues puzzle". In: *The Journal of finance* 50.1 (1995), pp. 23–51.

[95] Ulrike Malmendier and Stefan Nagel. "Depression babies: do macroeconomic experiences affect risk taking?" In: *The Quarterly Journal of Economics* 126.1 (2011), pp. 373–416.

[96] Ulrike Malmendier and Stefan Nagel. "Learning from inflation experiences". In: *The Quarterly Journal of Economics* 131.1 (2016), pp. 53–87.

[97] N Gregory Mankiw and Ricardo Reis. "Sticky information versus sticky prices: a proposal to replace the New Keynesian Phillips curve". In: *The Quarterly Journal of Economics* 117.4 (2002), pp. 1295–1328.

[98] Charles F Manski. "Survey Measurement of Probabilistic Macroeconomic Expectations: Progress and Promise". In: *NBER Macroeconomics Annual* 32 (Apr. 2018), pp. 411–471.

[99] Lilia Mukhlynina and Kjell G Nyborg. "The choice of valuation techniques in practice: education versus profession". In: (2016).

[100] Stefan Nagel and Zhengyang Xu. *Asset pricing with fading memory.* Tech. rep. National Bureau of Economic Research, 2019.

[101] Stefan Nagel and Zhengyang Xu. *Asset pricing with fading memory.* Tech. rep. National Bureau of Economic Research, 2019.

[102] Robert Novy-Marx. "The other side of value: The gross profitability premium". In: *Journal of financial economics* 108.1 (Apr. 2013), pp. 1–28.

[103] Robert Novy-Marx and Mihail Velikov. "A Taxonomy of Anomalies and Their Trading Costs". In: *The review of financial studies* 29.1 (Nov. 2015), pp. 104–147.

[104] Ricardo De la O and Sean Myers. "Subjective cash flow and discount rate expectations". In: (2020).

[105] James A Ohlson. "Financial ratios and the probabilistic prediction of bankruptcy". In: *Journal of accounting research* (1980), pp. 109–131.

[106] Ľuboš Pástor, Meenakshi Sinha, and Bhaskaran Swaminathan. "Estimating the Intertemporal Risk-Return Tradeoff Using the Implied Cost of Capital". In: *The Journal of finance* 63.6 (Dec. 2008), pp. 2859–2897.

[107] Ľuboš Pástor and Robert F Stambaugh. "Predictive Systems: Living with Imperfect Predictors". In: *The Journal of finance* 64.4 (2009), pp. 1583–1628.

[108] Ľuboš Pástor and Pietro Veronesi. “Was there a Nasdaq bubble in the late 1990s?” In: *Journal of financial economics* 81.1 (July 2006), pp. 61–100.

[109] Walter Pohl, Karl Schmedders, and Ole Wilms. “Asset pricing with heterogeneous agents and long-run risk”. In: *Journal of Financial Economics* (2021).

[110] Wang Renxuan. “Asset Prices When Investors Ignore Discount Rate Dynamics”. Available at SSRN: `https://ssrn.com/abstract=3740790` or `http://dx.doi.org/10.2139/ssrn.3740790`. Sept. 2020.

[111] Wang Renxuan. “Subjective Return Expectations”. Available at SSRN: `https://ssrn.com/abstract=3740792`. May 2020.

[112] Oleg Rytchkov. “Filtering out expected dividends and expected returns”. In: *The Quarterly Journal of Finance* 2.03 (2012), p. 1250012.

[113] Tano Santos and Pietro Veronesi. “Habit formation, the cross section of stock returns and the cash-flow risk puzzle”. In: *Journal of financial economics* 98.2 (Nov. 2010), pp. 385–413.

[114] Paul Schneider, Christian Wagner, and Josef Zechner. “Low-Risk Anomalies?” In: *The Journal of finance* 75.5 (Oct. 2020), pp. 2673–2718.

[115] Herbert A Simon. “Rational choice and the structure of the environment.” In: *Psychological review* 63.2 (1956), p. 129.

[116] R G Sloan. “Do stock prices fully reflect information in accruals and cash flows about future earnings?” In: *Accounting review* (1996).

[117] Robert F Stambaugh and Yu Yuan. “Mispricing Factors”. In: *The review of financial studies* 30.4 (Apr. 2017), pp. 1270–1315.

[118] Sheridan Titman, KC John Wei, and Feixue Xie. “Market development and the asset growth effect: International evidence”. In: *Journal of Financial and Quantitative Analysis* (2013), pp. 1405–1432.

[119] Edward Tower. “Are GMO’s Predictions Prescient? Using Them to Predict Vanguard’s Mutual Fund Returns, October 2010”. Oct. 2010.

[120] Amos Tversky and Daniel Kahneman. “Judgment under uncertainty: Heuristics and biases”. In: *science* 185.4157 (1974), pp. 1124–1131.

[121] J H Van Binsbergen and R S J Koijen. “Predictive regressions: A present-value approach”. In: *The Journal of finance* 65.4 (2010), pp. 1439–1471.

[122] Annette Vissing-Jorgensen. “Perspectives on behavioral finance: Does" irrationality" disappear with wealth? Evidence from expectations and actions”. In: *NBER macroeconomics annual* 18 (2003), pp. 139–194.

[123] Tuomo Vuolteenaho. “What drives firm-level stock returns?” In: *Journal of Finance* 57.1 (2002), pp. 233–264.

[124] Charles CY Wang. “Measurement errors of expected-return proxies and the implied cost of capital”. In: *Harvard Business School Accounting & Management Unit Working Paper* 13-098 (2015).

[125] Michael Weber. “Cash flow duration and the term structure of equity returns”. In: *Journal of financial economics* 128.3 (June 2018), pp. 486–503.

[126] Ivo Welch. *Simpler better market betas*. Tech. rep. National Bureau of Economic Research, 2019.

[127] Ivo Welch and Amit Goyal. “A comprehensive look at the empirical performance of equity premium prediction”. In: *The Review of Financial Studies* 21.4 (2008), pp. 1455–1508.

Appendix A: Appendix for Chapter 1: "Subjective Return Expectations"

A.1 More Details About Sell-Side Analysts Return Expectations Data

A.1.1 Measuring Analyst Return Expectations Using Analyst Price Targets

Firm- and market- level analyst return expectation are constructed using a bottom-up approach based on analyst-level return expectations per analyst issuance.

I collect single issuance of price targets from individual analyst's 12-month[1] price targets for individual firms from IBES unadjusted data base and match it with the closing price from CRSP on the date the price target is issued[2] to compute return expectation with price targets for individual firms. The expected returns are computed by dividing analyst's price targets by the daily closing price on the day the estimates was issued and subtracted by 1[3], or

$$\mu^A_{i,f,d} = \frac{P^{A,12}_{i,f,d}}{P_{f,d}} - 1$$

where $P^{A,12}_{i,f,d}$ is the price target of analyst i for firm f, issued at day d. The superscript 12 denotes the 12-month ahead estimates. Notice this methodology ensures there is no mechanical relation between mean estimated expected returns and the level of prices. On each issuing date the analyst has the freedom to pick her own price target since she observes the prices.

Firm-level return expectations are constructed together with the stop file provided by IBES to ensure individual estimates are not stale. IBES keeps track of the activeness of the

[1] Other horizons are available, though the coverage is poor.

[2] In case the issuance date is a weekend, the last Friday prices are used; In case the issuance is a holiday, the previous business day closing prices are used.

[3] The same formula is used in Brav and Lehavy (2003a) and Da and Schaumburg (2011)

individual estimates and provides a stop file for price targets[4]. I merge the point-in-time analyst-level expected return file with the stop-file on price targets to exclude estimates that analysts and IBES have confirmed to be no longer valid. Furthermore, to avoid stale estimates, I further restrict the estimates to be no older than 90 days when entering mean consensus estimates [5]

I construct weekly firm-level consensus expected returns by taking the mean of all active analyst-level forecasts, although using median makes no discernible difference for the main results. I drop analyst-level estimates that are greater than 5 standard deviation away from the mean estimates and I winsorize the entire analyst-level data base by 1% and 99% before calculating firm-level consensus. I take the mean of the available expected return estimates for each firm by the end of Saturday each week, or

$$\mu^A_{f,w} = \sum_i \mu^A_{i,f,w}/I_f$$

where I_f is the number of analyst for firm f at week w. For most of the application of the paper, I use firm-level return estimates based on monthly data, which is the consensus data on the last Saturday before each calendar month end.

Market level aggregate return expectations are constructed based on the SP500 universe. The aggregate market level return expectations for SP500 index is the firm market-cap ($M_{f,t}$) weighted average of firm-level return expectations at the end of the month t, or

$$\mu^A_{m,t} = \sum_f \frac{M_{f,t-1}}{\sum_f M_{f,t-1}} \mu^A_{f,t}$$

[4]According to IBES, this stop-file "includes stops applied to estimates that are no longer active. This can result from several events, e.g. an estimator places a stock on a restricted list due to an underwriting relationship or the estimator no longer covers the company. Prior to June 1993, actual stop dates did not exist in the archive files used to create the Detail History. An algorithm was developed to determine the date when an estimate became invalid if, for example, a merger between companies occurred or an analyst stopped working for a firm, etc. Estimate that are not updated or confirmed for a total of 210 days, the estimate is stopped."

[5]Engelberg, McLean, and Pontiff (2019b) allows the estimates to be at most 12 month old, in case the estimates are not covered by the stop-file, although the choice makes little difference for the main results, as verified in the Appendix.

In the Appendix, I also examined the results based on equal-weighted index. The results do not change qualitatively.

Additionally, the firm-level 12-month forward earnings to price ratio is constructed based on IBES analyst 1 fiscal year and 2 fiscal year ahead EPS estimates. Analyst level detailed unadjusted EPS estimates are multiplied with the number of shares outstanding at the date when analyst issued the EPS estimates to get total earnings. Subsequently, firm-level earnings estimates for 12-month ahead are linearly interpolation between the 1 year and 2 year ahead median earnings estimates for the firm at each month end. This methodology is consistent with how CRSP constructing their indices and is also used in O and Myers (2020)

A.1.2 Summary Statistics

The data set on analyst target prices has good and stable coverage for a large number of firms, especially when compared to surveys from CFOs and others that were studied in the literature. The coverage for SP500 is significantly better than the smaller firms, which is the reason why I choose SP500 universe as the venue for most of the empirical tests. Table A.1 shows the summary statistics for the variables used in this paper.

Panels (a) and (b) show the coverage of return expectations for the SP500 firms and all other firms. The number of analysts who filled survey far exceeds those of CFO (From Duke University), or retail investors (Shiller Individual), which has 390 and 81 respondents, respectively[6]. At a point in time, there are about 2700 analysts from 236 brokerage firms in the universe, among which 1410 analysts from 144 firms at a point in time cover SP500 firms, or 2.6 analysts per firm. The coverage deteriorates as the firm size becomes smaller, as shown in Panel (b), the number of analysts peer firm reduces to only 0.71 for the entire COMPUSTAT universe. For this reason, I use the SP500 universe as the main data set for analysis. On the other hand, the median analyst in the data set covers 35 firms, with a standard deviation of 22 firms. This is consistent with the practice of an analyst covering a

[6]These numbers are as reported in Table 1 of Adam, Matveev, and Nagel (2021a).

sector.

For SP500 firms, analysts revise their forecasts on average every 20 days, with a standard deviation of about 16 days. Notice that when constructing the sample, I exclude all estimates that are older than 60 days. In Appendix A.1.3, I describe the timing of the issuance in more detail. The existing surveys on CFOs and retail investors, are all in quarterly frequency.

Figure A.1: Coverage statistics SP500 over time

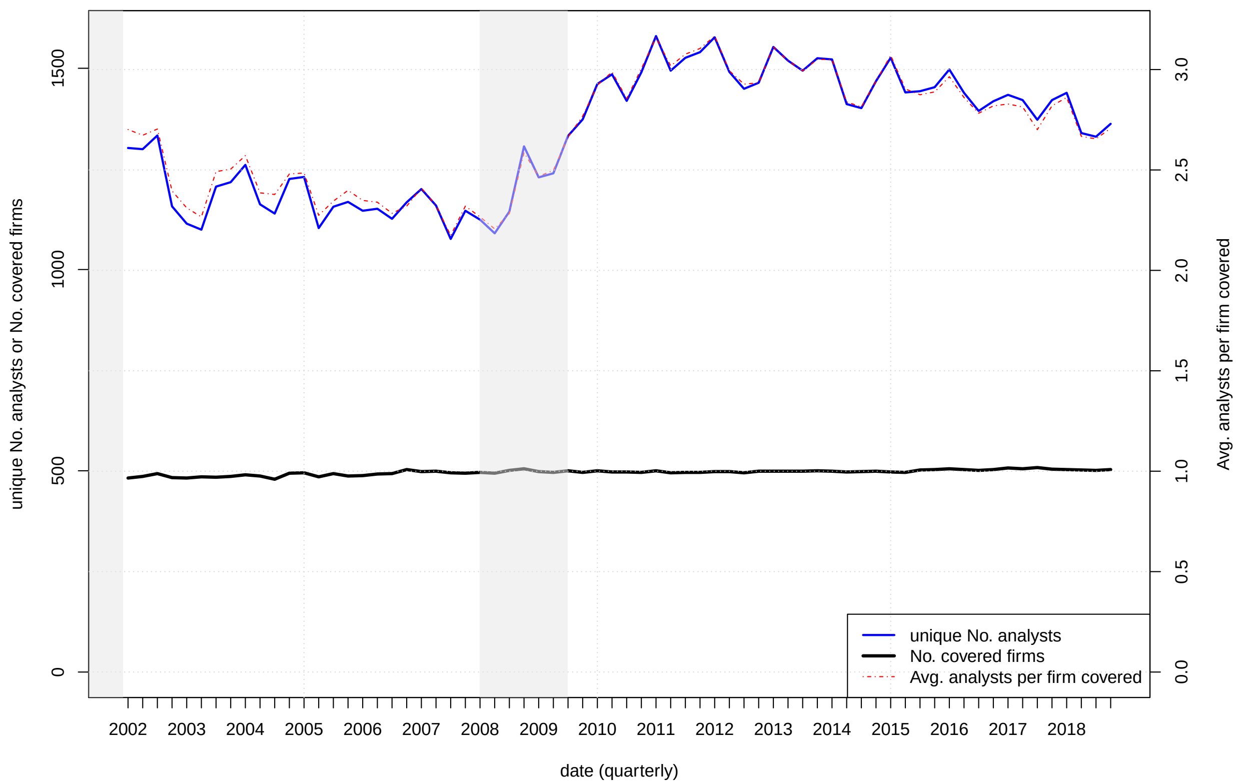


Figure A.1 and Figure A.2 show the coverage of analyst return expectations data is stable over time for both the S&P 500 and the CRSP-COMPUSTAT universe, respectively. For the SP500 firms, the number of analysts submitting price targets goes from around 1200 before 2008 to about 1500 in the last decade. The average number of firms covered stays very close to 500 over time. Figure A.2 shows coverage over time for all firms with analyst price targets, which shows stable coverage as well.

Figure A.2: Coverage statistics all firms over time

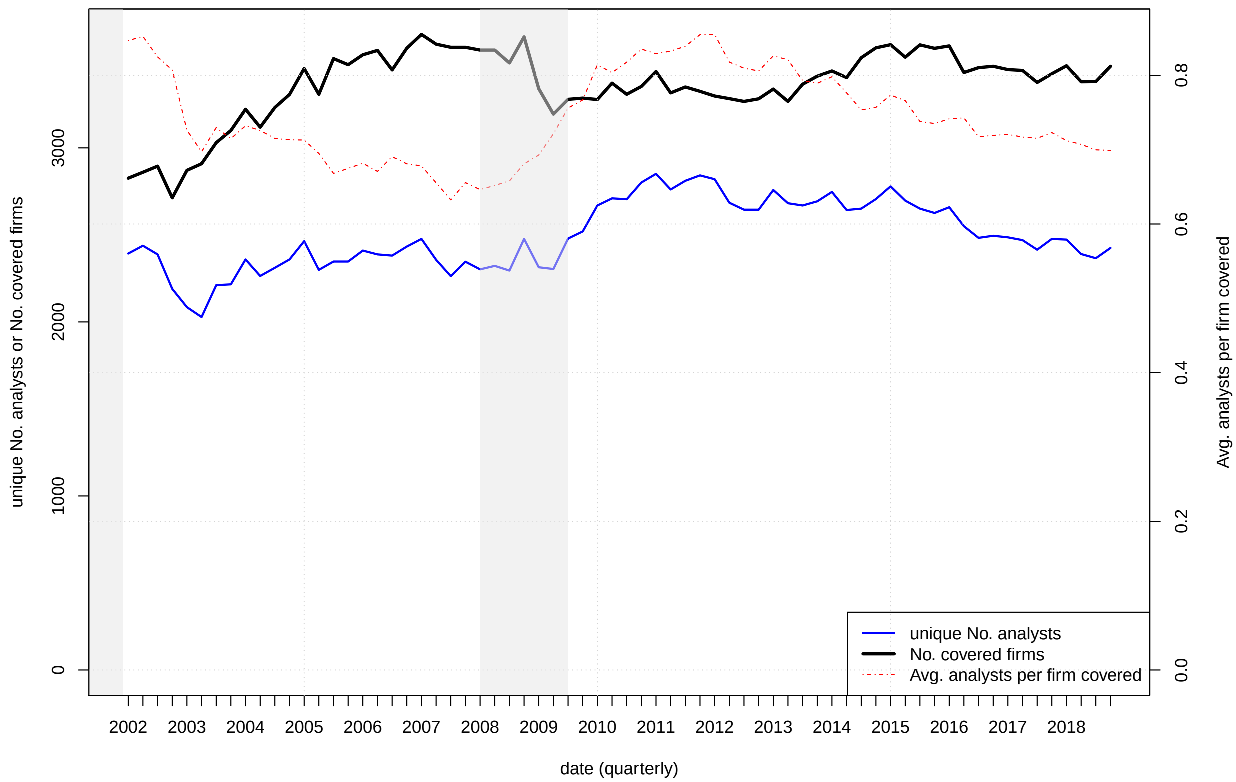


Table A.1: Summary Statistics

(a) Coverage Statistics SP500

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
nr.brokerage.firms	143.353	15.717	100	135.8	152	181
nr.analysts	1,334.926	151.938	1,076	1,167.8	1,461.8	1,580
nr.covered.firms	496.162	6.601	479	493.8	500.2	508
avg.days.since.last.revise	20.141	1.173	17.600	19.347	20.776	22.112
std.days.since.last.revise	16.374	0.852	14.633	15.697	16.953	18.422
avg.nr.analysts.per.firm	2.689	0.289	2.174	2.398	2.929	3.167

(b) Coverage Statistics All

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
nr.brokerage.firms	230.338	27.148	125	228.8	241.2	271
nr.analysts	2,495.632	195.426	2,028	2,354.5	2,669	2,851
nr.covered.firms	3,360.691	209.747	2,713	3,282.8	3,514.2	3,652
avg.days.since.last.revise	20.017	1.033	17.523	19.393	20.481	23.032
std.days.since.last.revise	15.928	0.661	14.602	15.479	16.364	17.625
avg.nr.analysts.per.firm	0.744	0.061	0.632	0.699	0.804	0.855

(c) SP500 Firm-level Analyst Expectation Data

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
ER.analyst	101,985	0.135	0.110	−0.107	0.066	0.185	0.549
fwd.12m.E/P	103,575	0.065	0.031	−0.038	0.048	0.080	0.162
LTG	99,094	0.115	0.080	−0.114	0.074	0.149	0.462

(d) Other SP500 firm-level data

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
tot.ret	103,676	0.008	0.085	−0.255	−0.037	0.054	0.266
mcap(in bil.)	103,693	27.927	51.410	0.083	6.570	26.041	1,099.436
ann.earnings(in bil.)	103,084	1.553	3.264	−53.557	0.322	1.465	56.518
ROE	102,099	0.183	0.233	−0.701	0.089	0.233	1.504
B/M	102,360	0.483	0.415	−0.034	0.211	0.629	2.440

Notes: The sample period is from 2002-01-01 to 2018-12-31. Monthly and quarterly data are measured at calendar month and quarter end, respectively. "ER.analyst" denotes the analyst return expectations; "fwd.12m.E/P" denotes forward 12 month earnings expectation, constructed using 1 and 2 fiscal year ahead earnings expectation, divided by the market cap; "LTG" are analyst long-term growth expectation; "ann.earnings" are annual actual earnings; "ROE" is actual annual earnings divided by total book value; "B/M" are book to market ratio.

The distributional statistics in Panel (c) of Table A.1 show that the average analyst return expectation for firms in SP500 universe is 13.4%. This is much higher than the realized average total return of about 9.6% per year, as shown in Panel (d). The positive bias documented here is consistent with those in the previous literature, such as Engelberg, McLean, and Pontiff (2019b) and Brav and Lehavy (2003a). Besides, the analyst earnings forecasts statistics are similar to those documented in Bordalo et al. (2019) and O and Myers (2020).

A.1.3 The Timing of Analyst's Price Target Issuance

I describe in more detail the timing of analyst's issuance of price targets. First, I examine the frequency at which an analyst issues a price target. Second, I investigate whether analysts issue more price targets during firms earnings announcement month.The results in this section show that the median analyst issues a new price target every 16 days. For analysts that issues a price target less infrequently, they tend to issue new estimates during earnings announcement month, and particularly on or one day after earnings.

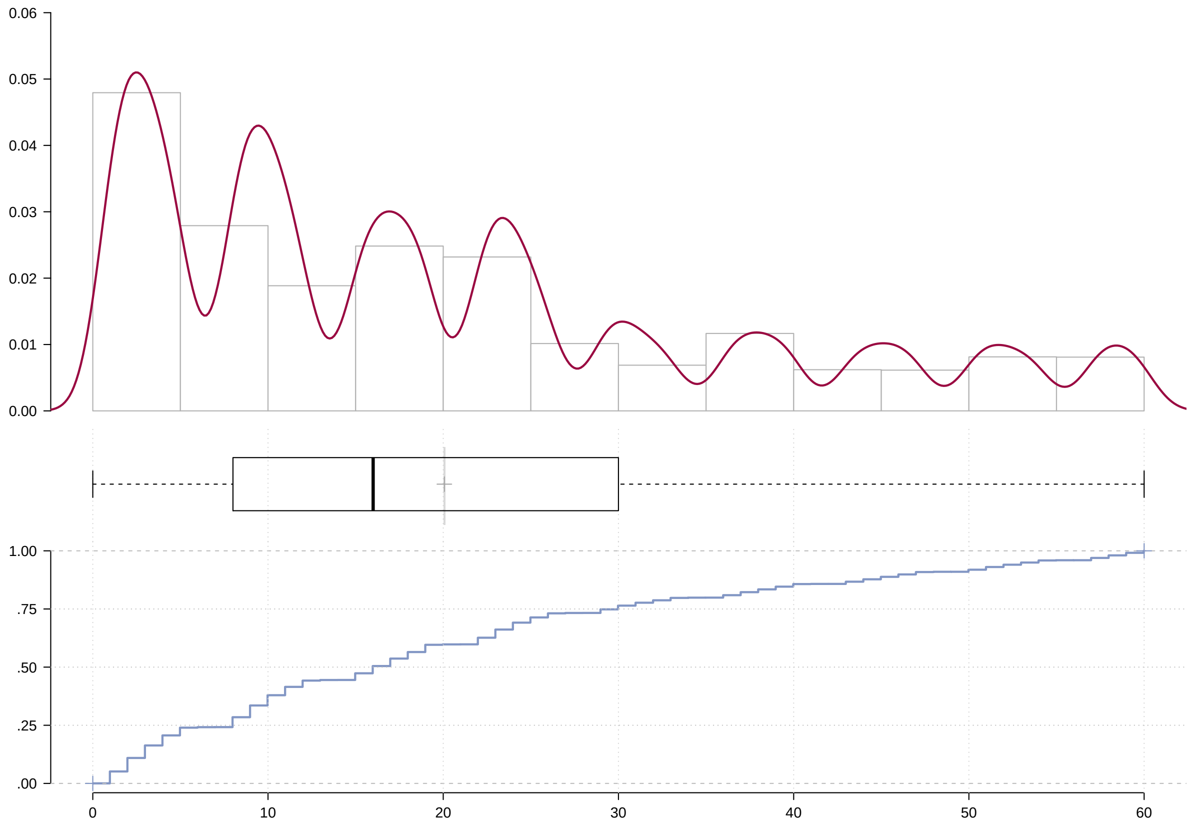
How Frequent an Analyst Issues a Price Targets

On average, a median analyst issues a new price target every 16 days for a particular firm he/she covers. Only 2% of these estimates are the same as price targets issued previously.

Figure A.3 plots the empirical distribution of the number of days between an analyst's newly issued price target and his/her previous issuance on the same firm. On average, a median analyst issues a new price target every 16 days, with a mean of 20 days. Furthermore, about 75% of the analysts would issue a new estimate each month, or within 30 days. Combining this information together with Figure A.4, those analysts who issues less frequently, say those who issue an new estimate every 60 days, will typically issue during the earnings month each quarter.

Another question is that, among these frequent updates, how many incidences analysts

Figure A.3: Empirical Distribution (Upper: PDF, Lower: CDF) of the No. of days between two subsequent issuance of price targets by the same analyst for the same firm



Note: Upper panel: Probability density function; Lower panel: Cumulative distribution function.

would maintain, or issue the same price targets as previous price targets? Table A.2 shows that only about 2% of all issued price targets are the same as the previous one. This percentage is much lower for return expectation. Over time, this percentage is also stable, varying at about 1% standard deviation per quarter.

Table A.2: Summary Statistics: Percentage of Analysts Who Issues the Same Price Target Each Quarter

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
prop.maintain.PTG	68	0.023	0.011	0.004	0.015	0.030	0.053
prop.maintain.ER	68	0.0004	0.001	0.000	0.000	0.0002	0.011

Note: "prop.maintain.PTG" is calculated as follows. First, for each analyst issued price target for a particular firm, the previous issuance for the same firm was compared, if available. Subsequently, for each calendar quarter, count the number of incidences where the current issuance is equal to the previous quarter and the number that they are not equal. "prop.maintain.PTG" is the proportion of the former (same price target) divided by the total number of analyst issuance. "prop.maintain.ER" is the same proportion but calculated using expected returns, instead of price target estimates.

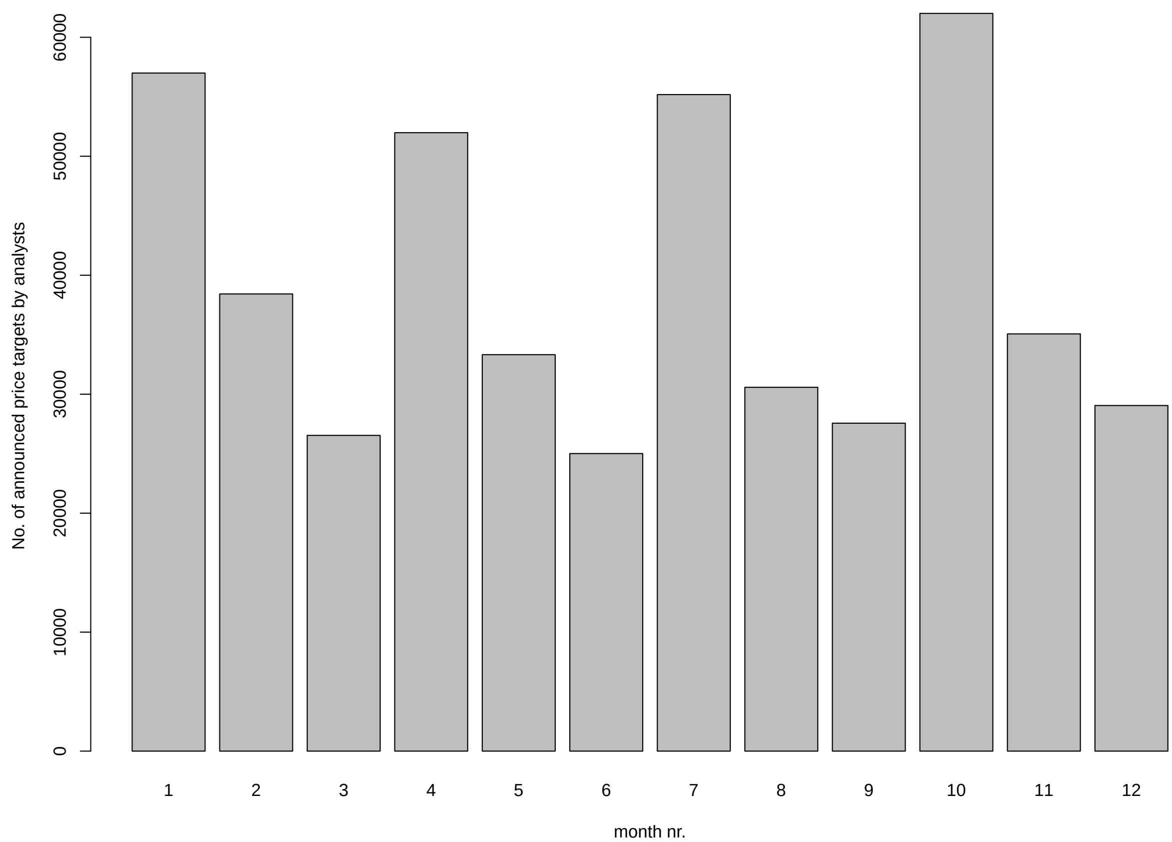
Price Targets Issuance and Firm Quarterly Earnings Announcements

Analysts are more likely to announce their price targets during the first month of each quarter, during which more firms announce their quarterly earnings. During firms' announcement months, analysts typically announce new price targets on or shortly following the announcement day.

Figure A.4 plots the number of announced price targets by all analysts in the sample of SP500 firms for the whole sample. The total price targets announced during Jan., Apr. Jul. and Oct. are about 48% of all announced, higher than the 33% if they are announced evenly through out the year. This is similar to the seasonal pattern of earnings announcements.

In fact, The seasonal pattern of firms' earnings announcements are much more dramatic compared to the announcements of analysts' price targets. As shown in Figure A.5, the number of earnings announcements are almost 8 times than those in Q2, Q3 and Q4. This

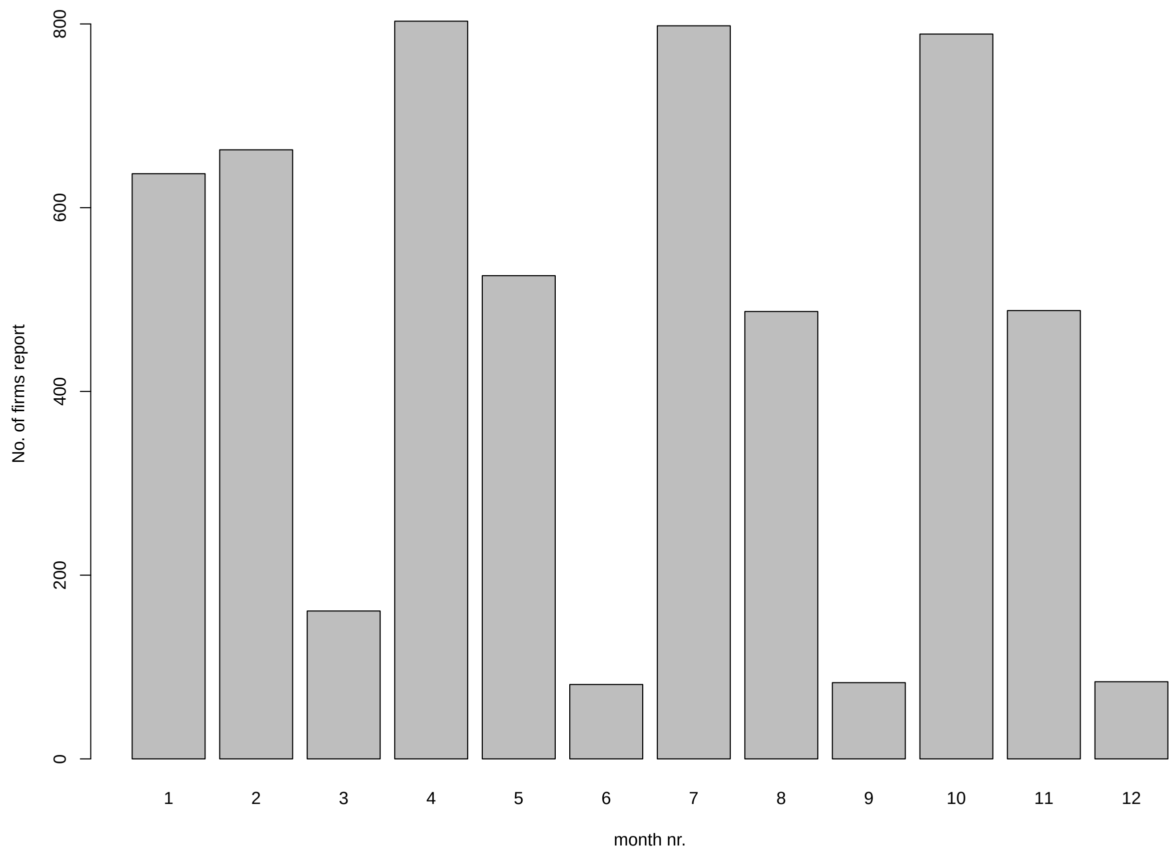
Figure A.4: Number of Price Targets Announced by Month, SP500 Universe



Note: Sample starts from 2002-01-01 to 2018-12-31. The No. of announced price targets by analysts count each single analyst submitted price targets for a particular firm during a month as one. Price targets that are issued longer than 60 days ago are excluded.

suggests that analysts' price targets change are not only driven by firms' earnings.

Figure A.5: Number of Earnings Announcements By Month, SP500 Universe



Note: Sample starts from 2002-01-01 to 2018-12-31. The No. of firms that report earnings count all firms that ever are in the SP500 index throughout the sample period and therefore can exceed 500.

During the month that firms do announce, about 50% of all new price target issuance is concentrated on or one day after the earnings day. The distribution of the number of days between announcement of price targets and the earnings announcement day is summarized in Table A.3.

Table A.3: Summary Distribution: Days Between Announcement of Price Targets and Nearest Earnings Announcements Within Each Month

Statistic	N	Mean	Median	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
days.after.earnings	282,806	−1.580	1	7.790	−30	−2	1	30

A.2 Details on the Sources of Return Expectation Data

As an overview, for analyst earnings and return forecast data, I use the IBES unadjusted file. Firm fundamentals, SP500 membership data are fro m COMPUSTAT and daily pricing data are from CRSP. CFO return expectation data are from Duke University CFO Global Business Outlook that are available on their website[7]. Retail investor return expectations are based on Robert Shiller's surveys[8] as well as the consumer survey conducted by University of Michigan[9].

For objective measures of expected returns, I construct aggregate price dividend ratio using the total returns and price returns on SP500 index. I document the detail of the construction in Appendix A.3. For the aggregate price to earning data, I used the data kindly provided by Prof. Robert Shiller on his website. The consumption-wealth ratio, or CAY is downloaded from Prof. Martin Lettau's website, which is constructed based on Lettau and Ludvigson (2004).Below are more details about these data sources.

Shiller Survey asks participants about their expected percentage increase on the Dow-Jones index over different horizons in the future. The participants consist of two groups, one retail investor, which is randomly selected U.S. wealthy individuals and the other is institution, or "the investment managers section of the Money Market Directory of Pension Funds and Their Investment Managers". I use the data set which aggregates the raw data and report the "the percent of the population expecting an increase in the Dow in the coming

[7]https://www.cfosurvey.org/

[8]Available on Yale University's website: https://som.yale.edu/faculty-research-centers/centers-initiatives/international-center-for-finance/data/stock-market-confidence-indices/united-states-stock-market-confidence-indices

[9]Available on https://data.sca.isr.umich.edu/data-archive/mine.php

year" from July 2001 to Dec 2020, during which the data was collected monthly and the moving average of the 6 month data are published. I also consider a data set that used and made public by Adam, Marcet, and Beutel (2017), which uses the raw averages of expected price growth from Shillar Survey, but only available from 2001-Q1 to 2012-Q4.

The Michigan survey asks about 500 households in the U.S. "What do you think is the percent chance that this one thousand dollar investment will increase in value in the year ahead, so that it is worth more than one thousand dollars one year from now?" and calculates the average across all responses. The survey is conducted monthly. I use data that starts from August, 2002 and end in December, 2018. I mainly use quarterly data at the end of each calendar quarter, to be consistent with the other survey data.

GMO 7-year Asset Class Forecasts is produced quarterly by GMO, which consists of return forecasts for 7-year ahead for different segments of equity and bond markets, including but not limited to U.S. large caps, international small caps or emerging market bonds. The data after 2017 are available directly on their website. For pre-2017 data, I hand-collected the data from the internet. This is possible because the company publishes the return forecasts dating back from the second quarter of 2000 and the publication is in the form of a standalone imagine, as shown in Figure A.6. On websites of advisers or consultants, they have cached historical figures of these snapshots. However, I was not able to collect the full history of their return forecasts, in quarterly frequency.

Figure A.6: A Snapshot of One GMO 7-year Asset Class Forecasts

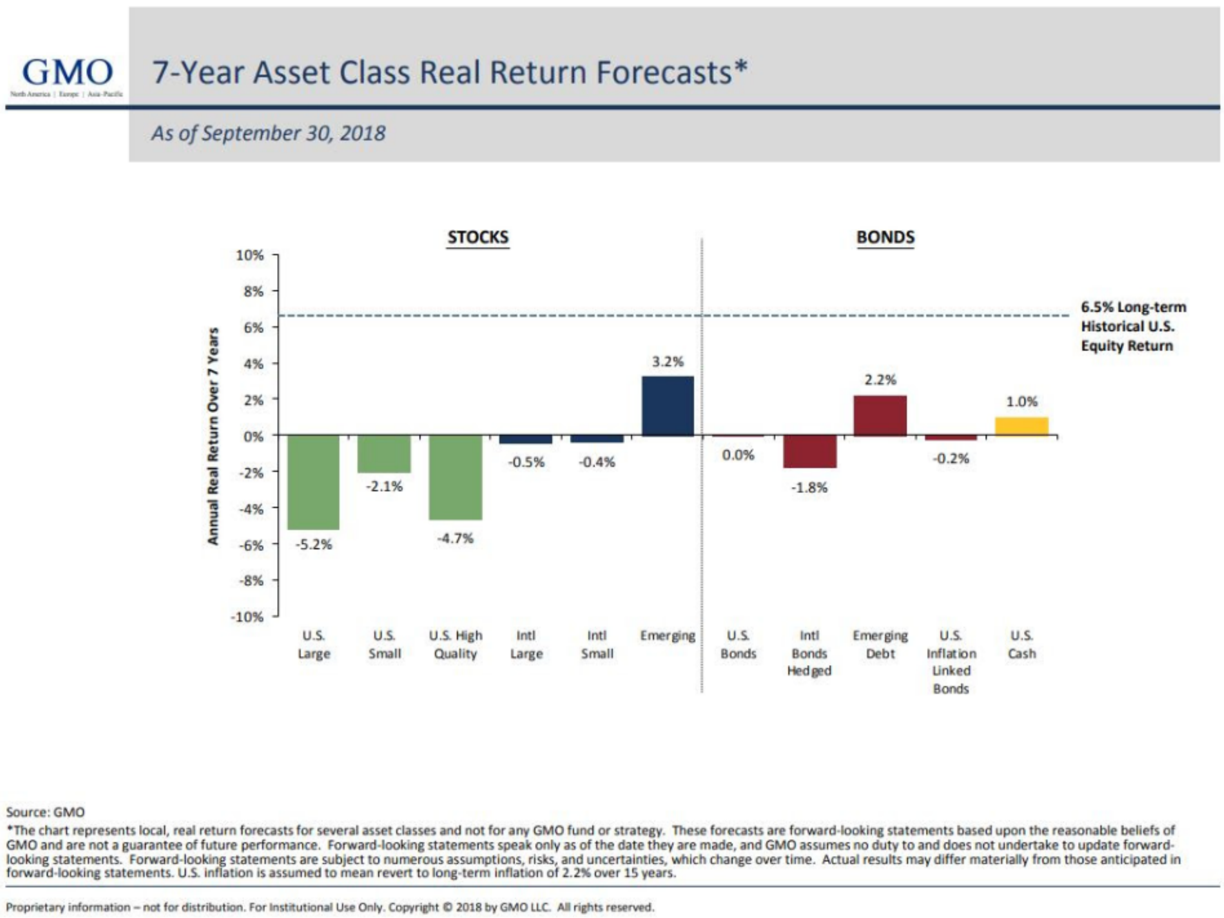


Table A.4: Summary Statistics Subjective Quarterly Expectations

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
ER.analyst	70	0.143	0.037	0.095	0.114	0.161	0.270
LTG.analyst	70	0.115	0.015	0.084	0.105	0.122	0.153
ER.CFO	69	0.056	0.014	0.022	0.046	0.065	0.091
ER.Shiller.12m	46	0.051	0.021	0.018	0.036	0.068	0.108
ER.GMO.7y	70	−0.0002	0.025	−0	−0.02	0.01	0
Shiller.Inst.Pct.Up	70	77.917	5.689	62.810	74.797	81.385	92.520
Shiller.Ind.Pct.Up	70	77.293	8.486	61.270	71.978	84.220	95.280
Michigan.Pct.Up	66	53.117	6.290	36.500	48.000	57.400	63.600
ER.SPF.Pct.10y	70	6.735	0.796	5.337	6.152	7.437	7.683
No.IPO	70	42.914	22.122	1	31	61.8	85
Equity.Net.Issuance	70	−0.048	0.033	−0.150	−0.071	−0.031	0.033
tot.ret.qtrly.SP500	70	0.020	0.078	−0.219	−0.013	0.063	0.159

Table A.5: Summary Statistics Quarterly Subjective Expectations

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
ER.analyst	70	0.143	0.037	0.095	0.114	0.161	0.270
ER.CFO	69	0.056	0.014	0.022	0.046	0.065	0.091
ER.consumer(raw)	66	53.117	6.290	36.500	48.000	57.400	63.600
ER.consumer(proj)	66	0.041	0.023	0.002	0.025	0.060	0.102
ER.buy.side	70	−0.0002	0.025	−0	−0.02	0.01	0
No.IPOs	70	42.914	22.122	1	31	61.8	85
past.6m.cum.ret	70	0.031	0.115	−0.316	−0.015	0.090	0.325

Table A.6: Correlations Between Different Subjective Expectations

	ER.analyst	LTG.analyst	ER.CFO	ER.Shiller.12m	ER.GMO.7y	Shiller.Inst.Pct.Up	Shiller.Ind.Pct.Up	Michigan.Pct.Up	ER.SPF.Pct.10y	No.IPO	Equity.Net.Issuance	tot.ret.qtrly.SP500
ER.analyst	1.00											
LTG.analyst	0.26**	1.00										
ER.CFO	-0.27**	0.29**	1.00									
ER.Shiller.12m	0.00	0.19	0.68***	1.00								
ER.GMO.7y	0.41***	-0.39***	-0.06	-0.20	1.00							
Shiller.Inst.Pct.Up	0.21*	0.03	0.20*	0.33**	0.26**	1.00						
Shiller.Ind.Pct.Up	0.37***	0.20*	0.47***	0.65***	0.34***	0.57***	1.00					
Michigan.Pct.Up	-0.68***	0.38***	0.38***	0.04	-0.69***	-0.37***	-0.27**	1.00				
ER.SPF.Pct.10y	0.45***	0.21*	0.35***	0.30**	0.53***	0.50***	0.81***	-0.31**	1.00			
No.IPO	-0.59***	0.11	0.20*	-0.22	-0.31***	-0.22*	-0.24*	0.66***	-0.12	1.00		
Equity.Net.Issuance	-0.01	0.05	0.45***	0.49***	0.01	0.16	0.29**	-0.12	0.04	-0.27**	1.00	
tot.ret.qtrly.SP500	-0.60***	-0.18	0.44***	0.16	0.06	-0.05	0.03	0.15	-0.05	0.29**	0.32***	1.00

A.3 Constructing Index-level Price Dividend Ratio

I follow Adam, Marcet, and Beutel (2017) to construct index level price dividend ratios on monthly frequency. Monthly data on the level of the S&P500 index denoted as P_t^{SP} , as well as the monthly holding returns on the index without dividend, or R_t^{ND} are from CRSP. The monthly total returns on S&P500 index including dividend, or R_t^D are from Global Insight[10].

The monthly total dividend is

$$D_t = (\frac{1 + R_t^D}{1 + R_t^{ND}} - 1)P_t^{SP}$$

and the annual dividend is the sum of total dividend in the last 12 month

$$D_t^A = \sum_{i=0}^{11} D_{t-i}$$

and the log price dividend ratio used in this study is

$$pd_t = log(PD_t) = log(\frac{P_t}{D_t^A})$$

Notice, since the return expectation from analysts are based on analyst's forecasts of *Price* of the stock instead of returns, the index return expectations by analysts correspond to

$$E_{t-1}^A(R_t^{ND}) = E_t^A(\frac{P_{t+1}^{SP}}{P_t^{SP}} - 1)$$

where the superscript on the expectation operator denote the analyst expectation.

[10]CRSP computes itself a value-weighted total returns including dividends and without dividend. However, upon examine the monthly implied dividend series, I found outliers. For example the November 2014 monthly dividend is almost 3 times the magnitude than that of the dividend in any other months in 2013 or 2014. Therefore, I used the global insights total returns. The implied dividend series does not have the irregular pattern throughout the entire sample from 1970 to 2019.

A.4 Detailed Analysis on Sell-Side Analysts' Contrarian Return Expectations

A.4.1 Contrarian Effects Across Different Horizons

Results from Column 2 and 3 of Table 1.3 also raise the question of which horizon of past returns matter most to analyst future return expectation. To investigate this question, I follow Greenwood and Shleifer (2014) to estimate the nonlinear regression of the form

$$\mu^A_{m,t} = a + b\sum_{j=0}^{k}\omega_j R_{m,t-2,t-j} + e_t \tag{A.1}$$

where $\omega_j = \frac{\lambda^j}{\sum_{i=0}^{k}\lambda^i}$ is the weight on past returns and λ measures how quickly past return die out in analysts' memory. A value of λ equal to 1 implies the returns of different horizons are equally important in influencing analyst future expectation while a value smaller than one means more recent past returns are more important than distant ones.

For the empirical implementation, I run the regression A.1 using monthly data based on cumulative quarterly returns that range from one-quarter (returns that are lagged by 3 to 6 months) to 12-quarters, so 16 regressors in total. I correct for the auto-correlations in the return expectations by using Newey-West standard errors with 12 month lags. The non-linear least squares estimates are presented in Table A.7.

The estimates of λ is 0.904, which shows that although analysts pay more attention to recent past returns. Compared to the result found in Greenwood and Shleifer (2014), who estimates the average value of λ to be 0.56 based on a host of other subjective expectations, distant returns have much more important impact for analysts. This estimate can also be contrasted with Malmendier and Nagel (2011), who find that *distant* but salient past history play a role in investor market participation decisions.

Table A.7: Analyst Return Expectations and Past Returns of Different Horizons

	a	b	λ
Estimate	0.152	-0.845	0.866
Std. Error	(0.009)	(0.372)	(0.064)

Notes: Nonlinear least squares monthly time-series regressions of analyst return expectations of market returns for the next year on cumulative quarterly returns that range from one-quarter (returns that are lagged by 3 to 6 months) to 12-quarters:

$$\mu^A_{m,t} = a + b \sum_{j=0}^{k} \omega_j R_{m,t-2,t-j} + e_t \tag{A.2}$$

where $\omega_j = \frac{\lambda^j}{\sum_{i=0}^{k} \lambda^i}$. Newey-West standard errors with twelve-month lags are shown in brackets. Sample starts from 2002-01-01 and ends in 2018-12-31, a total of 204 months.

A.4.2 Firm-level Results

The results in Table 1.3 show the analyst's contrarian views are present at the market level, which is a value-weighted average of firm-level variables. At the firm level, lower past returns are typically associated with an increase in valuation ratio, such as book to market ratio. An extensive literature has documented a positive relation between valuation ratios and future stock returns. This raises the question of whether the negative relationship between past returns and analyst expectations is merely an effect of analysts using firm's valuation ratios as a determinant in forming their expectation.

To answer this question, I run firm-level analyst return expectation on past returns together with a host of firm-level characteristics, including valuation ratios as control variables

$$\mu^A_{i,t} = \alpha_i + bR_{i,t-2,t-6} + cX_{i,t-2} + e_{i,t} \tag{A.3}$$

Since the paper focuses on the *time-variation* of analyst return expectation, I include a firm-effect in the panel regression.

First, the results show that analysts also hold strong contrarian view at the firm-level. In fact, when only including past 6-month returns (Column 1), the coefficient on analyst contrarian view has an estimate of -0.11, very close to the results on the aggregate, shown in Table 1.3.

Second, the coefficient on past returns changes very little when including other control variables, as shown in Column 3 and 4. Firm's valuation ratios such as book to market ratio does predict analyst's return expectation, as shown in Column 2, although the economic magnitude is much smaller, when compared to past returns.

Interestingly, Column 5 also show that analyst's own forecasts on future earnings, both one-year ahead and long-term, have a strong correlation with their own return expectation. This is consistent with the results documented in Da, Hong, and Lee (2016). Furthermore, the higher a firm's investment, the higher analysts would expect its future expected returns to be. To the best of my knowledge, there are not other prior literature documenting the effect of investment on subjective return expectation. However, this is not the focus of the current paper so I will not explore further.

In sum, the pattern that past returns are strong predictors for analyst's return expectations are robust at the firm level and are not due to analysts using firm's valuation ratios to make forecasts on firms' future returns.

A.4.3 Concerns with Stale Estimates? Return Expectations of Analyst First-time Issuance

One concern regarding the conclusion that analysts holding contrarian views is that analysts' stale price targets might be driving the negative relation between past returns and future analyst return expectation. To illustrate the concern more clearly, consider an extreme case where analysts never change their price targets. As prices go up, the expected returns go down mechanically and the contrarian conclusion follows. Although this might still be due to analyst intentionally holding slow-moving return expectations and thus appear to be contrarian, or come from analyst's limited attention or simply being lazy.

To eliminate such concerns, I show that the contrarian results are robust to a sample

Table A.8: Cross-Sectional Determinants of Analyst Return Expectations

	Dependent variable:			
	Analyst Return Expecations			
	(1)	(2)	(3)	(4)
lag.2m.cum.6m.ret	−0.111***		−0.102***	−0.094***
	(0.011)		(0.010)	(0.010)
lag.2m.CF/P		−0.014		
		(0.012)		
lag.2m.B/M		0.024***	0.013***	0.014***
		(0.005)	(0.004)	(0.004)
lag.2m.fwd.12m.E/P				0.519***
				(0.074)
lag.2m.LTG				0.351***
				(0.025)
lag.2m.Prof				−0.021***
				(0.006)
lag.2m.Inv				0.097***
				(0.014)
Constant	0.139***	0.123***	0.133***	0.059***
	(0.003)	(0.003)	(0.004)	(0.007)
Observations	99,716	91,202	92,735	86,184
R^2	0.049	0.012	0.052	0.138
Adjusted R^2	0.049	0.012	0.052	0.138
Residual Std. Error	0.105 (df = 99714)	0.108 (df = 91199)	0.105 (df = 92732)	0.099 (df = 86177)
Note:				*p<0.1; **p<0.05; ***p<0.01

Note: Firm-level analyst return expectation regressed on past returns together with other of firm-level characteristics together with control variables in $X_{i,t-2}$

$$\mu_{i,t}^{A} = \alpha_i + bR_{i,t-2,t-6} + cX_{i,t-2} + e_{i,t}$$

with a firm fixed effect α_i. Sample is based on SP500 firms from 2002-01-01 to 2018-12-31. Standard errors are clustered by firm and month. "lag.2m" means variables are lagged by 2 months. "cum.6m.ret" cumulative total returns for the firm in the past 6 months; "CF/P" Cash flow to market cap; "B/M": book to market ratio; "fwd.12m.E/P" analysts 1-year ahead forward earnings divided by market cap; "LTG" analyst long-term growth estimates; "Prof" Operating profitability defined as in Fama and French (2006); "Inv" annual asset changes divided by assets, as dfined in Fama and French (2006). Firm-level variables are winsorized at 1% and 99% over the entire sample.

containing only return expectations based on each analyst's first-time issuance ever for a particular firm in the entire IBES data base. Because the first-time issuance is always fresh and there is no potential staleness due to aggregation process, such results mean the negative correlation between analyst return expectations and past returns are mainly driven by analyst's contrarian views instead of staleness of analyst forecasts.

Table A.9 shows the results for the following regression

$$\mu^A_{j,i,t} = a + bR_{i,t-1,t-6} + cX_{i,t-1} + e_{j,t} \tag{A.4}$$

where j denotes an analyst and i denotes a firms. In particular, $\mu^A_{i,j,t}$ is the first estimate a particular analyst ever issued for a particular firm for both the EPS and price targets data bases[11]. Notice the analyst's issuance are recorded on the day of the issuance within each month and subsequently pushed to the end of the month to run on monthly data. Therefore, to avoid look-ahead bias, I require the independent variables to enter the regression with a one-month lag, so the predictive regression is entirely out of sample. I calculate standard errors by clustering by firm and month.

Results in Table A.9 show that the contrarian results documented at the aggregate and firm level also hold for the analyst level regression. Coefficients on the past returns are statistically negative for both the entire IBES and the SP500 universe. Furthermore, the magnitude of the coefficients across all specifications is very similar to those in the aggregate level and firm level regression, ranging from -0.12 to -0.17. Other firm-level controls do not drive away the contrarian effect.

In sum, these results confirm the staleness and other mechanical reasons are not the force driving the negative coefficients, and supports the conclusion that analysts' hold contrarian return expectations.

[11]EPS forecasts go back much further than price target data, which start to have good quality data from early 1980 and 2000, respectively. The reason for considering also EPS data base is to avoid cases in which an analysts have potentially stale price target estimates but are not reported in the price target data base. More details regarding how the two data bases are merged, see A.5.

Table A.9: Analyst-level first-time-issued return expectations vs. past past returns

	Dependent variable:			
	Analyst First Ever Issued Return Expectation For A Firm			
	(1)	(2)	(3)	(4)
past 6m.ret	-0.166^{***}	-0.148^{***}	-0.126^{***}	-0.125^{***}
	(0.022)	(0.020)	(0.022)	(0.023)
B/M		-0.013		-0.006
		(0.015)		(0.015)
Inv		0.207^{***}		0.222^{***}
		(0.037)		(0.042)
OpIB		-0.220^{***}		-0.033^{**}
		(0.017)		(0.014)
Constant	0.251^{***}	0.272^{***}	0.155^{***}	0.154^{***}
	(0.009)	(0.013)	(0.007)	(0.012)
Universe	All IBES	All IBES	SP500	SP500
Observations	9,282	8,310	2,887	2,795
R^2	0.032	0.107	0.029	0.052
Adjusted R^2	0.032	0.107	0.029	0.051
Residual Std. Error	0.335 (df = 9280)	0.299 (df = 8305)	0.189 (df = 2885)	0.187 (df = 2790)
Note:				*p<0.1; **p<0.05; ***p<0.01
				SEs are clustered by firm and month

Note: Sample period: 2002-01-01 to 2018-12-31. An analyst's first ever issued return expectation is regressed on the firm-level monthly variables including (lagged 1-month) cumulative past 6 month returns and other control variables. Analyst first-ever return expectation is based on analyst's first price targets ever issued in both EPS and price targets data in I/B/E/S data base. "lag.1m" denotes the variables are lagged by 1 month. "B/M", "Inv" and "OpIB" are book to market, investment and operating profitability variables defined as in Fama and French (2006). Independent variables are winsorized at 1% and 99%.

A.5 Merging Analyst Price Target Forecasts with EPS Forecasts

I construct an analyst-level historical coverage data set based on detailed analyst EPS and price target forecast data. For each analyst, before each date the analyst issues a price target, I trace all of the EPS and price target he/she has ever issued in the past. This set of firms are defined his/her coverage[12].

The EPS forecasts is the longest available analyst survey and has the best coverage, which goes back to 1980-01-01. The I/B/E/S database identify an analyst through a unique "analyst code", which I use to merge between the price target file and the EPS forecast file.

I first create a EPS-based coverage list in which all the firms for which an analyst has ever issued an EPS forecasts are included. The first ever announced EPS estimate of an analyst is considered as the start of the analyst's career. Additionally, the first and the last (or current) date on which he/she issues an eps estimate for a firm, is recorded as the start/end of his/her coverage for that particular firm. A similar coverage list is created for the price target data set. Empirically, the price target coverage is a subset of the coverage of EPS forecast for most of the analysts.

I make several filters to get rid of potential erroneous observations. First, I only include analyst's 1-, 2- fiscal year and 1-fiscal quarter ahead forecasts to make the EPS coverage list. The reason is that these periods are the most commonly surveyed horizons and are less prone to errors. Second, if an analyst stops appearing in the EPS file and reappears after 36-month, I count the restarting date as the analyst's career start. This is because only very few observations (6% of all observations) actually do reappear after 3 years. Analysts do update the forecasts quite often. The reason for not updating is mostly because of erroneous analyst identification code. Third, I delete analysts who cover more than 200 firms. The average number of firms covered by an analyst in the EPS data base is about 41 with a

[12] Admittedly, this coverage might not be complete. An analyst might be covering other firms and has not issued any EPS or price targets in the past for those firms. However, this potential under-estimation will not affect the results on the impact of past experience on the future price target forecasts, if the under-estimation is systematically correlated with the past experienced returns/earnings of the covered firms.

median of 33. Analysts who cover more than 200 firms are highly unusual, which amounts to less than 0.1% of all observations.

Table A.10: Summary statistics analyst-level historical coverage

Statistic	N	Mean	Median	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
no.firms.covered	14,352	18.224	10	22.099	1	3	26	189
no.firms.w.eps	14,352	17.160	9	21.782	0	2	24	186
no.firms.w.ptg	14,352	12.663	6	15.745	1	2	18	168
no.firms.w.eps.only	14,352	5.562	1	13.042	0	0	5	170
no.firms.w.ptg.only	14,352	1.065	0	3.911	0	0	1	122
no.firms.w.both	14,352	11.598	5	15.133	0	1	16	167
no.months.analyst.career	14,352	82.932	51.567	85.366	0.000	15.633	126.067	462.433

Note: Analyst-level historical coverage data set for analysts who issues at least 1 price target (ptg) during the entire sample period from 1999-01-01 to 2020-02-01. Analyst-level detailed unadjusted EPS data set is from 1980-01-01 to 2020-02-02 and price target data set is from 1999-01-01 to 2020-02-01. "no.firms.covered" is the number of unique firms that an analyst issues at least one price target and/or eps forecasts for. "no.months.analyst.career" is the number of months from the first ever price target or eps to the last time the analyst issues a ptg or eps.

The EPS-based coverage list is merged with analyst-level price target issuance to obtain the coverage history for each individual analyst that *ever issues at least one price target*.

Table A.11 summarizes the analyst-by-analyst coverage data set. The data set contains more than 14 thousand analysts. The average number of firms an analyst covers is about 18 firms, consistent with industry standard of about 10 to 30 names per analyst. The coverage is skewed to the left, with a median analyst only covers 10 firms.

Typically, an analyst submit more EPS forecasts than price targets for firms he/she covers. Among the 18 firms that an average analyst covers, more than 16 has an eps forecasts and only 1 firm has only price targets but no eps forecasts. The less price target forecasts as compared to EPS is consistent with the facts documented in previous literature, such as Da, Hong, and Lee (2016).

This data set contains the point-in-time data for all the firms an analyst has ever issued EPS or price target forecasts, before he/she issues a new price targets. Furthermore, it also has the first-ever issuance the analyst ever makes for either EPS or price targets, which I use to construct the number of years experience variable. This data is the basis for studying the impact of experience on analyst's return expectations.

This heterogeneity in the duration of career is helpful for the analysis on analyst's past experience. A median analyst typically last for about 4 years between his/her first issuance and last, with a standard deviation of about 7 years. Notice that there are some veterans who go on to have a career spanning almost four decade. One such analyst that I am able to trace online is Chuck Cerankosky from Northcoast Research, who have started his career at Rouston Research in 1979 and is still active. Notice that this sample only include the analysts who ever issues a price target. Compared to analysts who ever issues an EPS estimate only but not necessarily a price target, whose median career span is about 31 months, the analysts in this sample have substantially longer professional career.

Table A.11: Summary Statistics: Analyst Coverage History

(a) Point-in-time coverage statistics

Statistic	N	Mean	Median	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Coverage.eps.forecasts.per.analyst	55,093	13.084	12	9.407	1	5	19	116
Coverage.price.targets.per.analyst	55,093	9.110	7	7.496	1	3	14	81
Overlap.coverage.per.analyst	55,093	8.294	6	7.142	1	2	13	73

(b) No. firms covered over an analyst's career

Statistic	N	Mean	Median	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
No.firms.covered.EPS.analyst.career	12,058	18.639	11	21.625	1	3	26	184
No.firms.covered.PTG.analyst.career	12,058	13.048	7	15.062	1	2	19	127
Total.No.months.analyst.career	12,058	75.905	47.417	78.798	0.000	17.500	109.933	459.267

Note: "Coverage.eps.forecasts.per.analyst" ("Coverage.price.targets.per.analyst") is the number of unique firms an analyst issues at least one EPS (price target) forecast in a given calendar year; "Overlap.firms.per.analyst" is the number of firms an analyst issues at least one EPS forecast and one price target forecast in a given calendar year. "No.firms.covered.EPS.analyst.career" (No.firms.covered.PTG.analyst.career) is the unique number of firms an analyst has issued at least one EPS (price target) forecast over his/her career; "Total.No.months.analyst.career" is the total number of months between the first and the last time an analyst ever issues an EPS/Price Target forecasts in the data base.

The analyst-level point-in-time coverage data is then merged with daily stock returns from CRSP as well as quarterly firm-level earnings data to obtain the analyst-level experienced returns and earnings variables used in the analysis.

Table A.12 shows that the sample summary statistics [13]. A typical analyst would cover 6 firms and has an career length of a bit more than 4 years, or 50 months. Both number of firms covered and the career length has a large standard deviation, amounting to 22 firms and 84 months, respectively.

Table A.12: Summary statistics analyst-level historical coverage

Statistic	N	Mean	Median	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
no.firms.covered	14,352	18.224	10	22.099	1	3	26	189
no.firms.w.eps	14,352	17.160	9	21.782	0	2	24	186
no.firms.w.ptg	14,352	12.663	6	15.745	1	2	18	168
no.firms.w.eps.only	14,352	5.562	1	13.042	0	0	5	170
no.firms.w.ptg.only	14,352	1.065	0	3.911	0	0	1	122
no.firms.w.both	14,352	11.598	5	15.133	0	1	16	167
no.months.analyst.career	14,352	82.932	51.567	85.366	0.000	15.633	126.067	462.433

Note: Analyst-level historical coverage data set for analysts who issue at least 1 price target (ptg) during the entire sample period from 1999-01-01 to 2020-02-01. Analyst-level detailed unadjusted EPS data set is from 1980-01-01 to 2020-02-02 and price target data set is from 1999-01-01 to 2020-02-01. "no.firms.covered" is the number of unique firms that an analyst issues at least one price target and/or eps forecasts for. "no.months.analyst.career" is the number of months from the first ever price target or eps to the last time the analyst issues a ptg or eps.

A.6 Simulating Return Expectations

To show that the model proposed has the ability to capture the key empirical moments for both return predictability as well as the heterogeneous return expectation dynamics, I conduct simulation exercises.

I first simulate 500 quarters of data based on the system from 1.3 to 1.7. Panel A in Figure A.7 shows the simulated data over time, which plots the predictor x_t as dividend

[13] In Appendix A.5, I document in more details how the data set is constructed

price ratio against true expected return and realized returns.

Panel B of Table A.7 demonstrate the rational of using realized returns. In this sample, when combining realized returns with the observable predictor through Kalman Filter, the projected return out-of-sample is a much more accurate estimate for true expected returns compared to simple predictive regression using x_t. Of course, this result is based on the correct prior beliefs on the covariance structure. In fact, this technique of using Kalman filter to improve return forecasts are empirically verified in Van Binsbergen and Koijen (2010) and Pástor and Stambaugh (2009).

Figure 1.2 shows subjective return expectations generated from simulated data. These are annual series from 150th quarter onward in the simulation. $\widetilde{ER}_t^{DR}$ denotes the (annualized) return expectation formed based on the prior that $\widetilde{W}_t^{\mu} = 0.9687$, $\widetilde{\beta}$=0.9 and $\widetilde{E_r} = 0.03$, where $\widetilde{W}_t^{\mu} = \frac{\kappa_\mu \sigma_\mu}{\sigma_{v,t}}$ is defined in Section 1.2.3; $\widetilde{ER}_t^{CF}$ denotes the (annualized) return expectation data based on the prior that $\widetilde{W}_t^{\mu} = 0.01$ $\widetilde{\beta}$=0.96 and $\widetilde{E_r} = 0.02$. The other parameters in these two expectation series are calibrated to match moments of actual historical data of dividend yield and realized returns.

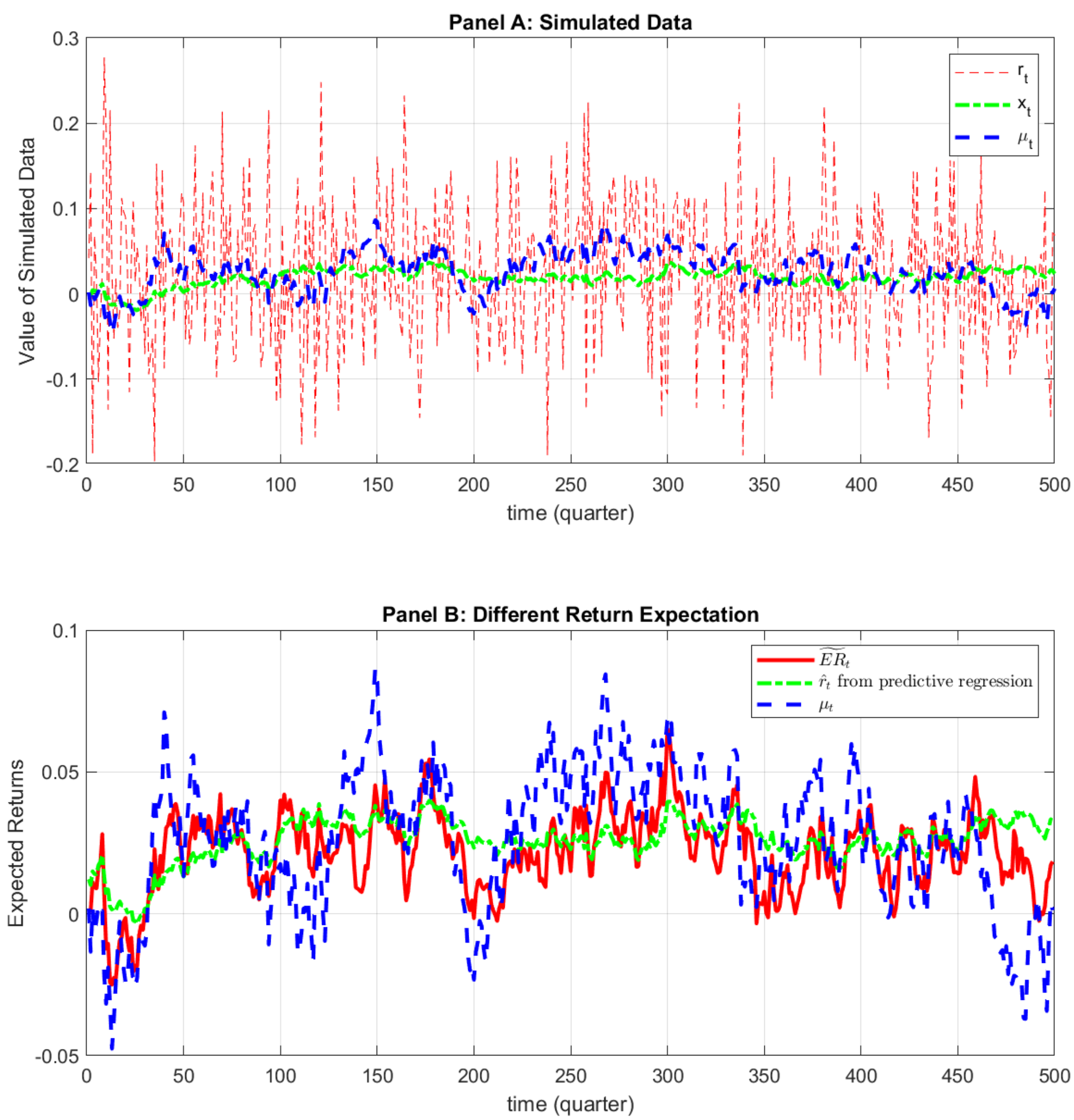
Figure 1.2 shows the model can match closely the heterogeneous return expectations graph 1.1 observed in the data, for a set of selected parameter values. As shown in the figure, even though the underlying data are the same, two forecasters can reach very different return expectation through the expectation formation model of 1.8, because of their different prior beliefs on how cash flows and discount rate interact and how important are cash flow process in driving asset prices.

A.7 Deriving Subjective Return Expectation Dynamics

I derive the expectation formation dynamics described in Equation (1.8). I further provide expressions for the steady-state parameters used in Section 1.2.3.

Investors infer the value of the unobservable expected returns μ_t based on information

Figure A.7: Returns, Predictors and Return Expectation in Simulated Data



Note: In Panel A, 500 quarters of data are simulated based on the system in 1.3 to 1.7. Observable parameters are calibrated based on actual data actual quarterly data on returns and dividend yield: $a = 0.97$, $Er = 0.02$, $Ex = 0.029$, $\sigma_r = 0.08$, $\rho_{rv} = \rho_{uv} = -0.89$. Additional parameters are chosen based on results from Pástor and Stambaugh (2009): $\beta = 0.9$, $\sigma_u = 0.78$, $\sigma_w = 0.0078$, ρ_{uw}=-0.71 and $\rho_{vw} = 0.5198$.

set $\mathcal{F}_t$ they observe from time 1 through time t

$$\mathcal{F}_t = (z_1, z_2, ..., z_t)$$
$$z_t = (r_t, x_t')'$$

Their prior belief, which is based on $\mathcal{F}_0$, is that the shocks follow multivariate normal:

$$\begin{bmatrix} u_{t+1} \\ v_{t+1} \\ \epsilon_{\mu,t+1} \end{bmatrix} |\mathcal{F}_0 \sim N \left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{uu} & \sigma_{uv} & \sigma_{u\mu} \\ \sigma_{uv} & \sigma_{vv} & \sigma_{v\mu} \\ \sigma_{u\mu} & \sigma_{v\mu} & \sigma_{\mu\mu} \end{pmatrix} \right] \tag{A.5}$$

where u_t and v_t are observable shocks to the returns and predictor vector x_t, respectively. $\epsilon_{\mu,t}$ is shock to the unobservable expected return process defined in Equation (1.5).[14] Since they also believe that the dynamics of expected returns, predictors follow (1.5) and (1.6), respectively, they consistently believe that

$$\begin{bmatrix} r_{t+1} \\ x_{t+1} \\ \mu_{t+1} \end{bmatrix} |\mathcal{F}_0 \sim N \left[\begin{pmatrix} E_r \\ E_x \\ E_r \end{pmatrix}, \begin{pmatrix} V_{rr} & V_{rx} & V_{r\mu} \\ V_{rx} & V_{xx} & V_{x\mu} \\ V_{r\mu} & V_{x\mu} & V_{\mu\mu} \end{pmatrix} \right] \tag{A.6}$$

where the parameters in the variance-covariance matrix is a function of the parameters in (A.5) and the persistent parameters.

The investors use Kalman Filter to form their expectations. For convenience, denote

$$a_t = \widetilde{E}(\mu_t|\mathcal{F}_{t-1}) \quad b_t = \widetilde{E}(\mu_t|\mathcal{F}_t) \quad f_t = \widetilde{E}(z_t|\mathcal{F}_{t-1})$$

as the conditional subjective expectations based on different information set. Furthermore,

[14] Notice here the shocks are following multivariate normal but the parameter values governing the variance-covariance matrix already are subject to investors' own prior beliefs.

let

$$P_t = \widetilde{Var}(\mu_t|\mathcal{F}_{t-1}), \quad Q_t = \widetilde{Var}(\mu_t|\mathcal{F}_t), \quad R_t = \widehat{Var}(z_t|\mu_t, D_{t-1})$$
$$S_t = \widehat{Var}(z_t|\mathcal{F}_{t-1}), \quad G_t = \widehat{Cov}(z, \mu_t|\mathcal{F}_{t-1})$$

be the subjective conditional variances that investors obtain after observing data.

Applying the updating algorithm of Kalman Filter,[15] we have

$$P_t = \beta^2 Q_{t-1} + \sigma_{\mu\mu} \tag{A.7}$$

$$S_t = \begin{pmatrix} Q_{t-1} + \sigma_{uu} & \sigma_{uv} \\ \sigma_{uv} & \sigma_{vv} \end{pmatrix} \tag{A.8}$$

$$G_t = \begin{pmatrix} \beta Q_{t-1} + \sigma_{u\mu} \\ \sigma_{v\mu} \end{pmatrix} \tag{A.9}$$

$$R_t = S_t - G_t P_t^{-1} G_t' \tag{A.10}$$

$$Q_t = P_t (P_t + G_t' R_t^{-1} G_t)^{-1} P_t \tag{A.11}$$

$$a_t = (1-\beta)E_r + \beta b_{t-1} \tag{A.12}$$

$$f_t = \begin{pmatrix} b_{t-1} \\ (I-A)E_x + Ax_{t-1} \end{pmatrix} \tag{A.13}$$

So the updated return expectation is

$$\begin{aligned} b_t &= a_t + G_t' S_t^{-1}(z_t - f_t) \\ &= a_t + m_t(r_t - b_{t-1}) + n_t' \left[x_t - \widetilde{E}_t(x_t|D_{t-1})\right] \\ &= a_t + m_t u_t + n_t' v_t \end{aligned} \tag{A.14}$$

[15]The internet appendix of Pástor and Stambaugh (2009) provides a similar derivation. See Durbin and Koopman (2012) for a more general treatment of Kalman filters and the state-space model in general.

and we arrive at Equation (1.8), since

$$\widetilde{ER}_{t|t} := b_t$$

and

$$\widetilde{ER}_{t|t-1} := a_t.$$

The parameters that govern the dynamics, m_t and n_t are dependent on the subjective prior beliefs of parameter values in (A.6). To see this, understand that the Kalman Filtering process from Equation (A.7) to (A.14) are recursive relations starting from $t = 2$, which depends on values of parameters in $t = 1$ when investors need to start from

$$\begin{aligned}
a_1 &= E_r \\
P_1 &= V_{\mu\mu} \\
f_1 &= E_z \\
S_1 &= V_{zz} \\
G_1 &= V_{z\mu} \\
R_1 &= S_1 - G_1 P_1^{-1} G_1' \\
Q_1 &= P_1 (P_1 + G_1' R_1^{-1} G_1)^{-1} P_1 \\
b_1 &= a_1 + G_1' S_1^{-1} (z_1 - f_1)
\end{aligned}$$

These values are based on the prior belief parameters and as analyzed in Section 1.2.3, not all of the parameters are identifiable through historical data, leaving room for heterogeneous expectation dynamics.

We know that

$$\begin{pmatrix} m_t & n_t' \end{pmatrix} = Cov(z_t', \mu_t | \mathcal{F}_{t-1}) Var(z_t | \mathcal{F}_{t-1})^{-1}$$

which is a function of Q_t defined in (A.11). The steady-state value of Q_t, computed from

(A.11), is

$$Q = \frac{\sqrt{\lambda_1^2 - 4\lambda_2} - \lambda_1}{2}$$

$$\lambda_1 = (1-\beta^2)Var(u|v) + 2\beta Cov(u, \epsilon_\mu|v) - Var(\epsilon_\mu|v)$$

$$\lambda_2 = Cov(u, \epsilon_\mu|v)^2 - Var(u|v)Var(\epsilon_\mu|v)$$

so the steady-state values of the m_t and n_t are

$$m = [\beta Q + Cov(u, \mu|v)]\,[Q + Var(u|v)]^{-1} \tag{A.15}$$

$$n = (\sigma_{\mu v} - m\sigma_{uv})\sigma_{vv}^{-1} \tag{A.16}$$

A.8 Understanding What Drives Differences in Return Expectations

I provide a more technical analysis to support the intuitive explanations in Section 1.2.3. Furthermore, I provide more details about how I make the plot of Figure 1.3.

The two prior beliefs can lead to either contrarian/extrapolative return expectations. To understand how, note that return expectations defined in (1.8) in the steady-state depend on the past returns through

$$\widetilde{m} = [\beta Q + Cov(u_t, \epsilon_{\mu,t}|v)] \,/ [Q + Var(u_t|v_t)]$$

where Q is the steady-state variance of $\tilde{E}(r_t|\mathcal{F}_t)$.[16] An investor will only appear to be contrarian if and only if $\widetilde{m} < 0$ or

$$Cov(u_t, \epsilon_{\mu,t}|v_t) < -\beta Q$$

[16]The expressions of Q is given in Appendix A.7

This condition is equivalent to

$$\rho_{d\mu} < -\frac{\beta Q}{\sigma_d \sigma_\mu} + \frac{\sigma_v - \rho_{uv}\sigma_u}{\sigma_d}\rho_{v\mu} \tag{A.17}$$
$$\approx -\frac{\beta Q}{\sigma_d \sigma_\mu} + 1.58\rho_{v\mu} \tag{A.18}$$

where the approximation (A.18) is due to the fact that investors will have a fairly accurate estimates from the data about $(\sigma_v - \rho_{uv}\sigma_u)/\sigma_d$.[17]

This condition (A.18) shows whether an investor appears contrarian depends largely on the value of $\rho_{v\mu}$: if $\rho_{v\mu}$ is very large and positive, investors will likely be a contrarian because a large set of values of $\rho_{d\mu}$ would lead to a negative $\widetilde{m}$.

Furthermore, based on the present value ration in Equation 1.13, the value of $\rho_{v\mu}$ is given by:[18]

$$\rho_{v\mu} = \frac{\kappa_\mu \sigma_\mu}{\sigma_v} - \rho_{g\mu}\frac{\kappa_g \sigma_g}{\sigma_v} \tag{A.19}$$

where σ_v is the volatility of the dividend price ratio. This condition shows that the value of $\rho_{v\mu}$ depends on 1). $W_\mu := \frac{\kappa_\mu \sigma_\mu}{\sigma_v}$, or how investors interpret the importance of discount rate news and 2). $\rho_{\mu,g}$: how they interpret expected cash flow news for future returns.

Notice that $W_\mu := \frac{\kappa_\mu \sigma_\mu}{\sigma_v}$ does not equal exactly to the conventional variance decomposition of dividend price ratio, or

$$V_{\mu,dp} = \frac{Var(\mu_t)}{Var(dp_t)}$$
$$= \frac{1-\phi^2}{1-\beta^2}\frac{1}{\Delta + \frac{1}{W_\mu^2}} \tag{A.20}$$

[17]This is because ρ_{uv} is easy to measure empirically so we have $\rho_{uv} \approx \rho_{r,dp} = -0.89$, which leads to $(\sigma_v - \rho_{uv}\sigma_u)/\sigma_d = 1.58$ and the last approximation obtains.

[18]Multiply both sides of the equation and taking expectation

where

$$\Delta := \frac{\kappa_\mu^2(\beta - \phi)^2}{1 - \beta^2}$$

the Equation (A.20) are obtained by taking unconditional variance on Equation (1.14) and (1.5) for the denominator and numerator, respectively.

Equation (A.20) shows that although W_μ and the $V_{\mu,dp}$ are not the same, W_μ does increase with $V_{\mu,dp}$ if the persistent parameters are held to be constant.

A.9 Estimating Expectation Formation Process

I estimate the system from (1.15) to (1.19) in three steps. First, I estimate the shocks $\epsilon_{\mu,t}$, $\epsilon_{g,t}$ and $\epsilon_{d,t}$ and parameters in the predictive system as captured in equation (1.15), (1.16) and (1.17) together with the parameters in the (1.2). To estimate this, I write the system into a State-Space Form and estimate the parameters using Kalman Filter based on maximum-llikelihood function. I write down the state-space form of the system in Section A.9.1. Second, the parameters in Equation (1.18) are estimated separately by Ordinary-Least-Squares together with the residuals. Finally, the correlations between innovations in predictors x_t and $\epsilon_{\mu,t}$, $\epsilon_{g,t}$ and $\epsilon_{d,t}$ are using these estimated series. [19]

A.9.1 A Simplified System With Return Expectations

$$\hat{r}_{t+1} = \frac{1}{L(\beta)}\hat{\mu}^A_{t+1} + \epsilon_{\Delta d,t+1} - \rho\kappa_\mu(\beta)\epsilon_{\mu,t+1} + \rho\kappa_g(\phi_g)\epsilon_{g,t+1} \tag{A.21}$$

$$\hat{dp}_{t+1} = \phi_g\hat{dp}_t + \frac{\kappa_\mu(\beta - \phi_g)}{L(\beta)}\hat{\mu}^A_{t+1} + \kappa_\mu\epsilon_{\mu,t+1} - \kappa_g\epsilon_{g,t+1} \tag{A.22}$$

$$\hat{\mu}^A_{t+1} = \beta\hat{\mu}^A_t + L(\beta)\epsilon_{\mu,t+1} \tag{A.23}$$

[19] These estimates are consistent estimates of the parameters. Potentially, I can use these estimates to re-estimate the entire system all together using Maximum-likelihood. The resulting estimates are similar to the three-step approach estimated here. Details of the estimation is provided in A.9.

where

$$L(\beta) = \sum_{k=0}^{3} \beta^k$$
$$\kappa_\mu(\beta) = \frac{1}{1-\rho\beta}$$
$$\kappa_g(\phi_g) = \frac{1}{1-\rho\phi_g}$$
$$B_{dp,\mu}(\beta, \phi_g) = \kappa_\mu(\beta - \phi_g)$$

and

$$\begin{pmatrix} \epsilon_{\Delta d,t+1} \\ \epsilon_{\mu,t+1} \\ \epsilon_{g,t+1} \end{pmatrix} \sim N \left[0, \begin{pmatrix} \sigma_d^2 & \sigma_{\mu d} & \sigma_{gd} \\ \sigma_{\mu d} & \sigma_\mu^2 & \sigma_{\mu g} \\ \sigma_{gd} & \sigma_{\mu g} & \sigma_g^2 \end{pmatrix} \right] \tag{A.24}$$

This system A.21 to A.24 is a linear system of underlying shocks and non-linear with respect to parameters

$$\theta = (\phi_g, \beta, \sigma_d, \sigma_\mu, \sigma_g, \sigma_{\mu d}, \sigma_{gd}, \sigma_{\mu g})'$$

Given the normality assumption in A.24, the system can be estimated by Maximum-likelihood Estimation. I write the system into state-space form.

The observable vector is

$$y_{t+1} = \begin{pmatrix} \hat{r}_{t+1} \\ \hat{dp}_{t+1} \\ \hat{\mu}^A_{t+1} \end{pmatrix}$$

and the latent processes

$$\alpha_t = \begin{pmatrix} \mu_{t+1} \\ \mu_t \\ \hat{dp}_{t+1} \\ \hat{dp}_t \\ \epsilon_{d,t+1} \\ \epsilon_{\mu,t+1} \\ \epsilon_{g,t+1} \end{pmatrix}$$

The dynamics of the measurement equations are captured by

$$\begin{pmatrix} \hat{r}_{t+1} \\ \hat{dp}_{t+1} \\ \hat{\mu}^A_{t+1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & \frac{1}{L(\beta)} & 0 & 0 & 1 & -\rho\kappa_\mu & \rho\kappa_g \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mu_{t+1} \\ \mu_t \\ \hat{dp}_{t+1} \\ \hat{dp}_t \\ \epsilon_{d,t+1} \\ \epsilon_{\mu,t+1} \\ \epsilon_{g,t+1} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$Z = \begin{pmatrix} 0 & \frac{1}{L(\beta)} & 0 & 0 & 1 & -\rho\kappa_\mu & \rho\kappa_g \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$d = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\eta_t = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$H = 0_{3\times 3}$$

and the state-equation is characterized by

$$\begin{pmatrix} \mu_{t+1} \\ \mu_t \\ \hat{dp}_{t+1} \\ \hat{dp}_t \\ \epsilon_{d,t+1} \\ \epsilon_{\mu,t+1} \\ \epsilon_{g,t+1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \beta & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{B_{dp,\mu}(\beta,\phi_g)}{L(\beta)} & 0 & \phi_g & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{\mu}_t^A \\ \hat{\mu}_{t-1}^A \\ dp_t \\ dp_{t-1} \\ \epsilon_{d,t} \\ \epsilon_{\mu,t} \\ \epsilon_{g,t} \end{pmatrix} + \begin{pmatrix} 0 & L(\beta) & 0 \\ 0 & 0 & 0 \\ 0 & \kappa_\mu(\beta) & -\kappa_g(\phi_g) \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_{d,t+1} \\ \epsilon_{\mu,t+1} \\ \epsilon_{g,t+1} \end{pmatrix}$$

$$T = \begin{pmatrix} \beta & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{B_{dp,\mu}(\beta,\phi_g)}{L(\beta)} & 0 & \phi_g & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$c = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

and

$$\epsilon_t = \begin{pmatrix} \epsilon_{d,t+1} \\ \epsilon_{\mu,t+1} \\ \epsilon_{g,t+1} \end{pmatrix}$$

$$R = \begin{pmatrix} 0 & L(\beta) & 0 \\ 0 & 0 & 0 \\ 0 & \kappa_\mu(\beta) & -\kappa_g(\phi_g) \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$Q = \begin{pmatrix} \sigma_d^2 & \sigma_{\mu d} & \sigma_{gd} \\ \sigma_{\mu d} & \sigma_\mu^2 & \sigma_{\mu g} \\ \sigma_{gd} & \sigma_{\mu g} & \sigma_g^2 \end{pmatrix}$$

where Σ_x has been pre-estimated using the OLS residuals. The system has an initial state vector

$$\alpha_{1|0} = (I_7 - T)^{-1} c$$

$$vec(P_{1|0}) = (I_{49} - T \otimes T)^{-1} vec(RQR')$$

A.9.2 A Simplified System With Dividend Expectations

$$\Delta \hat{d}_{t+1} = \frac{1}{L(\phi_g)} \hat{g}^A_{t+1} + \epsilon_{\Delta d,t+1} \tag{A.25}$$

$$\hat{dp}_{t+1} = \beta \hat{dp}_t + \frac{B_{dp,g}(\beta, \phi_g)}{L(\phi_g)} \hat{g}^A_{t+1} + \kappa_\mu \epsilon_{\mu,t+1} - \kappa_g \epsilon_{g,t+1} \tag{A.26}$$

$$\hat{g}^A_{t+1} = \phi_g \hat{g}^A_t + L(\phi_g) \epsilon_{g,t+1} \tag{A.27}$$

where

$$
\begin{aligned}
L(\phi_g) &= \sum_{k=0}^{3} \phi_g^k \\
\kappa_\mu(\beta) &= \frac{1}{1-\rho\beta} \\
\kappa_g(\phi_g) &= \frac{1}{1-\rho\phi_g} \\
B_{dp,\mu}(\beta,\phi_g) &= \kappa_g(\beta-\phi_g)
\end{aligned}
$$

and

$$
\begin{pmatrix} \epsilon_{\Delta d,t+1} \\ \epsilon_{\mu,t+1} \\ \epsilon_{g,t+1} \end{pmatrix} \sim N\left[0, \begin{pmatrix} \sigma_d^2 & \sigma_{\mu d} & \sigma_{gd} \\ \sigma_{\mu d} & \sigma_\mu^2 & \sigma_{\mu g} \\ \sigma_{gd} & \sigma_{\mu g} & \sigma_g^2 \end{pmatrix}\right] \tag{A.28}
$$

This system A.25 to A.28 is a linear system of underlying shocks and non-linear with respect to parameters

$$
\theta = (\phi_g, \beta, \sigma_d, \sigma_\mu, \sigma_g, \sigma_{\mu d}, \sigma_{gd}, \sigma_{\mu g})'
$$

Given the normality assumption in A.28, the system can be estimated by Maximum-likelihood Estimation. I write the system into state-space form.

The observable vector is

$$
y_{t+1} = \begin{pmatrix} \Delta\hat{d}_{t+1} \\ \hat{dp}_{t+1} \\ \hat{g}^A_{t+1} \end{pmatrix}
$$

and the latent processes

$$\alpha_t = \begin{pmatrix} \hat{g}^A_{t+1} \\ \hat{g}^A_t \\ \hat{dp}_{t+1} \\ \hat{dp}_t \\ \epsilon_{d,t+1} \\ \epsilon_{\mu,t+1} \\ \epsilon_{g,t+1} \end{pmatrix}$$

The dynamics of the measurement equations are captured by

$$\begin{pmatrix} \Delta\hat{d}_{t+1} \\ \hat{dp}_{t+1} \\ \hat{g}^A_{t+1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & \frac{1}{L(\phi_g)} & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{g}^A_{t+1} \\ \hat{g}^A_t \\ \hat{dp}_{t+1} \\ \hat{dp}_t \\ \epsilon_{d,t+1} \\ \epsilon_{\mu,t+1} \\ \epsilon_{g,t+1} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$Z = \begin{pmatrix} 0 & \frac{1}{L(\phi_g)} & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$d = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\eta_t = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$H = 0_{3\times 3}$$

and the state-equation is characterized by

$$\begin{pmatrix} \hat{g}^A_{t+1} \\ \hat{g}^A_t \\ \hat{dp}_{t+1} \\ \hat{dp}_t \\ \epsilon_{d,t+1} \\ \epsilon_{\mu,t+1} \\ \epsilon_{g,t+1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \phi_g & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{\kappa_g(\beta-\phi_g)}{L(\phi_g)} & 0 & \beta & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{g}^A_t \\ \hat{g}^A_{t-1} \\ \hat{dp}_t \\ \hat{dp}_{t-1} \\ \epsilon_{d,t} \\ \epsilon_{\mu,t} \\ \epsilon_{g,t} \end{pmatrix} + \begin{pmatrix} 0 & 0 & L(\phi_g) \\ 0 & 0 & 0 \\ 0 & \kappa_\mu(\beta) & -\kappa_g(\phi_g) \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \epsilon_{d,t+1} \\ \epsilon_{\mu,t+1} \\ \epsilon_{g,t+1} \end{pmatrix}$$

$$T = \begin{pmatrix} \phi_g & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{\kappa_g(\beta-\phi_g)}{L(\phi_g)} & 0 & \beta & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$c = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

and

$$\epsilon_t = \begin{pmatrix} \epsilon_{d,t+1} \\ \epsilon_{\mu,t+1} \\ \epsilon_{g,t+1} \end{pmatrix}$$

$$R = \begin{pmatrix} 0 & 0 & L(\phi_g) \\ 0 & 0 & 0 \\ 0 & \kappa_\mu(\beta) & -\kappa_g(\phi_g) \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$Q = \begin{pmatrix} \sigma_d^2 & \sigma_{\mu d} & \sigma_{gd} \\ \sigma_{\mu d} & \sigma_\mu^2 & \sigma_{\mu g} \\ \sigma_{gd} & \sigma_{\mu g} & \sigma_g^2 \end{pmatrix}$$

where Σ_x has been pre-estimated using the OLS residuals. The system has an initial state vector

$$\alpha_{1|0} = (I_7 - T)^{-1} c$$

$$vec(P_{1|0}) = (I_{49} - T \otimes T)^{-1} vec(RQR')$$

A.10 Consistency in Analysts' Return and Cash Flow Expectations

In the estimation framework presented in the previous section, the cash flow expectation is backed out through the present value relation based on return expectation and valuation ratio similar to the the VAR framework developed by Campbell (1991). Inherently, this framework assumes that subjective return and cash flow expectations of market participants are consistent with the present value model and the mean-reverting expectation dynamics. To see if analysts' own expectations are indeed consistent, I compare analysts' own cash flow expectations, which is observable, with the implied cash flow expectations based on return expectations.

In summary, I find that directly observed analysts' cash flow expectations are broadly consistent with the implied cash flow expectations based on return expectations and price fundamental ratios, although there exists finer nuances in the observed expectation dynamics directly reported by analysts.

Table A.13 shows the correlations between various directly observed cash flow expectations (shocks) with implied expectations (shocks). First, implied cash flow expectations (shocks) are strongly, although imperfectly correlated with analysts' own cash flow expectations (shocks). In particular, as shown in Table A.13a and A.13b, the implied dividend growth expectations and analysts' earnings growth expectations are 82% correlated and the shocks are about 50% correlated. The imperfect correlation could potentially be due to three reasons: first, measurement errors on the expectations series; second, the expectations on returns or cash flows are more complicated than a simple AR(1) process, for example, it contains a term structure of cash flow expectations; third, analysts' expectations processes are not perfectly consistent. As a result of the imperfect correlation, the correlation between the shocks on return expectations and cash flow expectations are less strong than estimated through the model, as shown in Table A.13b. Finally, the multi-variable regression results in Table A.13c shows that the implied cash flow expectations are correlated by more than one observed cash flow expectation measure, including earnings and dividend expectations.

Table A.14a shows the variance decomposition based on directly used cash flow and return expectations measures by regressing the expectation measures on the dividend price ratio. The coefficients on the dividend price ratio from the regression can be interpreted directly as the variance explained by the expectation measure. Despite the implied cash flow growth expectations being imperfectly correlated with a particular observed cash flow expectations, the results show that cash flow expectations alone explains 99% of price dividend ratio variation, while return expectations only explain about 10%. These results are

Table A.13: Correlations Between Implied Dividend Growth Expectation and Analysts' Cash Flow Expectations

(a) Pair-wise correlations between implied and directly reported cash flow expectations

	g	Analyst.Div.Growth	Analyst.E.Growth	Analyst.Payout	LTG	Actual.Div.Growth
g	1.00					
Analyst.Div.Growth	0.79***	1.00				
Analyst.E.Growth	0.82***	0.81***	1.00			
Analyst.Payout	-0.03	-0.06	-0.20	1.00		
LTG	0.51***	0.46***	0.54***	0.11	1.00	
Actual.Div.Growth	0.67***	0.82***	0.73***	-0.30**	0.37***	1.00

(b) Pair-wise correlations between implied and directly reported cash flow expectation shocks

	implied.shock.ER	implied.shock.g	implied.shock.d	shock.analyst.earnings	shock.analyst.div.growth	shock.analyst.LTG
implied.shock.ER	1.00					
implied.shock.g	-0.54***	1.00				
implied.shock.d	-0.38***	-0.11	1.00			
shock.analyst.earnings	-0.04	0.50***	-0.24*	1.00		
shock.analyst.div.growth	-0.15	0.47***	-0.06	0.80***	1.00	
shock.analyst.LTG	-0.19	0.36***	0.00	0.51***	0.51***	1.00

(c) Linear Regressions With Multiple Regressors

	Dependent variable:		
	g		
	(1)	(2)	(3)
Analyst.Div.Growth	0.618***		0.268**
	(0.079)		(0.132)
Analyst.E.Growth		0.738***	0.479***
		(0.158)	(0.125)
Analyst.Payout		0.312	0.223
		(0.195)	(0.216)
LTG		0.496	0.477
		(1.852)	(1.398)
Constant	−0.073***	−0.252	−0.219
	(0.025)	(0.190)	(0.190)
Observations	62	62	62
R^2	0.626	0.689	0.728
Adjusted R^2	0.619	0.673	0.708
Residual Std. Error	0.076 (df = 60)	0.070 (df = 58)	0.066 (df = 57)
F Statistic	100.284*** (df = 1; 60)	42.768*** (df = 3; 58)	38.058*** (df = 4; 57)
Note:			*p<0.1; **p<0.05; ***p<0.01

Note: g is the implied cash flow expectation from the model estimation. All statistics are calculated from the sample between 2003-Q1 to 2019-Q4, where all series have data available.

consistent with the findings presented in the previous section, confirming the model are broadly consistent with the data. The small and insignificant coefficient on analyst's long-term growth expectation is also interesting, which means that most of the variation of price dividend ratios, from analyst's own perspective, is due to short-term cash flow expectations, consistent with the findings of O and Myers (2020). In their framework, O and Myers (2020), however, they use the CFO expectations as the measure for return expectations. As shown in the right-most column of Table A.14a, such assumption can result in different conclusions regarding how return expectations are related to prices and cash flow expectations, which is demonstrated by the different signs of the regression coefficients between analyst and CFO return expectations.

Finally, Table A.14b shows that variance decomposition using the subjective expectations based on directly observed cash flow expectations are very different from that using the VAR framework based on dividend price ratio employed in Cochrane (2011a). The implied long-run coefficient on log(D/P) in the return equation from this framework is 0.678, which means that based on the same sample period, an econometrician would conclude that discount rate variation is the main driver behind the variation of price dividend ratio, instead of short-term cash flow growth. These results highlight the difference between the subjective expectations and objective expectation, when making inference on variance decomposition of fundamental to price ratios.

Overall, the results in this section show that analysts' expectations about future returns and cash flows are broadly consistent with the simple present value model described above. However, the reported expectation data might have nuances that is not captured by the simple mean-reverting process that worth further research. Potentially, there might be a term structure on both return and cash flow expectations that are not considered in the current model. However, this is outside the scope of the current paper so I leave to future researchers.

Table A.14: Variance Decomposition of Log Dividend Price Ratios

(a) Variance Decomposition Subjective Expectations Analysts

	Dependent variable:			
	Div.Growth.1y	LTG	ER.Analyst	ER.CFO
	(1)	(2)	(3)	(4)
log(DP)	0.992***	0.049	0.098***	−0.063***
	(0.214)	(0.031)	(0.032)	(0.021)
Adjusted R^2	0.615	0.209	0.183	0.403
Note:				*p<0.1; **p<0.05; ***p<0.01

(b) Variance Decomposition VAR as in Cochrane (2011a)

	Dependent variable:	
	Next.Year.Excess.Ret	Next.Qtr.Log(DP)
	(1)	(2)
Log(DP)	0.362***	0.839***
	(0.128)	(0.101)
Implied Long Run Coefficent	0.678	
Adjusted R^2	0.088	0.700
Note:		*p<0.1; **p<0.05; ***p<0.01

Appendix B: Appendix for Chapter 2: "Asset Prices When Investors Ignore Discount Rate Dynamics"

B.1 The Return Expectation Biases Due to CDR Assumption

I demonstrate why assuming a constant discount rate could result in a bias in return expectations and analyze how the bias is related to firm-level characteristics.

B.1.1 The Setup

We start by considering a general discounted cash flow model with potentially time-varying discount rates and expected cash flows. Let V_0 be the be the value of an equity that pays c_t, $t = 1, 2, ..., \infty$ into the future. Further denote M_t as the expected return, or discount rate known at the beginning of period t, for the cash flow to be paid on $t+1$. For the convenience of exposition, let $\mu_t = log(M_t)$. We have

$$\begin{aligned} V_0 &= E_0 \left[\sum_{t=0}^{\infty} (\prod_{s=0}^{t} e^{-\mu_s}) c_{t+1} \right] \\ &= \sum_{t=0}^{\infty} E_0 \left[(\prod_{s=0}^{t} e^{-\mu_s}) c_{t+1} \right] \end{aligned} \tag{B.1}$$

If one ignores the dynamics of discount rate and instead assumes a constant discount rate, $\pi_t = log(\Pi_t)$, he/she would valuate the stock using

$$\begin{aligned} \widetilde{V}_0 &= E_0 \left[\sum_{t=0}^{\infty} e^{-t\pi_0} c_{t+1} \right] \\ &= \sum_{t=0}^{\infty} \left[e^{-t\pi_0} E_0 \left(c_{t+1} \right) \right] \end{aligned} \tag{B.2}$$

Equation (B.2) represents the valuation formula taught in a typical undergraduate or MBA class: first project future cash flows to obtain $E_0\left(c_{t+1}\right)$ and subsequently apply a discount rates, either using weighted average cost of capital (WACC) or a CAPM model to obtain a value for π_t. This valuation formula is also a log version of the commonly used Discounted Cash Flow Models (DCF) as in popular valuation textbooks, such as Damodaran (2012).

To understand the implication of the constant discount rate assumption more precisely, I follow Hughes, Liu, and Liu (2009) to assume the dynamics of the discount rates μ_t and c_t:

$$\mu_t = r_f + \beta_t \lambda \tag{B.3}$$

$$\beta_t = \bar{\beta} + \sigma_\beta \epsilon_{\beta,t} \tag{B.4}$$

$$c_{t+1} = c_t exp\left[g + \sigma_c(\rho\epsilon_{\beta,t+1} + \sqrt{1-\rho^2}\epsilon_{c,t+1})\right] \tag{B.5}$$

$$\begin{pmatrix} \epsilon_{\beta,t} \\ \epsilon_{c,t+1} \end{pmatrix} \sim N\left(0, \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)$$

The discount rate dynamics specified in (B.3) is different from a constant discount rate for two reasons. First, it has own volatility, which leads to uncertainty in prices with respect to discount rate itself. More specifically, this dynamics in this particular specification is due to the conditional β_t [1]. The volatility of the risk premium is therefore

$$\sigma_\mu := \lambda \sigma_\beta.$$

This specification is consistent with a version of the conditional CAPM model[2]. If investors ignore the dynamics of β_t and instead use a static CAPM, he/she would use $\pi_t = \bar{\beta}\lambda$ instead. Second, the discount rate is correlated with stochastic discount cash flows, which impacts the prices through ρ. [3].

[1]The analytical difference to be presented is invariant if the risk premium λ is stochastic, as shown in Hughes, Liu, and Liu (2009).

[2]See for example Jagannathan and Wang (1996) for conditional CAPM.

[3]Notice that here the discount rates have a simple term-structure. In the case of discount rates have a term structure. μ_t follow the same dynamics with respect to the horizon t. This makes the analysis less complicated. In a fully specified model such as those of Ang and Liu (2004), discount rates also have a potential term structure. Besides, the discount rate shocks are i.i.d., Ang and Liu (2004) also consider cases

The cash flow process, as specified in Equation (B.5), has a constant growth g and an interaction with the discount rates, through ρ. Notice that the cash flow shocks are permanent growth shocks, while the discount rate shocks are temporary. This is reasonable as in general cash flows grow at a positive rate in the long run while discount rate should be stationary in the long run. The specifications also means that we can interpret a firm's discount rate volatility as mainly due to its systematic risk through β_t, while the cash flow shocks are mostly idiosyncratic, which drives most of the idiosyncratic volatility in stock returns.

B.1.2 The Biases

Under the specifications in Equation (B.3) through to (B.5), the rational "fair-value" of the equity should be valued, according to (B.1) at

$$V_0 = c_0 \frac{exp(g + \frac{1}{2}\sigma_c^2)}{exp(\mu_0)\left\{1 - exp\left[-\left(r_f + \lambda\bar{\beta} - g\right) - \frac{1}{2}(\rho\sigma_c - \sigma_\mu)^2 - \frac{1}{2}(1-\rho^2)\sigma_c^2\right]\right\}} \tag{B.6}$$

On the other hand, an investor who values the stock using a constant discount rate, or (B.2), would arrive at

$$\widetilde{V}_0 = c_0 \frac{exp(g + \frac{1}{2}\sigma_c^2)}{exp(\pi_0) - exp(g + \frac{1}{2}\sigma_c^2)} \tag{B.7}$$

To understand the impact of dynamic discount rate in valuating a stock, considering the case where $\mu_t = \bar{\mu}$, Equation B.6 becomes the familiar Gordon-Growth formula with uncertain cash flows

$$A_0 = \frac{c_0}{exp(\mu_0 - g - \frac{1}{2}\sigma_c) - 1} \tag{B.8}$$

which makes clear the impact of discount rates being stochastic: it adds the volatility of the discount rates, σ_μ and the interaction between discount rates and cash flow ρ into the

where the discount rate and cash flow processes are persistent. For the sake of simplicity, the current paper only uses the simple case with analytical solutions.

valuation formula.

By equating Equation (B.6) and (B.7), we have the relationship between the two expected returns:

$$\Pi_0 = M_0 - exp\left\{\mu_0 - \left[\left(r_f + \lambda\bar{\beta} - g\right) - \frac{1}{2}(\rho\sigma_c - \sigma_\mu^2) - \frac{1}{2}(1-\rho^2)\sigma_c^2\right]\right\} + exp(g + \frac{1}{2}\sigma_c^2) \tag{B.9}$$

The equation also implies that Π_0 will equal M_0 if $\mu_0 = \bar{\mu}$. The formula provides an analytical expression for the bias, b_t,

$$b_t^i = -M_t^i e^{-\Delta^i} + exp(g^i + \frac{1}{2}(\sigma_c^i)^2)$$

where

$$\Delta^i = \left(r_f + \lambda\bar{\beta}^i - g^i\right) - \frac{1}{2}(\rho\sigma_c^i - (\sigma_\mu^i)^2) - \frac{1}{2}(1 - (\rho^i)^2)(\sigma_\mu^i)^2$$

and I added the superscript i, which runs across different firms, to stress that the biases are related to the characteristics of different firms. The biases are dynamic because it depends on the realization of the discount rate M_t^i. Besides, the bias is related to firm-level fundamental characteristics, such as expected growth and volatility of the growth. Confirming our intuition, the bias is higher for firms with higher growth rates and uncertainty.

We are more interested in the unconditional expectation of the bias, which is given by

$$b^i = E(b_t^i) = \delta^i exp(g^i + \frac{1}{2}(\sigma_c^i)^2) \tag{B.10}$$

where

$$\delta^i = 1 - exp\left[\sigma_\mu^i(\sigma_\mu^i - \rho^i\sigma_c^i)\right] \tag{B.11}$$

B.1.3 A Discussion: The Sign and Magnitude of the Biases

Equation (B.10) relates the bias to firm-specific characteristics, and therefore will have implication for the cross-section. However, is the channel suggested by the CDR important enough to have any impact on the cross-section of stock returns and is it plausible to explain any cross-sectional anomalies? I discuss the plausibility of the channel based on empirical findings in the literature before examining the data in Section (2.2.2).

Analytically, the relationship between b^i and characteristics depends on the sign of δ^i. In the case that $\delta^i > 0$ (< 0), we have $b^i > 0$ (<0). Furthermore, b^i depends on fundamentals of the firm such as expected growth rates g^i and σ_c^i.

The sign of δ^i potentially differs on the market and firm-level. On the market level, δ^i has been shown to be negative, leading to a negative market-level bias b^m and the magnitude of the bias has been estimated in the literature. This negative bias is mainly due to the fact that discount rate dominates (for example Cochrane (2011b)) on the market level and that aggregate cash flows and discount rate news are weakly negatively correlated (for example Lochstoer and Tetlock (2020), Campbell (1990)). In fact, the negative b^m is directly supported by the empirical literature on implied cost of capital, which assumes a constant expected returns in the model. Claus and Thomas (2001) and Pástor, Sinha, and Swaminathan (2008) estimates that the market level implied risk premium (Π_t^m-R^f) is around 3% or less using the constant discount rate assumption, significantly less than estimates of market premium, which are typically above 5%.[4] Hughes, Liu, and Liu (2009) calibrates the magnitude of b^m and shows that the magnitude is at -2.3%.[5] Given the robustness of these empirical findings, in my empirical tests, I directly use $b^m = -0.023$ and show my results are not sensitive to different choice of b^m.

[4] Avdis and Wachter (2017) is the latest in the literature estimating the market risk premium. Their estimate for the market risk premium in the U.S. is at 5.1%, lower than the ones in the previous literature, which typically was above 6%. However, this estimate is still more than 2% above those based on the constant discount rate assumption.

[5] Their calibration is based on the parameter sets of $\sigma_c = 0.15$, $g = 0.05$, $\sigma_\mu = 0.14$ and $\rho = -0.1$, which translates into a $b^m = -0.023$

On the firm level, which is the focus of this paper, the sign of δ^i is likely positive, although the magnitude is unclear. This is because cash flow news likely dominates on the firm level and the shocks between discount rate and cash flows are positive, as shown in Vuolteenaho (2002) and Cohen, Polk, and Vuolteenaho (2009), for example. Compare to the market level evidence, no direct estimates are provided in the literature for the firm level biases δ^i.[6] Therefore, I empirically verify the sign and magnitude of different components of b^i in Section (2.2.2).

Depending on the sign of δ^i on the firm level, the bias relates to a firm's fundamental characteristics, g^i and σ_c^i. In the case of a positive δ^i, firms with higher future cash flow growth and/or cash flow volatility have higher biases.

B.2 Biased Return Expectations and Equilibrium Asset prices: A More Formal Analysis

I study a multi-asset economy in which some investors with biased return expectations (CDR investors) trade with risk-averse rational investors (arbitrageurs). CDR investors take up $\theta \in (0,1)$ share of the economy, so arbitrageurs are left with $1-\theta$. Both of these investors live for two periods and they invest in the first period into the risky securities and risk-free rate r_f to maximize their terminal wealth. There are N risky assets, each of which pays a dividend of D_t^i for asset i in the next period. The number of shares outstanding of these risky assets are $x^* = (x^1, x^2, ..., x^N)'$ and risk free assets are in unlimited supply.

Both CDR investors and arbitrageurs have the same utility function with the same risk-aversion coefficient, γ. The key difference is that the CDR investors have a subjective return expectations, $\widetilde{E}(.)$ that are biased, or

$$\tilde{E}_t(R_{t+1}^i) = E_t(R_{t+1}^i) + b_t^i \tag{B.12}$$

[6]Gode and Mohanram (2003) regress firm-level cost of capital on firm characeteristics and found the estimated implied cost of capital are positively related to analysts LTG estimates, leverage, earnings volatility, using data from 1984-1998. They consider three different ICC models. Their results support a positive firm-level δ^i.

In particular, the CDR investors solve the problem

$$\max_{\omega} \sum_{i=1}^{N} \omega^i P_t^i \left[\tilde{E}_t(R_{t+1}^i) - R_f \right] - \frac{\gamma}{2} \omega' \Sigma_t \omega \tag{B.13}$$

where

$$R_{t+1}^i = \frac{P_{t+1}^i + D_{t+1}^i}{P_t^i}$$

While the arbitrageurs solve the problem

$$\max_{y} \sum_{i=1}^{N} y^i P_t^i \left[E_t(R_{t+1}^i) - R_f \right] - \frac{\gamma}{2} y' \Sigma_t y \tag{B.14}$$

Denote $\omega^* = (\omega^1, \omega^2, ..., \omega^N)'$ and $y^* = (y^1, y^2, ..., y^N)'$ the optimal demand of the CDR investors and the arbitrageurs, respectively. Market clears and we have

$$\theta \omega^* + (1 - \theta) y^* = x^* \tag{B.15}$$

The equilibrium asset prices and expected returns are outlined in Proposition B.1.

Proposition B.1. *In the multi-asset economy featuring biased investors and arbitrageurs, whose return expectations are governed by Equation* (B.12)*, and solve optimization problems in* (B.13) *and* (B.14)*, respectively. With market clearing conditions* (B.15)*, the equilibrium asset price for asset i is*

$$P_t^i = \frac{1}{1 + R_f - \theta b_t^i} \left[E_t(P_{t+1}^i + D_{t+1}^i) - \gamma e^{i'} \Sigma_t x^* \right] \tag{B.16}$$

where e^i is a vector of zeros with 1 on the i^{th} entry. The expected return of asset i is

$$E_t(R_{t+1}^i) - R_f = \theta(-b_t^i + \beta_t^i b_t^M) + \beta_t^i \left[E_t(R_{t+1}^M) - R_f \right] \tag{B.17}$$

where $b_t^M = \sum_{i=1}^{N} \frac{x^i P^i}{\sum_j x^j P^j} b_t^i$ is the market-level expectation bias of CDR investors, $\beta_t^i =$

$\frac{Cov_t(R_t^i, R_t^M)}{Var_t(R_t^M)}$ *the CAPM-beta in its usual definition and* $R_t^M = \sum_{i=1}^N \frac{x^i P^i}{\sum_j x^j P^j} R_t^i$ *is the value-weighted market-returns.*

Proof. See Appendix B.3. □

Results in Proposition B.1 confirm the earlier intuition about how biases in the return expectation could cause mispricing in equilibrium. As shown in Equation (B.16), the more CDR investors in the economy, i.e. the higher value of θ, the more serious the mispricing potentially becomes. Furthermore, fixing the share of CDR investor, the higher the bias the CDR investors have for the return expectation of an asset, the higher its price and the lower its expected return, as shown in Equation (B.17). This is intuitive as the CDR investors will demand more of such an asset, leading to a lower expected returns.

Equation (B.17) reveals that the return expectation bias on the asset level as well as market level together contribute to the non-zero CAPM-alpha. This is intuitive as the CDR investors irrational demand on the asset level would also lead to an equilibrium impact on the market level.

B.3 A Proof of Proposition B.1

Solving the first-order-condition of (B.13) and (B.14), we have the optimal demands given by

$$\omega^* = \frac{1}{\gamma}\Sigma_t^{-1}\left[E(P_{t+1} + D_{t+1}) + B_t P_t - P_t(1 + R_f)\right] \tag{B.18}$$

where B_t is a diagonal matrix with biases b_t^i being on the i^{th} row and i^{th} column, and

$$y^* = \frac{1}{\gamma}\Sigma_t^{-1}\left[E(P_{t+1} + D_{t+1}) - P_t(1 + R_f)\right] \tag{B.19}$$

respectively.

Market clearing conditions imply that

$$\theta\omega^* + (1-\theta)y^* = x^*$$

or

$$\theta\frac{1}{\gamma}\Sigma_t^{-1}\left[E(P_{t+1}+D_{t+1}) + B_tP_t - P_t(1+R_f)\right] + (1-\theta)\frac{1}{\gamma}\Sigma_t^{-1}\left[E_t(P_{t+1}+D_{t+1}) - P_t(1+R_f)\right] = x^*$$
$$\theta B_tP_t + E_t(P_{t+1}+D_{t+1}) - P_t(1+R_f) = \gamma\Sigma x^*$$

$$\left[(1+R_f)I - \theta B_t\right]P_t = E_t(P_{t+1}+D_{t+1}) - \gamma\Sigma_t x^*$$

which leads to

$$P_t^i = \frac{1}{1+R_f-\theta b^i}\left[E_t(P_{t+1}^i + D_{t+1}^i) - \gamma e^{i'}\Sigma_t x^*\right]$$

which is Equation (B.16) of Proposition (B.1).

The expected returns follow

$$\begin{aligned}
E_t(R_{t+1}^i) - R_f &= -\theta b_t^i + \gamma\frac{1}{P_t^i}e_i'\Sigma_t x^* \\
&= -\theta b_t^i + \gamma\frac{1}{P_t^i}e_i'Cov_t(P_{t+1}+D_{t+1}, P_{t+1}+D_{t+1})x^* \\
&= -\theta b_t^i + \gamma Cov_t(R_{t+1}^i, P_{t+1}+D_{t+1})x^* \\
&= -\theta b_t^i + \gamma Cov_t(R_{t+1}^i, (P_{t+1}+D_{t+1})'x^*) \\
&= -\theta b_t^i + \gamma Cov_t(R_{t+1}^i, R_{t+1}^M)P_t'x^* \qquad \text{(B.20)}
\end{aligned}$$

Now define the market-cap weight for asset i as

$$\omega_M^i = \frac{x^iP_t^i}{\sum_j x^jP_t^j}$$

and pre-multiply Equation (B.20) by the weights and sum over different assets to obtain

$$R^M_{t+1} - R_f = -\theta b^M_t + \gamma Var_t(R^M_{t+1}) P'_t x^*$$

which gives

$$\begin{aligned} \gamma Var_t(R^M_{t+1}) P'_t x^* &= R^M_{t+1} - R_f + \theta b^M_t \\ P'_t x^* &= \frac{E_t(R^M_{t+1} - R_f)}{\gamma Var_t(R^M_{t+1})} \end{aligned} \tag{B.21}$$

Substituting Equation (B.21) into (B.20) and we have

$$\begin{aligned} E_t(R^i_{t+1}) - R_f &= -\theta b^i_t + \gamma Cov_t(R^i_{t+1}, R^M_{t+1}) \frac{E_t(R^M_{t+1} - R_f)}{\gamma Var_t(R^M_{t+1})} \\ &= \theta(-b^i_t + \beta^i_t b^M_t) + \beta^i_t \left[E_t(R^M_{t+1}) - R_f \right] \end{aligned}$$

the last equation is the Equation (B.17) in Proposition (B.1).

B.4 Detailed Data Descriptions

In sum, the estimation of firm-level equity requires 5 firm-level variables, 1 industry-level variable and 1 aggregate variable. The firm-level variables are:3 analyst's consensus forecasts for a firm's earnings of current fiscal year (FY1), the next fiscal year (FY2) and the fiscal year thereafter (FY3); 1 analyst's consensus long-term forecast (LTG); 1 payout ratio, which is based on the firm's previous year total dividend to firm's net income. The industry-level variable is the average LTG based on 48 Fama-French industry classification. The aggregate variable is the long-term average of gdp growth, which goes down from 7% to 6% over the 35 years in the sample. Based on these 5 inputs, I compute the implied cost of capital $q_{i,t}$ and the entire term structure of a firm's payout ratio $PB_{i,t+s}$ based on (B.22), which is a function of the last year's payout ratio and aggregate gdp growth rate and the $q_{i,t}$.

In the IBES monthly summary history file, I use analyst EPS estimates for fiscal year 1, fiscal year 2 and fiscal year 3 (fpi = 1, 2, 3) and the long-term growth estimates to take full advantage of the term structure of analyst forecast[7]. Furthermore, I require both fiscal year 1 and fiscal year 2 consensus to be based on no less than 3 analyst estimates available and at least 2 estimates for FY3 estimates[8] in order to be included in the sample. I only use the latest monthly consensus estimates within each calendar quarter: March, June, September and December to obtain firm-quarter consensus estimates. In addition, the firms included in the sample need to be a US firm whose reporting currency is in US dollars. For the base case, I consider the median estimates as the consensus estimate, but my results do not change when using the mean estimates. The EPS estimates are multiplied by shares outstanding from daily CRSP data as of the date the EPS estimates were announced to obtain estimates for total dollar earnings. In addition, I also adjust for stock splits for the shares outstanding data. To merge the IBES data base with the CRSP data base, I first match them by 8-digit historical CUSIP. Additionally, I match firms whose ticker and/or company names are the same and those who have the same 6-digit historical CUSIP. In terms of timing, I match the quarterly IBES data with the monthly CRSP-COMPUSTAT merged by calendar quarter. In all asset pricing tests, I require the analyst estimates from the IBES summary file to be announced at least 1 quarter before the date returns are observed. Since the IBES summary file's statistical period is in the middle of each month, this means that the analyst expectation information is lagged about 3 month and 2 weeks.

To compute the payout ratio, I collect the common dividends (DVC), net income (IBCOM) as well as firms historical industry SIC code from COMPUSTAT. If a firm's net income is negative, I replace it with 6% of asset value (AT). I winsorize the payout ratio so that they ar also between 0 and 1. For other fundamental data and the price related variables I use the CRSP-COMPUSTAT Merged (Annual) . I include common shares (share codes 10

[7]Further horizons are available, but the coverage is much poorer

[8]The reason for using 2 FY3 estimates is because the coverage for FY3 is considerably poorer. My results are actually stronger when requiring 3 FY3 estimates, but the average number of firms covered will be only 60% of the sample in the base case.

and 11) in the CRSP database traded on NYSE/AMEX, and NASDAQ exchanges with the beginning-of-month prices above \$1. I follow the convention in the literature (for example Fama and French (2015)), lagging the annual fundamental information of each firm for at least 6 month and assume information on all the firms' fundamental data are observed by end of June each year when forming portfolios based on fundamental variables. Annual and monthly stock returns, as well as market prices gross and net of dividends are obtained from CRSP and are adjusted for stock delistings. The market capitalization (ME) of a stock is its price times the number of shares outstanding, adjusted for stock splits, using the cumulative adjustment factor provided by CRSP, which is also used to compute a firms total expected earnings and actual earnings.

B.5 Measuring Analyst Return Expectations Using Analyst Price Targets

Firm- level analyst return expectation are constructed using a bottom-up approach based on analyst-level return expectations per analyst issuance.

I collect single issuance of price targets from individual analyst's 12-month[9] price targets for individual firms from IBES unadjusted data base and match it with the closing price from CRSP on the date the price target is issued[10] to compute return expectation with price targets for individual firms. The expected returns are computed by dividing analyst's price targets by the daily closing price on the day the estimates was issued and subtracted by 1[11], or

$$\mu^A_{i,f,d} = \frac{P^{A,12}_{i,f,d}}{P_{f,d}} - 1$$

where $P^{A,12}_{i,f,d}$ is the price target of analyst i for firm f, issued at day d. The superscript 12 denotes the 12-month ahead estimates. Notice this methodology ensures there is no mechanical relation between mean estimated expected returns and the level of prices. On

[9] Other horizons are available, though the coverage is poor.

[10] In case the issuance date is a weekend, the last Friday prices are used; In case the issuance is a holiday, the previous business day closing prices are used.

[11] The same formula is used in Brav and Lehavy (2003a) and Da and Schaumburg (2011)

Table B.1: Returns and Alphas of the Universe with Available Estimates of Misvaluation (Analyst's Forecasts)

Sample period 1985-07 to 2018-12. Monthly value-weighted excess returns of the universe with available firm misvaluation measure $\hat{\alpha}_t^i$, or "vw.mkt.rf.analyst" are regressed on a constant (Column 1), value-weighted excess returns of market based on CRSP universe (Column 2) and Fama-French five factor returns downloaded from Ken French's webiste (Column 3).

	Dependent variable:		
		vw.mkt.rf.analyst	
	avg.ex.ret	CAPM.alpha	FF5.alpha
	(1)	(2)	(3)
mkt.rf		1.016***	1.028***
		(0.005)	(0.006)
smb			0.001
			(0.009)
hml			0.041***
			(0.011)
cma			0.0003
			(0.016)
rmw			0.028**
			(0.011)
Constant	0.005**	−0.001***	−0.002***
	(0.002)	(0.0002)	(0.0002)
Observations	402	402	402
R^2	0.000	0.989	0.990
Adjusted R^2	0.000	0.989	0.990
Residual Std. Error	0.045 (df = 401)	0.005 (df = 400)	0.005 (df = 396)
F Statistic		35,728.420*** (df = 1; 400)	7,837.528*** (df = 5; 396)
Note:			*p<0.1; **p<0.05; ***p<0.01

each issuing date the analyst has the freedom to pick her own price target since she observes the prices.

Firm-level return expectations are constructed together with the stop file provided by IBES to ensure individual estimates are not stale. IBES keeps track of the activeness of the individual estimates and provides a stop file for price targets[12]. I merge the point-in-time analyst-level expected return file with the stop-file on price targets to exclude estimates that analysts and IBES have confirmed to be no longer valid. Furthermore, to avoid stale estimates, I further restrict the estimates to be no older than 90 days when entering mean consensus estimates [13]

I construct weekly firm-level consensus expected returns by taking the mean of all active analyst-level forecasts, although using median makes no discernible difference for the main results. I drop analyst-level estimates that are greater than 5 standard deviation away from the mean estimates and I winsorize the entire analyst-level data base by 1% and 99% before calculating firm-level consensus. I take the mean of the available expected return estimates for each firm by the end of Saturday each week, or

$$\mu^A_{f,w} = \sum_i \mu^A_{i,f,w} / I_f$$

where I_f is the number of analyst for firm f at week w. For most of the application of the paper, I use firm-level return estimates based on monthly data, which is the consensus data on the last Saturday before each calendar month end.

[12]According to IBES, this stop-file "includes stops applied to estimates that are no longer active. This can result from several events, e.g. an estimator places a stock on a restricted list due to an underwriting relationship or the estimator no longer covers the company. Prior to June 1993, actual stop dates did not exist in the archive files used to create the Detail History. An algorithm was developed to determine the date when an estimate became invalid if, for example, a merger between companies occurred or an analyst stopped working for a firm, etc. Estimate that are not updated or confirmed for a total of 210 days, the estimate is stopped."

[13]Engelberg, McLean, and Pontiff (2019a) allows the estimates to be at most 12 month old, in case the estimates are not covered by the stop-file, although the choice makes little difference for the main results, as verified in the Appendix.

B.6 Estimating Implied-Cost-Capital

Methodology: The ICC Model of Pástor, Sinha, and Swaminathan (2008)

I follow the ICC model of Pástor, Sinha, and Swaminathan (2008) in estimating the implied cost of capital. Chen, Da, and Zhao (2013) details the way they calculate the ICC model in the cross section and I therefore follow the procedure outlined in their Appendix to estimate the ICC on the stock level.

1. Collect firm-level analyst earnings projections from IBES monthly summary file. Include firm-level earnings projections at the end of March, June, September and December for the current fiscal year (the next annual reporting date), the next fiscal year and the long-term-growth forecast (LTG);

2. Estimate firm-level Implied-Cost-Capital (ICC) model. This involves assuming a firm-level long-term growth rate as well as a plow-back rate (or 1- payout rate):

   (a) Assuming

   $$P_t = \sum_{k=1}^{15} \frac{FE_{t+k}(1-b_{t+k})}{(1+q_t)^k} + \frac{FE_{t+16}}{q_t(1+q_t)^{15}} = f(c^t, q_t) \qquad \text{(B.22)}$$

   where P_t is the stock price, FE_{t+k} is the earnings forecast k years ahead, b_{t+k} is the plowback rate $(1 - payout)$, and q_t is the ICC.

   (b) Estimate FE_{t+k} :

      i. FE_{t+1} and FE_{t+2} are proxied by the current fiscal year and the next fiscal year IBES analyst summary file data. $FE_{t+3} = FE_{t+2}(1 + LTG_t)$

         A. Assuming the individual firm-level earnings growth rates to revert to industry growth forecast (LTG_t^{Ind}) by year $t+16$:

$$g_{t+k} = g_{t+k-1} \times exp[log(LTG_{t+3}^{ind}/LTG_{t+3})/13]$$
$$\forall 4 \leq k \leq 15$$
$$g_{16} = g_t^{GDP},$$
$$FE_{t+k} = FE_{t+k-1}(1+g_{t+k}) \quad \forall 4 \leq k \leq 16$$

g_t^{GDP} is the GDP growth rate using an expanding rolling window since 1947

(c) Estimate b_{t+k} :

i. b_{t+1} and b_{t+2} are estimated from the most recent net payout ratio for each firm. The net payout ratio ratio of common dividends (DVC in compustat) to net income (item IBCOM). If net income is negative, replace it by 6% of assets[14].

ii. b_{t+k},$3 \leq k \leq 16$ is assumed to

$$b_{t+k} = b_{t+k-1} - \frac{b_{t+2} - b_t^{ss}}{15} \tag{B.23}$$
where $b_t^{ss} = g_t^{GDP}/q_t$

(d) The q_t is then backed out numerically by solving Eq. (B.22) and (B.23) together numerically. When there exists multiple roots, choose the root that is closest to the risk-free rate. Exclude any stock whose price is below $1. Winsorize the

[14]Notice that about 50% of the firms do not pay dividend in the last year. As a result, the first two years of plowback ratio is 1. This does not mean that the projected earnings for the first two years have no impact on the estimation of the implied cost of capital q_t. Since FE_{t+k} are first calculated using the first two to three years of earnings projections together with the firm and industry level LTG, as long as any path during the first 15 years contain a non-zero payout ratio, the first three years of projections will have an impact on the estimation of the ICC.

sample at 1% and 99%. Notice that by assuming the steady-state plowback ratio, we implicitly impose the constraint that

$$q_t \geq g_t^{GDP}$$

since in steady-state, the plowback ratio can not exceed 1.

Summary Statistics of ICC and Input Variables

B.7 Robustness Checks

B.7.1 Equal-Weighted Portfolio Sorts

Table B.2: Summary Statistics

(a) Empirical distributions of variables

	statistic	Pi	b_1_2	pb_7	EP_1	EP_2	EP_3	LTG	g_ind
1	mean	0.130	0.990	0.848	0.070	0.095	0.114	0.169	0.168
2	std	0.058	0.057	0.073	0.113	0.151	0.187	0.089	0.055
3	std cs	0.056	0.055	0.072	0.111	0.149	0.184	0.086	0.047
4	std ts	0.030	0.031	0.042	0.065	0.065	0.076	0.054	0.026
5	min	0.068	0.589	0.603	-0.153	0.004	0.014	0.040	0.048
6	p25	0.094	1	0.807	0.033	0.045	0.054	0.110	0.130
7	median	0.116	1	0.848	0.053	0.066	0.076	0.150	0.159
8	p75	0.144	1	0.896	0.076	0.090	0.104	0.200	0.195
9	max	0.428	1	0.993	0.876	1.216	1.512	0.500	0.351

(b) AR(1) coefficients

	variable	Pi	b_1_2	pb_7	EP_1	EP_2	EP_3	LTG	g_ind
1	AR(1)	0.920	0.943	0.882	0.897	0.950	0.956	0.893	0.946
2	std	0.005	0.009	0.004	0.008	0.005	0.005	0.005	0.010

(c) Correlations between variables

	Pi	pb_7	b_1_2	EP_1	EP_2	EP_3	LTG	g_ind
Pi	1	-0.746	0.023	0.611	0.747	0.784	0.409	0.336
pb_7		1	0.461	-0.352	-0.417	-0.434	-0.351	-0.284
b_1_2			1	-0.006	0.006	0.010	0.079	0.102
EP_1				1	0.873	0.821	-0.136	-0.107
EP_2					1	0.970	-0.070	-0.063
EP_3						1	-0.030	-0.042
LTG							1	0.511
g_ind								1

Note: Statistics are calculated over the whole sample. Firm-level variables are winsorized at 1% and 99% for the whole sample. Variable definitions: “Pi”: implied constant discount rate (ICC); “b_1_2” is the plow-back ratio from the last year; “pb_7” is the implied plow back ratio in Year 7. “Ek/P”, k = 1, 2, 3 are the fiscal year k earnings consensus estimates divided by the current market capitalization; “LTG”: long-term growth forecasts; “g_ind” industry long-term growth estimates where industries are defined based on 48 Fama-French classification. In Panel B.2a, “std” denotes the standard deviations for the variables over the entire sample, “std cs” and “std ts” are the average cross-sectional standard deviation over time and the time-series standard deviation over different firms, respectively. AR(1) coefficients are estimated by regressing the current value of the variable on its respective one-quarter lagged value based on the whole sample. Standard errors for the AR(1) coefficients are clustered by firm-quarter.

Table B.3: Pre-estimated Mis-valuation ($\widehat{\alpha}_t^i$) Sorted Portfolios and Realized Average Stock Returns (1986-06 to 2018-12, value-weighted)

All returns, alphas and their standard errors are expressed in percentages. Stocks are into quantile portfolios based on the misvaluation measure $\widehat{\alpha}^i$ at the end of June each year, using the available information up to that point. Portfolios are rebalanced with equal weights every month. "Low" denote the portfolio with lowest $\widehat{\alpha}_t^i$. "High-Low" are excess returns of a portfolio that goes long on stocks with the highest $\widehat{\alpha}_t^i$ and short those with the lowest $\widehat{\alpha}_t^i$. "SE" are standard errors which are shown in brackets. "Mean ex.ret" are monthly returns over 3 month treasury rates; "SR" are monthly Sharpe Ratios. "FF-5 alpha" denote Fama-French 5 factor alphas. "num_stocks" are average number of stocks included in the portfolio over time. "Ex-Ante Misvaluation" are value-weighted portfolio $\widehat{\alpha}_t^i$ measured at each end of June. Their standard errors are measured using Newey-West methods based on 4 lagbs ("SE (NW-4)").

stats	Low	2	3	4	High	High - Low
Ex-Ante Misvaluation	-1.8	-0.88	-0.67	-0.5	-0.3	1.5
SE (NW-4)	(0.19)	(0.32)	(0.21)	(0.18)	(0.15)	(0.12)
CAPM alpha	-0.63	-0.26	0.07	0.2	0.38	0.98
SE CAPM alpha	(0.22)	(0.19)	(0.13)	(0.11)	(0.11)	(0.17)
mean ex.ret	0.19	0.48	0.76	0.87	1.12	0.94
SE ex.ret	(7.22)	(6.28)	(5.56)	(5.23)	(5.91)	(3.31)
SR	0.03	0.08	0.14	0.17	0.19	0.28
CAPM beta	1.33	1.15	1.12	1.09	1.24	-0.08
SE CAPM beta	(0.05)	(0.04)	(0.03)	(0.02)	(0.03)	(0.04)
FF-5 alpha	-0.49	-0.29	-0.05	0.07	0.33	0.79
SE FF-5 alpha	(0.15)	(0.14)	(0.07)	(0.06)	(0.08)	(0.14)
num_stocks	456.68	453.78	456.09	456.09	453	