Recovering Investor Expectations from Demand for Index Funds

Recovering Investor Expectations from Demand for Index Funds*

Mark Egan
Harvard University
and NBER§

Alexander MacKay
Harvard University†

Hanbin Yang
Harvard University‡

July 27, 2021

Abstract

We use a revealed-preference approach to estimate investor expectations of stock market returns. Using data on demand for index funds that follow the S&P 500, we develop and estimate a model of investor choice to flexibly recover the time-varying distribution of expected future returns across investors. Our analysis is facilitated by the prevalence of leveraged funds that track the same underlying asset: by choosing between higher and lower leverage, investors trade off higher return against less risk. Our estimates indicate that investor expectations are heterogeneous, extrapolative, and persistent. Following a downturn, investors become more pessimistic on average, but there is also an increase in disagreement among participating investors due to the presence of contrarian investors.

Keywords: Stock Market Expectations, Demand Estimation, Exchange-Traded Funds (ETFs)
JEL Classification: D12, D81, D84, G11, G50, L0

*We thank John Campbell, Robin Greenwood, Sam Hanson, Ian Martin, Jesse Shapiro, Andrei Shleifer, Adi Sunderam and the seminar participants at the Columbia University Discussion Group on Intermediation, Harvard Business School, Notre Dame, Princeton University, Swiss Finance Institute Lugano, the University of Chicago Booth Household Finance Conference, the University of Copenhagen, and the University of Washington. We also thank John Graham and Campbell Harvey for providing the Duke CFO Global Business Outlook Survey data.

†Harvard University, Harvard Business School. Email: amackay@hbs.edu.
‡Harvard University, Harvard Business School and Department of Economics. Email: hyang1@g.harvard.edu.
§Harvard University, Harvard Business School. Email: megan@hbs.edu.

1 Introduction

Understanding investor beliefs, how these beliefs are formed, and the dynamics of these beliefs is critical for explaining investment and saving behavior and may have profound macroeconomic implications. For example, beliefs that diverge from rational expectations may affect the distribution of wealth across households or exacerbate credit cycles (Bordalo et al., 2018). Thus, a better understanding of investor beliefs can inform macroeconomic policy and the regulation of financial markets. A number of surveys have been designed to elicit beliefs about the future performance of the stock market from households, investment professionals, and managers. While recent evidence suggests that these surveys produce consistent and valuable information, surveys can be criticized for being noisy and sensitive to interpretation (Greenwood and Shleifer, 2014; Giglio et al., 2019). We propose an alternative method to recover investor beliefs about future returns based on observed investment decisions. In contrast to other methodologies used to recover investor beliefs, our methodology uses data on investment flows, rather than asset prices, which allows us to recover the distribution of beliefs across investors.

Specifically, we develop a parsimonious model of demand for exchange-traded funds (ETFs) that reflect the performance of the stock market. By revealed preference, estimation of the model allows us to recover the underlying distribution of investor expectations of future returns. For each ETF purchase, we recover the beliefs about future returns and risk aversion that rationalizes the purchase given the fees and risk associated with the ETF. The key feature of our data for identification is that investors choose investment options from a menu of several ETFs with different risk/return profiles and fee structures. Identification in our setting is conceptually related to Barseghyan et al. (2013), who show how beliefs and risk aversion can be separately identified in the context of insurance choice.

This paper has two main empirical contributions. First, we use our framework to construct a time series of stock market expectations. At each point in time, we recover the distribution of expectations across investors. We find that heterogeneity in expectations is meaningful and varies over time. Second, we examine how investor expectations are formed. We confirm prior findings that average beliefs are extrapolative and violate full-information rational expectations. However, our estimated distribution also suggests that a large fraction of investors are contrarian. In particular, periods with increased disagreement suggest the simultaneous presence of both contrarian and extrapolative investors. These results indicate that understanding dispersion in beliefs is important, and our framework allows us to study how this dispersion evolves over time.

To implement the approach, we apply a model of investor choice to observed market shares for investments linked to the performance of the S&P 500. Our data on market shares comes from monthly purchases of ETFs by retail (non-institutional) investors. ETFs are passive investment funds designed to track another underlying asset. Collectively, ETFs linked to the S&P 500 average $240 billion in assets under management during our sample, and they provide varying levels

of leverage for the same benchmark.[1] The ETFs are designed to (a) track the return of the S&P 500, (b) provide leveraged return (2x or 3x return) of the S&P 500, or (c) provide inverse leveraged return (-3x, -2x, or -1x) of the S&P 500. Leveraged ETFs are popular investment products among retail investors. Relative to all S&P 500 linked ETFs held by retail investors, leveraged ETFs accounted for roughly one quarter of assets under management (AUM) and almost half of retail trading volume during the financial crisis. In each month, we observe the fraction of investors purchasing S&P 500 linked ETFs in each leverage category.

Studying leveraged index funds offers a clean setting for identifying investor expectations of stock market returns. By choosing among different levels of leverage exposure to the same underlying asset, the investor reveals information about her expectations for the future performance of the asset and her risk preferences. With higher leverage, an investor increases the expected mean return, but also the risk associated with the investment. We model this decision and estimate the model to recover a time-varying distribution of investor expectations of stock market returns that rationalize aggregate choices.

Identification of the model works conceptually as follows. Consider an investor who elects to purchase a 2x leveraged ETF, and for simplicity, assume the investor has no other wealth or investments. Compared to a 1x ETF, the investor has doubled the expected return and taken on twice the risk. Thus, the investor's purchase indicates that the investor is either more optimistic about the return of the stock market or more risk tolerant than an investor that chooses a 1x ETF. Because the investor could have further increased the mean return and the risk by purchasing a 3x ETF but chose not to, we have a second restriction on the investor's expectation and risk aversion. Full nonparametric identification can be achieved with empirical variation in fees or perceived risk, as these inform the mean expectation and risk aversion, respectively.

Using maximum likelihood, we estimate a flexible, time-varying distribution of expectations at a quarterly frequency over the period 2008-2018. Our framework allows us to quantify those expectations in terms of the expected annualized return of the stock market. The results suggest that accounting for belief heterogeneity across investors is of first-order importance, as in Meeuwis et al. (2018) and Brunnermeier et al. (2014). For example, we find that, while the median investor in December 2009 expected a market risk premium of 5 percent, roughly 10 percent of investors expected the stock market to fall by more than 10 percent. To validate our results, we compare our estimates to widely used surveys of investor expectations (e.g., the Shiller Index, Gallup, etc.) that are commonly used in the literature (Greenwood and Shleifer, 2014; Nagel and Xu, 2019). Despite the fact that these two approaches draw on different populations and are collected with different methods, we find that our estimates are positively correlated with existing surveys.

Consistent with the survey data results, we interpret our revealed-preference estimates as investors' beliefs about the expected future return of the stock market. However, this interpretation has two important caveats. First, we do not observe investors' portfolios; we only observe pur-

[1] Hortaçsu and Syverson (2004) develop and estimate a sequential search model to understand price dispersion within the 1x leverage funds designed to track the S&P 500. We broaden the set of funds to include leveraged ETFs in order to study the "first-stage" decision of which leverage category to invest in.

chases of S&P 500 ETFs. If investors use leveraged S&P 500 ETFs to hedge other investments, our estimated risk parameter would capture a mix of risk aversion and hedging demand. To address this, we estimate an extension of our model where investors account for the risk of the ETF both in terms of the variance of the ETF and the covariance of the ETF with the rest of the investors' wealth (i.e., hedging demand). We find little evidence of hedging demand, and our estimated time series of beliefs remain similar. Second, we are studying a subset of retail investors who choose to invest in leveraged ETFs. Even though the market for leveraged ETFs is quite large,[2] one may be concerned that our estimated beliefs are not representative of a general population. As discussed above, we find that our estimates of investor expectations are highly correlated with survey estimates, which suggests that ETF investors in our sample have similar expectations to broader groups of market participants. We conclude that our parsimonious model generates reasonable estimates of investor beliefs. Our methodology can be particularly useful when high-frequency survey or micro data is unavailable. Further, survey expectations are often criticized for being potentially unreliable, unrelated to the portfolio decisions of households, and not linked to marginal investors (Cochrane, 2011). In contrast, our methodology extracts investor beliefs directly from their investment decisions.

We use our estimates to examine how the distribution of investor expectations evolves over time. Our estimates suggest that the mean, dispersion, and skewness of the belief distribution evolve in response to past returns. Following a period of negative stock market performance, investor beliefs become more pessimistic on average, more dispersed, and more negatively skewed. Though the mean expectation is extrapolative, we observe an increase in the median belief following negative returns, suggesting that a large share of investors display contrarian tendencies. Both extrapolative and contrarian tendencies appear stronger in response to negative returns compared to positive returns. These asymmetric responses lead to increased disagreement after stock market busts, while periods of positive returns generate greater consensus.[3] We also find that expectations are persistent: one month of poor stock market performance impacts investor expectations up to two years in the future. Lastly, we find evidence that average forecast errors are predictable, which suggests that investor expectations violate full information rational expectations, consistent with much of the existing literature.[4]

We also consider an extension of the model where we allow risk aversion, in addition to beliefs, to vary over time. Consistent with Jensen (2018), our estimated risk aversion parameter is low at the start of the financial crisis and increases over time in our sample. One caveat is that our estimated risk aversion parameter captures variation in both risk aversion and risk perceptions, both of which may be time-varying and cyclical.

[2]For example, the retail market share of leveraged S&P 500 ETFs was roughly the same as tracker (1x leverage) S&P 500 ETFs during the financial crisis (after adjusting for trading volume). We use "leveraged ETFs" to describe both ETFs with positive leverage (2x, 3x) and inverse ETFs with negative leverage (-1x, -2x, -3x).

[3]Using data from the Survey of Professional Forecasters, Ilut and Schneider (2014) also find that forecast dispersion is counter-cyclical.

[4]For example, see Bacchetta et al. (2009); Coibion and Gorodnichenko (2012, 2015); Amromin and Sharpe (2014); Greenwood and Shleifer (2014); Gennaioli et al. (2016); Bordalo et al. (2019) among others.

Next, we use our estimated distribution of investor beliefs to understand the relationship between investor beliefs and asset prices. Consistent with the previous literature, we find that the average expected return across investors does not forecast future returns (Greenwood and Shleifer, 2014; Amromin and Sharpe, 2014). In contrast, we find evidence suggesting that dispersion in beliefs is negatively correlated with future returns, and that this effect is amplified when a larger fraction of investors potentially face short-selling constraints. This finding is consistent with the literature that argues that belief disagreement and short-selling constraints can generate equilibrium asset prices that are higher than the valuation of the average investor, leading to lower future returns.[5]

While the bulk of our analysis focuses on S&P 500 linked ETFs and investor expectations of stock market returns, our approach readily extends to other asset classes. We use our model to recover the time-varying distribution of investor expectations for gold, oil, European equities, emerging markets equities, US real estate, medium-term (7-10 year) Treasury, and long-term (20+ year) Treasury using ETFs linked to the primary benchmarks in these asset classes.

We recover the distribution of investor beliefs at monthly and quarterly frequencies from 2008 through 2018. However, our methodology, in principle, allows us to recover the distribution of beliefs at the same frequency as ETF flow data. As an extension, we use daily ETF flow data to recover the distribution of investor beliefs at a daily level during the COVID-19 pandemic. Our results suggest that news of the first recorded death in the US in late February coincided with a substantial increase in the dispersion of investor beliefs that peaked in mid-March and persisted through June. The average retail investor appeared cautiously optimistic until US states started imposing lockdowns in mid-to-late March, after which the average expected return fell by more than 20 percent. Negative mean expectations lasted until mid-April, when infection rates in the US started to fall. These results highlight the potential usefulness and scope of our methodology; we recover the distribution of investor beliefs in real time using readily available data.

The paper proceeds as follows: Section 2 describes the data used in our analysis. Section 3 introduces our model of investor choice and describes how variation in leverage within the choice set allows us to nonparametrically identify the distribution of beliefs. Section 4 details the parameterization of our empirical model, describes the estimation routine, and presents the results along with a comparison to survey data. We analyze the formation of investor expectations in Section 5 and additional extensions to our method in Section 6. Section 7 concludes.

Related Literature

Our paper builds on the demand estimation literature at the intersection of industrial organization and finance.[6] On a conceptual level, our paper complements the recent work of Koijen and Yogo

[5]The theoretical literature dates back to Miller (1977) and is supported by an empirical literature that includes Chen et al. (2002) and Diether et al. (2002).

[6]Demand estimation has recently been used in a number of other financial applications such as demand for bank deposits (Dick, 2008; Egan, Hortaçsu, and Matvos, 2017; Egan, Lewellen, and Sunderam, 2017; Wang, Whited, Wu, and Xiao, 2018), bonds (Egan, 2019), credit default swaps (Du et al., 2019), insurance (Koijen and Yogo, 2016, 2018),

(2019a). Koijen and Yogo (2019a) develop an equilibrium asset pricing model where investors have heterogeneous preferences, and each investor's portfolio is generated from a Berry et al. (1995) type demand system. This type of demand-side approach to asset pricing uses the revealed preferences of investors, by focusing on quantities rather than prices or returns. It is conceptually related to the approaches that Shumway et al. (2009) and Berk and van Binsbergen (2016) use to study mutual fund flows and Heipertz et al. (2019) uses to study French banks.

We build on the idea of estimating preference heterogeneity across investors by studying investor expectations. To this end, our paper also relates to Ross (2015) and the corresponding literature that uses asset prices to separately recover risk preferences and transition probabilities, which correspond to beliefs about returns in our context. These other methodologies, such as Ross (2015) or Ghosh and Roussellet (2020), use asset prices across products that span different states to recover beliefs for a representative investor. By contrast, we use data on investment flows, rather than asset prices, to recover the distribution of investor expectations. In other words, we recover heterogeneity across investors, whereas the previous methods typically focus on a hypothetical, homogeneous investor. We find heterogeneity across investors to be very important empirically.[7] Further, in the context of leveraged ETFs, we observe plausibly exogenous variation in the cost of leverage, allowing us to recover the distribution of beliefs without making restrictive assumptions about the structure of asset prices, beliefs, or preferences.[8]

In other contexts, Calvet et al. (2019) and Martin (2017) also focus on recovering expectations and risk preferences. Using household level data from Sweden, Calvet et al. (2019) calibrates a life-cycle model to recover the distribution of risk aversion in the population under the maintained assumption that investors hold common expectations of returns. Lastly, Martin (2017) derives a lower bound on the equity premium using data from index option prices. Although we use a very different approach and data set, our time-varying estimates of the equity premium are similar to the estimates in Martin (2017).

Our paper also relates to the growing work on structural behavioral economics (see DellaVigna, 2018, for a discussion of the literature). Our paper relates most closely to Barseghyan et al. (2013).

---

mortgages (Benetton, 2018) and investments more generally (Koijen and Yogo, 2019a,b; Koijen et al., 2019).

[7] There is also a related strand of literature that focuses on robust identification of investor beliefs (Chen et al., 2020; Ghosh and Roussellet, 2020; Ghosh et al., 2020). Given a stochastic discount factor in a candidate asset pricing model and a statistical measure of divergence between probabilities, the methods developed in this literature can recover beliefs that are consistent with observed prices within a certain distance of divergence from rational expectation. Ghosh et al. (2020) extends the methodology to allow for two different investor types that have potentially different preferences and beliefs. Identification of beliefs using these methodologies depends on the underlying asset pricing model chosen by the researcher and beliefs that do not stray too far from rational expectations.

[8] By comparison, Ross (2015) uses state prices computed from options to back out state transition probabilities that are assumed to be time-homogeneous and Markov. Previous research points out that such assumptions are strong and rule out persistent shocks to the stochastic discount factor in realistic asset pricing models (Borovička et al., 2016). Jensen et al. (2019) relax the time-homogeneous and Markov assumptions. Their approach requires prices of securities with different maturities, where the number of maturities is greater or equal to the fixed number of states. In addition, Martin and Ross (2019) investigate the recovery theorem in fixed income markets where the stationary and Markov assumptions are more likely to hold. Using a similar approach, d'Arienzo (2020) tests for the rationality of recovered belief and show that non-rational beliefs could provide explanations for bond market puzzles.

In our baseline specification, we assume that investors have mean/variance preferences; however, in principle our methodology allows for richer investor preferences.

Barseghyan et al. (2013) develops a demand-side framework that shows how belief distortions can be separately identified from risk preferences using data on insurance choice. We use a similar framework and identification strategy to recover the distribution of beliefs, and then use the corresponding estimates of the belief distribution to better understand the evolution of beliefs and the corresponding behavioral frictions.

Our demand estimation framework and estimated beliefs complement the findings of Vissing-Jorgensen (2003), Ben-David et al. (2013), Amromin and Sharpe (2014), Greenwood and Shleifer (2014), and Nagel and Xu (2019). These papers use survey evidence to better understand investor expectations. Using very different data and empirical approach, we find similar patterns of investor expectations. While we use trading activity data to infer investor beliefs, recent account-level evidence from Giglio et al. (2019) suggests that an investor's beliefs, as measured by surveys, are reflected in the direction and magnitude of her trading decisions. Given this finding, it is not surprising that our measure of beliefs during the COVID-19 pandemic also matches the patterns Giglio et al. (2020) find in their follow-up study on the pandemic. Our finding that beliefs are heterogeneous and persistent is consistent with one of the main empirical facts (Fact 3) that Giglio et al. (2019) document using survey data.

Our findings also relate to the literature linking beliefs to asset prices. One theme of this literature is the importance of the interrelationship between belief heterogeneity and short-selling constraints. A long theoretical literature suggests that belief disagreement in conjunction with short-selling constraints can lead to elevated equilibrium asset prices and lower future returns.[9] Restrictions on short selling constrain pessimistic investors from selling assets to the more optimistic investors in the market. In equilibrium, the most optimistic investors will hold the asset and the price will reflect the beliefs of the marginal-optimistic buyer. If the valuation of the average buyer is correct in expectation, the asset will have lower future returns. Consistent with the theoretical and corresponding empirical literature (Chen et al., 2002; Diether et al., 2002), we find evidence that increased dispersion in beliefs is associated with lower future returns and that the effect is magnified when more investors face short-selling constraints.

One of our key findings is that investor beliefs appear extrapolative across a number of asset classes. This finding complements the literature that uses survey evidence to document extrapolation in the stock market (Benartzi, 2001; Greenwood and Shleifer, 2014), the housing market (Case et al., 2012), risk taking (Malmendier and Nagel, 2011), investment decisions (Gennaioli et al., 2016), and inflation markets (Malmendier and Nagel, 2015). In our setting, we are able to study extrapolation and beliefs simultaneously across several common asset classes. A novel finding is that, while beliefs are extrapolative for the average investor, they do not appear extrapolative for all investors. For example, following downturns, the average investor becomes more pessimistic, but a substantial fraction of investors become more optimistic. This empirical finding of heterogeneous extrapolation is a feature of theoretical asset pricing models with extrapolative

[9] See, e.g., Miller (1977); Harrison and Kreps (1978); Morris (1996); Chen et al. (2002); Scheinkman and Xiong (2003); and Hong et al. (2006). Geanakoplos (2010) and Simsek (2013) extend this idea further by adding credit and endogenous borrowing constraints, which can strengthen the negative relationship between disagreement and returns.

beliefs (Cutler et al., 1990; De Long et al., 1990; Barberis and Shleifer, 2003; Barberis et al., 2015, 2018). Thus, our findings provide additional evidence for understanding the formation of beliefs. A recent literature documents that such adaptive expectations could have profound impacts on the macroeconomy (Bordalo et al., 2018; Gennaioli and Shleifer, 2018; Bordalo et al., 2018).

2 Data

2.1 Overview of Leveraged ETFs

Leveraged ETFs provide investors a menu of different exposures to an underlying asset. Leveraged ETFs cover many asset categories, including broad indices (S&P 500) and commodity prices (oil). They offer discrete leverage categories of 2x or 3x on the long side and -1x, -2x, and -3x on the short side. These products provide active retail investors access to leveraged exposure with limited liability as an alternative to more complicated derivative contracts, which require margins and specialty knowledge.[10]

Despite easy access to leverage, these products mostly attract short-term investors. Leveraged ETFs typically rebalance either daily or monthly in order to maintain a constant ratio of leverage relative to the linked asset. Though this rebalancing provides constant leverage in the short run, the return on these ETFs over a longer period (i.e., more than one year) may diverge from the short-run leverage target.[11] Because ETF trades predominantly reflect shorter horizons, this feature does not affect the typical ETF investor. In our data, the average holding period for ETFs is less than one month.

2.2 ETF Data Sources

We assemble ETF data from Bloomberg, ETF Global, and CRSP. Bloomberg reports monthly data on ETF AUM, net asset value, trading volume, and quarterly data on ETF institutional ownership. We rely on benchmark and fund descriptions in ETF Global accessed via WRDS to identify the choice sets of S&P 500 ETFs with leverage categories from -3x to 3x. Lastly, CRSP Mutual Fund Database also accessed through WRDS provides ETF expense ratios. When missing in CRSP, we use expense ratios from ETF Global. Our panel ranges from 2008 to 2018.

We aggregate individual ETFs to their leverage categories, so that the primary unit of observation in our analysis is at the month-by-leverage level. Our main focus is understanding investor expectations and risk aversion, so we focus on investors' choice of leverage (i.e., 1x vs 2x leverage) rather than individual ETFs (i.e., ProShares Ultra 2x S&P 500 ETF vs. Direxion Bull 2x S&P

[10] In addition to ETFs, retail investors can also buy leveraged mutual funds and exchange-traded notes (ETNs). We focus on ETFs primarily because of better data quality and for comparability reasons. The structures of leveraged mutual funds and ETNs are the same as ETFs, so our model could also be applied to these products.

[11] Thus, investors holding ETFs for longer periods are subject to additional risk from changes in leverage. The difference between nominal and realized leverage is commonly described as "volatility drift." See "Important information about leveraged, inverse, and commodity exchange-traded products," Vanguard.com. https://investor.vanguard.com/investing/leveraged-inverse-etf-etn. Accessed November 12, 2019.

500 ETF). Implicitly, we assume that investors choose leverage and issuer separately. We consider this approach reasonable because the risk and return profiles of ETFs are homogeneous within a leverage category, similar to the maintained assumption in Hortaçsu and Syverson (2004). To aggregate our data from ETF to leverage level, we sum the market shares across ETFs and take the market-share weighted average expense ratio. We detail our construction of ETF-specific market shares below.

2.3 Constructing Market Shares from Leverage Choice

A key input in our empirical model is the quantity of ETFs purchased by retail investors. We measure quantities as the dollar amount retail investors purchase during each month. This flow measure reflects the investor expectations at the time of transaction, as opposed to stock measures that reflect past purchase decisions.[12] Stock measures such as AUM will place greater weight on passive investors whose holdings do not reflect contemporaneous information. This distinction is important in our context, because a large fraction of AUM in trackers (1x) comes from passive investors, whereas trading in leveraged ETFs is dominated by active investors. In addition, we remove demand from institutional investors. Institutions are major investors in trackers, but they rarely buy leveraged ETFs because they have access to more cost-effective leveraged contracts such as futures and swaps.

We construct our measure of ETF purchases from data on trading volume and net fund flows each month, with the ultimate goal of measuring the quantity of ETFs purchased by active retail investors in a given month. To calculate the quantity purchased, we assume that every month a fraction of retail investors consider investing a fixed portion of their wealth in ETFs that follow a particular asset (e.g., the S&P 500). Investors may set aside this wealth in the process of actively rebalancing their portfolios in that month. If an investor does not choose to invest this wealth in an ETF, the investment is put into a money market account.

When investors purchase ETFs, they purchase from a market maker who has previously purchased the ETF from another investor. Thus, measures of trading volume ($Trading\,Volume_{jt}$) count investor purchases and investor sales as separate transactions. When purchases exceed sales, new shares are issued by the market maker to satisfy excess demand. These new shares are measured as net flows ($Net\,Flow_{jt}$). When sales exceed purchases, $Net\,Flow_{jt}$ is negative and represents redemptions.

We construct the quantity of ETF j purchased by retail investors at time t as

$$Quantity_{jt} = Retail_j \times [(Trading\,Volume_{jt} - Net\,Flow_{jt})/2 + Net\,Flow_{jt}].$$

To measure purchases only, we first subtract net flows from trading volume, capturing trades of existing shares that have both a purchase and sale. We divide this measure in half to get a measure

[12] Account-level evidence from Giglio et al. (2019) indicates that, conditional on trading, investor trading decisions are highly correlated with beliefs.

of purchases. We add back in net flows to adjust for shares created or redeemed. To adjust for retail demand, we scale this measure by the average retail ownership of each ETF in our sample, $Retail_j$.

As in most demand estimation exercises, we do not directly observe investors that consider investing in S&P 500 ETFs but instead choose a risk-free option (0x leverage). To construct shares for this outside option, we measure flows into retail money market accounts for retail investors considering S&P 500 ETFs. Investing in a money market account is a natural risk-free option for most retail investors. First, we obtain the total amount of assets invested in retail money market funds from FRED. We scale this total by the fraction of AUM in S&P 500 ETFs out of the AUM in all retail investment vehicles (including all ETFs and retail mutual funds). This constructs a proxy for the share of money market AUM corresponding to S&P 500 investors. To convert this stock measure into a flow measure of investor purchases, we scale this proxy by the ratio of retail quantity (defined above) to retail AUM. We calculate the ratio as the average across all S&P 500 ETFs within each month of our sample. As a robustness check, instead of using this measure, we estimate the share in the outside option as a free parameter; the estimation results are not materially different. We discuss this and other robustness checks in Section 4.

Table 1 compares market shares based on our demand measure with shares of raw AUM, which includes holdings of passive or institutional investors. Because institutional investors hold a disproportionate share of tracker funds, the shares in trackers are on average 88 percent under AUM but only 57 percent according to our market share definition.[13]

2.4 Summary Statistics and Trends

The market for S&P 500 linked ETFs and leveraged ETFs grew dramatically over the period 2008-2018. Figure 1 displays total AUM held in S&P 500 linked ETFs by retail investors and the associated trading volumes over the period 2008-2018. As of 2018, retail investors held around $180 billion in S&P 500 linked ETFs.

The primary unit of observation in our analysis is the market share of each leverage category at the monthly level. Figure 2a displays the market share of each leverage category over the period 2008-2018. While S&P 500 tracker funds (1x leverage) are the most commonly held product on average, during the financial crisis leveraged ETFs collectively became more popular than tracker ETFs.

Table 1 shows a breakdown of leverage categories, with average AUM and expense ratios. As discussed above, leveraged ETFs are smaller in AUM compared to trackers. Leveraged ETFs also charge substantially higher fees, and ETFs with more leverage tend to be marginally more expensive. Figure 2b shows the trends in ETF fees. ETF fees are relatively stable over time, though the

[13]We have run additional specifications where we define market shares based on AUM instead of trading volume (Appendix B.5). The estimated distributions of beliefs are similar to our baseline estimates, though there are some differences, which is to be expected given that previous research shows that an investor's beliefs are more correlated with her trading decisions than her holdings (Giglio et al., 2019). For example, a buy-and-hold investor may not update her portfolio in response to a change in beliefs.

average fee for 1x trackers has been declining since 2013.

3 Empirical Framework

3.1 Demand for ETFs

We model an investor's decision as a discrete-choice problem. Each investor i has a fixed amount of wealth to allocate to ETFs that are benchmarked to the performance of the S&P 500 Index. Investors choose an ETF leverage category $j \in \{-3, -2, -1, 0, 1, 2, 3\}$ with corresponding leverage $\beta_j = j$, where $j = 0$ represents the outside option of placing their money in a retail money market account.

Investor i's indirect utility from choosing leverage j is given by

$$u_{ij} = \beta_j \mu_i - p_j - \frac{\lambda}{2} {\beta_j}^2 \sigma^2. \tag{1}$$

The term μ_i reflects investor i's expectation of future stock market returns. Investors have heterogeneous expectations that are distributed $\mu_i \sim F(\cdot)$. If an investor chooses $\beta_j = 2$, she will realize twice the return of the S&P 500 Index. Collectively the term $\beta_j \mu_i - p_j$ captures the investor's (subjective) expected return as a function of leverage β_j and net of ETF fee p_j. Without any loss in generality, we normalize preferences with respect to the annualized ETF fee p_j to one. Because ETF fee p_j is measured as annualized percentage of AUM, this allows us to interpret μ_i as the annualized return in excess of the risk-free rate offered by a money market account.

Risk aversion is additively separable, following the second-order Taylor expansion used in Barseghyan et al. (2013) and the Arrow-Pratt approximation of risk premium. The parameter λ is the investor's coefficient of risk aversion, and can be interpreted to represent constant relative risk aversion.[14] The term $\beta_j^2 \sigma^2$ measures the volatility of leverage j, where σ^2 is the volatility of the S&P 500 Index. Thus, the combined term $-\frac{\lambda}{2} {\beta_j}^2 \sigma^2$ captures the (time-varying) risk penalty for leverage category j.

In our baseline analysis, we assume that risk aversion is constant across investors and over time. We later extend the model to allow for heterogeneous risk aversion where $\lambda_i \sim G(\cdot)$ and time-varying risk aversion where λ potentially varies over time. Though we label λ_i as a risk aversion parameter, it may also capture heterogeneous beliefs over the volatility of the stock market. Thus, λ_i may be interpreted as $\lambda_i \frac{\sigma_i^2}{\sigma^2}$, where σ_i^2 is investor i's expectation of stock market volatility. In this more general interpretation, λ_i captures heterogeneity in both risk aversion and risk perceptions.

Another feature of our model is that we treat an investor's ETF investment choice independently from her more general portfolio allocation problem. To address this, we consider an extension of the model where investors account for how the ETF covaries with their wealth/portfolio, and we allow ETF choice to potentially hedge against wealth/portfolio risk. For a derivation of this

[14] Another way to derive equation (1) is to assume a utility function with constant absolute risk aversion and normally distributed returns. In this case λ represents the constant absolute risk aversion coefficient multiplied by wealth.

model and the corresponding estimates, see Appendix B.1. Neither extension has a first-order effect on our estimated belief distribution. Moreover, our estimates suggest that hedging demand plays a minimal role in retail investment in ETFs.[15] For these reasons, we proceed with the more parsimonious model to develop our main results.

The investor's problem is to choose the leverage category that maximizes her indirect utility

$$\max_{j\in\{-3,-2,-1,0,1,2,3\}} \beta_j\mu_i - p_j - \frac{\lambda}{2}{\beta_j}^2\sigma^2. \tag{2}$$

An investor chooses leverage j if and only if it maximizes her subjective risk-adjusted return relative to the other leverage choices $k \neq j$.[16] So an investor who chooses j prefers leverage j to leverage $j-1$ such that

$$u_{ij} = \beta_j\mu_i - p_j - \frac{\lambda}{2}{\beta_j}^2\sigma^2 > \beta_{j-1}\mu_i - p_{j-1} - \frac{\lambda}{2}{\beta_{j-1}}^2\sigma^2 = u_{ij-1}.$$

This inequality can be re-written to provide a lower bound on investor i's expectation of future stock market returns:

$$\mu_i > \frac{\lambda}{2}\left(\beta_j^2 - \beta_{j-1}^2\right)\sigma^2 + p_j - p_{j-1}, \tag{3}$$

noting that $\beta_j - \beta_{j-1} = 1$. Intuitively, investor i must believe that the stock market return μ_i is sufficiently high to offset the incremental change in risk $\frac{\lambda}{2}\left(\beta_j^2 - \beta_{j-1}^2\right)\sigma^2$ and fees $p_j - p_{j-1}$ associated with leverage j over leverage $j-1$. Similarly, an investor who chooses j prefers leverage j to leverage $j+1$ such that

$$u_{ij} = \beta_j\mu_i - p_j - \frac{\lambda}{2}{\beta_j}^2\sigma^2 > \beta_{j+1}\mu_i - p_{j+1} - \frac{\lambda}{2}{\beta_{j+1}}^2\sigma^2 = u_{ij+1},$$

generating an upper bound on investor i's expectation of future stock market returns:

$$\mu_i < \frac{\lambda}{2}\left(\beta_{j+1}^2 - \beta_j^2\right)\sigma^2 + p_{j+1} - p_j. \tag{4}$$

In words, the above inequality implies that investor i's expectation of future stock market returns is not sufficiently high to offset the incremental change in risk and fees to justify purchasing leverage category $j+1$ over j.

Inequalities (3) and (4) imply that an investor's optimal leverage choice is simply a function

[15] The estimated mean expectation of stock returns with hedging demand is highly correlated with the estimated mean expectation in our main results with a correlation coefficient of 0.98.

[16] We assume that investors are aware of all of the available leverage categories $j \in \{-3,-2,-1,0,1,2,3\}$. In practice, it is possible that some investors were not aware of all of the available leverage categories, especially the -3x and 3x categories that were introduced in 2009 (after the introduction of the 2x and -2x categories). As an additional robustness check in Appendix B.4, we estimate an alternative version of the model where we combine the 2x and 3x leverage categories and the -2x and -3x leverage categories. While this approach inevitably "throws away" some data because we are not using information about the investors who chose a 3x ETF over a 2x ETF or a -3x ETF over a -2x ETF, it is useful for understanding our results. We find that the estimated beliefs and risk aversion parameters from this alternative specification are quantitatively similar to our baseline estimates.

of her expectation μ_i. We assume that every leverage category j is optimal for some investors, i.e., there exists some μ_i that satisfies both (3) and (4) for all j.[17] Therefore, an investor chooses leverage category j if and only if

$$\frac{\lambda}{2}\left(\beta_{j+1}^2 - \beta_j^2\right)\sigma^2 + p_{j+1} - p_j > \mu_i > \frac{\lambda}{2}\left(\beta_j^2 - \beta_{j-1}^2\right)\sigma^2 + p_j - p_{j-1}.$$

Given the distribution of beliefs $F(\cdot)$, the share of investors purchasing leverage j, s_j, is then

$$s_j = F\left(\frac{\lambda}{2}\left(\beta_{j+1}^2 - \beta_j^2\right)\sigma^2 + p_{j+1} - p_j\right) - F\left(\frac{\lambda}{2}\left(\beta_j^2 - \beta_{j-1}^2\right)\sigma^2 + p_j - p_{j-1}\right). \tag{5}$$

The above market share equation captures the probability that any given investor would purchase leverage j. This relationship is at the heart of our estimation strategy described below. Given market share data s_j and product characteristics p and σ, we can recover the preference parameter λ and the distribution of expectations $F(\cdot)$.[18]

3.2 Identification

We now describe how risk aversion λ and the distribution of expectations $F(\cdot)$ are nonparametrically identified using aggregate market share and product characteristic data. As described in Section 4, we allow F to vary over time when we estimate beliefs. Here, we provide conditions to identify the distribution that applies in each relevant period. We then briefly discuss the empirical variation in our data to illustrate how identification works in our setting.

3.2.1 Nonparametric Identification Using Cross-Sectional and Within-Quarter Variation

Identification is obtained by using two sources of variation. The first source of variation comes from the menu of choices facing investors. We choose parameters so that the implied shares from the model match the shares chosen by investors, relying only on cross-sectional variation in investor expectations. The second source of variation comes from time series variation in fees and volatility. For our baseline results, we allow fees and volatility to vary at a higher frequency than the parameters governing the distribution of investor expectations. This modeling assumption allows us to use variation in these observables in estimation.

We start by describing in more detail the role of cross-sectional variation in leverage choice. By revealed preference, an investor that chooses a leverage category of 2x has a higher expected return than an investor that chooses a 1x ETF, and a lower expected return than an investor that

[17] In other words, we assume that no leverage is dominated by another leverage. This can be tested empirically for any set of parameters. Because $\beta_{j+1}^2 - \beta_j^2 = 2j + 1$ and $\beta_j^2 - \beta_{j-1}^2 = 2j - 1$, this assumption can be written as the condition $\lambda\sigma^2 > (p_j - p_{j-1}) + (p_j - p_{j+1})$ for interior j ($j \neq \{-3, 3\}$). Intuitively, prices for leverage j can not be too high relative to the nearby leverage categories.

[18] In our baseline setup, we assume that investors' trading decisions are perfectly sensitive to their own beliefs. This assumption is consistent with one of the main facts in Giglio et al. (2019) that, conditional on trading, an investor's trading decisions are highly correlated with her beliefs. If investors are not perfectly sensitive to their own beliefs, then our methodology would recover attenuated beliefs, i.e., we would underestimate the true dispersion in beliefs.

chooses a 3x return. By observing the market shares of purchases in each leverage category, we can pin down features of the distribution of investor expectations.

Formally, the distribution of expectations is semi-parametrically identified by the shares of investors in each leverage category, similar to identification in an ordered probit or logit model. For notational convenience, let S_j denote the cumulative share of investors purchasing all leverage categories $k \leq j$: $\sum_{k=-3}^{j} s_k$, where j ranges from -3 to 3. We can add up the shares from equation (5) to obtain a system of equations satisfying

$$S_j = F\left(\frac{\lambda}{2}(2j+1)\sigma^2 + p_{j+1} - p_j\right), \tag{6}$$

where $2j+1 = \beta_{j+1}^2 - \beta_j^2$ for all $j < 3$. S_3 is always equal to 1 and is not informative. The right-hand side elements depend on the observed characteristics σ, p_{j+1}, and p_j, as well as the unknown parameter λ and the distribution F. Because we observe six unique cutoff points in our data, $\{S_j\} = \{S_{-3}, S_{-2,}S_{-1}, S_0, S_1, S_2\}$, we have a system of six equations, allowing us to identify up to six parameters in each period. The parameters of interest include the risk aversion parameter λ and a period-specific distribution for F. The coefficient on fees is known to be 1, as utility is normalized to annual returns. Thus, the distribution of F can be estimated as a flexible distribution of up to five parameters. For example, if F is parameterized as normal, then F has two degrees of freedom (mean and variance) to attempt to fit the six observed values of $\{S_j\}$.

Our second source of variation, which allows us to obtain full nonparametric identification, comes from time series variation in fees and volatility. Changes in these variables change the cutoff points between leverage classes that are captured in equation (6). Intuitively, by observing how changes in fees affect the choice of leverage relative to volatility, we can pin down the scale of risk aversion. We provide a formal argument below.

Changes in fees alone can be sufficient to identify risk aversion and obtain nonparametric identification of the distribution of expectations. Formally, suppose that there exist two realizations of the data (σ, p_j, p_{j+1}) and $(\tilde{\sigma}, \tilde{p}_j, \tilde{p}_{j+1})$ for which $\tilde{\sigma}^2 = \sigma^2$ and $\tilde{S}_k = S_j$ for $k \neq j$. Then it must be that $\frac{\lambda}{2}(2j+1)\sigma^2 + p_{j+1} - p_j = \frac{\lambda}{2}(2k+1)\sigma^2 + \tilde{p}_{k+1} - \tilde{p}_k$. Therefore, we have $\frac{\lambda}{2}\sigma^2(2j-2k) = (\tilde{p}_{k+1} - \tilde{p}_k) - (p_{j+1} - p_j)$, or $\lambda = \frac{(\tilde{p}_{k+1} - \tilde{p}_k) - (p_{j+1} - p_j)}{\sigma^2(j-k)}$. The risk aversion coefficient is identified from the data. Because the coefficient on price is normalized to 1, we have identified the distribution at the quantile $F^{-1}(S_j)$. In other words, once we have identified λ, both S_j and the argument of F are known, so we have nonparametric identification of the distribution F at S_j. Furthermore, we only have to identify λ once, so two realizations of the data can provide identification at all quantiles $\{S_j\} \cup \{\tilde{S}_j\}$, which are the cutoff values in equation (6).

Likewise, changes in volatility can aid in identification. Intuitively, if we observe the same realization of market shares from the same belief distribution, but prices have changed, then it must be the case that changes in volatility have exactly offset the changes in prices for the marginal investor. Formally, consider two different realizations of the data (σ, p_j, p_{j+1}) and $(\tilde{\sigma}, \tilde{p}_j, \tilde{p}_{j+1})$ for which $S_j = \tilde{S}_j$. Then, it must be that $F^{-1}(S_j) = F^{-1}(\tilde{S}_j)$, or $\frac{\lambda}{2}(2j+1)\sigma^2 + p_{j+1} - p_j =$

$\frac{\lambda}{2}(2j+1)\tilde{\sigma}^2 + \tilde{p}_{j+1} - \tilde{p}_j$. The risk aversion coefficient is then $\lambda = 2\frac{(\tilde{p}_{j+1}-\tilde{p}_j)-(p_{j+1}-p_j)}{(2j+1)(\sigma^2-\tilde{\sigma}^2)}$. More generally, this exactness can be relaxed by using a local approximation to estimate how leverage market shares vary with respect to variation in prices, $\frac{\partial S_j}{\partial p_j}$, and volatility $\frac{\partial S_j}{\partial \sigma^2}$. Because $\frac{\partial S_j}{\partial p_j} = -f(\cdot)$ and $\frac{\partial S_j}{\partial \sigma^2} = \frac{\lambda}{2}(2j+1)f(\cdot)$, we can recover λ as $\lambda = \frac{-2}{2j+1}\frac{\frac{\partial S_j}{\partial \sigma^2}}{\frac{\partial S_j}{\partial p_j}}$.

By assumption, we hold the distribution of investor expectations fixed over a certain period.[19] This is a key assumption, as it implies that within-period changes in fees and volatility are independent of $F(\cdot)$. If—contrary to our model—the expectation distribution had within-period changes, and if these changes were also correlated with changes in fees and volatility, then this could bias our results. Since we estimate each period separately, correlation between the distribution of expectations and fees/volatility *across* periods is not problematic. Thus, the only concern is potential within-period endogeneity.[20]

In the data, we find that ETF fees are unlikely to generate this concern. We observe that fees are relatively fixed in the short run, with the largest changes occurring at an annual frequency. Further, our sample period coincided with a "cut-throat price war" in the ETF industry, which leads us to believe that much of the variation in fees is driven by supply-side factors—i.e., greater competition and increased scale—over this period.[21] These observations help alleviate concerns that ETF issuers are endogenously changing fees at a high frequency in response to changes in investor expectations.

An alternative approach that eliminates the within-period endogeneity concern is to estimate F at a monthly frequency, leveraging only the cross-sectional variation in investor expectations. We pursue this approach and provide an alternative set of estimates in Appendix A. The alternative estimates closely resemble our main results, though they are not identical. The differences we do observe between the alternative estimates and our baseline results indicate that variation in fees and volatility does play some role in pinning down investor expectations.

3.2.2 Illustration Using Empirical Variation

Our main empirical results use both sources of variation described above. We estimate the belief distribution at the quarterly level, allowing monthly variation in prices and volatility to assist in identification. In each month, we observe the market shares of each of the leverage categories, allowing us to construct the cumulative shares, $\{S_j\}$. In addition, we observe the corresponding fees, $\{p_j\}$, and implied volatility, σ^2. Given the risk aversion parameter, λ, we can compute the investor expectations needed to rationalize the shares as per equation (6). If S_j investors purchased

[19] We do not require individual investor expectations to be fixed over the period, only the aggregate distribution.

[20] A modification to our identification argument with exogenous fees shows how this could bias our estimates. For example, consider the case where fees systematically increase for leverage categories that see increased demand from changes in expectations. Formally, consider two different realizations of the data (σ, p_j, p_{j+1}) and $(\tilde{\sigma}, \tilde{p}_j, \tilde{p}_{j+1})$ for which volatility is constant and $\tilde{S}_k = S_j$ for $k = 1$ and $j = 2$. Fees increase for $j = 2$, and, at the same time, investors become more optimistic such that $\tilde{F}$ first-order stochastically dominates F. Then $F^{-1}(S_2) < \tilde{F}^{-1}(\tilde{S}_1)$, implying $\lambda < \frac{(\tilde{p}_2-\tilde{p}_1)-(p_3-p_2)}{\sigma^2}$ instead of $\lambda = \frac{(\tilde{p}_2-\tilde{p}_1)-(p_3-p_2)}{\sigma^2}$. In this example, our estimates of risk aversion would be biased upward, because we would infer that investors are less sensitive to fees (relative to volatility) than they actually are.

[21] For example, see https://www.ft.com/content/10238d8e-b320-4667-944d-d463e7311213.

leverage category j or lower, this implies that S_j fraction of investors believe that the stock market will increase by less than $\frac{\lambda}{2}(2j+1)\sigma^2 + p_{j+1} - p_j$. Otherwise, they would have purchased a higher leverage category. Thus, each cutoff point provides information about F, the distribution of expected returns.

Figure 3 illustrates how both sources of identification work empirically. Each panel corresponds to the cross section of expectations across investors in a given quarter. The *y*-axis represents the value of cumulative shares, $\{S_j\}$, and the *x*-axis represents the investor expectations needed to rationalize the leverage choice given some risk aversion parameter, $\left(\frac{\lambda}{2}(2j+1)\sigma^2 + p_{j+1} - p_j\right)$. Cumulative shares ($\{S_j\}$), implied volatility ($\sigma$), and fees ($p$) are directly observed. Risk aversion can be inferred because different values of λ will change the implied beliefs needed to rationalize investors' choices, shifting where the observations fall on the *x*-axis. Observations in the figure are at the monthly level, which is the unit of observation in our baseline analysis. The plotted points correspond to one quarter of data. Each of the three months in a quarter is represented by a different color. In each month, we have six observations, corresponding to the $J-1$ leverage categories and the cutoff points implied by leverage, fees, and volatility. Each cutoff value is represented by a different shape.

The estimation objective is to choose parameters for risk aversion and expectations that best allow us to fit a CDF for F through the 18 observed points in the quarter. As discussed in the previous paragraph, different guesses for the risk aversion parameter lead to different cutoff points, shifting the horizontal locations of the markers in Figure 3. Different guesses for parameters of the CDF change the curvature of the black line. We jointly choose risk aversion and the distribution of beliefs (F) to best fit the data.

Panel (a) in the figure shows a quarter in which the cross-sectional (within-month) variation across leverage categories acts as a primary source of identification in pinning down the CDF. Because few investors choose the highest-leverage categories, the cutoff points in a given month span most of the distribution. Variation in fees and leverage provide small, localized variation because cutoff points and cumulative shares are stable across months in that quarter. Alternatively, panel (b) provides an example quarter in which within-quarter variation in fees and volatility plays a more substantial role in identification. In this case, larger within-quarter variation in fees or volatility causes the observations for each leverage category to be more dispersed across months, especially along the *x*-axis. A few observations overlap for different values of leverage, approximately meeting our condition for nonparametric identification of risk aversion. An example from panel (b) corresponds to $\tilde{S}_{-1} = S_{-2}$, where $\tilde{S}_{-1}$ is the third blue marker (solid blue circle) from the left and S_{-2} is the second purple marker from the left (hollow purple square). When $\tilde{S}_{-1} = S_{-2}$, this implies that investors at the cutoff for $\tilde{S}_{-1}$ must have the same beliefs as investors at the cutoff for S_{-2}. This pins down risk aversion. Recall that once risk aversion is identified, the distribution of beliefs F is nonparametrically identified as per equation (6).

Panels (c) and (d) further illustrate how the risk aversion parameter is pinned down by our model. In panel (c), we replicate the observations and estimated CDF from panel (b) in black.

Note that the scale differs from panel (b). In red, we plot a set of observations and CDF that would be obtained with a risk aversion parameter that is five times greater. In panel (d), we again replicate the observations and estimated CDF from panel (b) in black. In blue, we plot the observations corresponding to a risk aversion parameter that is five times smaller. We cannot fit a CDF for expected returns through the observations in panel (d) because the expected return cutoff points are not monotonically increasing with leverage. At that level of risk aversion, our model predicts that no investor would choose a 2x ETF over a 3x ETF. Thus, the smaller risk aversion parameter is inconsistent with what we observe in the data. The larger risk aversion parameter in panel (c) admits a CDF, but the fit is poor compared to the estimated risk aversion parameter. The fit is poor because the larger risk aversion parameter dramatically increases the dispersion in the cutoff points for a given leverage category within a quarter despite there being little change in market shares (i.e., the solid red square markets). The variation in the relative horizontal location of each observation illustrates how our approach pins down the risk aversion parameter.

3.3 Heterogeneous Risk Aversion

The main objective of this paper is to estimate a parsimonious model that allows us to recover the distribution of investor beliefs. However, we also consider the natural extension of the model where investors have heterogeneous risk aversion $\lambda_i \sim G(\cdot)$.

$$u_{ij} = \beta_j \mu_i - p_j - \frac{\lambda_i}{2} {\beta_j}^2 \sigma^2.$$

Here, we assume that investors agree over the volatility of the S&P 500 Index but have heterogeneous risk aversion. This framework corresponds to the random coefficients and latent class/mixture ordered choice models,[22] and also relates more generally to the random coefficients models commonly used in the demand estimation literature (Berry et al., 1995). As discussed in Section 3.1, one could recast the model of heterogeneous risk preferences into an empirically equivalent model where investors have heterogeneous beliefs over the volatility of the stock market.

With heterogeneity in risk aversion, the share of investors choosing leverage j is

$$s_j = \int \left[F\left(\frac{\lambda_i}{2} \left(\beta_{j+1}^2 - \beta_j \right) \sigma^2 + p_{j+1} - p_j \right) - F\left(\frac{\lambda_i}{2} \left(\beta_j^2 - \beta_{j-1}^2 \right) \sigma^2 + p_j - p_{j-1} \right) \right] dG(\lambda_i).$$

Identification of heterogeneity in risk preferences comes from variation in the substitution patterns with different levels of volatility similar to identification in Berry et al. (1995).[23] In the above section, we showed that two realizations from the data are sufficient to pin down a single risk aversion parameter. If we observe more than two realizations of the data that generate the same quantile, then we have multiple measures of the risk aversion parameter. These can be used as overidentifying restrictions to reject a model of homogeneous risk aversion, or, intuitively, these additional

[22] See Chapter 8 of Greene and Hensher (2010) for a discussion of the literature.
[23] See Cunha et al. (2007) for further discussion of ordered choice models.

realizations can be used to pin down properties of the distribution of risk aversion coefficients.

3.4 Discussion and Alternative Interpretations

Our model makes a few key assumptions that merit discussion. First, we assume that investor expectations about future stock market performance can be collapsed into a single expected return. We do not view this assumption as particularly problematic. Investor uncertainty will be absorbed by the risk aversion parameter in our model. Implicitly, the parameter captures both market-level uncertainty and investor-specific uncertainty, as described above. Investor-specific uncertainty may reflect both forecast uncertainty and beliefs about volatility.

Second, we assume that the investor is making a discrete decision to invest a certain amount of wealth in these ETFs. The discrete choice assumption rules out behavior where an investor splits their wealth between two different leverage categories. The way we justify this assumption is the standard approach in empirical discrete choice models: we allow, in theory, individual investors to have multiple realizations from the distribution $F(\cdot)$. Thus, μ_i may represent different perspectives within a single individual, without any modification to the model.

Third, we assume investors only focus on financial characteristics of ETFs summarized by expected return and volatility. Non-financial characteristics such as fund issuer marketing, distribution channels, and brand recognition are ruled out. Issuers of leveraged ETFs offer almost the entire menu of leverage choices, and so they are unlikely to steer investors toward a specific leverage. By omitting ETF-specific demand shocks, we could potentially overstate the expected return needed for investors to shift from a tracker to a 2x ETF if investors have a brand preference for the three (more well-known) issuers that only offer trackers.

Finally, we do not make any assumptions about an investor's investment horizon, nor does our method require one. As discussed previously, without any loss in generality, we normalize an investor's disutility from fees p to one, which allows us to interpret investor expectations over stock market growth μ_i in term of annual returns. Note that this does not mean that we assume that the investor intends to hold the ETF for a year. Rather, it means that the investor expects the S&P 500 to grow at rate of μ_i, where μ_i is measured in annualized returns.[24] We do not directly observe each investor's investment horizon. In the data, investors in leveraged and inverse-leveraged ETFs tend to have similar holding periods, which are less than one month on average.

4 Estimation

4.1 Empirical Model

Following our framework in Section 3, we develop and estimate an empirical model of investor leverage choice. We allow the distribution of investor expectations to vary over time, estimating

[24] In other words, our methodology allows us to recover an investor's expectations about the growth rate of the S&P 500 rather than the investor's beliefs about what the level of the S&P 500 should be.

F_s for each set of periods s. The subscript s indexes time-varying distributions and also the set of months T_s for which the distribution applies, i.e., the distribution F_s applies to any period $t \in T_s$. In our baseline specification, we estimate the model using monthly data and allow the distribution of expectations to vary at the quarterly level such that $|T_s| = 3$. We estimate the expectation distribution via maximum likelihood. The likelihood contribution of an investor who chooses leverage j is $F_s(x_{jt}) - F_s(x_{(j-1)t})$, where F_s is the distribution of expectations and x_{jt} is the cutoff utility value corresponding to the expected return that renders an investor indifferent between choice j and choice $j+1$. Let a_i denote the leverage choice for investor i and N_t denote the number of potential investors in period t. Then, the likelihood component for F_s is

$$\prod_{t \in T_s} \prod_{i \in N_t} \prod_{j \in J} \left(F_s(x_{jt}) - F_s(x_{(j-1)t})\right)^{\mathbf{1}[a_i = j]} \tag{7}$$

and the log-likelihood is

$$\sum_{t \in T_s} \sum_{i \in N_t} \sum_{j \in J} \mathbf{1}\left[a_i = j\right] \ln\left(F_s(x_{jt}) - F_s(x_{(j-1)t})\right). \tag{8}$$

We observe market share data, rather than individual choices. We sum over the (latent) individuals in each period and scale by N_t to obtain the following expression for the log-likelihood

$$\sum_{t \in T_s} \sum_{j \in J} s_{jt} \ln\left(F_s(x_{jt}) - F_s(x_{(j-1)t})\right). \tag{9}$$

The parameter vector, θ, characterizes the time-varying distribution F_s and risk aversion λ. Our estimate $\hat{\theta}$ is chosen to maximize the log-likelihood. We parameterize F_s as a skewed t distribution with four parameters. The parameters correspond to location, dispersion, skewness, and kurtosis; these are further described in Table 2.[25] The four-parameter skewed t distribution is a flexible distribution that nests other common distributions such as the Normal and Cauchy distributions. We estimate location, dispersion, and skewness separately for each three-month period, while holding kurtosis fixed for the entire sample. As discussed in Appendix A, we also re-estimate the model where we allow the location, dispersion, and skewness to vary at the monthly rather than quarterly level, and we find quantitatively similar results.

x_{jt} is the utility index and is parameterized as

$$x_{jt} = \frac{\lambda}{2}\left(\beta_{j+1}^2 - \beta_j^2\right)\sigma_t^2 + p_{(j+1)t} - p_{jt}$$

where ETF leverage (β_j) and fees (p_{jt}) are directly observed in the data, and we calculate implied volatility (σ_t^2) based on the VIX. In our baseline specification, we hold λ constant over time. Section

[25] We parameterize the skewed t distribution as $F(x_{jt}) = \tilde{F}\left(\frac{x_{jt}-a}{b}\right)$. The corresponding density $\tilde{f}$ is given by $\tilde{f}(x) = \frac{2}{c+1/c} g(cx)$ for $x < 0$ and $\tilde{f}(x) = \frac{2}{c+1/c} g(\frac{x}{c})$ for $x \geq 0$, where c is the skewness parameter and g is the density of a t distribution with degree of freedom d. In estimation, we use the skewt package in R for calculating $\tilde{F}$.

4.4 includes an extension of the model where we allow risk aversion to vary annually.

Thus, we estimate three parameters in each quarter, corresponding to the time-varying distribution of expectations, plus the kurtosis parameter and an additional parameter to capture risk aversion. Since we have 11 years and 44 quarters of data, we estimate 134 parameters in total. In alternative specifications, we allow λ_i to be heterogeneous across investors and vary over time.

4.2 Baseline Results

Our estimates for investor expectations are plotted in Figure 4. Panel (a) shows the distribution of time-varying expectations in each quarter. The mean expectation is plotted with red dots and the median is plotted with a solid red line. Dashed lines show the 25th and 75 percentiles, and dotted lines show the 10th and 90th percentiles. The estimated time-varying parameters that characterize the distribution are displayed in the other three panels. Panels (b), (c), and (d) plot the estimates for the location, dispersion, and skewness parameters, respectively. 95 percent confidence intervals are displayed with dashed lines and are calculated using the maximum likelihood formula for asymptotic standard errors. Here, we describe and interpret our baseline estimates of investor expectations. In Section 5 we further study the evolution of and the factors driving investor expectations.

Our estimates of investor expectations in Figure 4a suggest that investors became substantially more pessimistic surrounding the 2008 financial crisis and that pessimism persisted for several years after the crisis. During the crisis, the average investor's expectation of the market risk premium fell by over 20 percent and remained below zero for the following two years. Over our whole sample the average expected market risk premium of the median investor in our sample is roughly 3 percent, which is slightly smaller than other estimates in the literature (Welch, 2000; Graham and Harvey, 2008).

We find that there is a large variation in the dispersion of expectations across investors over time. The changing dispersion in investor expectations is captured by our dispersion parameter, shown in panel (c), which is roughly analogous to the standard deviation. Investors have greater disagreement during the crisis, as can be seen in the large differences between the 90th and 10th percentile of expectations from 2008 to 2011. At the most extreme, our estimated mean expectation in 2008 Q4 is an annualized return of less than -20 percent. In this quarter, we estimate that 10 percent of investors thought the return on the S&P 500 would be worse than -67 percent. The results suggest that disagreement tends to rise in times of crisis. As illustrated in Figure 4a, there is also a substantial increase in disagreement among investors surrounding the 2011-2012 European Sovereign Debt Crisis and the 2015-2016 Chinese stock market turbulence. From 2016 to 2018, we estimate that investors had much less disagreement about the future return of the stock market. The expectation distribution has remained more stable with tighter bands between the 90th and 10th percentiles.

We estimate that the distribution tends to have a negative skew. In panel (d), this corresponds to $c_s < 1$. This affects the overall distribution by lowering the mean relative to the median, which

can be seen in panel (a). Skewness has the greatest effect on the mean in 2008 Q4, when the dispersion in expectations is highest. This suggests that a mass of investors became particularly pessimistic during the financial crisis.

While the mean expectation fell during the financial crisis, we find that the median investor became slightly more optimistic during the financial crisis. As we explore further in Section 5.1, this suggests that a large fraction of investors are extrapolative and became substantially more pessimistic during the financial crisis, while an even larger fraction are contrarian and became more optimistic as the stock market fell. This is consistent with the evidence in Luo et al. (2020) that many retail investors tend to trade as contrarians.

We summarize our estimated parameters in Table 3. For our time-varying parameters, we report the median value and the corresponding standard errors. We report our time-invariant parameter for kurtosis, which reflects how much of the distribution lies in the tails. Our estimated kurtosis parameter of 1.26 implies fat tails that are roughly in line with the Cauchy distribution.[26] Our estimated risk aversion parameter of $\lambda = 0.98$ implies that investors are willing to pay an additional 39 basis points in fees for a one standard deviation reduction in volatility.[27] One caveat to interpreting the risk aversion parameter is that our implied volatility measure VIX includes both physical volatility and a variance risk premium term, which is usually positive. When we scale the VIX-based estimates of λ by the average ratio of VIX to realized volatility (a factor of 1.77), we obtain an average parameter value 1.73. To put these numbers in perspective, our estimate of λ is lower than other risk aversion estimates traditionally found in the literature. For example, using life cycle models Fagereng et al. (2017) estimate relative risk aversion of 7.3, Calvet et al. (2019) estimate relative risk aversion of 5.8, and Meeuwis (2019) estimate relative risk aversion of 5.4. These differences are not necessarily surprising, given our distinct population and the fact that our parameter may capture additional uncertainty. As described in Section 3.1, estimates of risk aversion coefficients may not be directly comparable to the extent that investors have heterogeneity in beliefs about volatility and forecast uncertainty.

4.3 Heterogeneous Risk Aversion

In our baseline specification, we hold the risk aversion parameter fixed for all investors. We also estimate a specification in which investors have heterogeneous preferences for risk. As discussed above, this assumption is isomorphic to a model in which investors have heterogeneous beliefs about the volatility of the stock market.

Formally, we assume that $\lambda_i \sim G(\cdot)$, where λ_i is independent from investor expectations μ_i. We parameterize G as a uniform distribution. As above, we estimate our model using maximum likelihood, while integrating out the distribution for λ_i. The estimated parameters are summarized

[26] Technically, our estimates imply that moments higher than the mean are not defined. Hence, we talk about a dispersion parameter rather than a standard deviation. For convenience, we use the terms skewness and kurtosis, whose corresponding moments are not defined.

[27] This is computed as $\frac{\lambda}{2} sd(\sigma)^2$, where $sd(\sigma)$ denotes the standard deviation of VIX in our sample.

in Table 3. We report our estimate of G in terms of its midpoint and dispersion, where dispersion captures the distance from the midpoint to the upper and lower bounds.

Incorporating heterogeneity in risk aversion makes little difference to our overall estimates. We estimate a risk aversion distribution of $\lambda_i \sim U[0.71, 1.00]$. Thus, the midpoint of 0.85 is slightly lower than the constant risk aversion parameter estimate of 0.98. The other parameters are only slightly affected by the change. Figure 5 provides a comparison of the two specifications. The top three panels correspond to the specification with fixed risk aversion, and the bottom three panels correspond to the specification with heterogeneous risk aversion. Panels (a) and (d) show the distribution of investor expectations, which track each other closely. Panels (c) and (f) show the fit of log shares, where the x-axis represents the log shares in the data and the y-axis represents the fitted shares in the model.

For a more specific comparison, we plot the distribution of investor expectations for a single period in panels (b) and (e). These panels show the PDF of expectations in September 2009, which is plotted in yellow. The vertical blue lines correspond to the cutoff points of indifference between leverage categories, in terms of excess return. The area under the yellow line between two vertical blue lines corresponds to the model-predicted shares for a particular leverage. For example, investors with expectations between $\mu_i = 11$ and $\mu_i = 16$ would choose 2x leverage. Comparing panel (b) to panel (e), we see that incorporating heterogeneity in risk aversion compresses the cutoff points toward zero, though this effect is small. For example, the implied expectation to choose 1x leverage over the outside option is $\mu_i = 3.3$ in our baseline specification and $\mu_i = 2.9$ with risk aversion heterogeneity.

4.4 Time-Varying Risk Aversion

We also estimate an extension of our baseline model where we allow the risk aversion coefficient to vary annually. We rely on the menu of choices and the empirical variation in fees and volatility within each calendar year to identify the risk aversion of that year.

Figure 6a displays our estimates for investor expectations with time-varying risk aversion, which are qualitatively similar to the distribution of expectations in our baseline model. The estimated mean expected returns in these two models are highly correlated with a correlation coefficient of 0.96. The average risk aversion is lower but similar to the level in our baseline model (0.61 vs 0.98). The most notable difference from the baseline model is that the dispersion of expectation becomes smaller during periods with high volatility, especially in 2008 and 2009. As shown in Figure 6b, we estimate lower risk aversion during those periods, and hence we need less dispersion in beliefs to rationalize leverage choices in the data.

One interesting pattern in the data is that the risk aversion parameter λ appears to be procyclical. Our estimated risk aversion parameter λ is relatively low at the start of the financial crisis and then increases over time in our sample. One important caveat, as discussed in Sections 3.1 and 3.3, is that it is difficult to separately identify risk aversion from variation in risk perceptions more generally. While this does not impact the identification of the belief distribution, it does impact the

interpretation of the parameter λ. When λ reflects risk perceptions more generally, rather than just risk aversion, there are a handful of intuitive explanations for why λ might have been low during the financial crisis and increased over time. For example, as a result of either salience (Bordalo et al., 2012) or prospect theory (Kahneman and Tversky, 1979), it is possible that during the financial crisis investors became less risk averse as they entered the loss domain and then, during the subsequent stock market boom, became more risk averse as they entered the gain domain.[28] Relatedly, Lian et al. (2019) and Jensen (2018) suggest that, as a result of saliency, an increase in an investor's expected returns relative to the risk-free rate could lead the investor to become more risk tolerant. Recent evidence from Gormsen and Jensen (2020) suggests that tail risk is perceived to be larger in good times, and evidence from Jensen (2018) suggests risk aversion is procyclical. These results are consistent with our findings.

In Appendix B.2, we also consider several additional model extensions where we allow the parameter λ to be both time-varying and heterogeneous, and allow investor beliefs to be correlated with λ across investors and over time. The estimated distribution of beliefs in these alternative models is qualitatively similar to our baseline specification where risk aversion is homogeneous across investors and constant over time. The correlations between the mean belief from our baseline model and the mean beliefs from these extensions are more than 0.90. While these extensions slightly improve the fit of the model, they do so at the expense of adding many more parameters. Consequently, our more parsimonious baseline model ranks higher than these extensions in terms of common criterion used for model selection (e.g., BIC, AIC). See Table A1 for a summary of goodness-of-fit for different specifications.

4.5 Comparison with Survey Data

We examine how our estimates of investor beliefs compare with survey responses, which have been previously used to understand the formation of beliefs (Vissing-Jorgensen, 2003; Ben-David et al., 2013; Greenwood and Shleifer, 2014; Nagel and Xu, 2019). We examine the following surveys/indices that are commonly used in the literature: the Duke CFO Global Business Outlook, the Wells Fargo/Gallup Investor and Retirement Optimism Index, the University of Michigan Survey of Consumers, the American Association of Individual Investors (AAII) Sentiment Survey, the Shiller U.S. Individual One-Year Confidence Index, and the Survey of Professional Forecasters. An advantage of surveys is that they can be constructed to be representative of a desired target population of individuals; conversely, the advantage of our revealed preference approach is that it is based on the actual decisions of individuals, albeit from a specific subset of the population.

Each survey asks different questions to elicit investor beliefs about the stock market. For example, the Duke CFO Global Business Outlook asks survey respondents to report what they believe the stock market will return over the course of the next year, while the Shiller U.S. Individual One-Year

[28] According to the prospect theory (Kahneman and Tversky, 1979), individuals have an S-shaped value function such that they are risk averse with respect to gains but risk loving with respect to losses. Saliency theory (Bordalo et al., 2012) generates similar types of framing effects that are driven by the saliency of the payouts rather than the curvature of the S-shaped value function.

Confidence Index measures the percentage of respondents who expect the stock market to increase over the upcoming year.

Because we recover the full distribution of expectations, we can use our estimates to calculate the implied responses to each survey question. For example, our estimated mean corresponds to a survey that asks for expected return, whereas our estimated fraction of investors taking positive leverage corresponds to investors who think the stock market will increase. In principle, we can simulate survey statistics quite flexibly. Overall, the survey responses implied by the estimated distribution of beliefs from our model are statistically and positively correlated with the survey data. For additional details of each survey and scatterplots of the relationship, see Section D in the Appendix.

Table 4 evaluates the relationships between the survey data and our estimated distribution of expectations. All seven of the chosen measures from surveys are positively and significantly correlated (at the 1 percent level) with the corresponding statistic from our estimated distribution. Each survey captures a slightly different subset of investors in the population, so comparing the correlations across surveys potentially provides insight into the underlying ETF investors. We find that our estimated distribution of beliefs is more correlated with surveys that capture the beliefs of average retail investors (i.e., Wells Fargo/Gallup Investor and Retirement Optimism Index and the University of Michigan Survey) than with surveys that capture the beliefs of more sophisticated investors (i.e., Duke CFO Global Business Outlook Survey, Shiller Index, and the AAII Survey).[29] While these results suggest that our estimates of investor beliefs might be picking up the beliefs of average retail investors, we also find that our measure of beliefs is highly correlated with professional GDP forecasts. This is not necessarily surprising given that the University of Michigan and Gallup indices are highly correlated with professional GDP forecasts, which suggests that retail investor forecasts of stock market returns are highly correlated with expectations of economic growth.

A number of these surveys also report measures of dispersion in investor expectations. In Table 5, we report the relationship between the dispersion in our estimated beliefs and the corresponding survey measures. The results in column (1) indicate that there is a positive relationship between our estimated dispersion in investor expectations and dispersion in CFO expectations as measured by the Duke CFO Global Business Outlook Survey. Similarly, we find a positive relationship between the interquartile range of expected returns across investors from our model and the interquartile range of GDP forecasts across professional forecasters in the Survey of Professional Forecasters. Last, in columns (3) and (4) we examine the relationship between our estimated dispersion in investor expectations and other measures of uncertainty and dispersion used in the literature: the

[29] Each survey varies with respect to the underlying populations. In terms of sophistication, we might expect the Duke CFO survey, American Association of Individual Investors (AAII) survey, and Cash Shiller survey to reflect wealthier and more sophisticated investors as detailed in the Appendix. At the other end of the spectrum, we might expect the University of Michigan Survey, which covers the broad US population, to reflect less sophisticated investors. The Wells Fargo/Gallup Investor and Retirement Optimism Index is constructed based on a nationally representative survey of U.S. investors with $10,000 or more invested in stocks, bonds, and mutual funds, and the underlying survey population likely falls somewhere in the middle in terms of investor sophistication.

Equity Market Volatility Tracker from Baker et al. (2019) and the Financial Uncertainty Index from Ludvigson et al. (2020).[30] We find positive relationships between the dispersion in investor expectations and both of these alternative measures.

Overall, the results displayed in Tables 4 and 5 help shed light on the external validity of our estimates. The expectations we recover from demand for S&P 500 linked ETFs are highly and significantly correlated with the investor expectations measured in six different surveys. The results also indicate that the relationship between our estimated beliefs and beliefs elicited from survey data extends to other moments of the distribution beyond the mean.

4.6 Robustness Checks

We find that allowing for skewness and kurtosis, as we do in our baseline specification, provides estimates that best fit the data. However, for robustness, we also estimate the model using a normal distribution for expectations, where we allow the mean and standard deviation (the location and dispersion parameters) to vary over time. Using a normal distribution maintains several of the qualitative features of our baseline specification, but the model fit is worse. The normal distribution does a poor job fitting the fat tails of the expectation distribution, and it cannot account for skewness.

To test the sensitivity of our results to our definition of the outside option, we consider two alternative measures. In one specification, we scale the outside share by a factor of 5 rather than the average ratio of purchase volume to AUM, with the idea that outside options may not trade at the same frequency as the inside goods. We also consider a specification where we estimate the share choosing the outside option as a free parameter, rather than bringing in the data. Neither assumption makes a meaningful difference in our estimates. The resulting expectation distributions and the plots of model fit are displayed in Figure A1 of the Appendix.

We also present four sets of results discussed earlier as robustness checks. Appendix A provides results using only within-menu variation in choices, at both quarterly and monthly frequencies. Appendix B.1 provides a discussion of our more general model where we allow an investor's ETF choice to incorporate hedging demand as part of the investor's broader portfolio allocation problem. In this extension, investors account for how the ETF investment covaries with their wealth/portfolio, and we estimate this covariance term for each investor as a random coefficient. The corresponding estimated time-varying distribution of investor expectations is similar to the estimated distribution in our baseline specification. Appendix B.3 shows that when we use realized volatility instead of VIX, we obtain very similar estimates with time-varying risk aversion. Although VIX includes both physical volatility and also the variance risk premium, it is unlikely that variance

[30]The Equity Market Volatility tracker (Baker et al., 2019) is constructed based on monthly counts of newspaper articles that include keywords related to economy, market, or volatility across eleven major U.S. newspapers. Ludvigson et al. (2020) construct the financial uncertainty index based on a wide range of financial series, including valuation ratios, dividend and price growth rates, treasury and corporate bond yields, industry portfolio equity returns, and Fama and French risk-factors. For each of these series, the authors compute the conditional volatility of the unforecastable component and aggregate these individual conditional volatilities into an uncertainty index.

risk premium drives our time-varying risk aversion estimates. Lastly, Appendix B.4 presents results where we combine the market shares of positive leveraged (2x and 3x) and negative leveraged (-2x and -3x). The corresponding estimates are very similar to our baseline specification. This suggests that the underlying variation driving the belief distribution is mostly coming from an investors' choice to be positively leveraged or negatively leveraged rather than the specific leverage class they choose. The only change in the choice set of leverage categories in our sample is the introduction of 3x and -3x leverage in 2009. This extension shows that any lack of awareness for these new leverage categories is unlikely to bias our estimates for the distribution of beliefs.

5 Understanding Investor Expectations

In this section, we use our estimates to contribute to the understanding of how investors form expectations. First, we confirm a previous finding that, on average, investors extrapolate recent stock market returns when forming expectations. We contribute to the literature by showing how extrapolation impacts not only the mean expectations but also the variance and skewness. In other words, we show how historical returns are correlated with investor disagreement and pessimism. Second, we examine the persistence of beliefs and find that a one-time -10 percent return shock impacts investors' beliefs for up to two years into the future. Third, we show that investor forecast errors are predictable and violate full-information rational expectations, which is consistent with the vast evidence documenting the predictability of forecast errors. Last, we explore whether investor expectations of returns forecast future returns. While we find that the average expectation is uncorrelated with future returns, we find some evidence suggesting that dispersion in expectations is negatively correlated with future returns.

5.1 Determinants of Investor Expectations: Extrapolated Beliefs

There is a long theoretical and empirical literature highlighting the role of extrapolation in the formation of investor beliefs. We examine the relationship between past stock market returns and the expectations we recover from our model. An advantage of our model is that we recover the full distribution of beliefs, rather than just the mean or median, which allows us to examine how other moments, such as the standard deviation and skewness of beliefs, change in response to historical stock market returns.

We examine the relationship between the estimated mean expected excess return versus the previous year-over–year excess return of the stock market in the following regression

$$\overline{E[R]}_q = \alpha + \beta AnnualRet_q + \epsilon_q \tag{10}$$

where $\overline{E[R]}_q$ is the mean expected return from our model and $AnnualRet_q$ is the past one year excess return of S&P 500.[31] Observations are at the quarterly level.

[31]Here, we look at extrapolation based on the annual returns because they are potentially the most salient, but we

We report the corresponding estimates in column (1) of Table 6. Due to potential autocorrelation of the error term, we report t-statistics based on Newey and West (1987) with four lags. The results in column (1) indicate that a one percentage point increase in historical returns is correlated with a 0.11 percentage point increase in investor beliefs about the stock market return. The results also indicate that historical returns explain 58 percent of the variation in the mean expected return, suggesting that recent returns are first-order in explaining investor expectations.

Building on these results, we examine how other moments of the expectation distribution covary with recent stock market returns. Column (2) of Table 6 displays the regression estimates corresponding to the dispersion of expected returns across investors versus historical returns. The two series are negatively and significantly correlated. The estimates reported in column (2) indicate that a one percentage-point decrease in the past 12-month excess return of the stock market is correlated with a 2.6 percent increase in the dispersion parameter (which is analogous to the standard deviation of a normal distribution). The results suggest that there is a substantial increase in disagreement following negative returns, while investor beliefs become more homogeneous following positive returns.

Part of the increase in dispersion appears to be driven by the presence of contrarian investors. In column (3) of Table 6, we examine the relationship between median beliefs and historical returns. We estimate a small negative and statistically significant relationship between the median belief and the previous 12-month excess return of the stock market. The results indicate that a one percentage point increase in historical returns is correlated with an 0.04 percentage point decrease in the median belief. These results suggest that while the mean belief is extrapolative, the median investor is contrarian.

Understanding changes in the skew of the distribution helps reconcile the differences between the mean and median belief. Column (4) of Table 6 illustrates how the skewness of the distribution varies with recent stock market returns. The results indicate that investor expectations become more positively skewed following positive past returns. Conversely, investor expectations become more negatively skewed following negative returns. The results indicate that a one percentage-point increase in recent historical returns is correlated with a 0.26 percent increase in the skewness parameter. Combined, these findings suggest that there exists a mass of behavioral investors that become very pessimistic after a market downturn, making the belief distribution more negatively skewed and decreasing the mean expectation.[32]

find similar results if we look at shorter horizons.

[32]An advantage in our setting is that it is straightforward to apply our method across asset classes. We construct a panel of the distributions of investor beliefs across eight different asset classes at a quarterly level. Using this panel data set provides additional statistical power and insight into the formation of investor beliefs, allowing us to exploit cross-sectional variation in asset returns. In Appendix C, we exploit the panel structure of our data to show that the extrapolative behavior we document among S&P 500 investors, in terms of the mean, dispersion, and skewness of investor expectations, holds more generally across asset classes.

5.2 Persistence of Beliefs

Figure 4 panel (a) suggests that the financial crisis had a large and persistent impact on investor beliefs. After the decline in stock market in the late fall of 2008, the mean and skewness of investor expectations become more negative, and there is also a large increase in disagreement. As illustrated in the figure, these effects persist for up to two years.

We examine how the belief distribution evolves by estimating how the location, dispersion and skewness parameters of the distribution evolve as an AR(1) process.

$$\begin{aligned} Location_q &= \alpha_a + \beta_a MonthlyRet_q + \rho_a Location_{q-1} + \epsilon_{aq} \\ Dispersion_q &= \alpha_b + \beta_b MonthlyRet_q + \rho_b Dispersion_{q-1} + \epsilon_{bq} \\ Skewness_q &= \alpha_c + \beta_c MonthlyRet_q + \rho_c Skewness_{q-1} + \epsilon_{cq} \end{aligned} \tag{11}$$

Observations in equation (11) are at the quarterly level over the period 2008-2018. We examine how each parameter evolves as a function of the parameter value from the previous quarter, and also $MonthlyRet_q$, the previous monthly excess return of the stock market averaged across the current quarter. We report the corresponding estimates in Table 7. The results indicate that there is strong persistence in the belief distribution over time, as the AR(1) component of each parameter estimate is positive and significant. Consistent with our previous estimates, we also continue to find evidence that beliefs are extrapolative and impact multiple moments of the distribution.

Figure 7 displays the impulse response of how the expectation distribution evolves in response to large swings in the stock market. Panel (a) displays how investor expectations respond following a one-time one-month return of negative 10 percent. As illustrated in the figure, the mean expectation across investors immediately falls, and it remains negative for almost two years. One interesting feature, which is consistent with our previous results, is that the mean and the median move in opposite directions. While the mean expectation across investors falls, the median expectation increases in response to a negative return. This disagreement in measures of central tendency is possible because a negative stock market return has a large impact on the skewness and dispersion of the distribution of beliefs. Following the negative return, there is substantial disagreement among investors, and the interquartile range of investor expectations almost doubles. In response to the negative return shock, the 10th and 25th percentile of investors become dramatically more pessimistic while the median investor becomes more optimistic. The expected return among investors in the 10th percentile falls by more than 10 percent.

Panel (b) of Figure 7 shows how the expectation distribution evolves in response to a one-time one-month return of positive 10 percent. The average investor's expectation of future stock market returns jumps up and remains elevated for the next 1-2 years. In sharp contrast to the effect of a negative return, investor expectations become less dispersed in response to positive news about the stock market. Expectations among investors at the 25th and 75th percentiles of the distribution converge to the median in response to the recent positive stock market return. The interquartile

range among investor beliefs falls by half.

Our results suggest that investor beliefs are extrapolative and persistent, such that a change in recent returns has a profound impact on the mean, variance, and skewness of investor beliefs over the following two years. Our findings suggest that extrapolative and contrarian tendencies are stronger in response to poor returns, leading to increased disagreement during and after crises. Conversely, investors tend to agree more after periods of positive returns.

5.3 Are Beliefs Rational? Predictability of Forecast Errors

We explore the rationality of investor expectations by examining the predictability of forecast errors. The unpredictability of forecast errors is a necessary condition for rational forecasts. We construct forecast errors as

$$\epsilon_{q+4} \equiv AnnualRet_{q+4} - \overline{E[R]}_q,$$

where $AnnualRet_{q+4}$ is the annual excess return of the S&P 500 from q to $q+4$ and $\overline{E[R]}_q$ is the average expected return across investors from our model estimates at time q.[33] We test the forecastability of errors in the following linear regression model:

$$\epsilon_{q+4} = \alpha_0 + \alpha_1 X_q + \eta_{q+4} \tag{12}$$

The vector X_q consists of macroeconomic variables known to investors at the time of the forecast that are potentially correlated with stock market returns and investor expectations. We also include lagged parameters from our belief distribution (dispersion and skew) in vector X_q. Under the null hypothesis that expectations are rational and incorporate all relevant information, the macroeconomic variables known at time q should be orthogonal to the forecast error at $q+4$, or in other words, $\alpha_1 = 0$.

We report the corresponding estimates in Table 8. Our macroeconomic control variables include past 12-month excess returns of S&P 500 and one quarter lagged values of: VIX, log price-dividend ratio, and the consumption wealth ratio (cay) of Lettau and Ludvigson (2001). We also examine whether four quarters lagged forecast errors ϵ_q and one quarter lagged parameters or moments from the belief distribution (dispersion, skew, and interquartile range) predict forecast errors. We find evidence that forecast errors are predictable by the dispersion of the belief distributions, implied volatility, log price-dividend ratio, and the consumption wealth ratio.[34] We also find some evidence, albeit statistically insignificant, that forecast errors are negatively correlated with past returns. This is consistent with our finding that beliefs are extrapolative and suggests that such extrapolation is irrational. Similarly, we find that forecast errors are positively correlated with past investor disagreement. Together with the evidence that disagreement is negatively correlated with future returns (Section 5.4), this suggests the average investor belief is overly pessimistic during times of heightened disagreement. Given that disagreement tends to increase following negative

[33] In untabulated results we find that forecast errors are predictable at other horizons (i.e., at three or six months).

[34] In the Appendix, we show that the median forecast error is also predictable.

returns (Table 6) this again suggests that investors become overly pessimistic following negative returns.

5.4 Beliefs and Asset Prices: Forecasting Future Returns

5.4.1 Expectations and Future Returns

Lastly, we explore whether investor expectations of returns can forecast future returns. We regress the future 3-month excess return of the S&P 500 on the estimated mean expected return from our sample and report the corresponding results in column (1) of Table 9. Rather than predicting excess returns, the estimated mean expected returns instead have a weakly negative correlation with future returns. Our evidence is consistent with the findings in Greenwood and Shleifer (2014) and Amromin and Sharpe (2014) that investor expectations do not forecast future returns.[35]

In contrast, Greenwood and Shleifer (2014) and a long previous literature show that model-based measures can forecast future returns. We examine how our estimates of investor expectations about future returns vary with model expected returns. First, following Greenwood and Shleifer (2014) we use the dividend price ratio as a proxy for expected returns, second, we use the consumption wealth ratio (cay) of Lettau and Ludvigson (2001) as a proxy for expected returns, and third, we use the lower bound on the equity premium that Martin (2017) constructs using prices on S&P 500 index options. Columns (2)-(4) of Table 9 report the regression results corresponding to the regressions of the dividend-price ratio, cay, and the lower bound on the equity premium on our estimate of the mean expected return. The results indicate that model expected returns are negatively and significantly correlated with our estimate of the mean expected return. This evidence is consistent with the findings from Greenwood and Shleifer (2014) and Martin (2017) that investor expected returns are negatively correlated with model-based measures of expected returns.

The lower bound on equity premium in Martin (2017) binds for a representative investor with log utility who is unconstrained, rational, and fully invested in the market. In our framework, we also estimate a range of expected returns such that investors with these expectations find it optimal to hold the market (i.e., choose 1x leverage). Figure A2 in the Appendix shows that our estimated range almost always contains the equity premium in Martin (2017). Although using different methods based on different products, we generate very similar implications about the beliefs of investors holding the market.

In addition, Martin and Papadimitriou (2019) highlight the difference between the median and average beliefs in a model of heterogeneous investors. In their model, investors who happen to be correct in hindsight become wealthier, and the wealth-weighted average belief can be extrapolative even if no investors change their beliefs in response to past returns. The divergence between average and median beliefs from our estimates reflects similar patterns. This further shows how heterogeneity can complicate attempts to generate a single measure of market sentiments.

[35] We find similar results when we examine future 6-month and 12-month excess returns, which we report in Table A6 in the Appendix.

5.4.2 Dispersion in Beliefs and Future Returns

As detailed in the introduction, there is a long theoretical literature and modest empirical literature documenting how belief heterogeneity in conjunction with short-selling constraints can lead to inflated asset prices and lower future returns. The simple intuition for this mechanism from Miller (1977) is that short-selling constraints keep pessimistic investors out of the market such that prices reflect the valuation of more optimistic investors. Assuming the average valuation across all investors reflects the true value of an asset, short-selling constraints will lead to inflated prices and lower subsequent returns.

Building on this literature, we examine the relationship between our estimated dispersion in beliefs and future returns by regressing the future 3-month excess return of the S&P 500 on our measures of dispersion in beliefs. We measure dispersion in beliefs using both the levels and changes in the interquartile range of beliefs and the dispersion parameter b from our estimated belief distribution, which corresponds to the standard deviation of beliefs. In our specification we also control for the average expected return across investors. Columns (1)-(4) of Table 10 display the corresponding results. We estimate a negative relationship between dispersion in beliefs and returns in each specification, and the estimates are statistically significant in two out of the four specifications. The results in column (1) indicate that a one percentage point increase in the interquartile range is associated with a 0.39pp decrease in quarterly future returns, which is consistent with the predictions from the literature.

The theoretical literature has pointed to the interaction of short-selling constraints and dispersion in beliefs as the key factor driving the negative relationship between dispersion in beliefs and returns. We explore this potential mechanism by examining the relationship between returns and the fraction of constrained pessimistic investors in the following regression:

$$R_{q+1} = \beta_1 Constrained_q + \beta_2 Dispersion_q + \beta_3 Constrained_q \times Dispersion_q + \beta_4 \overline{E[R]}_q + \nu_q. \quad (13)$$

The dependent variable measures the 3-month return of the S&P 500 from period q to $q+1$. One of the key independent variables of interest is the variable $Constrained_q$ which captures the fraction of constrained pessimistic investors and is measured as the share of investors who purchase the lowest available leverage category. For the early part of our sample, the lowest available leverage was -2x which changed to -3x in 2009 after the introduction of new ETFs. The other independent variable of interest is the interaction term $Constrained_q \times Dispersion_q$. Much of the theoretical literature indicates that the combination of disagreement and short-selling constraints leads to lower future returns such that $\beta_3 <0$.

We report the corresponding estimates in columns (5)-(7) of Table 10. In column (5) we control for the share of constrained pessimistic investors ($Constrained_q$) and find a negative and significant relationship between the share of constrained pessimist investors and future returns. The results indicate that a one percentage point increase in the share of investors in the lowest available leverage category (constrained pessimistic investors) is correlated with a 2.25pp decrease in

returns. In columns (6)-(7) we include the interaction term $Constrained_q \times Dispersion_q$. where we measure dispersion using the interquartile range in Column (6) and using the dispersion parameter b in column (7). Consistent with the theoretical literature we estimate a negative coefficient on the interaction term, indicating that returns are lower when both short-selling constraints bind and dispersion in beliefs is high. We also add the caveats that our measure of constrained short sellers is likely an imperfect proxy for short-selling constraints and the investors we study reflect only a subset of the market. Furthermore, both short-selling constraints and dispersion in beliefs are potentially endogenous. However, the results suggest that the interaction of short-selling constraints and dispersion in beliefs may play a role in explaining the negative relationship between belief dispersion and returns.[36]

6 Extensions

6.1 Extending the Methodology to Other Assets

It is straightforward to extend our approach to other asset classes. We extend our analysis to estimate time-varying investor expectations for gold, oil, European equities, emerging market equities, US real estate, medium-term (7-10 year) Treasury, and long-term (20+ year) Treasury. In Appendix C, we describe the ETFs corresponding to each asset class in detail and report corresponding market shares. We follow the same methodology as above, using maximum likelihood to recover time-varying distribution of expectations separately for each asset class. For oil and US real estate, we have less empirical variation in choices, so we restrict the skewness parameter to be 1 (no skew) throughout the sample.[37]

Figure 8 panels (a)-(g) plot the estimated expected return distribution over time across the seven different asset classes. We capture time-varying expectations that seem reasonable and are consistent with intuition. For example, following the 2008 financial crisis, investor expectations over the real estate sector fall dramatically and then rebound in 2010 and 2011 (8e). Similarly, the negative effects of the European sovereign debt crisis on investor beliefs are immediately apparent in Figure 8c, as the average investor expected a decline in equity prices and there was an increase in disagreement across investors.

We estimate different risk aversion parameters for each asset class, because the sample of investors trading each asset class may be different. For example, we find that investors in gold are slightly less risk averse than those in S&P 500 ($\lambda = 0.78$ vs. $\lambda = 0.98$). We estimate that investors in oil are much less risk averse, with a risk aversion parameter of 0.28. One caveat is that the interpretation of these estimates as risk aversion depends on the strict interpretation of the model. If investors have heterogeneity in beliefs about volatility, this could be reflected in the estimated

[36] We find that dispersion is similarly predictive of future 6-month returns. It is less predictive of 12-month returns, though we have less statistical power at that horizon. We replicate Table 10 with 6-month and 12-month horizons in Tables A7 and A8 in the Appendix.

[37] If we relax this constraint, we do not estimate the skewness parameter to be significantly different from 1.

parameter. The differences in estimated risk aversion could also vary because the size of the ETF investment relative to the investor's portfolio varies across asset classes.

6.2 Using Higher-Frequency Data: Evidence from the COVID-19 Pandemic

Our methodology for recovering investor beliefs allows us to calculate the distribution of beliefs at the same frequency as the ETF flow data. From 2008 through 2018, we estimate the distribution of beliefs at monthly and quarterly frequencies. As an extension, we exploit daily ETF flow data to calculate the distribution of beliefs at a daily level during the COVID-19 pandemic.

We estimate the distribution of beliefs over the period January 2020 through June 2020 following the same procedure described in Section 4.1. We use the high-frequency data to estimate the time-varying parameters of the distribution—location, dispersion, and skewness—at a daily frequency. For this shorter time series, we set the static parameters for risk aversion and kurtosis equal to our baseline estimates. Figure 9 panel (a) displays our daily estimates of investor beliefs and panel (b) displays the incidence of COVID-19 and the S&P 500 Index level. Panel (a) shows that there was a sharp increase in disagreement among investors in mid-February around the time the first COVID-19 death was reported in the US. The increase in disagreement peaked in late March, but remained elevated through June. Part of the increase in dispersion was driven by an increase in pessimism, but it was also driven by an increase in optimism among contrarian investors. The results suggest that the average investor became optimistic about returns starting in February until states in the US started imposing lockdowns in mid-to-late March. Following the lockdowns in the US, the average expected return fell by more than 20 percent from its peak and did not recover until mid-to-late April when infection rates in the US started to fall. The patterns we document are consistent with the survey evidence in Giglio et al. (2020). Using monthly survey data from February through April 2020, Giglio et al. (2020) find that in mid-March the average investor became more pessimistic and that there was a persistent increase in disagreement. Overall, the results displayed in Figure 9 highlight the potential usefulness and scope of our methodology; we recover the distribution of investor beliefs at a high frequency in real time using readily available data.

7 Conclusion

We use a revealed-preference approach to estimate investor expectations of stock market returns. We apply our methodology to the market for S&P 500 ETFs. ETF investors face a fixed menu of investment alternatives, each with a different fee structure and risk/return profile. Measuring how investors trade-off risk/return among a fixed choice set allows us to separately identify investor expectations of returns and risk aversion.

Our framework allows us to recover the full distribution of investor beliefs at a quarterly frequency over the period 2008-2018. Our empirical estimates of investor expectations are highly correlated with the leading survey measures of investor expectations that are commonly used in

the literature (Greenwood and Shleifer, 2014). Because we recover the distribution of investor expectations, we are able to provide new insights into the drivers of investor beliefs. Consistent with the literature, we find evidence of extrapolative beliefs: mean expectations about future returns are highly and positively correlated with recent historical returns. In addition, we find that the distribution of beliefs becomes more dispersed and more negatively skewed following a period of negative stock market returns. As a stark illustration of heterogeneity in beliefs, we find that mean and median expectations move in opposite directions after negative S&P 500 returns.

Our framework is straightforward to apply to other asset classes. While we study the market for ETFs for tractability reasons, this type of demand framework could be used to provide insight into investor expectations and risk preferences in other settings going forward, and could be particularly useful when survey or micro data is unavailable.

References

Amromin, G. and S. A. Sharpe (2014). From the horse's mouth: Economic conditions and investor expectations of risk and return. *Management Science 60*(4), 845–866.

Bacchetta, P., E. Mertens, and E. Van Wincoop (2009). Predictability in financial markets: What do survey expectations tell us? *Journal of International Money and Finance 28*(3), 406–426.

Baker, S. R., N. Bloom, S. J. Davis, and K. J. Kost (2019). Policy news and stock market volatility. *NBER Working Paper*.

Barberis, N., R. Greenwood, L. Jin, and A. Shleifer (2015). X-capm: An extrapolative capital asset pricing model. *Journal of financial economics 115*(1), 1–24.

Barberis, N., R. Greenwood, L. Jin, and A. Shleifer (2018). Extrapolation and bubbles. *Journal of Financial Economics 129*(2), 203–227.

Barberis, N. and A. Shleifer (2003). Style investing. *Journal of Financial Economics 68*(2), 161–199.

Barseghyan, L., F. Molinari, T. O'Donoghue, and J. C. Teitelbaum (2013). The nature of risk preferences: Evidence from insurance choices. *American Economic Review 103*(6), 2499–2529.

Ben-David, I., J. R. Graham, and C. R. Harvey (2013). Managerial miscalibration. *The Quarterly Journal of Economics 128*(4), 1547–1584.

Benartzi, S. (2001). Excessive extrapolation and the allocation of 401 (k) accounts to company stock. *The Journal of Finance 56*(5), 1747–1764.

Benetton, M. (2018). Leverage regulation and market structure: A structural model of the UK mortgage market. Working paper.

Berk, J. B. and J. H. van Binsbergen (2016). Assessing asset pricing models using revealed preference. *Journal of Financial Economics 119*, 1–23.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica*, 841–890.

Bordalo, P., N. Gennaioli, Y. Ma, and A. Shleifer (2018). Over-reaction in macroeconomic expectations. Working paper, National Bureau of Economic Research.

Bordalo, P., N. Gennaioli, R. L. Porta, and A. Shleifer (2019). Diagnostic expectations and stock returns. *The Journal of Finance 74*(6), 2839–2874.

Bordalo, P., N. Gennaioli, and A. Shleifer (2012). Salience theory of choice under risk. *The Quarterly journal of economics 127*(3), 1243–1285.

Bordalo, P., N. Gennaioli, and A. Shleifer (2018). Diagnostic expectations and credit cycles. *The Journal of Finance 73*(1), 199–227.

Borovička, J., L. P. Hansen, and J. A. Scheinkman (2016). Misspecified recovery. *The Journal of Finance 71*(6), 2493–2544.

Brunnermeier, M. K., A. Simsek, and W. Xiong (2014). A welfare criterion for models with distorted beliefs. *The Quarterly Journal of Economics 129*(4), 1753–1797.

Calvet, L. E., J. Y. Campbell, F. J. Gomes, and P. Sodini (2019). The cross-section of household preferences. Working paper.

Case, K. E., R. J. Shiller, and A. K. Thompson (2012). What have they been thinking? Homebuyer behavior in hot and cold markets. *Brookings Papers on Economic Activity*, 265.

Chen, J., H. Hong, and J. C. Stein (2002). Breadth of ownership and stock returns. *Journal of financial Economics 66*(2-3), 171–205.

Chen, X., L. P. Hansen, and P. G. Hansen (2020). Robust identification of investor beliefs. *National Bureau of Economic Research*.

Cochrane, J. H. (2011). Presidential address: Discount rates. *The Journal of finance 66*(4), 1047–1108.

Coibion, O. and Y. Gorodnichenko (2012). What can survey forecasts tell us about information rigidities? *Journal of Political Economy 120*(1), 116–159.

Coibion, O. and Y. Gorodnichenko (2015). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review 105*(8), 2644–78.

Cunha, F., J. J. Heckman, and S. Navarro (2007). The identification and economic content of ordered choice models with stochastic thresholds. *International Economic Review 48*(4), 1273–1309.

Cutler, D. M., J. M. Poterba, and L. H. Summers (1990). Speculative dynamics and the role of feedback traders. *American Economic Review 80*(2), 63–68.

d'Arienzo, D. (2020). Maturity increasing overreaction and bond market puzzles. *Available at SSRN 3733056*.

De Long, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann (1990). Positive feedback investment strategies and destabilizing rational speculation. *The Journal of Finance 45*(2), 379–395.

DellaVigna, S. (2018). Structural behavioral economics. In *Handbook of Behavioral Economics: Applications and Foundations 1*, Volume 1, pp. 613–723. Elsevier.

Dick, A. A. (2008). Demand estimation and consumer welfare in the banking industry. *Journal of Banking & Finance 32*(8), 1661–1676.

Diether, K. B., C. J. Malloy, and A. Scherbina (2002). Differences of opinion and the cross section of stock returns. *The Journal of Finance 57*(5), 2113–2141.

Du, W., S. Gadgil, M. B. Gordy, and C. Vega (2019). Counterparty risk and counterparty choice in the credit default swap market. Working paper, SSRN 2845567.

Egan, M. (2019). Brokers vs. retail investors: Conflicting interests and dominated products. *The Journal of Finance 74*(3), 1217–1260.

Egan, M., A. Hortaçsu, and G. Matvos (2017). Deposit competition and financial fragility: Evidence from the us banking sector. *American Economic Review 107*(1), 169–216.

Egan, M., S. Lewellen, and A. Sunderam (2017). The cross section of bank value. Working paper, National Bureau of Economic Research.

Fagereng, A., C. Gottlieb, and L. Guiso (2017). Asset market participation and portfolio choice over the life-cycle. *The Journal of Finance 72*(2), 705–750.

Geanakoplos, J. (2010). The leverage cycle. *NBER macroeconomics annual 24*(1), 1–66.

Gennaioli, N., Y. Ma, and A. Shleifer (2016). Expectations and investment. *NBER Macroeconomics Annual 30*(1), 379–431.

Gennaioli, N. and A. Shleifer (2018). *A crisis of beliefs: Investor psychology and financial fragility*. Princeton University Press.

Ghosh, A., A. G. Korteweg, and Q. Xu (2020). Recovering heterogeneous beliefs and preferences from asset prices. *Working Paper*.

Ghosh, A. and G. Roussellet (2020). Identifying beliefs from asset prices. *Working Paper*.

Giglio, S., M. Maggiori, J. Stroebel, and S. Utkus (2019). Five facts about beliefs and portfolios. Working paper, National Bureau of Economic Research.

Giglio, S., M. Maggiori, J. Stroebel, and S. Utkus (2020). Inside the mind of a stock market crash. Technical report, National Bureau of Economic Research.

Gormsen, N. J. and C. S. Jensen (2020). Higher-moment risk. *Available at SSRN 3069617*.

Graham, J. R. and C. R. Harvey (2008). The equity risk premium in 2008: evidence from the global cfo outlook survey. Working paper, SSRN 1162809.

Graham, J. R. and C. R. Harvey (2011). The Duke/CFO business outlook survey. *Duke CFO Global Business Outlook. http: www. cfosurvey. org*.

Greene, W. H. and D. A. Hensher (2010). *Modeling ordered choices: A primer*. Cambridge University Press.

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

Harrison, J. M. and D. M. Kreps (1978). Speculative investor behavior in a stock market with heterogeneous expectations. *The Quarterly Journal of Economics 92*(2), 323–336.

Heipertz, J., A. Ouazad, and R. Rancière (2019). The transmission of shocks in endogenous financial networks: A structural approach. Working paper, National Bureau of Economic Research.

Hong, H., J. Scheinkman, and W. Xiong (2006). Asset float and speculative bubbles. *The journal of finance 61*(3), 1073–1117.

Hortaçsu, A. and C. Syverson (2004). Product differentiation, search costs, and competition in the mutual fund industry: A case study of S&P 500 index funds. *The Quarterly Journal of Economics 119*(2), 403–456.

Ilut, C. L. and M. Schneider (2014). Ambiguous business cycles. *American Economic Review 104*(8), 2368–99.

Jensen, C. (2018). *Essays on Asset Pricing*. Ph. D. thesis, Denmark.

Jensen, C. S., D. Lando, and L. H. Pedersen (2019). Generalized recovery. *Journal of Financial Economics 133*(1), 154–174.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*(2), 263–292.

Koijen, R. S., R. J. Richmond, and M. Yogo (2019). Which investors matter for global equity valuations and expected returns? Working paper, SSRN 3378340.

Koijen, R. S. and M. Yogo (2016). Shadow insurance. *Econometrica 84*(3), 1265–1287.

Koijen, R. S. and M. Yogo (2018). The fragility of market risk insurance. Working paper, National Bureau of Economic Research.

Koijen, R. S. and M. Yogo (2019a). A demand system approach to asset pricing. *Journal of Political Economy 127*(4), 1475–1515.

Koijen, R. S. and M. Yogo (2019b). Exchange rates and asset prices in a global demand system. Working paper, SSRN 3383677.

Lettau, M. and S. Ludvigson (2001). Consumption, aggregate wealth, and expected stock returns. *The Journal of Finance 56*(3), 815–849.

Lian, C., Y. Ma, and C. Wang (2019). Low interest rates and risk-taking: Evidence from individual investment decisions. *The Review of Financial Studies 32*(6), 2107–2148.

Ludvigson, S., S. Ma, and S. Ng (2020). Uncertainty and business cycles: Exogenous impulse or endogenous response? *American Economic Journal: Macroeconomics forthcoming*.

Luo, C., E. Ravina, M. Sammon, and L. M. Viceira (2020). Retail investors' contrarian behavior around news and the momentum effect. *Available at SSRN 3544949*.

Malmendier, U. and S. Nagel (2011). Depression babies: Do macroeconomic experiences affect risk taking? *The Quarterly Journal of Economics 126*(1), 373–416.

Malmendier, U. and S. Nagel (2015). Learning from inflation experiences. *The Quarterly Journal of Economics 131*(1), 53–87.

Martin, I. (2017). What is the expected return on the market? *The Quarterly Journal of Economics 132*(1), 367–433.

Martin, I. and D. Papadimitriou (2019). Sentiment and speculation in a market with heterogeneous beliefs. Working paper, SSRN 3428381.

Martin, I. and S. Ross (2019). Notes on the yield curve. *Journal of Financial Economics 134*(3), 689–702.

Meeuwis, M. (2019). Wealth fluctuations and risk preferences: Evidence from U.S. investor portfolios. Working paper.

Meeuwis, M., J. A. Parker, A. Schoar, and D. I. Simester (2018). Belief disagreement and portfolio choice. Working paper, National Bureau of Economic Research.

Miller, E. M. (1977). Risk, uncertainty, and divergence of opinion. *The Journal of finance 32*(4), 1151–1168.

Morris, S. (1996). Speculative investor behavior and learning. *The Quarterly Journal of Economics 111*(4), 1111–1133.

Nagel, S. and Z. Xu (2019). Asset pricing with fading memory. Working paper, National Bureau of Economic Research.

Newey, W. K. and K. D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 55*(3), 703–708.

Ross, S. (2015). The recovery theorem. *The Journal of Finance 70*(2), 615–648.

Scheinkman, J. A. and W. Xiong (2003). Overconfidence and speculative bubbles. *Journal of political Economy 111*(6), 1183–1220.

Shumway, T., M. Szefler, and K. Yuan (2009). The information content of revealed beliefs in portfolio holdings. Working paper, University of Michigan.

Simsek, A. (2013). Belief disagreements and collateral constraints. *Econometrica 81*(1), 1–53.

Vissing-Jorgensen, A. (2003). Perspectives on behavioral finance: Does "irrationality" disappear with wealth? evidence from expectations and actions. *NBER macroeconomics annual 18*, 139–194.

Wang, Y., T. M. Whited, Y. Wu, and K. Xiao (2018). Bank market power and monetary policy transmission: Evidence from a structural estimation. Working paper.

Welch, I. (2000). Views of financial economists on the equity premium and on professional controversies. *The Journal of Business 73*(4), 501–537.

Tables

Table 1: Summary Statistics Across S&P 500 Leverage Categories

	Adj. Share (%)		Raw Share (%)		Raw AUM (Billion)		Purchases (Billion)		Retail Fraction		Expense Ratio (bps.)	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
-3x	2.87	0.70	0.33	0.13	0.75	0.26	2.50	1.35	0.91	0.03	93.48	1.26
-2x	6.59	5.48	1.23	1.08	1.98	0.89	7.76	10.83	0.86	0.05	89.99	1.10
-1x	1.06	0.48	0.83	0.50	1.66	0.59	1.18	0.80	0.72	0.03	90.74	2.34
1x	56.98	10.47	88.41	4.37	230.77	139.91	233.13	73.38	0.25	0.05	8.66	0.36
2x	4.85	2.08	0.99	0.63	1.93	0.62	5.41	4.36	0.79	0.04	91.24	2.09
3x	3.49	0.62	0.37	0.08	1.04	0.66	3.19	1.30	0.82	0.08	95.27	0.58
Total	75.26	5.55	92.09	2.76	237.96	140.44	252.67	83.12	0.27	0.04	29.31	9.01

Notes: Table 1 shows summary statistics at month $\times$ leverage category level. *Adj. Share* and *Raw Share* compare market shares based on our adjusted purchase volume outlined in Section 2.3 and the raw AUM in the data. *Raw AUM* and *Purchases* display the original AUM and our adjusted purchase volume in billion dollars. Lastly, *Retail Fraction* shows the retail ownership and *Expense Ratio* shows the fee charged by ETFs. The last row corresponds to the means and standard deviations of monthly total adjusted market share, AUM share, AUM, and purchase volume across all leverage categories, monthly average retail ownership and monthly average expense ratio weighted by market share.

Table 2: Parameters for Time-Varying Belief Distribution F_s

Parameter	Function	Interpretation
a_s	Location	Corresponds to mean and median with no skew ($c = 1$)
b_s	Dispersion	Multiplicative scale; corresponds to standard deviation when ($d = \infty, c = 1$)
c_s	Skewness	More extreme negative values ($c < 1$) or positive values ($c > 1$)
d_s	Kurtosis	Special cases are Cauchy ($d = 1, c = 1$) and Normal ($d = \infty, c = 1$)

Table 3: Estimated Parameters

	Constant Risk Aversion		Heterogeneous Risk Aversion		Time-Varying Risk Aversion	
	Coef	SE	Coef	SE	Coef	SE
Expected Return						
Location (Median)	2.856	(0.695)	2.543	(0.667)	2.053	(0.802)
Dispersion (Median)	1.080	(0.424)	1.005	(0.456)	0.910	(0.395)
Skewness (Median)	0.766	(0.343)	0.756	(0.243)	0.732	(0.313)
Kurtosis	1.262	(0.133)	1.322	(0.151)	1.510	(0.193)
Risk Aversion: Time Invariant						
Mean	0.982	(0.020)	0.853	(0.072)		
Dispersion			0.145	(0.080)		
Risk Aversion: Time Varying						
Min					0.255	(0.076)
Median					0.617	(0.045)
Max					1.164	(0.407)
Model Fit						
R^2	0.921		0.923		0.935	
Log Likelihood	-168.4		-168.3		-166.7	
AIC	2565.0		2565.5		2560.7	
BIC	3208.5		3213.8		3252.2	

Notes: Table 3 shows estimation results with constant and heterogeneous risk aversion. The first panel displays parameters for the expected return distributions. Location, dispersion and skewness parameters are allowed to vary over time, and we estimate one set of coefficients for each quarter. We display the median location, dispersion, and skewness coefficients, as well as their corresponding standard errors. The next panel shows mean risk aversion and the dispersion (half length of the range) when it follows uniform distribution. Standard errors are computed using the inverse of numerical Hessian. Next, we compute the implied mean expected return in each quarter and display the quantiles of the across-time distribution of mean expectations. The last two rows show R^2 and log likelihood of each specification.

Table 4: Comparison with Surveys

Survey	Gallup	Mich.	AAII	Shiller	Duke CFO		SPF
Survey Measure	Index	Index	Index	Index	Mean	Share Neg.	Mean
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Est. Measure	0.71***	0.77***	0.45***	0.55***	0.32***	0.68***	0.88***
	(0.10)	(0.075)	(0.099)	(0.17)	(0.11)	(0.12)	(0.096)
Observations	40	44	44	44	44	44	44
R-squared	0.511	0.585	0.202	0.307	0.100	0.462	0.779

Notes: Table 4 shows the correlation between the estimated expectations from our model and six additional surveys: (1) the Gallup Investor and Retirement Optimism Index, (2) the University of Michigan Survey of Consumers, (3) the American Association of Individual Investors (AAII) Sentiment Survey, (4) the Shiller U.S. Individual One-Year Confidence Index,(5-6) the Duke CFO Global Business Outlook Survey (CFO), and (7) the Survey of Professional Forecasters (SPF). The estimated belief distribution corresponds to our baseline model estimates reported in column (1) of Table 3. Observations are at the quarterly level over the period 2008-2018. For details on these surveys, see Section 4.5. Each column displays the correlation between each survey measure and the analogous measure from our model. Surveys in columns (1)-(3) are compared to the relative share of investors preferring positive to negative leverage, based on our estimated distribution of expectations. The Shiller index in column (4) is compared to the fraction of investors choosing positive leverage (greater than 1x). The CFO mean expected return in column (5) is compared to the mean estimated expected return from our model. The fraction of CFOs who expect returns to be negative next year is compared to the fraction of investors choosing negative leverage in column (6). The SPF average GDP growth forecast in column (7) is compared to the mean estimated expected return from our model. We winsorize the mean and standard deviation of expected returns from our model at the 5% level to account for outliers during the financial crisis. Winsorizing the data does not change inference on the relationship between the corresponding series. *** p<0.01, ** p<0.05, * p<0.10.Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.10.

Table 5: Comparison with Other Measures of Dispersion and Uncertainty

Measure	Duke CFO	SPF	BBDK	LMN
	(1)	(2)	(3)	(4)
Est. Measure	0.56***	0.72***	0.77***	0.82***
	(0.15)	(0.19)	(0.067)	(0.14)
Observations	44	44	44	44
R-squared	0.310	0.512	0.586	0.678

Notes: Table 5 shows the correlation between the estimated dispersion in beliefs from our model and four additional measures surveys and the literature: (1) the Duke CFO Global Business Outlook Survey (CFO), (2) the Survey of Professional Forecasters (SPF), (3) the US Equity Market Volatility Tracker constructed by Baker, Bloom, Davis, and Kost (2019) (BBDK), and (4) the Financial Uncertainty Index constructed by Ludvigson, Ma, and Ng (2020) (LMN). The estimated belief distribution corresponds to our baseline model estimates reported in column (1) of Table 3. Observations are at the quarterly level over the period 2008-2018. For details on these surveys, see Section 4.5. Each column displays the correlation between each alternative measure and the analogous measure from our model. Column (1) displays the correlation between standard deviation of expected returns across CFOs and standard deviation parameter from our model. Column (2) displays the correlation between th interquartile range of GDP forecasts across professional forecasters in the SPF and the interquartile range of estimated expected returns from our model. Column (3) displays the correrlation between the US Equity Market Volatility Tracker by Baker, Bloom, Davis, and Kost (2019) and the interquartile range of estimated expected returns from our model. Column (4) displays the correrlation between the Financial Uncertainty Index by Ludvigson, Ma, and Ng (2020) and the interquartile range of estimated expected returns from our model. We winsorize the standard deviation of expected returns from our model at the 5% level to account for outliers during the financial crisis. Winsorizing the data does not change inference on the relationship between the corresponding series. *** p<0.01, ** p<0.05, * p<0.10. Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.10.

Table 6: Expected Returns versus Past 12-Month Returns

	Mean	ln(Dispersion)	Median	ln(Skewness)
	(1)	(2)	(3)	(4)
Annual Return	0.11***	-0.026***	-0.038**	0.0026***
	(0.030)	(0.0076)	(0.014)	(0.00089)
Observations	44	44	44	44
R-squared	0.582	0.310	0.230	0.184

Notes: Table 6 displays the regression of different moments of the estimated expected returns distribution on the past 12-month excess return of the S&P 500. Observations are at the quarterly level over the period 2008-2018. The dependent variable in each column corresponds to different moments/parameters of the estimated expected returns distribution corresponding to our baseline estimates reported in column (1) of Table 3. The dependent variable in column (1) is the mean and is measured in percentage points, in column (2) is the dispersion parameter in logs, in column (3) is the median and is measured in percentage points, and in column (4) is the skewness parameter in logs. The independent variable Annual Return is measured in percentage points. We winsorize all independent and dependent variables at the 5% level to account for outliers during the financial crisis. Winsorizing the data does not change inference on the relationship between the corresponding series. Newey-West based standard errors are in parentheses with four lags. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 7: Evolution of the Parameters of the Expectation Distribution: Vector Autoregressions

	Location	Dispersion	Skewness
Lag Parameters	0.55***	0.65***	0.41***
	(0.13)	(0.085)	(0.15)
Monthly Return	-0.65***	-0.63***	0.0015
	(0.23)	(0.20)	(0.0038)
Const	2.32***	1.33***	0.45***
	(0.69)	(0.41)	(0.12)
Observations	43	43	43
R-squared	0.726	0.791	0.195

Notes: Table 7 displays the regression results to three linear regression models (eq. 11). Observations are at the quarterly level over the period 2008-2018. The dependent variable in each column corresponds to different moments/parameters of the estimated expected returns distribution corresponding to our baseline estimates reported in column (1) of Table 3. The dependent variable in column (1) is the mean parameter, in column (2) is the standard deviation parameter, and in column (3) is the skew parameter. We include the lag dependent variable in each regression as a control variable. The independent variable Monthly Return is the previous monthly excess return of the S&P 500 averaged across the quarter and is measured in percentage points. Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table 8: Predictability of Forecast Errors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lagged Forecast Error	-0.23							
	(0.16)							
Annual Return		-0.12						
		(0.15)						
Lagged ln(Dispersion)			4.08*					
			(2.34)					
Lagged IQR				0.63***				
				(0.12)				
Lagged ln(Skewness)					-3.57			
					(20.8)			
Lagged VIX						0.46**		
						(0.19)		
Lagged ln(Price/Dividend)							41.4***	
							(8.13)	
Lagged cay								272*
								(151)
Observations	40	44	43	43	43	43	43	39
R-squared	0.077	0.028	0.076	0.162	0.001	0.119	0.235	0.122

Notes: The table displays the results corresponding to a linear regression model (eq. 12). Observations are at the quarterly level. The dependent variable is the forecast error constructed as the difference between the 12-month return of the S&P 500 from period q to $q+4$ and the mean expected return at time q. The forecast error is scaled by 100 such that it is measured in percentage points. Independent variables include four quarters lagged values of forecast errors, past 12-month S&P 500 excess returns, one quarter lagged value of parameters or moments from the belief distribution (dispersion, skew, and interquartile range), and one quarter lagged value of VIX, log price-dividend ratio, and the consumption wealth ratio (cay) of Lettau and Ludvigson (2001). Newey-West based standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

Table 9: Comparison with Future Returns and Model Returns

	Future Returns	ln(Div/Price)	cay	Martin (2017)
	(1)	(2)	(3)	(4)
Mean of Exp. Returns	-0.13	-0.047***	-0.0026***	-0.014***
	(0.58)	(0.0054)	(0.00090)	(0.0036)
Observations	44	44	39	44
R-squared	0.003	0.749	0.200	0.630

Notes: Table 9 displays the regression of future and model-based return measures on the mean belief from our estimated distribution of investor beliefs. Observations are at the quarterly level over the period 2008-2018. The dependent variable in column (1) is the future 3-month excess return of the S&P 500. The dependent variable in column (2) is the log dividend-pice ratio. The dependent variable in column (3) is cay from Lettau and Ludvigson (2001). Lastly, the dependent variable in column (4) is the lower bound on the equity premium in Martin (2017). The independent variable Mean of Exp. Returns corresponds to our baseline model estimates reported in column (1) of Table 3. We winsorize the Mean of Exp. Returns and the future 3-month excess return of the stock market variables at the 5% level to account for outliers during the financial crisis. Newey-West based standard errors are in parentheses with one lag. *** p<0.01, ** p<0.05, * p<0.10.

Table 10: Dispersion in Beliefs and Returns

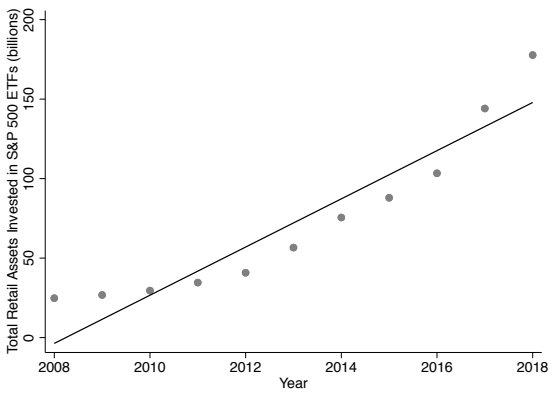
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
IQR	-0.39					0.77**	
	(0.28)					(0.35)	
Δ IQR		-0.44***					
		(0.047)					
Dispersion (b)			-0.65				1.64**
			(0.69)				(0.70)
Δ Dispersion				-1.00***			
				(0.13)			
Constrained					-2.25***	-1.69**	-1.66**
					(0.63)	(0.81)	(0.79)
Constrained$\times$IQR						-0.064***	
						(0.021)	
Constrained$\times$Dispersion							-0.14***
							(0.044)
Observations	44	43	44	43	44	44	44
R-squared	0.065	0.222	0.039	0.217	0.242	0.338	0.340

Notes: The table displays the results corresponding to a linear regression model (eq. 13). Observations are at the quarterly level. The dependent variable is the 3-month return of the S&P 500 from period q to $q+1$ and scaled by 100 such that it is measured in percentage points. The variable IQR measures the interquartile range of the estimated belief distribution at time q and is scaled by 100 such that it is measured in percentage points. The variable $Dispersion$ corresponds to the estimated dispersion parameter (b) from our belief distribution. The variable $Constrained$ is share of investors who purchase the lowest available leverage category and is scaled by 100 such that it is measured in percentage points. In each specification we also control for the estimated mean belief at time q. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

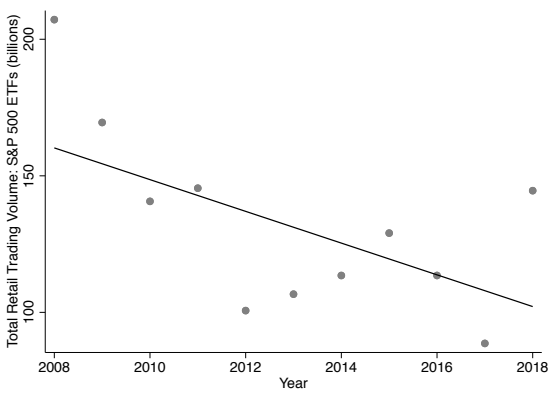
Figures

Figure 1: S&P 500 ETFs

(a) Assets Under Management (Retail Investors)



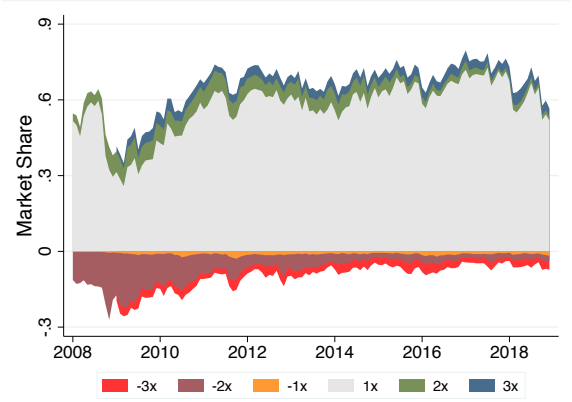
(b) Trading Volume (Retail Investors)



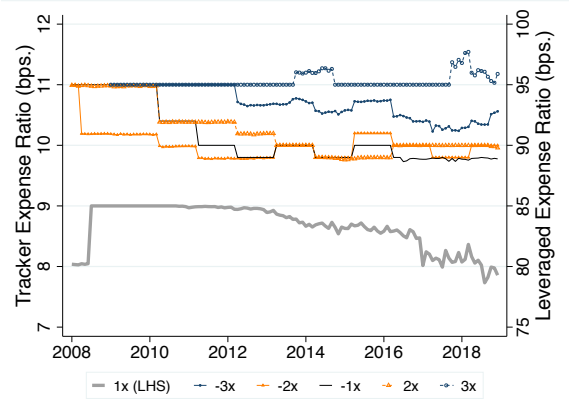
Notes: Figure 1 shows binned scatters at annual frequency along with the linear fitted lines for total retail AUM in panel (a) and total retail trading volume in panel (b) of ETFs that track S&P 500. Retail AUM is computed as $Retail_j \times AUM_{jt}$ and trading volume is computed as $Retail_j \times TradingVolume_{jt}$ according to the market share construction discussed in Section 2.3.

Figure 2: Data at Leverage Category Level (S&P 500)

(a) Market Share

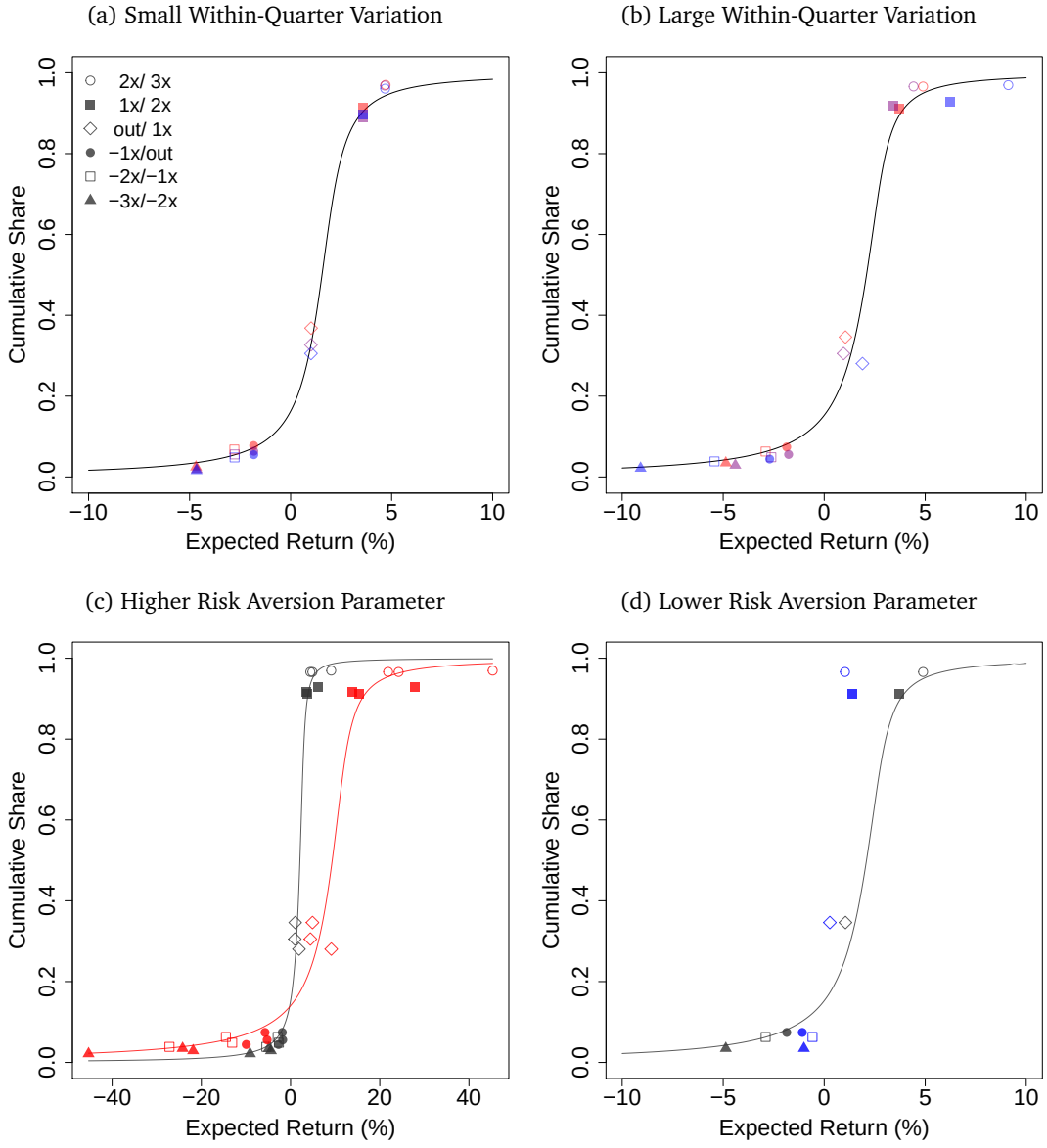


(b) Expense Ratio



Notes: Figure 2 top panel plots adjusted market share for each leverage category. The bottom panel plots market share weighted average expense ratio in each leverage category.

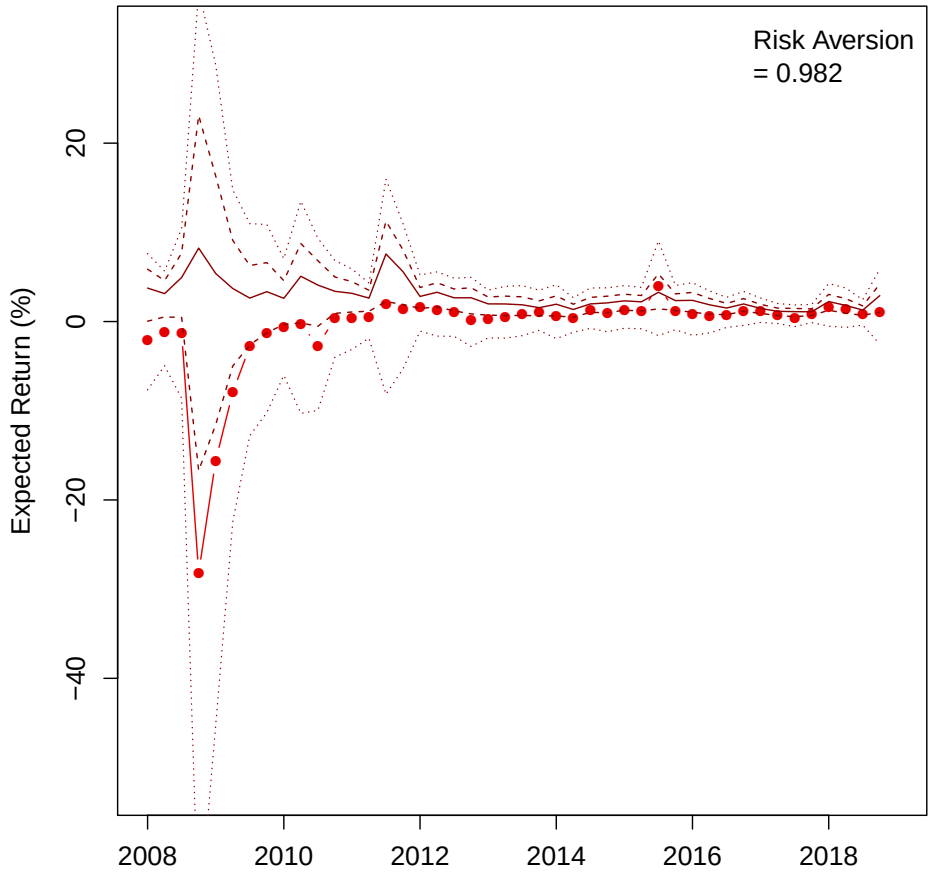
Figure 3: Identification Using Cross-Sectional and Within-Quarter Variation



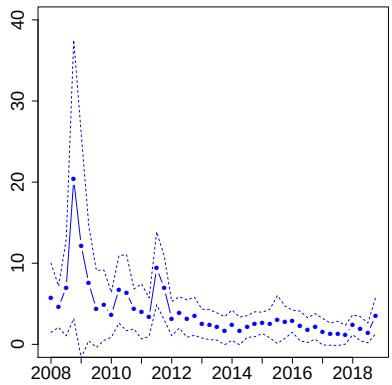
Notes: Figure 3 plots model-implied cutoff points of expected returns against the cumulative shares to illustrate how the distribution of expectations and risk aversion are identified. Each plotted line corresponds to an estimated CDF of investor expected returns. The parameters of the CDF are chosen to best fit the plotted points. Panel (a) shows a quarter in which the cross-sectional (within-month) variation acts a primary source of identification in pinning down the CDF. Panel (b) provides an example quarter in which within-quarter variation in fees and volatility play a more substantial role in identification. In panel (b), the overlap in cutoff points for different months within a quarter—i.e., the hollow purple square and the solid blue circle, corresponding to S_{-2} in one month and S_{-1} in another—illustrates our condition for nonparametric identification of risk aversion. Panels (c) and (d) further illustrate how the risk aversion parameter is identified. In each panel, the cutoff points and the estimated CDF from panel (b) are plotted in black. In panel (c), we plot a set of observations and CDF that would be obtained with a risk aversion parameter that is five times greater (in red). In panel (d), we plot the observations corresponding to a risk aversion parameter that is five times smaller (in blue). We cannot fit a CDF for expected returns through these observations because the cutoff points are not monotonically increasing with leverage. At that level of risk aversion, our model predicts that no investor would choose a 2x ETF over a 3x ETF. Thus, the smaller risk aversion parameter is inconsistent with what we observe in the data. The larger risk aversion parameter admits a CDF, but the fit is poor compared to the estimated risk aversion parameter.

Figure 4: Time-Varying Investor Expectations

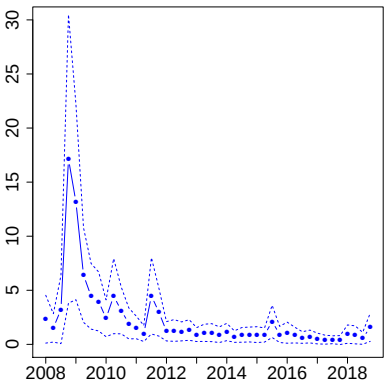
(a) Estimated Distribution



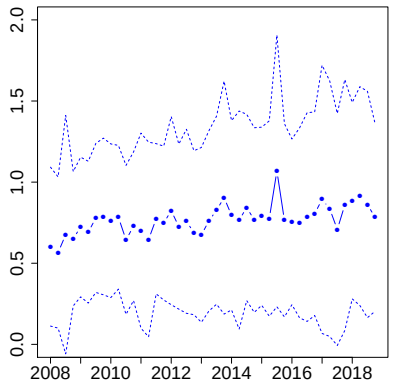
(b) Location Parameter (a_s)



(c) Dispersion Parameter (b_s)

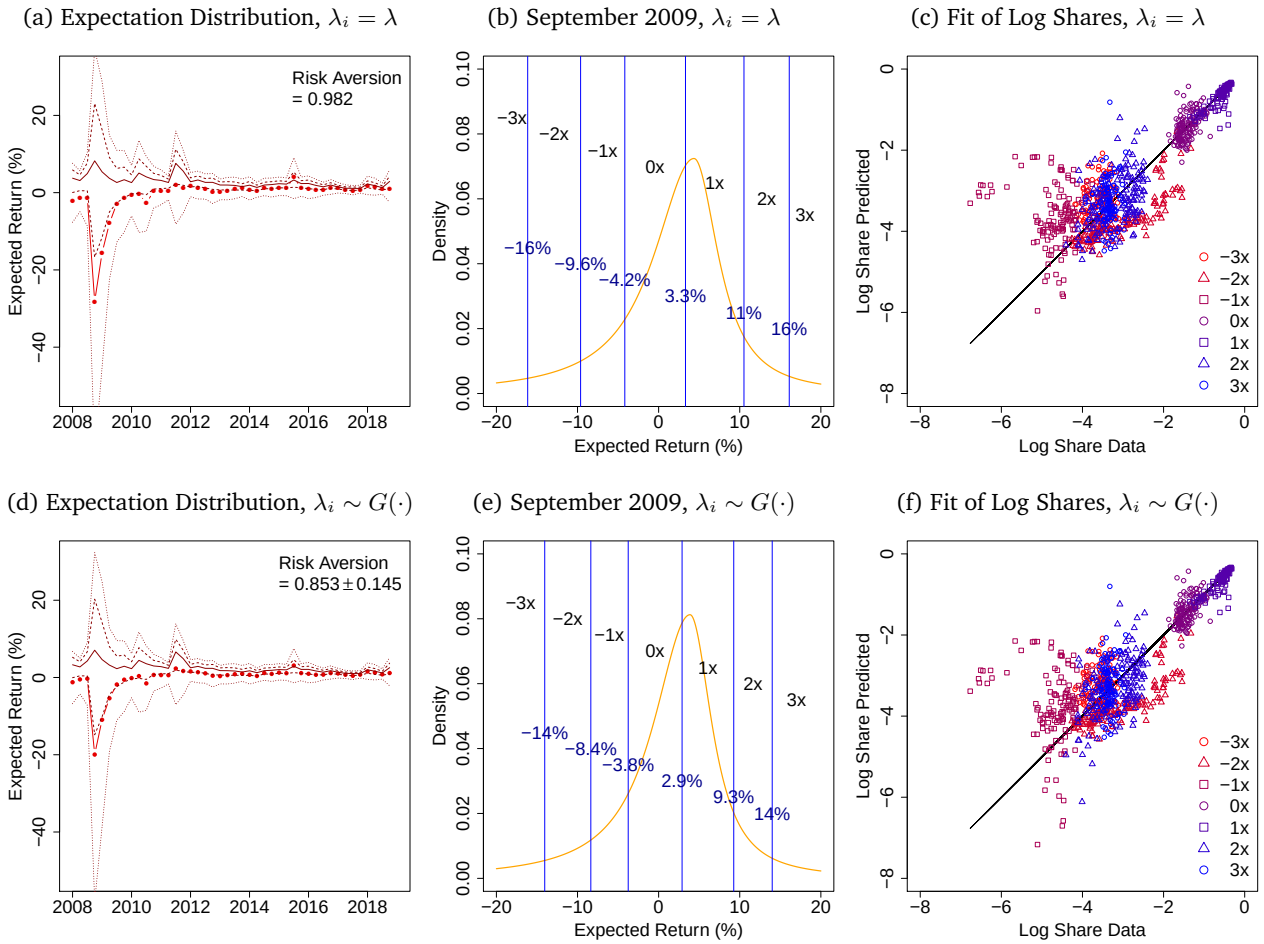


(d) Skewness Parameter (c_s)



Notes: Figure 4 panel (a) plots the estimated distribution of investor expectations over time. Red dots represent mean expected return. Solid dark red lines indicate median, dashed lines indicate 25th and 75th percentiles, and dotted lines indicate 10th and 90th percentiles. Panels (b) to (d) show estimated time-varying location, dispersion, and skewness parameters for expectation distribution in blue dotted lines, and the 90 percent confidence intervals in blue dashed lines.

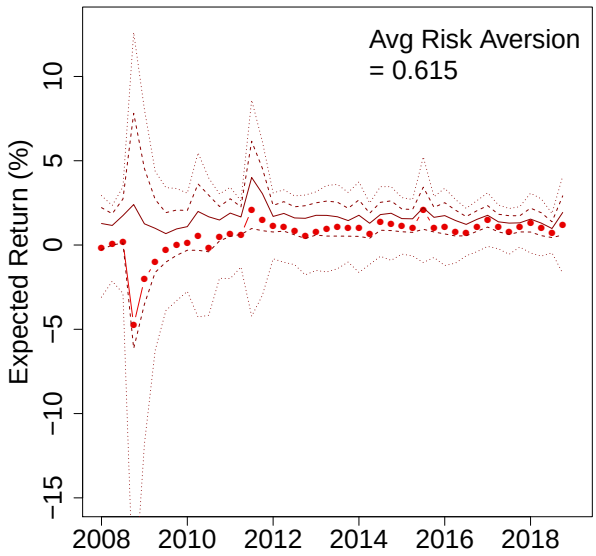
Figure 5: Expectations and Model Fit: Baseline and Heterogeneous Risk Aversion (S&P 500)



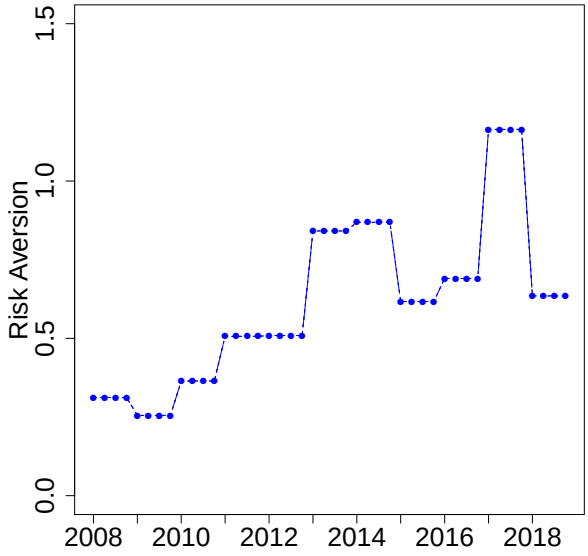
Notes: Figure 5 top panels correspond to the baseline specification with constant risk aversion. Bottom panels allow for heterogeneous risk aversion. Left panels plot the estimated distribution of investor expectations over time. Red dots represent mean expected return. Solid dark red lines indicate median, dashed lines indicate 25th and 75th percentiles, and dotted lines indicate 10th and 90th percentiles. Middle panels display the density of expectations for a given month (September 2009) and cutoff points corresponding to the expected return where investors are indifferent between two adjacent leverage categories. Right panels plot fit in terms of log market shares of each leverage. The *x*-axis corresponds to log market shares in the data, and *y*-axis corresponds to predicted log market shares. Color red to blue represents each leverage from -3x to 3x. The solid black lines correspond to the 45 degree line.

Figure 6: Time-Varying Risk Aversion

(a) Expectations



(b) Risk Aversion

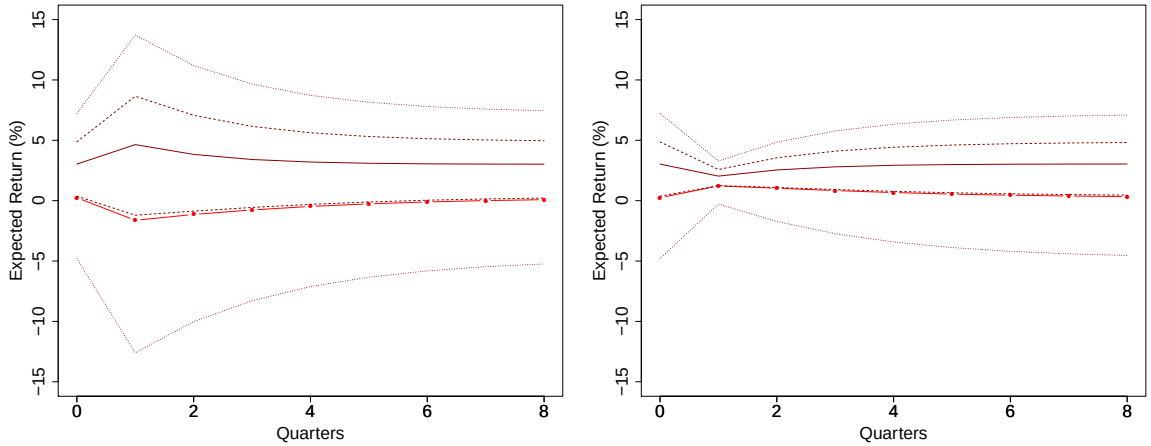


Notes: Figure 6 panel (a) plots the estimated distribution of expectations over time where we allow risk aversion to vary at annual level. Red dots represent mean expected return. Solid dark red lines indicate median, dashed lines indicate 25th and 75th percentiles, and dotted lines indicate 10th and 90th percentiles. Panel (b) plots the estimated risk aversion which varies at the annual level.

Figure 7: Impulse Response

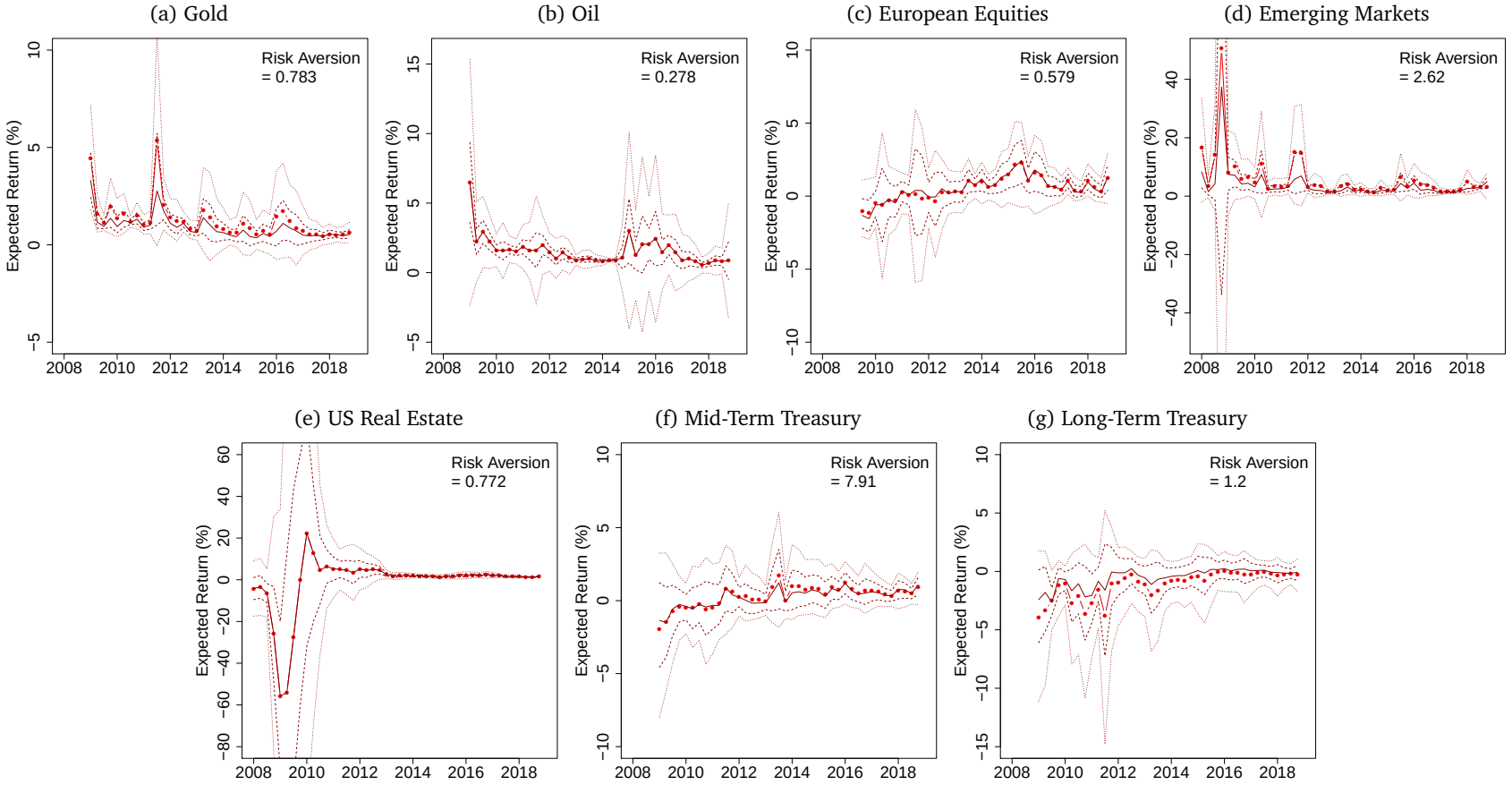
(a) Response to -10% One-Month Return

(b) Response to 10% One-Month Return



Notes: Figure 7 displays the impulse responses of a one-month S&P 500 return of -10 percent at $t = 1$ in the left hand side panel and a one-month S&P 500 return of 10 percent at $t = 1$ in the right hand side panel. Time t is measured in quarter. We predict each parameter separately using their lagged value in the previous quarter and the previous monthly S&P 500 excess return averaged cross the current quarter as reported in Table 7. The initial values are kept at steady state mean of each parameter. We assume averaged S&P 500 monthly returns are 0.84% for $t = 0$ and $t > 1$. Red dots correspond to analytical mean. Solid dark red line shows median, and dashed dark red lines show 10, 25, 75, and 90th percentiles.

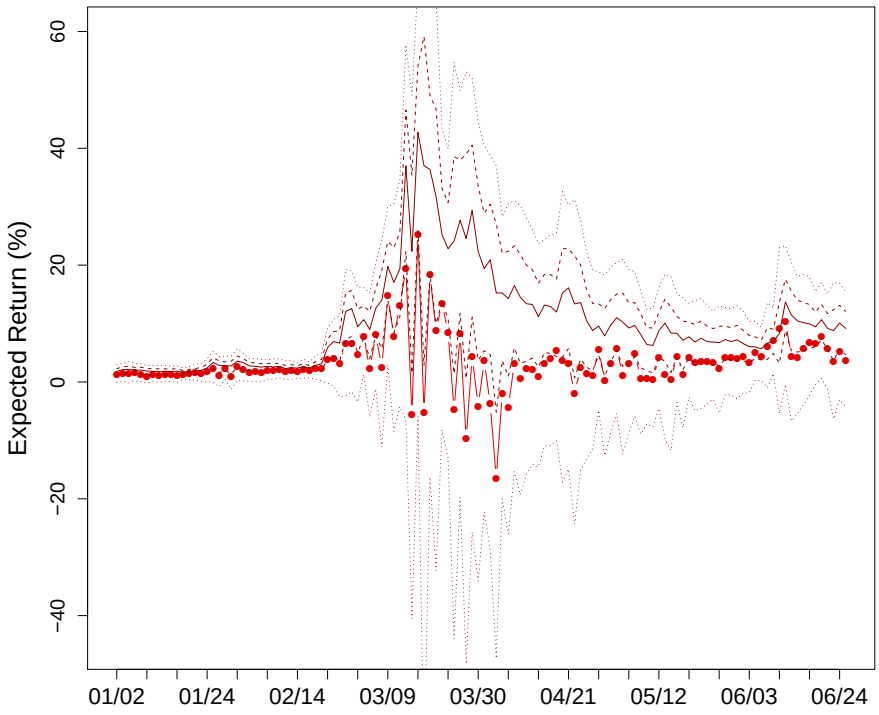
Figure 8: Expectations: Other Asset Classes



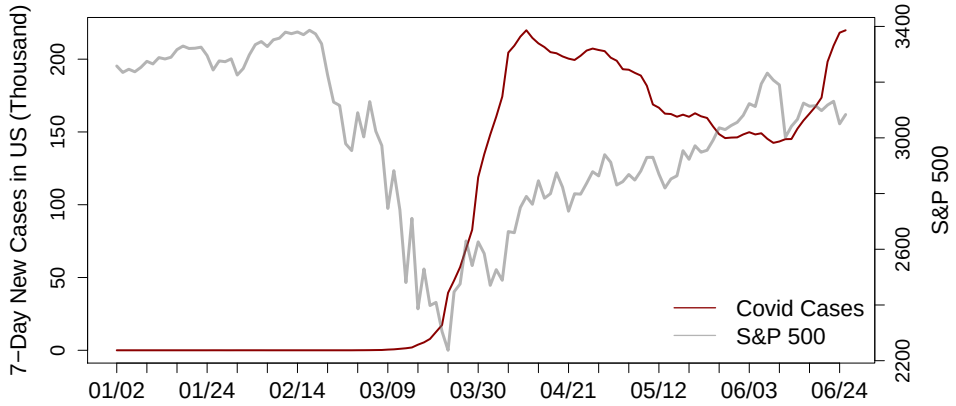
Notes: Figure 8 panels (a)-(g) displays the estimated expectation distribution corresponding to gold, oil, European equities, emerging market equities, US real estate, medium-term (7-10 year) Treasury, and long-term (20+ year) Treasury. Red dots represent mean expected return, solid dark red lines indicate median, dashed lines indicate 25th and 75th percentiles, and dotted lines indicate 10th and 90th percentiles.

Figure 9: Daily Expectations During COVID-19 Pandemic

(a) Estimated Distribution



(b) COVID-19 Cases and S&P 500 Index



Notes: Figure 9 panel (a) plots the estimated distribution of investor expectations at daily frequency during the first half of 2020. Panel (b) plots new COVID-19 cases over the past 7 days in the US in dark red with axis on the left hand side, and S&P 500 index levels in gray with axis at the right hand side. COVID cases are downloaded from the COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University

Appendix

A Alternative Estimation Approach

In this appendix, we provide an alternative set of estimates for our time-varying belief distribution. Our baseline estimates, which are presented in the text, make use of two sources of variation for identification. The first source of variation is in the choice of leverage facing investors. The second source is empirical variation in prices and volatility. How these sources provide identifying power are described in more detail in Section 3.

If we rely only on the first source of variation—the choices facing investors—then we can leverage the model to estimate beliefs at a higher frequency, as we would not require within-period variation in prices and volatility. For our alternative estimates, we follow this approach. Because we observe six unique points in the distribution in each period, corresponding to $\{S_j\} = \{S_{-3}, S_{-2,}S_{-1}, S_0, S_1, S_2\}$, we can identify, in principle, up to six period-specific parameters for the distribution F and risk aversion λ. Thus, even with this high degree of flexibility in the time series, our model has sufficient identifying restrictions.

For our alternative estimates, we use nonlinear least squares to estimate parameters that vary at the monthly level. As in our main results, we hold the risk aversion parameter (λ) and the kurtosis parameter fixed over the sample, allowing month-specific values for location, dispersion, and skewness. On advantage of the approach is computational efficiency. We estimate only a subset of the parameters with a nonlinear search and the rest are recovered by ordinary least squares.

Our estimation routine works as follows: in an outer loop, we choose the risk aversion parameter ($\hat{\lambda}$) and the kurtosis parameter ($\hat{d}$), which we hold fixed across periods. Then, in each period, we pick a value for the skewness parameter $\hat{c}_t$. We use the estimated skewness and kurtosis parameters to invert the cumulative share equation, obtaining

$$F^{-1}(S_{jt}; \hat{c}_t, \hat{d}_t) = \frac{1}{\hat{b}_t}\left(\frac{\hat{\lambda}}{2}(2j+1)\,\sigma_t^2 + p_{(j+1)t} - p_{jt} - \hat{a}_t\right) + \zeta_{jt},$$

where $\hat{a}_t$ and $\hat{b}_t$ are the period-specific location and dispersion parameters, and ζ_{jt} is a residual. We then run a period-specific regression of $F^{-1}(S_{jt}; \hat{c}_t, \hat{d}_t)$ on $(\frac{\hat{\lambda}}{2}(2j+1)\,\sigma_t^2 + p_{(j+1)t} - p_{jt})$ for all $j < 3$. As the coefficient on the combined term is normalized to 1, the regression coefficient provides us an estimate of the dispersion parameter $\frac{1}{\hat{b}_t}$. The constant is equal to $-\frac{\hat{a}_t}{\hat{b}_t}$ and provides us an estimate of the location parameter. We iterate over the outer-loop parameters $\hat{\lambda}$ and $\hat{d}$ until we find the value of all parameters that minimize $\sum_t \sum_j \hat{\zeta}_{jt}^2$.

Our monthly estimates using this procedure are displayed in Figure A3. These estimates track our main results fairly closely, though the skewness is somewhat less extreme during the crisis. This may be due to the fact that this alternative approach has residuals that allow the model to fit the shares exactly. Thus, extreme beliefs that may imply skewness in the distribution can be instead captured with a residual.

Figure A4 provides a more detailed comparison of the different estimates. Panels (a) and (e) report our baseline time series, which is based on maximum likelihood estimation, and the model fit. The alternative time series is shown in panel (d), and the fit, after removing the residuals, is shown in panel (h). Recall that the model fits the data perfectly when the residuals are accounted for.

To assist in comparison with the alternative estimates, we provide monthly maximum likelihood estimates in columns (b) and (f), where we allow the parameters of the belief distribution to vary at the monthly level. These estimates also rely only on variation in the choices and do not make use of empirical variation in fees and volatility. Likewise, we provide quarterly estimates for the alternative approach in panels (c) and (g).

The alternative estimates, which are obtained using different identifying restrictions and using a different objective function in estimation (least squares instead of maximum likelihood), return similar qualitative patterns to our baseline results. These alternative estimates show that our general approach is not sensitive to any single assumption.

B Robustness Checks

B.1 Heterogeneous Portfolios and Hedging Demand

In this appendix, we allow for portfolio hedging in our demand estimation. In our baseline specification, investor utility is given by

$$u_{ij} = \beta_j \mu_i - p_j - \frac{\lambda}{2} {\beta_j}^2 \sigma^2,$$

which specifies that the additional risk of adding an ETF to an investor's portfolio is ${\beta_j}^2\sigma^2$. This model does not account for how the ETF investment decisions may covary with the investor's wealth. If an investor considers the risk of her wealth not invested in ETFs, then she may prefer to pick ETFs that are negatively correlated with her other wealth to reduce her overall risk.

Formally, if an investor's wealth ω_i is correlated with the underlying ETF asset, the additional variance of investing a fraction of her wealth δ in ETF j is given by $\delta^2\beta_j^2\sigma^2 + 2\delta\beta_j\beta_{\omega_i}\sigma^2$, where β_{ω_i} is the market beta of the investor's portfolio. The term $\delta^2\beta_j^2\sigma^2$ reflects the variance of the ETF investment, and the term $2\delta\beta_j\beta_{\omega_i}\sigma^2$ reflects how the ETF investment changes the variance of the investor's existing portfolio.

To see this, consider an investor who has wealth W_0 exposed to market risk and sets aside δW_0 in active investment following S&P 500 ETFs. The total value of her wealth and ETF investment is $W = W_0(1 + \beta_{\omega_i} R) + \delta W_0(1 + \beta_j R)$, where R denotes S&P 500 returns and we assume there is no alpha in the wealth return. Taking a second-order Taylor Expansion of expected utility with respect to deviation from $W_0(1 + \delta)$ obtains

$$
\begin{aligned}
E[u(W)] &\approx u(W_0(1+\delta)) + u'(W_0(1+\delta)W_0 E[\beta_{\omega_i} R + \delta\beta_j R] + \frac{1}{2}u''(W_0(1+\delta))W_0^2 E[(\beta_{\omega_i} R + \delta\beta_j R)^2] \\
&\approx u(W_0(1+\delta)) + u'(W_0(1+\delta))W_0(\beta_{\omega_i}\mu + \delta\beta_j\mu) + \frac{1}{2}u''(W_0(1+\delta))W_0^2\sigma^2(\beta_{\omega_i}^2 + \delta^2\beta_j^2 + 2\delta\beta_j\beta_{\omega_i}) \\
&\propto \delta\beta_j\mu + \frac{1}{2}\frac{u''(W_0(1+\delta))}{u'(W_0(1+\delta))}W_0\sigma^2(\delta^2\beta_j^2 + 2\delta\beta_j\beta_{\omega_i}) \\
&= \delta\beta_j\mu - \frac{\lambda}{2}(\delta^2\beta_j^2 + 2\delta\beta_j\beta_{\omega_i})\sigma^2 \propto \beta_j\mu - \frac{\lambda}{2}\beta_j^2\sigma^2(\delta + \frac{2\beta_{\omega_i}}{\beta_j})
\end{aligned}
$$

where, in the second line, we plug in the definition of mean and variance of return and assume that $E[\beta_{\omega_i} R + \delta\beta_j R]^2 \approx 0$. In the third line, we drop terms unrelated to β_j and divide by $u'(W_0(1+\delta))W_0$. In the fourth line, we define risk aversion as constant relative risk aversion scaled by the fraction invested in ETF: $\lambda = -\frac{u''(W_0(1+\delta))W_0(1+\delta)}{u'(W_0(1+\delta))}\frac{1}{1+\delta}$. Also, note that purchasing an ETF will yield diversification benefits for the investor if and only if $\text{sgn}(\beta_j) \neq \text{sgn}(\beta_{\omega_i})$.

Thus, the indirect utility of leverage j for an investor whose wealth has market risk β_{ω_i} is given by

$$
u_{ij} = \beta_j\mu_i - p_j - \frac{\lambda}{2}{\beta_j}^2\sigma^2\delta - \lambda\beta_j\beta_{\omega_i}\sigma^2.
$$

For an average β_ω, the cumulative probability of purchasing leverage $k \leq j$ becomes

$$
S_j = F\left(\frac{\lambda\delta}{2}(2j+1)\sigma^2 + \lambda\beta_\omega\sigma^2 + p_{j+1} - p_j\right)
$$

With this extension, we identify $\lambda\delta$ and $\lambda\beta_\omega$. The coefficient on the first term inside the bracket captures risk aversion multiplied by the fraction of active ETF investment $\lambda\delta$. The second term corresponds to hedging and gives us an estimate for the average wealth market risk multiplied by risk aversion $\lambda\beta_\omega$. This specification considers the overall risk contribution of an ETF leverage choice, including its covariance with the investor's wealth in addition to its own variance. If the investor's wealth is positively correlated with the market ($\beta_\omega > 0$), $\lambda\beta_\omega$ shows that positive leverage has an additional risk of increasing the investor's overall market exposure while negative leverage yields an additional hedging return. Though we can no longer isolate risk aversion λ, we can still recover the distribution of investor expectations F.

We estimate specifications with fixed portfolio risk β_ω and investor-specific portfolio risk β_{ω_i}, where we allow β_{ω_i} to follow a normal distribution. We integrate out these unobserved preferences as random coefficients. We present our estimates for the model with heterogeneity in risk aversion and portfolio risk, but the results are similar when we do not allow for heterogeneity.

Figure A5 displays estimates for investor expectations with hedging, which are close to our baseline results in Figure 4. The estimated mean expected returns in these two models are highly correlated with a correlation coefficient of 0.97. Our estimates suggest that portfolio demand is not meaningful in the context of S&P 500 ETFs. While recover an average value of $\beta_\omega/\delta = -0.788$. Though we cannot separately identify δ, it is unlikely that active investment in S&P 500 ETFs makes up a significant fraction of investors' wealth. To provide an "upper bound" estimate of the effects of

hedging demand, assume that investors place a significant fraction of their wealth—10 percent—into ETFs. If on average $\delta \approx 0.1$, then $\beta_\omega = -0.0788$. This suggests investors behave as if their wealth is nearly uncorrelated with the market. In other words, this hedging term is close to zero, suggesting that investors behave as if there is little hedging consideration against market risk.

In estimation, we might pick up collinearity between $(2j+1)\sigma^2$ and σ^2, so we hesitate to interpret this result strongly. But this exercise shows that our method is capable to account for multiple sources of risks in more general settings.

B.2 Time-Varying and Heterogeneous Risk Aversion

In the main text, we show results where the perceptions of risk vary across investors in Section 4.3 and vary over-time in Section 4.4. This Appendix section presents three extensions. First, we allow perceptions of risk to vary both across time and individuals. Then, we show two extensions where we allow idiosyncratic risk aversion draws to be correlated with expectations.

With two dimensions of heterogeneity that vary over time, it becomes harder to evaluate distributions analytically as in our baseline estimation procedure in Section 4.1. Therefore, we use simulated maximum likelihood for estimations in this Appendix section. Specifically, we take 10,000 random draws for expectations and another 10,000 draws for risk aversion using the Sobol sequence,[38] and simulate choice probability of leverage categories based on these random draws.

Consider the extension of the model where investor utility is given by:

$$u_{ijt} = \beta_j \mu_{it} - p_{jt} - \frac{\lambda_{it}}{2} {\beta_j}^2 \sigma^2$$

where we assume risk aversion λ_{it} follows the time-varying distribution $G_t(\cdot)$. In our estimation procedure we allow the distribution of risk-aversion to vary at the annual level, and we initially assume that beliefs are independent of risk aversion, which we relax in additional specifications.

We present the estimated distributions of investor beliefs and risk aversion in Figure A6. The estimated distribution of investor beliefs is qualitatively similar to our baseline specification where risk aversion is homogeneous across investors and constant over time. The correlation between the mean belief from our baseline model and the mean belief from our extended model is 0.92. Figure A6b displays the distribution of investor risk aversion. In general, the dispersion in risk aversion is relatively small, ranging from 0.03 to 0.31. Consistent with results in the main text, we find that risk aversion, or alternatively investor perceptions about risk, were low heading into and the start of the financial crisis.

We also consider two additional specifications where we allow an investor's perceptions about risk to be time varying and heterogeneous as well as potentially correlated with the investor's beliefs about stock market returns μ_{it}. We use a Gaussian copula to model the correlation. Let $z_{\mu,it}$

[38] Sobol numbers fill the parameter space more evenly than (pseudo-)random numbers. Using Sobol sequences results in a faster convergence and more stable estimates. We draw Sobol random numbers between (0,1), and then use the cumulative distribution functions to invert expectations and risk aversions.

and $z_{\lambda,it}$ denote the CDF for expected return and risk aversion. The Gaussian copula assumes that their standard normal inverses follow a bivariate normal with correlation ρ.

$$\Phi_\rho(\Phi^{-1}(z_{\mu,it}), \Phi^{-1}(z_{\lambda,it}))$$

where Φ^{-1} denotes standard normal inverse and Φ_ρ a bivariate normal with correlation ρ.

It is perhaps more intuitive to consider correlation between risk aversion and a measure of how extreme the expected returns are. Investors with highly positive and highly negative beliefs may have low perception of risk, whereas investors who believe return will be around zero may be more risk averse. Hence, we also consider an additional specification where we allow for correlation between μ_{it}^2 and risk aversion. The square of a random variable from the skewed t distribution does not follow a standard distribution. For computational tractability, we use the F distribution (which captures the squared t distribution) to approximate the CDF of μ_{it}^2. Thus, we ignore the estimated skewness in μ_{it} when generating correlation between extremeness in beliefs and risk aversion.

We report the estimated distributions of beliefs about stock market returns and risk aversion in Figure A7. We estimate the correlation between perception of risk and expected return to roughly 0.04, and the correlation with expected return squared to be roughly 0.03. Note that these are different from the estimated parameter ρ in the Gaussian copula. The correlation between the mean beliefs from Figure A6 and these two specifications is 0.997 and 0.990. Thus, allowing for correlation between risk aversion and expectations has little impact on our estimates. The estimated distributions of investor beliefs are also similar to our baseline results. The correlation between the mean belief from our baseline model and the mean belief when we allow for correlation between perceptions about risk and expected return or expected return squared is 0.927 and 0.943. Using standard goodness-of-fit criteria (AIC and BIC in Table A1), we would select our baseline model with homogeneous risk aversion over any of these extensions.

B.3 Implied versus Realized Volatility

In the main text, we measure volatility using the VIX, which is a measure of the implied volatility and includes the variance risk premium. To the extent that the variance risk premium is constant and positive, our risk aversion or risk perception parameter λ is underestimated by a scalar factor because the VIX would overstate the true underlying volatility in the market. The variance risk premium also likely varies over time and with the market, which would impact our estimates of risk aversion in a potentially more meaningful way, especially when we allow risk aversion to vary over time.

As a robustness check, we estimate our model where we measure volatility using 90-day realized volatility rather than the VIX. Figure A8 shows that the estimated distribution of beliefs and risk aversion follow very similar patterns to the time varying risk aversion results when we use the VIX in Figure 6. The correlations of mean beliefs and risk aversions are 0.95 and 0.94 between the

estimates using VIX and the estimates using realized volatility. If we scale the VIX-based estimates of λ by the average ratio of VIX to realized volatility (a factor of 1.77), we obtain an average parameter value 1.1, which is similar to the average value of 1.3 displayed in Figure A8. These results give us some confidence that the variation in the variance risk premium is not driving the patterns of beliefs that we capture. Because realized volatility may often be noisier than expected volatility, we use VIX as our preferred measure of expected volatility.

B.4 Awareness of Leverage Categories

In the main text, we assume that investors are aware of all of the available leverage categories $j \in \{-3, -2, -1, 0, 1, 2, 3\}$. However, it is possible in practice that some investors are not aware of all of the available leverage categories, especially the -3x and 3x categories which were introduced in 2009 after the -2x and 2x categories. In this section, we estimate an alternative version of the model where we combine the 2x and 3x leverage categories and the -2x and -3x leverage categories.

We re-estimate the model where we combine the market shares of positive leveraged (2x and 3x) and negative leveraged (-2x and -3x) and update our likelihood equation. In our baseline estimation approach, we observe six different points of the belief distribution $F(\cdot)$ at any given moment in time. After combining the positive and negative leverage categories, we only observe four moments of the belief distribution $F(\cdot)$. Because of the reduced number of moments, we estimate the belief distribution with the skewed t distribution while holding fixed the parameter governing kurtosis (the heaviness of the tails) at our baseline estimate of 1.262.

We display the estimated distribution of beliefs in Figure A9. The distribution of beliefs is quantitatively similar to our baseline estimates (the correlation of mean beliefs is 0.98), and the estimated risk aversion parameter λ is also quite similar. This suggests that a lot of the underlying variation driving the belief distribution is coming from an investors' choice to be positively leveraged or negatively leveraged rather than the specific leverage category they choose.

The most obvious potential bias to our results is that some investors who purchase the 1x ETF products are unaware of leveraged options. To the extent this happens in the data, our methodology will understate the number of optimistic investors. If one knew how many investors were unaware of these products (i.e., 30% of S&P 500 Index ETF investors are unaware of leveraged ETFs), one could, in principle, adjust our estimated distribution by distributing the 30% of investors who purchase the 1x category across the 1x, 2x, and 3x leveraged ETFs.

In addition, the leveraged ETF market for S&P 500 ETFs was well established by the start of our sample period. Figure 1 shows that leveraged ETFs were actually the most popular in relative terms at the start of our sample. That time period obviously coincides with the global financial crisis, but it suggests that investors were familiar with these products.

A related concern is that investors may be differentially aware of positively and negatively leveraged ETFs. For example, if investors tend to be more familiar with positively leveraged ETFs than negatively leveraged ETFs, then we would systematically underestimate the number of pessimistic investors. Furthermore, awareness could vary over time and could be related to the past perfor-

mance of an ETF. While we do not have any data on consumer awareness of leveraged products, ETF providers tend to launch positively and negatively leveraged ETFs at the same time. To the extent that consumers learn about new products over time, the fact that positively and negatively leveraged ETFs are launched and marketed simultaneously suggests that investors are likely to be equally aware of both types of products.

B.5 Market Shares Based on AUM

In the main text, we measure market shares based on monthly trading activity. This definition allows us to capture investor expectations at the time of transaction. Alternatively, one could measure market shares based on the stock of holdings (AUM), which would also reflect past purchase decisions. Previous research shows that an investor's beliefs are more correlated with her trading decisions than her holdings (Giglio et al., 2019), motivating our baseline definition. As a robustness check, we also estimate the model where we construct market shares based on AUM.

Here, we construct the quantity of ETF j purchased by retail investors at time t as

$$Quantity_{jt}^{Alt} = \begin{cases} Retail_j \times AUM_{jt} & \text{leveraged ETFs} \\ Retail_j \times AUM_{jt} \times \omega & \text{1x trackers,} \end{cases}$$

where $Quantity_{jt}^{Alt}$ is our alternative measure. We first take AUM for each ETF at the end of each month. We scale AUM by the average retail ownership of each ETF in our sample, $Retail_j$, following the same procedure in our baseline definition. To account for the fact that trackers are often held by passive buy-and-hold investors (e.g., individuals saving for retirement), we scale by tracker AUM by the fraction tracker investors that are active (i.e., non buy-and-hold investors). Specifically, we assume that a fraction ω of tracker ETFs is held by active investors, while the remaining $1 - \omega$ fraction is held by passive buy-and-hold investors who never trade. We assume that all leveraged ETFs are held by active investors. Under the assumption that all active investors have the same probability of trading in a given month, we calculate the active fraction among trackers as:

$$\omega = \frac{\frac{Trading\,Volume_{tracker}}{AUM_{tracker}}}{\frac{Trading\,Volume_{leveraged}}{AUM_{leveraged}}}.$$

Thus, ω is equal to the ratio of the average trading propensity of tracker ETFs to the average trading propensity of leveraged ETFs, where trading propensity is defined as purchase volume during each month over AUM at the end of the month. In the data, we calculate ω using a time-invariant measure of average trading propensity that pools all tracker ETFs or leveraged ETFs over time. Note that because ω and $Retail_j$ are constant over time, the time-series variation in market shares will be driven entirely by time-series variation in AUM.

Figure A10 shows the distribution of expectations based on the alternative market share and the fit of log shares. The distribution of beliefs are similar to our baseline estimates. The estimated

location and dispersion parameters have correlation coefficients of 0.84 and 0.97 with the estimated parameters using our baseline market share definition. The mean belief has a weaker correlation of 0.27. The most notable difference is that the mean belief does not decline sharply during the last quarter in 2008, which suggests that holdings were slower to adjust to during the crisis and is consistent with the idea that trading decisions are potentially more reflective of beliefs than holdings.

C More Asset Classes

In addition to S&P 500, we also consider other asset classes including gold, oil, European equities, emerging market equities, US real estate, medium-term Treasury, and long-term Treasury. For each asset class, we include ETFs tracking the following indices:

- Gold: Bloomberg Gold Subindex, NYSE Arca Gold Miners Index, MSCI ACWI Select Gold Miners Investable Market Index, and the spot price of gold. Our data for gold ETFs starts in Q1 2009.
- Oil: Bloomberg WTI Crude Oil Subindex, WTI Crude Oil and Brent Crude Oil futures. Data starts in Q1 2009.
- European equities: FTSE Developed Europe Index. Data starts in Q3 2009.
- Emerging market equities: MSCI Emerging Market Index. Data starts in Q1 2008.
- US real estate: Dow Jones US Real Estate Index and MSCI US REIT Index. Data starts in Q1 2008.
- Medium-term Treasury: Barclays US Treasury 7-10 Year Index, ICE US Treasury 7-10 Year Bond Index, Merrill Lynch 7-15 Year US Treasury Index, and NYSE 7-10 Year Treasury Bond Index. Data starts in Q1 2009.
- Long-term Treasury: Barclays US Treasury 20+ Year Index, NYSE 20+ Year Treasury Bond Index. Data starts in Q1 2009.

Figures 8 and A11 plot estimated expectation distribution and market share for these seven other asset classes. Europe and US Real Estate show different peaks of dispersion corresponding to the sovereign debt crisis in Europe and the subprime mortgage crisis in the US. Long-term Treasury exhibits large dispersion in 2013, possibly due to speculation that Federal Reserve might start to wind down its quantitative easing program (tapering). Mid-term Treasury has a few discrete spikes in 2014, most likely corresponding to some idiosyncratic trading of institutions that we are unable to filter out using average retail ownership across time.[39]

[39]Different from other asset classes, leveraged ETFs in treasuries have reasonably low fees, so there is larger institutional demand.

We also compare the correlations across these asset classes. Table A2 shows that the mean expectations we recover generate reasonable correlation patterns across asset classes. US stock market comoves positively with European and emerging market equities in general, and is also positively correlated with real estate. US stock is positively correlated with Treasury and negatively correlated with commodities (gold and oil). On the other hand, emerging market is positively correlated with commodities and negatively correlated with Treasury.

C.1 Investor Expectations Across Asset Classes

We examine the extrapolative nature of beliefs in our panel setting using the following regression

$$\overline{E[R]}_{iq} = \beta AnnualRet_{iq} + \mu_i + \mu_q + \epsilon_{iq}. \tag{14}$$

The dependent variable $(\overline{E[R]}_{iq})$ is the average expected return of asset i at time q, and the key independent variable of interest is the corresponding past one year return $(AnnualRet_{iq})$. Observations are at the quarter by asset class level. The panel setting allows us to control for asset and time fixed effects.

Table A3 displays the corresponding regression estimates. Consistent with our previous results for the S&P 500, the results suggest investor beliefs are extrapolative across asset classes. The results in column (1) indicate that a one percentage point increase in historical returns is correlated with a 5 basis point increase in the mean expected return. We also find that the dispersion and skewness of investor beliefs are correlated with past returns across asset classes. Negative returns are correlated with an increase in investor disagreement; a one percentage point decrease in returns is correlated with a one percent increase in the dispersion parameter (column 2). We also find evidence that skewness of beliefs is positively correlated with past returns.

We also test whether investor expectations predict future returns across asset classes in the following regression specification

$$FutureRet_{iq} = \overline{E[R]}_{iq} + \mu_i + \mu_t + \epsilon_{iq} \tag{15}$$

The dependent variable $(FutureRet_{iq})$ measures the realized annual return of asset i from time q to $q+4$. The independent variable $(\overline{E[R]}_{iq})$ is the average expected return of asset i at time q. Observations are at the quarter by asset class level. We report the corresponding estimates in Table A4. The results in column (1) suggest that expected returns are negatively correlated with future returns; however, the effect becomes much smaller and statistically insignificant once we control for asset and time fixed effects. Consistent with our results for the S&P 500, investor beliefs do not forecast future returns across the eight major asset classes.

D Details on Survey Data

In this section, we provide additional details on the survey data and scatter plots showing the relationships between the survey measures and our estimated distribution of expectations.

Duke CFO Global Business Outlook: The Duke CFO Global Business Outlook surveys CFOs at a quarterly frequency about their views on the stock market and macroeconomic outlook. As part of the survey, CFOs are asked to report their expectations of the market risk premium over the upcoming year. The organizers of the survey report both the mean and standard deviation of the expected market risk premium across survey respondents, as well as the fraction with a negative outlook (Graham and Harvey, 2011). We examine how these moments of the distribution of the expected market risk premium across CFOs compare with the estimated moments from our model. This survey provides a nice demonstration of how we can construct statistics that map our model to survey results.

Figure A12 panels (a)-(c) display binned scatter plots, comparing the moments from the survey to our estimated moments. Each panel is constructed using quarterly data over the period 2008-2018 from the CFO survey and our estimates. Figure A12a displays a binned scatter plot of the estimated mean expected market risk premium across ETF investors versus the mean expected market risk premium across CFO survey respondents. The two series are positively and significantly correlated, exhibiting a correlation of 0.38. Figure A12b compares the standard deviation of expected returns across the two series. The standard deviation of the expected market risk premium across ETF investors is significantly and positively correlated (0.41) with the corresponding standard deviation across CFOs. The Duke CFO survey also reports the fraction of respondents expecting a negative market return over the course of the next year. We construct an analogous measure in our ETF data by examining the fraction of investors who prefer negative leveraged ETFs. Figure A12c displays a binned scatter plot of the share of CFO respondents versus the share of ETF investors with a negative market outlook. Again the two series are positively and significantly correlated with each other (0.65). It is also worth noting that the magnitudes are remarkably similar.[40] Overall, the results suggest that the distribution of investor beliefs about the stock market recovered from our model is similar to the distribution of investor beliefs reported in the Duke CFO Global Business Outlook.

Wells Fargo/Gallup Investor and Retirement Optimism Index: The Gallup Investor and Retirement Optimism Index is constructed using a nationally representative survey of U.S. investors with $10,000 or more invested in stocks, bonds, and mutual funds.[41] The index is designed to capture a broad measure of U.S. investors' outlook on their finances and the economy based on their survey

[40] A regression of the share of CFO respondents with a negative market outlook on the share of ETF investors who purchase negative leveraged ETFs yields a coefficient of 0.80 and is statistically indistinguishable from 1.

[41] The data is calculated from the figures reported online from https://news.gallup.com/poll/231776/investor-optimism-stable-strong.aspx. A full description of the index is available online https://www.gallup.com/207062/wells-fargo-gallup-investor-retirement-optimism-index-work.aspx.

responses and Gallup's proprietary index construction methodology. Given that we are unable to directly construct an analogous index, we construct a measure of "optimism" using the fraction of investors choosing positive leverage versus those choosing negative leverage. Specifically, we use the following measure

$$M = \frac{\sum_{j=\{1,2,3\}} \hat{s}_j}{\sum_{j=\{1,2,3\}} \hat{s}_j + \sum_{k=\{-3,-2,-1\}} \hat{s}_j} \tag{16}$$

where $\hat{s}_j$ is the predicted share from our model.[42] This measure is similar to the percent bullish minus percent bearish measure used in Greenwood and Shleifer (2014) and helps capture information about the beliefs of the median ETF investor.

Figure A13 displays the relationships between additional surveys and analogous measures from our ETF measurements, corresponding to quarterly time series from 2008 to 2018. Panel (a) presents a binned scatter plot of our measure of optimism compared to the Gallup Investor and Retirement Optimism Index. The two series are positively and significantly correlated (0.70) in the time series. In other words, there is a positive relationship between investor outlook measured by Gallup, and the relative share of investors preferring positive leverage to negative leverage based on their estimated expectations. Though we omit the results for brevity, the Gallup index is also positively and significantly correlated with our estimates of expected mean returns.

University of Michigan Surveys of Consumers: The University of Michigan Surveys of Consumers asks consumers about the probability that the stock market increases. Specifically, the survey asks a set of nationally representative of US consumers to report the percent chance that a "one thousand dollar investment in the stock market will increase in value a year ahead." Constructing an analogous measure using our model is challenging because the subjective belief about the probability of a stock market increase depends both on the expected stock market return and also the beliefs of the distribution of returns. Similar to our analysis with the Gallup index, we compare the University of Michigan index to the relative share positive versus negative from equation (16).

Figure A13 panel (b) shows that stock market beliefs from Michigan Surveys and our estimates are significantly and positively correlated (0.77). This correlation suggests that our ETF data and model estimates mirror the beliefs of consumers more broadly. The University of Michigan index is also positively and significantly correlated with our estimates of expected mean returns, though, as above, we omit the results for brevity.

American Association of Individual Investors (AAII) Sentiment Survey: The American Association of Individual Investors surveys its members each week about their sentiment towards the stock market over the next 6 months. Specifically, the survey asks respondents whether they believe the stock market over the next six months will be up (bullish), no change (neutral), or down

[42] Note that the predicted shares correspond closely to the shares in the data as we obtain a high degree of model fit.

(bearish).[43] Because the percent bullish and percent bearish are highly correlated in the survey, we construct a single measure, $\frac{bullish}{bullish+bearish}$, which corresponds closely to the relative share positive versus negative from equation (16). Comparing each response separately to analogous measures from our estimates yields similar results.

Panel (c) in Figure A13 displays the relationship between the AAII survey and our estimates. The plot shows the relative share bullish compared to our measure of relative share positive (omitting neutrals). The correlation between the two measures of sentiment is positive and significant (0.33), which indicates relatively more investors purchase positive leverage when AAII respondents have a more positive outlook on the market.

Shiller U.S. Individual One-Year Confidence Index: The Shiller US Individual One-Year Confidence Index measures the percentage of individual investors who expect the stock market (Dow Jones Industrial) to increase in the coming year.[44] Survey respondents, who are comprised of wealthy individual investors, are asked to provide their expected increase in the stock index over the upcoming year, and the confidence index measures the percentage of investors who report a positive expected increase in the stock market. For this survey, we produce a proxy measure using the fraction of investors who would choose positively leveraged ETFs, i.e., $\sum_{j=\{2,3\}} \hat{s}_j$. Panel (d) of Figure A13 displays a binned scatter plot of the share of investors purchasing positively leveraged ETFs and the One Year Confidence Index. The two series are positively and significantly correlated (0.47), indicating that the preferences revealed through leveraged ETF purchases line up well with the analogous Shiller survey measure.

Survey of Professional Forecasts: The Philadelphia Federal Reserve surveys professional forecasters each quarter about their views regarding economic growth as part of the Survey of Professional Forecasts (SPF). We focus on the forecast of annual real GDP growth since the SPF does not include stock market forecasts. Panel (e) in Figure A13 displays the relationship between the estimated mean expected stock market risk premium across ETF investors versus professional GDP forecasts. The two series are positively and significantly correlated (0.82). The SPF also reports the interquartile range of GDP forecasts. Panel (f) displays the interquartile range of GDP forecasts across investors versus the interquartile range of stock market beliefs across ETF investors. As with the mean belief/forecast, the two series are positively and significantly correlated (0.86).

Overall, the results displayed in Figures A12 and A13 help shed light on the external validity of our estimates. The expectations we recover from demand for S&P 500 linked ETFs are highly and significantly correlated with the investor expectations measured in six different surveys. Our

[43]The typical AAII member is a male in his mid-60s with a bachelor's or graduate degree. AAII members tend to be affluent with a median portfolio in excess of $1 million. The typical member describes himself as having a moderate level of investment knowledge and engaging primarily in fundamental analysis. For further details see https://www.aaii.com/journal/article3/is-the-aaii-sentiment-survey-a-contrarian-indicator [accessed 11/17/2019]

[44]Data are available online at https://som.yale.edu/faculty-research-centers/centers-initiatives/international-center-for-finance/data/stock-market-confidence-indices/united-states-stock-market-confidence-indices [Accessed 10/31/2019]

estimates of investor beliefs help complement the survey data. While the survey data are intended to capture the beliefs of distinct populations, our belief estimates come from the actual investment decisions of investors.

Table A1: Goodness of Fit Across Specifications, Ranked by BIC

		Goodness of Fit								Median Parameters of Belief Distribution			
μ	λ	AIC	BIC	log likelihood	r^2	Number of Parameters	Correlation of mean belief with baseline	Mean Risk Aversion	Median Risk Aversion	Mean	Location	Dispersion	Skewness
Skewed t	Homogeneous, time invariant (baseline)	2,565.0	3,208.5	-168.4	0.921	134	-	0.98	0.98	0.77	2.86	1.08	0.77
Skewed t	Uniform, time invariant	2,565.5	3,213.8	-168.3	0.923	135	0.999	0.85	0.85	0.72	2.54	1.01	0.76
Skewed t	Homogeneous, varying annually	2,560.7	3,252.2	-166.7	0.935	144	0.958	0.61	0.62	0.89	2.05	0.91	0.73
Normal	Homogeneous, time invariant	2,864.0	3,291.4	-197.0	0.617	89	0.567	1.76	1.76	3.03	3.03	6.15	-
Skewed t	Uniform, varying annually	2,574.2	3,318.5	-166.0	0.938	155	0.919	0.58	0.56	0.92	2.02	0.85	0.73
Skewed t	Uniform, varying annually, correlated with μ^2	2,575.8	3,325.0	-166.0	0.939	156	0.943	0.59	0.56	0.83	2.00	0.88	0.74
Skewed t	Uniform, varying annually, correlated with μ	2,576.2	3,325.3	-166.0	0.936	156	0.927	0.58	0.56	0.80	1.97	0.90	0.72
Skewed t	Homogeneous, varying quarterly	2,612.6	3,462.7	-165.6	0.943	177	0.790	0.55	0.51	0.86	1.84	0.89	0.73
Skewed t	Uniform, varying quarterly	2,695.3	3,756.6	-165.2	0.943	221	0.784	0.54	0.50	0.96	1.86	0.88	0.71

Notes: Table A1 displays goodness of fit measures and estimation results across different specifications. The baseline specification with homogeneous and time invariant risk aversion is reported in the first row. The first five columns show measures of goodness of fit including AIC, BIC, log-likelihood and R-squared, as well as the number of parameters estimated. The sixth column shows the correlation of mean belief from different specifications with the mean belief of the baseline specification. The next two columns display estimated risk aversion. When risk aversion is heterogeneous, we focus on the average risk aversion. When risk aversion is time-varying, these two columns present the median and mean estimate across time. The last four columns show the median of mean beliefs, and the median of location, dispersion, and skewness parameters across quarters.

Table A2: Correlation of Mean Expectation Across Asset Classes

	S&P 500	Europe	Emerging Mkt	Real Estate	Treasury 7-10	Treasury 20+	Gold	Oil
S&P 500	1							
Europe	0.641***	1						
	(0.000)							
Emerging Mkt	-0.790***	-0.282**	1					
	(0.000)	(0.002)						
Real Estate	0.686***	-0.00523	-0.362***	1				
	(0.000)	(0.956)	(0.000)					
Treasury 7-10	0.762***	0.719***	-0.251**	0.570***	1			
	(0.000)	(0.000)	(0.006)	(0.000)				
Treasury 20+	0.590***	0.531***	-0.533***	0.462***	0.647***	1		
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)			
Gold	-0.472***	-0.306***	0.717***	-0.340***	-0.313***	-0.679***	1	
	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)		
Oil	-0.761***	-0.128	0.391***	-0.653***	-0.571***	-0.472***	0.573***	1
	(0.000)	(0.173)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	
Observations	132							

Notes: Table A2 displays the correlation of mean expectation each quarter across all asset classes. Standard errors are shown in paranthesis. *** p<0.01, ** p<0.05, * p<0.10.

Table A3: Expected Returns versus Past 12-Month Returns Across All Asset Classes

	Mean	ln(Dispersion)	Median	ln(Skewness)
	(1)	(2)	(3)	(4)
Annual Return	0.049**	-0.0095*	0.033*	0.0014***
	(0.022)	(0.0054)	(0.018)	(0.00052)
Time Fixed Effects	X	X	X	X
Asset Fixed Effects	X	X	X	X
Observations	330	330	330	330
R-squared	0.351	0.619	0.256	0.684

Notes: Table A3 displays the regression of different moments of the estimated expected returns distribution on the past 12-month excess return of the corresponding asset class (eq. 14). Observations are at the asset class by quarter level over the period 2008-2018. See Appendix C for a further description of the data. The dependent variable in each column corresponds to different moments/parameters of the estimated expected returns distribution. The dependent variable in column (1) is the mean and is measured in percentage points, in column (2) is the dispersion parameter in logs, in column (3) is the median and is measured in percentage points, and in column (4) is the skewness parameter in logs. The idependent variable Annual Return is measured in percentage points. We winsorize all independent and dependent variables at the 5% level within each asset class to account for outliers during the financial crisis. Winsorizing the data does not change inference on the relationship between the corresponding series. Driscoll-Kraay based standard errors are in parentheses with four lags and are grouped by asset class. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table A4: Future Annual Returns vs. Expected Returns Across All Asset Classes

	(1)	(2)	(3)
$AverageExpectedReturn$	-0.34**	-0.20	0.14
	(0.15)	(0.24)	(0.21)
Asset Fixed Effects		X	X
Time Fixed Effects			X
Observations	330	330	330
R-squared	0.009	0.043	0.404

Notes: Table A4 displays the regression of different moments of the estimated expected returns distribution on the past 12-month excess return of the corresponding asset class (eq. 15). Observations are at the asset class by quarter level over the period 2008-2018. See Appendix C for a further description of the data. The dependent variable measures the realized return of the asset over the next twelve months. The independent variable, Average Expected Return, corresponds to the average expected return from our model. We winsorize all independent and dependent variables at the 5% level within each asset class to account for outliers during the financial crisis. Winsorizing the data does not change inference on the relationship between the corresponding series. Driscoll-Kraay based standard errors are in parentheses with four lags and are grouped by asset class. *** p<0.01, ** p<0.05, * p<0.10.

Table A5: Predictability of Median Forecast Errors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lagged Forecast Error	-0.20*							
	(0.11)							
Annual Return		0.036						
		(0.13)						
Lagged ln(Dispersion)			0.88					
			(1.93)					
Lagged IQR				0.27**				
				(0.13)				
Lagged ln(Skewness)					14.8			
					(21.6)			
Lagged VIX						0.14		
						(0.17)		
Lagged ln(Price/Dividend)							24.6***	
							(7.25)	
Lagged cay								99.7
								(133)
Observations	40	44	43	43	43	43	43	39
R-squared	0.069	0.003	0.004	0.034	0.020	0.012	0.098	0.019

Notes: The table displays the results corresponding to a linear regression model (eq. 12). Observations are at the quarterly level. The dependent variable is the forecast error constructed as the difference between the 12-month return of the S&P 500 from period q to $q+4$ and the median expected return at time q. The forecast error is scaled by 100 such that it is measured in percentage points. Independent variables include four quarters lagged values of forecast errors, past 12-month S&P 500 excess returns, one quarter lagged value of parameters or moments from the belief distribution (dispersion, skew, and interquartile range), and one quarter lagged value of VIX, log price-dividend ratio, and the consumption wealth ratio (cay) of Lettau and Ludvigson (2001). Newey-West based standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

Table A6: Comparison with Future Returns

	3-Month Future Returns	6-Month Future Returns	12-Month Future Returns
	(1)	(2)	(3)
Mean of Exp. Returns	-0.13	-0.77	-0.73
	(0.58)	(0.59)	(0.66)
Observations	44	44	44
R-squared	0.003	0.059	0.025

Notes: Table A6 displays the regression of future returns on the mean belief from our estimated distribution of investor beliefs. Observations are at the quarterly level over the period 2008-2018. The dependent variables are the future 3-month, 6-month, and 12-month excess return of the S&P 500. The independent variable Mean of Exp. Returns corresponds to our baseline model estimates reported in column (1) of Table 3. We winsorize the Mean of Exp. Returns and the future excess returns of the stock market variables at the 5% level to account for outliers during the financial crisis. Newey-West based standard errors are in parentheses with one lag. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Table A7: Dispersion in Beliefs and 6-Month Returns

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
IQR	-0.16					1.05*	
	(0.37)					(0.57)	
Δ IQR		-0.33***					
		(0.086)					
Dispersion (b)			-0.12				2.16*
			(0.89)				(1.12)
Δ Dispersion				-0.74***			
				(0.24)			
Constrained					-2.58***	-2.35**	-2.30**
					(0.72)	(0.99)	(0.97)
Constrained$\times$IQR						-0.062**	
						(0.030)	
Constrained$\times$Dispersion							-0.13**
							(0.062)
Observations	44	43	44	43	44	44	44
R-squared	0.065	0.158	0.060	0.153	0.234	0.286	0.289

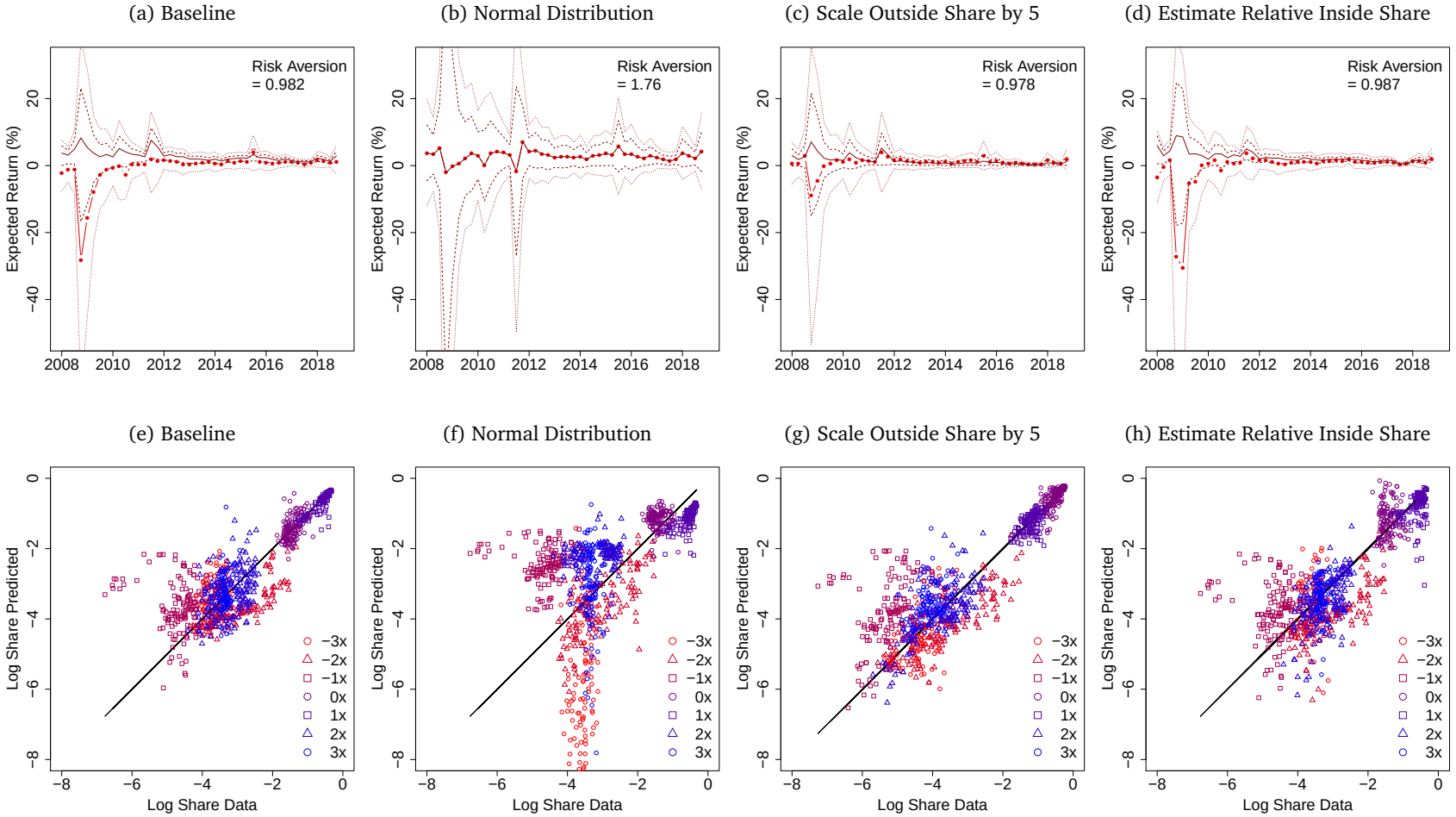
Notes: The table displays the results corresponding to a linear regression model (eq. 13). Observations are at the quarterly level. The dependent variable is the 6-month return of the S&P 500 from period q to $q+2$ and scaled by 100 such that it is measured in percentage points. The variable IQR measures the interquartile range of the estimated belief distribution at time q and is scaled by 100 such that it is measured in percentage points. The variable $Dispersion$ corresponds to the estimated dispersion parameter (b) from our belief distribution. The variable $Constrained$ is share of investors who purchase the lowest available leverage category and is scaled by 100 such that it is measured in percentage points. In each specification we also control for the estimated mean belief at time q. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

Table A8: Dispersion in Beliefs and 12-Month Returns

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
IQR	0.57					1.23	
	(0.41)					(0.74)	
Δ IQR		0.0026					
		(0.12)					
Dispersion (b)			1.53				2.51*
			(1.00)				(1.48)
Δ Dispersion				-0.00061			
				(0.30)			
Constrained					-2.59	-3.54*	-3.43*
					(1.65)	(1.85)	(1.81)
Constrained$\times$IQR						-0.016	
						(0.042)	
Constrained$\times$Dispersion							-0.016
							(0.089)
Observations	44	43	44	43	44	44	44
R-squared	0.059	0.049	0.076	0.049	0.106	0.201	0.211

Notes: The table displays the results corresponding to a linear regression model (eq. 13). Observations are at the quarterly level. The dependent variable is the 12-month return of the S&P 500 from period q to $q+4$ and scaled by 100 such that it is measured in percentage points. The variable IQR measures the interquartile range of the estimated belief distribution at time q and is scaled by 100 such that it is measured in percentage points. The variable $Dispersion$ corresponds to the estimated dispersion parameter (b) from our belief distribution. The variable $Constrained$ is share of investors who purchase the lowest available leverage category and is scaled by 100 such that it is measured in percentage points. In each specification we also control for the estimated mean belief at time q. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

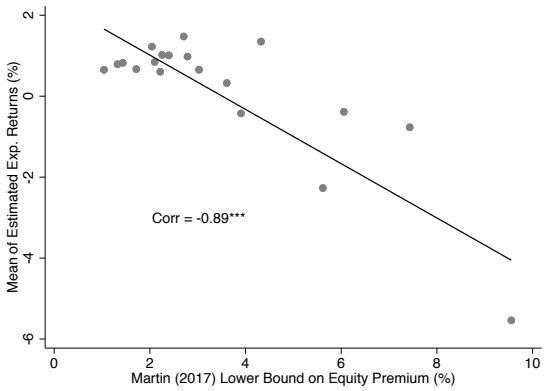
Figure A1: Expectations and Model Fit: Robustness Checks



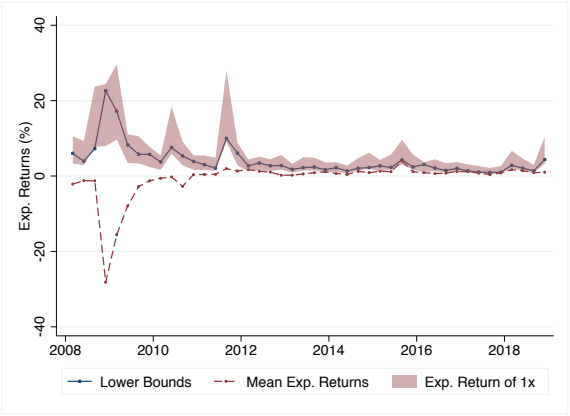
Notes: Figure A1 panel (a) and (e) correspond to the baseline estimates. In (b) and (f), we fit data assuming expectation follows normal distribution. In (c) and (d), we scale the outside share of our baseline definition by a factor of 5. In (d) and (h), we fit relative inside share only without using the share of outside option. For the top panels, red dots represent mean expected return, solid dark red lines indicate median, dashed lines indicate 25th and 75th percentiles, and dotted lines indicate 10th and 90th percentiles. The bottom panels plot fit in terms of log market shares of each leverage category. The x-axis corresponds to log market shares in the data, and y-axis corresponds to predicted log market shares. Color red to blue represents each leverage from -3x to 3x. The solid black lines correspond to the 45 degree line.

Figure A2: Comparision with Lower Bound on Equity Premium

(a) Mean Expected Return vs. Lower Bound

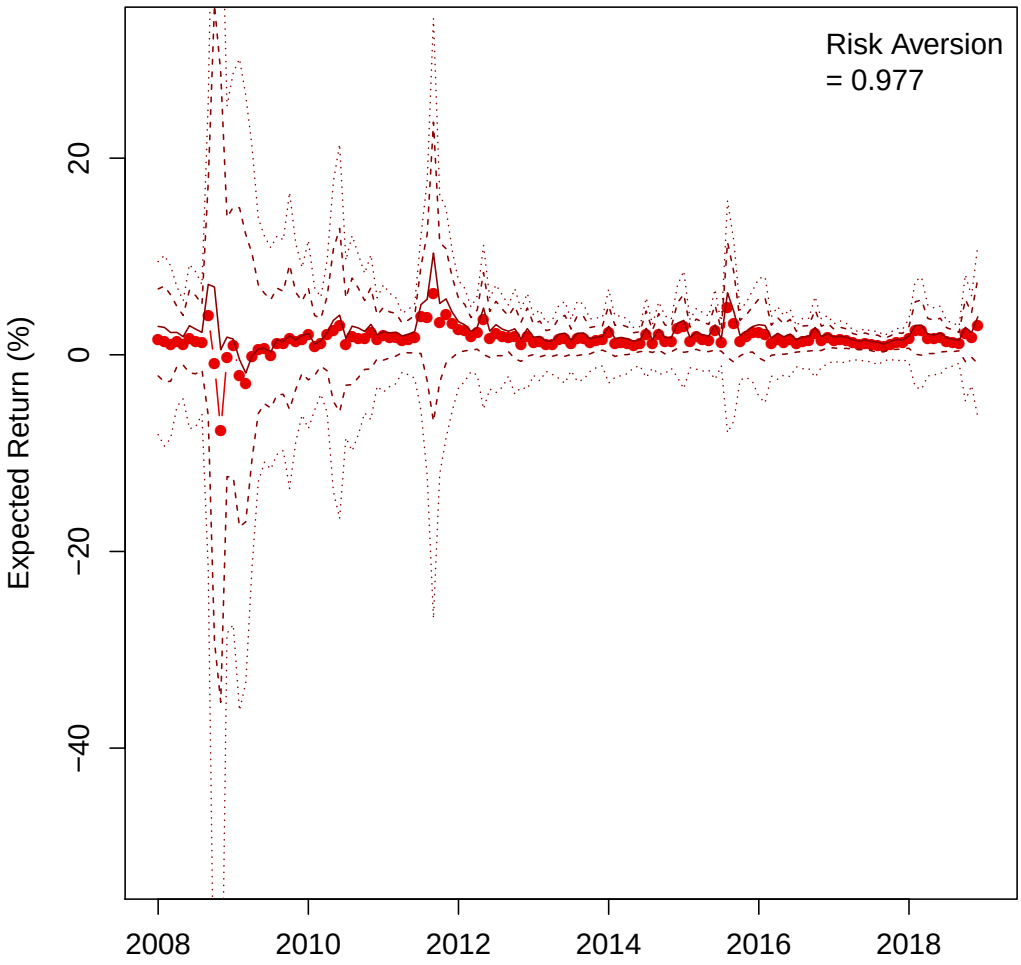


(b) Expected Returns of 1x vs. Lower Bound



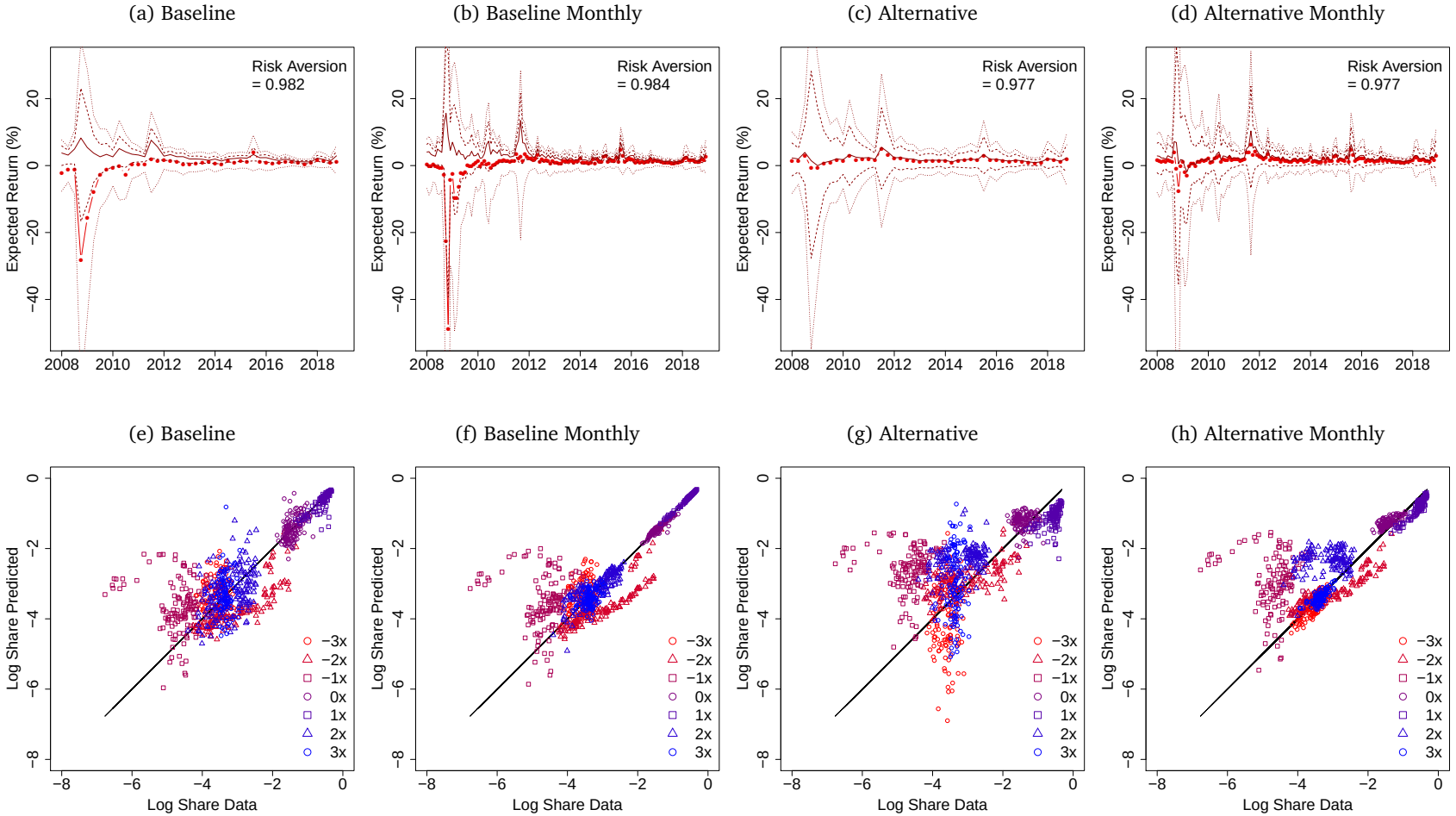
Notes: Figure A2 displays the relationship between the estimated expected returns from our model and our replication of lower bound on equity premium in Martin (2017). The estimated expected returns correspond to our baseline model estimates reported in column (1) of Table 3. Observations in each panel are at the quarterly level over the period 2008-2018. Panel (a) displays a binned scatter plot of the mean estimated expectation from our model versus the lower bound on equity premium. We winsorize the mean of expected returns from our model at the 5% level to account for outliers during the financial crisis. *** p<0.01, ** p<0.05, * p<0.10. Panel (b) plots the time-series of the lower bound on equity premium, the estimated mean expected returns, and the estimated range of expected returns consistent with choosing the 1x leverage.

Figure A3: Time-Varying Investor Expectations: Alternative Estimates



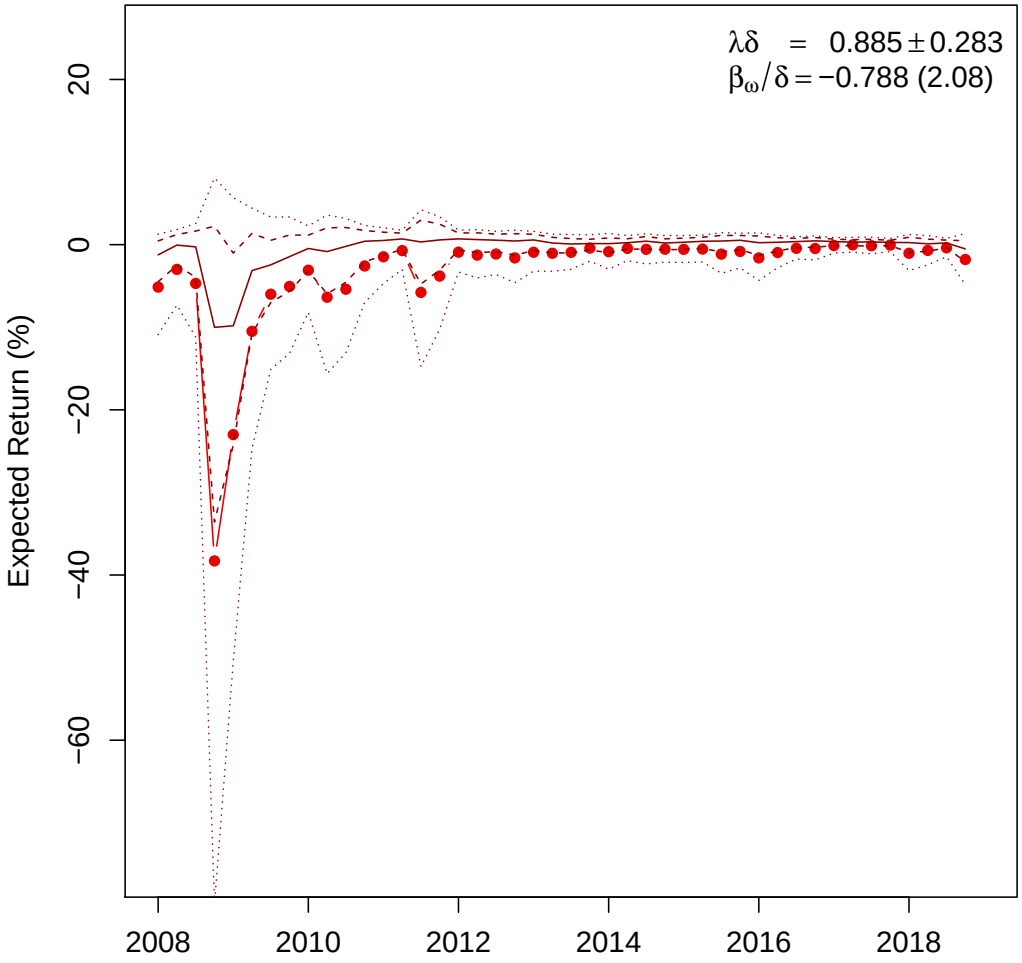
Notes: Figure A3 plots the estimated distribution of investor expectations over time in each month, using the alternative approach described in Appendix A. These estimates use only variation in the choices facing investors to recover the time-variation distribution of expecations. Red dots represent mean expected return. Solid dark red lines indicate median, dashed lines indicate 25th and 75th percentiles, and dotted lines indicate 10th and 90th percentiles.

Figure A4: Expectations and Model Fit: Comparison of Baseline and Alternative Estimates



Notes: Figure A4 panel (a) and (e) correspond to the baseline estimates. Panel (b) and (f) are based on the alternative approach described in Appendix A. These estimates use only variation in the choices facing investors to recover the time-variation distribution of expecations. Panel (c) and (d) are based on the alternative method in Appendix A. Panel (d) and (h) are based on monthly estimates uding the alternative method. For the top panels, red dots represent mean expected return, solid dark red lines indicate median, dashed lines indicate 25th and 75th percentiles, and dotted lines indicate 10th and 90th percentiles. The bottom panels plot fit in terms of log market shares of each leverage category. The x-axis corresponds to log market shares in the data, and y-axis corresponds to predicted log market shares. Color red to blue represents each leverage from -3x to 3x. The solid black lines correspond to the 45 degree line.

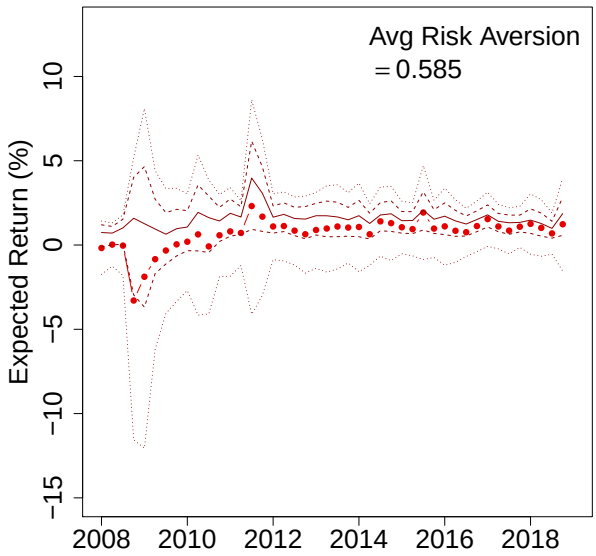
Figure A5: Time-Varying Investor Expectations: Hedging



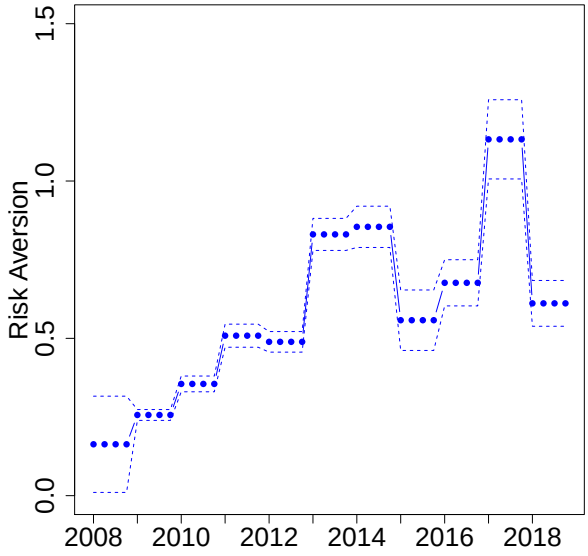
Notes: Figure A5 plots the estimated distribution of expectations over time for investors with hedging considerations. We display the risk aversion coefficient, which follows a uniform distribution, and the market risk beta of investors' wealth, which follows a normal distribution. These coefficients are scaled by the share of wealth invested in ETFs, which we cannot separately identify. Red dots represent mean expected return. Solid dark red lines indicate median, dashed lines indicate 25th and 75th percentiles, and dotted lines indicate 10th and 90th percentiles.

Figure A6: Time-varying Heterogeneous Risk Aversion

(a) Expectations



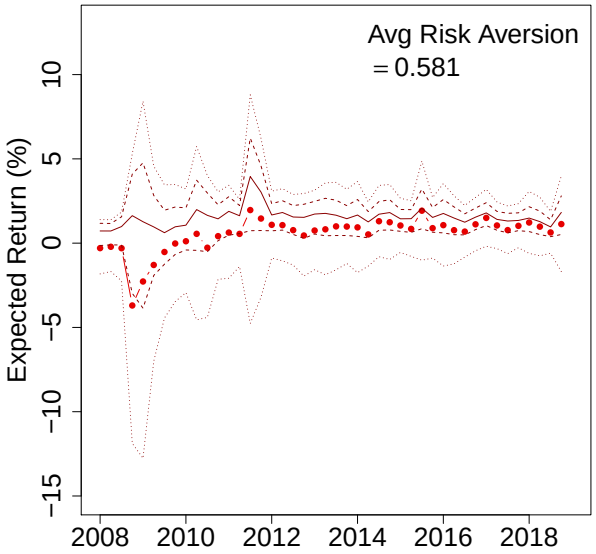
(b) Risk Aversion



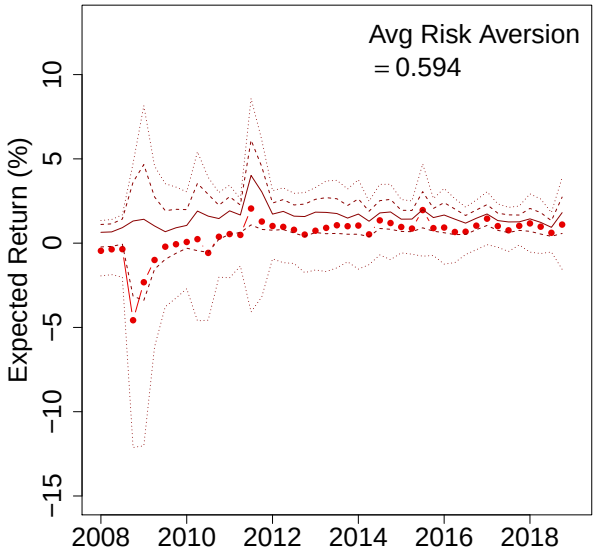
Notes: Figure A6 panel (a) plots the estimated distribution of expectations over time where we allow risk aversion and the dispersion to vary at annual level. Red dots represent mean expected return. Solid dark red lines indicate median, dashed lines indicate 25th and 75th percentiles, and dotted lines indicate 10th and 90th percentiles. Panel (b) plots the estimated distribution of risk aversion, which varies at annual level. The blue dots represent the mean risk aversion. Blue dashed lines show lower and upper bounds of risk aversion.

Figure A7: Time-varying Heterogeneous Risk Aversion where Beliefs are Correlated with Risk Perceptions

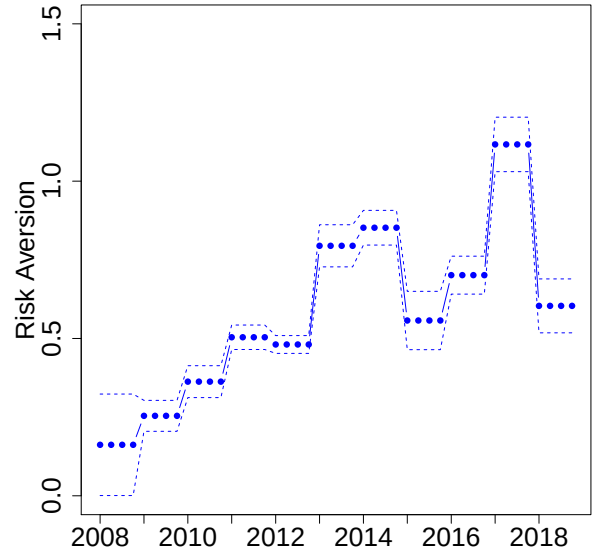
(a) Expectations, Correlation with μ



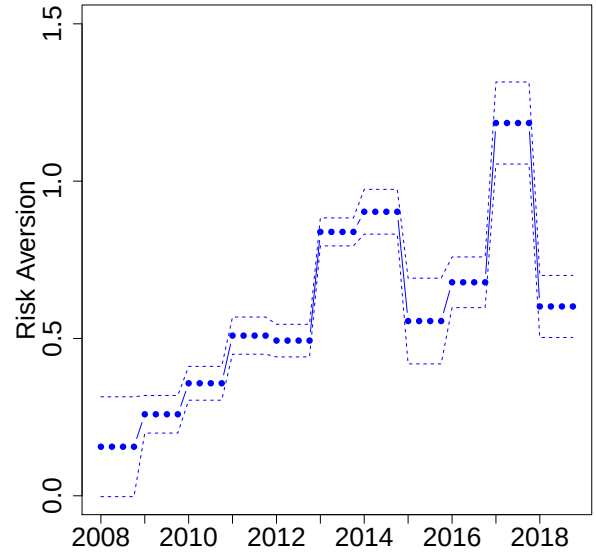
(b) Expectations, Correlation with μ^2



(c) Risk Aversion, Correlation with μ



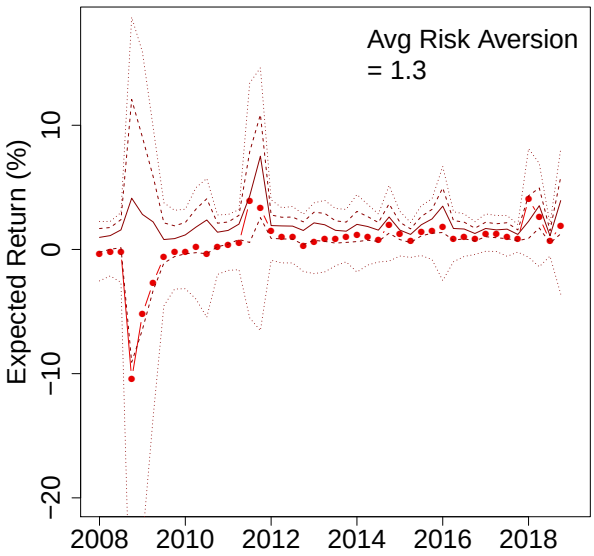
(d) Risk Aversion, Correlation with μ^2



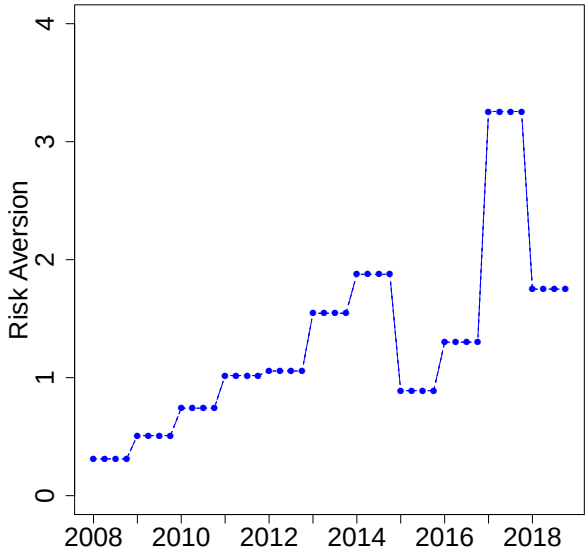
Notes: Figure A7 panels (a) and (b) plot the estimated distribution of expectations over time where we allow risk aversion and the dispersion to vary at annual level. Red dots represent mean expected return. Solid dark red lines indicate median, dashed lines indicate 25th and 75th percentiles, and dotted lines indicate 10th and 90th percentiles. Panels (c) and (d) plot the estimated distribution of risk aversion, which varies at annual level. The blue dots represent the mean risk aversion. Blue dashed lines show lower and upper bounds of risk aversion. In panels (a) and (c), we allow correlation between expectation and risk aversion. In panels (b) and (d), we allow correlation between expectation squared and risk aversion.

Figure A8: Estimation with Realized Volatility

(a) Expectations



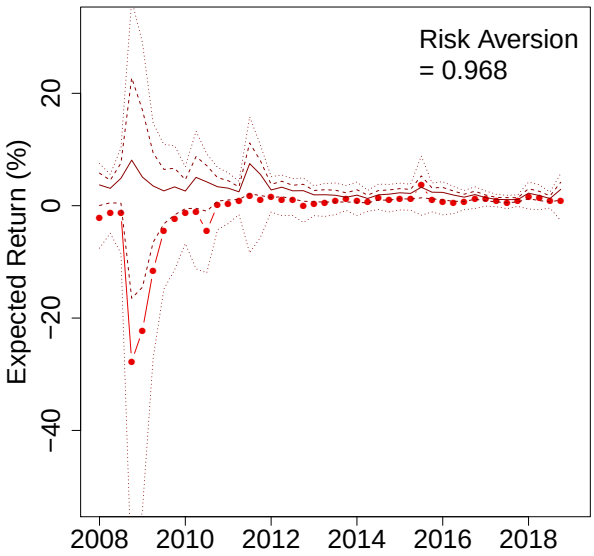
(b) Risk Aversion



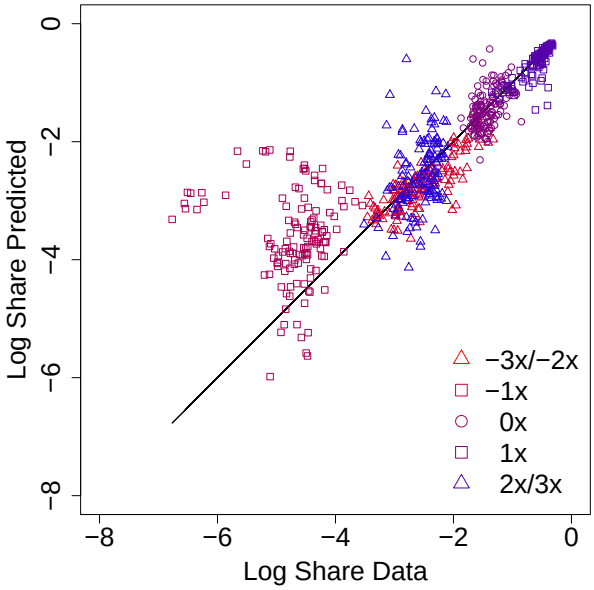
Notes: Figure A8 panel (a) plots the estimated distribution of expectations over time where we allow risk aversion to vary at annual level and use 90 days realized volatility as the proxy for the level of risk. Red dots represent mean expected return. Solid dark red lines indicate median, dashed lines indicate 25th and 75th percentiles, and dotted lines indicate 10th and 90th percentiles. Panel (b) plots the estimated risk aversion, which varies at annual level.

Figure A9: Combining -2x & -3x, and 2x & 3x

(a) Expectations



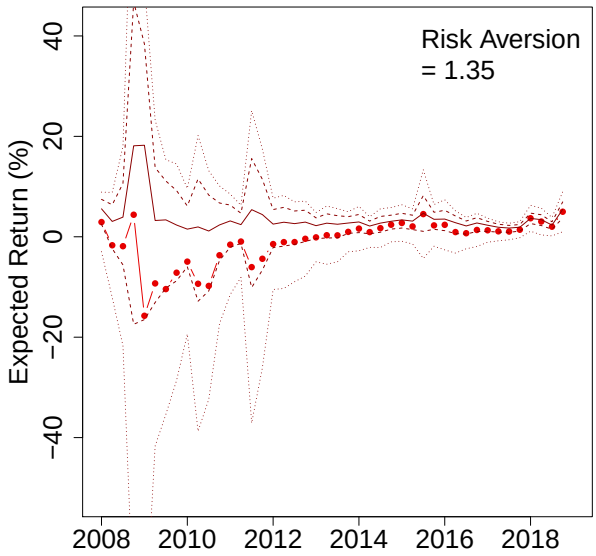
(b) Fit of Log Shares



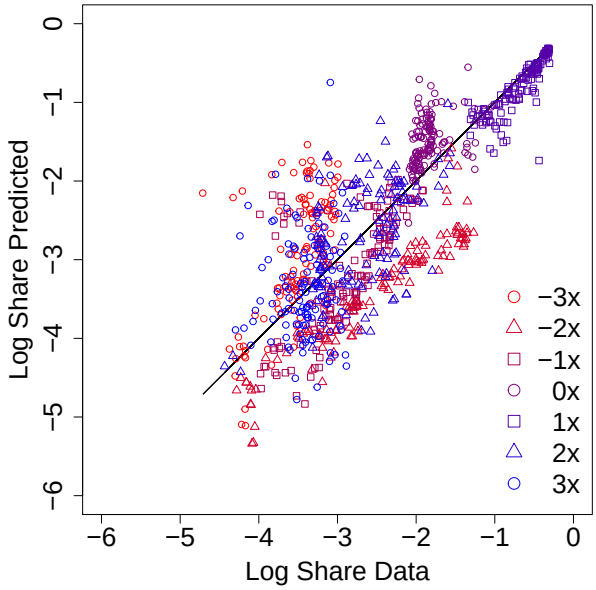
Notes: Figure A9 panel (a) plots the estimated distribution of expectations over time where we combine -2x & -3x into one leverage category and 2x & 3x into one leverage category. Red dots represent mean expected return. Solid dark red lines indicate median, dashed lines indicate 25th and 75th percentiles, and dotted lines indicate 10th and 90th percentiles. Panel (b) plots fit in terms of log market shares of each leverage. The x-axis corresponds to log market shares in the data, and y-axis corresponds to predicted log market shares. Color red to blue represents each leverage. The solid black lines correspond to the 45 degree line.

Figure A10: Expectations with Market Shares Based on AUM

(a) Expectation Distribution

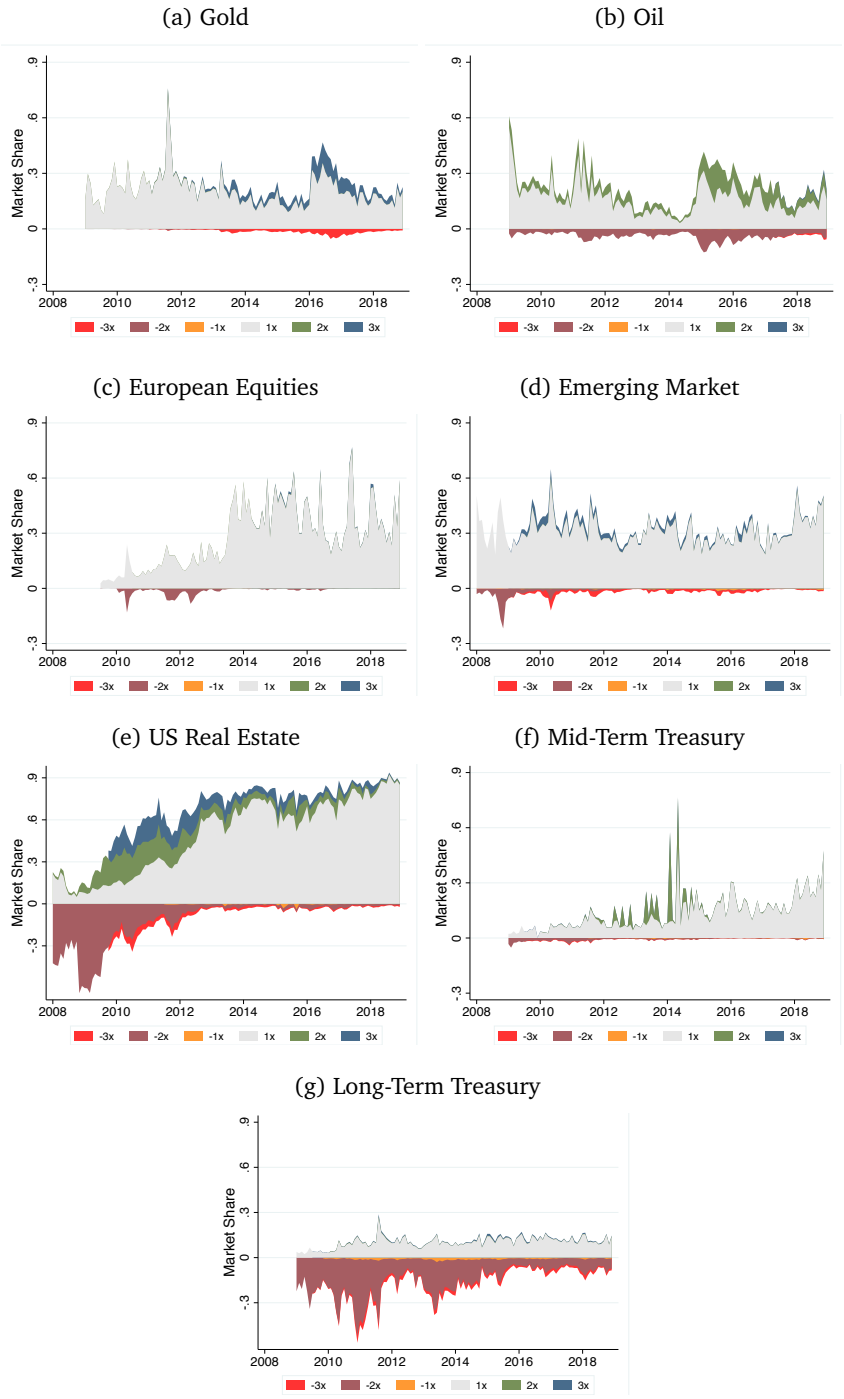


(b) Fit of Log Shares



Notes: Figure A10 plots the estimated distribution of expectations over time where we use assets under management (AUM) to calculate market shares. AUM for 1x trackers is scaled by a constant ratio of the average trading propensity of trackers to the average trading propensity of leveraged ETFs. Red dots represent mean expected return. Solid dark red lines indicate median, dashed lines indicate 25th and 75th percentiles, and dotted lines indicate 10th and 90th percentiles. Panel (b) plots fit in terms of log market shares of each leverage. The x-axis corresponds to log market shares in the data, and y-axis corresponds to predicted log market shares. Color red to blue represents each leverage. The solid black lines correspond to the 45 degree line.

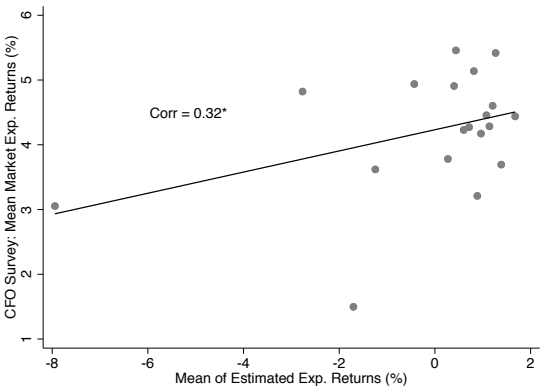
Figure A11: Market Shares: Other Asset Classes



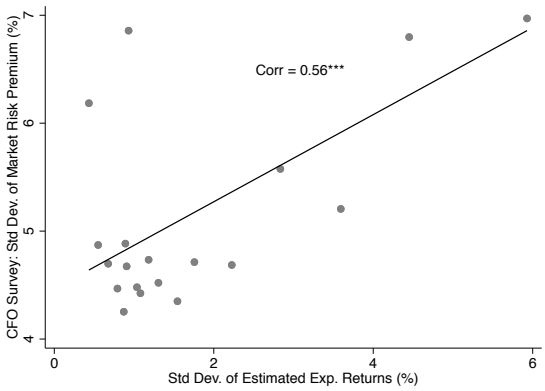
Notes: Figure A11 shows market share of each leverage for each asset class.

Figure A12: Comparison with Duke CFO Global Business Outlook Survey

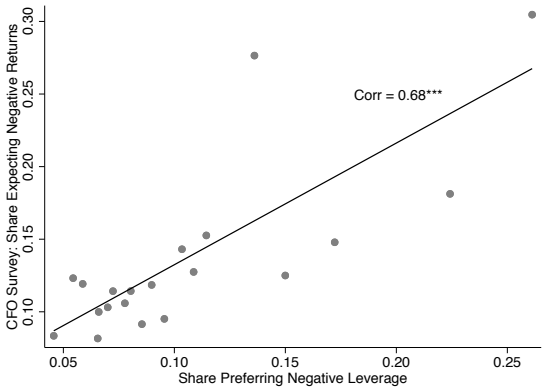
(a) Mean Expected Return



(b) Std. Dev. of Expected Return

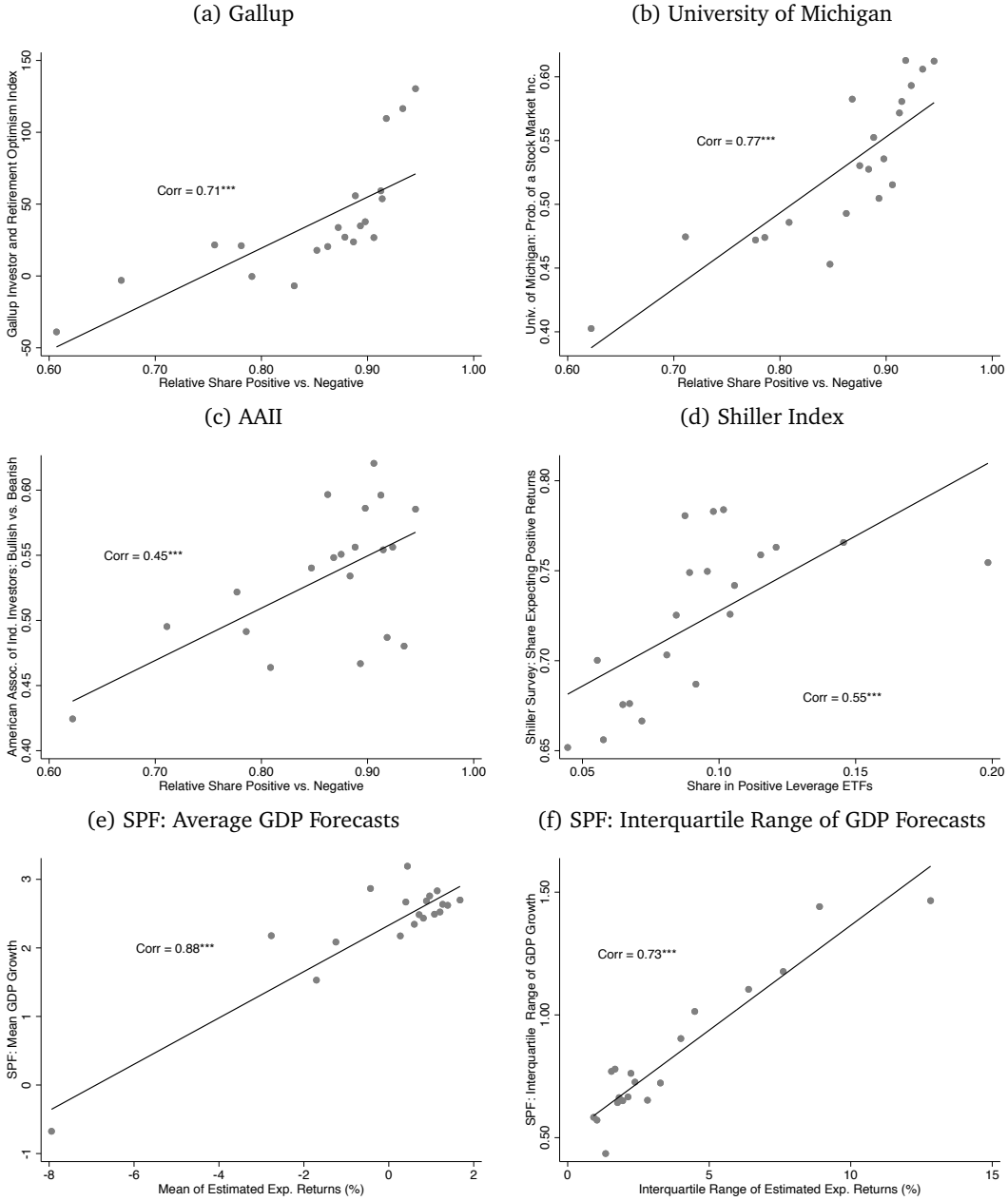


(c) Share with Negative Outlook



Notes: Figure A12 panels (a)-(c) display binned scatter plots of our estimated belief distribution versus results from the Duke CFO Global Business Outlook Survey. The estimated belief distribution corresponds to our baseline model estimates reported in column (1) of Table 3. Observations in each panel are at the quarterly level over the period 2008-2018. Panel (a) displays the relationship between the mean estimated expected return from our model versus the mean expected return from the Duke CFO survey. Panel (b) displays the relationship between the estimated standard deviation of expected returns across investors from our model versus the standard deviation of expected returns across CFOs as reported in the Duke CFO survey. Panel (c) displays the relationship between the market share of negative leveraged ETFs versus the share of CFOs who expect S&P 500 Returns to be negative next year. We winsorize the mean and standard deviation of expected returns from our model at the 5% level to account for outliers during the financial crisis. Winsorizing the data does not change inference on the relationship between the corresponding series. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Figure A13: Comparison with Surveys



Notes: Figure A13 displays the relationship between the estimated expectations from our model and five additional surveys: (a) the Gallup Investor and Retirement Optimism Index, (b) the University of Michigan Survey of Consumers, (c) the American Association of Individual Investors (AAII) Sentiment Survey, (d) the Shiller U.S. Individual One-Year Confidence Index, and (e)-(f) the Survey of Professional Forecasters (SPF). The estimated belief distribution corresponds to our baseline model estimates reported in column (1) of Table 3. Observations in each panel are at the quarterly level over the period 2008-2018. For details on these surveys, see Section 4.5. Panels (a)-(f) display binned scatter plots comparing each survey to an analogous measure from our model. Surveys in panels (a)-(c) are compared to the relative share of investors preferring positive to negative leverage, based on our estimated distribution of expectations. The Shiller index in panel (d) is compared to the fraction of investors choosing positive leverage (greater than 1x). The SPF average GDP growth forecast in panel (e) is compared to the mean estimated expected return from our model. The interquartile range of GDP forecasts across professional forecasters in the SPF in panel (f) is compared to the interquartile range of estimated expected returns from our model. In panels (e) and (f) we winsorize the mean and interquartile range of expected returns from our model at the 5% level to account for outliers during the financial crisis. Winsorizing the data does not change inference on the relationship between the corresponding series. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.