NBER WORKING PAPER SERIES

RECOVERING INVESTOR EXPECTATIONS FROM DEMAND FOR INDEX FUNDS

Mark L. Egan
Alexander MacKay
Hanbin Yang

Working Paper 26608
http://www.nber.org/papers/w26608

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2020, Revised March 2020

We thank John Campbell, Robin Greenwood, Sam Hanson, Ian Martin, Jesse Shapiro, Andrei Shleifer, Adi Sunderam and the seminar participants at Harvard Business School and the University of Chicago Booth Household Finance Conference. We also thank John Graham and Campbell Harvey for providing the Duke CFO Global Business Outlook Survey data. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Recovering Investor Expectations from Demand for Index Funds
Mark L. Egan, Alexander MacKay, and Hanbin Yang
NBER Working Paper No. 26608
January 2020, Revised March 2020
JEL No. D12,D81,D84,G11,G50,L0

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 returns. 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. Although generated from a different method (realized choices) and a different population, our quarterly estimates of investor expectations are positively and significantly correlated with the leading surveys used to measure stock market expectations. Our estimates suggest 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. Because investors have heterogeneous beliefs, we estimate meaningful ex ante gains from leverage variety.

Mark L. Egan
Harvard Business School
Baker Library 365
Soldiers Field
Boston, MA 02163
and NBER
megan@hbs.edu

Alexander MacKay
Morgan 242
Harvard Business School
Soldiers Field
Boston, MA 02163
amackay@hbs.edu

Hanbin Yang
Harvard University
hanbin.v.yang@gmail.com

1 Introduction

We propose a revealed-preference approach to estimate investor beliefs about the future performance of the stock market. Understanding investor beliefs, how these beliefs are formed, and the dynamics of these beliefs is critical to explaining the investment and saving behavior of consumers 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); a better understanding of beliefs can inform macroeconomic policy and the regulation of financial markets. A growing number of surveys have been designed to elicit such beliefs 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; Cochrane, 2011). To complement this literature, we develop a parsimonious model of demand for exchange-traded funds (ETFs) that allows us to recover the distribution of investor expectations of stock market returns from observed investment decisions.

Our framework builds on the industrial organization literature on estimating demand with heterogeneous investors. In the context of demand for ETFs, heterogeneity in beliefs about the future performance of an underlying asset will lead to different investment decisions. By modeling choices over different funds, we are able to recover the distribution of expected returns across investors. Identification in our setting is conceptually related to Barseghyan et al. (2013), who show how beliefs can be separately and nonparametrically identified from risk aversion in the context of insurance choice. Similar to Barseghyan et al. (2013), the key feature of our data for identification is that investors choose investment options from a menu of several (more than two) ETFs with different risk/return profiles and fee structures.

This paper has three empirical contributions. First, we use our framework to construct a time series of expected stock market returns. At each point in time, we recover the distribution of expectations across investors rather than just the average expectation. We find that heterogeneity in expectations is meaningful and varies over time. Our estimates are aligned with the survey evidence commonly used in the literature (Greenwood and Shleifer, 2014; Nagel and Xu, 2019). Second, we examine how investor expectations are formed. We confirm a prior finding, based on survey evidence, that beliefs are extrapolative. Further, because we recover the entire distribution of expectations, we shed new light on how the dispersion of beliefs, or disagreement, evolves over time. Lastly, we use counterfactual simulations from the model to show the value of leverage choice to investors with different beliefs. We find that investors realize substantial ex ante benefits from leverage choice; these gains were highest during the financial crisis when disagreement was greatest.

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 exchange-traded funds (ETFs) by retail (non-institutional) investors. ETFs are passive investment funds designed to track another underlying asset. In our sample, ETFs linked to the S&P 500 average $71 billion in assets under management, 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 leverage exposures 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 mean (expected) return and taken on twice the risk. Thus, the investor's purchase indicates that the investor is either relatively more optimistic about the return of the stock market or relatively more risk tolerant compared to 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, providing information on both objects. Full nonparametric identification can be facilitated by 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 expected market risk premium for the median investor in December 2009 was 3%, roughly 10% of investors expected the stock market to fall by more than 10%. To validate our results, we compare our estimates to widely-used surveys of investor sentiment (e.g., the Shiller index). Despite the fact that these two approaches draw on different populations and are collected with

[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.

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 the investor's beliefs about the expected future return of the stock market. However, this interpretation has two important caveats. First, we do not observe an investor's portfolio; we only observe purchases 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. Therefore, we proceed with our more parsimonious model for our main results. 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.

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.

Next, we examine how the distribution of investor expectations evolves over time. Our results suggest that the mean expected return is extrapolative, based on past stock market returns. In addition, we find that the dispersion in expectations, or the extent of disagreement among investors, also reflects past returns. Following a period of negative stock market performance, investor beliefs become more pessimistic on average, more dispersed, and more negatively skewed. This suggests that a subset of investors become very pessimistic following negative returns. In contrast, disagreement across investors tends to decline following periods of high stock market returns. In other words, investors tend to agree during stock market booms and disagree during stock market busts.[3] Further, we find that expectations are persistent: one month of poor stock market performance impacts investor expectations up to two years in the future.

In a counterfactual exercise, we use our estimated beliefs to measure the value of the leveraged funds to investors. We show that limiting investors to only 1x trackers or the outside

[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 (-1, -2x, -3x).

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

option costs investors on average 3.7 percentage points in ex ante return. A restricted choice set would be most harmful to pessimistic investors. In our data, 10.5 percent of retail investors choose negative leverage on average, and 25 percent choose negative leverage during the financial crisis; these options would not have been available in the counterfactual. Thus, the availability of these products provides value to investors with divergent beliefs. Our counterfactual is further motivated by a recent ban by Vanguard on these leveraged ETFs for users on the Vanguard platform.[4]

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. Section 6 provides our analysis of the value of the choices in the market and the cost of a ban on leveraged ETFs. Section 7 concludes.

Related Literature:

Our paper builds on the demand estimation literature at the intersection of industrial organization and finance.[5] On a conceptual level, our paper relates closely to the recent work of Koijen and Yogo (2019b). Koijen and Yogo (2019b) 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.[6] We build on the idea of estimating preference heterogeneity across investors, but focus on how we can recover the expectations and risk preferences. To this end, our paper relates closely to Barseghyan et al. (2013), Calvet et al. (2019), Ross (2015), and Martin (2017). 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. 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. Ross (2015) uses state

[4]Vanguard's stated motive for the ban was to protect investors that tend to buy and hold, as the realized leverage may differ from the nominal leverage over periods longer than the stated target period (typically 1 day or 30 days). In our data, the average investor holds a leveraged ETF for less than one month, which suggests that this may be less of a concern. Despite this, we check for the impact on buy-and-hold investors in Appendix E. We find that, after the crisis, ex post leverage tracks nominal leverage well even for long investment horizons. Even during the crisis, long-term investors benefited overall, despite the divergence between realized and nominal leverage. See Ivanov and Lenkey (2014) for a further discussion of concerns about these products.

[5]Demand estimation has recently been used in a number of other financial applications such as demand for bank deposits (Dick (2008); Egan et al. (2017); Egan et al. (2017); Wang et al. (2018); and Xiao (2019)), bonds (Egan (2019)), annuities (Koijen and Yogo (2016)), credit default swaps (Du et al. (2019)), and investments more generally (Koijen and Yogo (2019a) and Koijen and Yogo (2019b)).

[6]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 similar to the approaches 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.

prices computed from options, and backs out a distribution of physical beliefs by imposing a transition-independent assumption on the stochastic discount factor. Martin (2017) derives a lower bound on the equity premium using data from index option prices.

We use the demand estimation framework to recover and better understand investor expectations of stock market returns and risk preferences. Our work complements the findings of Vissing-Jorgensen (2003), Ben-David et al. (2013), Amromin and Sharpe (2013), Greenwood and Shleifer (2014), and Nagel and Xu (2019), which use survey evidence to better understand investor expectations. Using a very different data and empirical approach, we find similar patterns of investor expectations.

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). 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 finding potentially provides additional evidence for understanding the formation of beliefs. A recent literature documents that such extrapolative beliefs 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.[7] ProShares and Direxion are the two largest providers of leveraged ETFs, with almost $8 billion AUM at the end of 2018.

Despite easy access to leverage, these products are not considered suitable for buy-and-hold investors and typically attract short-term purchases.[8] Leveraged ETFs are intended to hit their nominal leverage after a short period—typically 1 day or 30 days—and the realized leverage

[7] 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.

[8] In a letter to the SEC, Direxion estimated that its shareholders hold triple-leveraged funds for between one and four days. (https://www.wsj.com/articles/sec-moves-to-curb-leveraged-etfs-1465205401)

may diverge from this target over a longer holding period. We explore this additional risk for longer-term investors in Appendix E.

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. Our panel ranges from 2008 to 2018.[9]

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 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 over the course of each month. This flow measure reflects the investor expectations at the time of transaction, as opposed to stock measures that reflect past purchase decisions. 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

[9] Although the first leveraged ETF was launched in 2006, we drop earlier periods due to data limitations.

a fraction of ETF investors become "active" and rebalance their portfolios. When rebalancing, active investors set aside a fixed portion of their wealth to invest in the ETF market.

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 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, $Retail_j$.[10]

As in most demand estimation exercises, we do not directly observe investors that consider investing in S&P 500 ETFs but who choose a risk-free option instead (0x leverage). To construct shares for this outside option, we calculate flows into retail money market accounts for retail investors considering S&P 500 ETFs, as 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 relevant for S&P 500 investors. To convert this stock measure into a flow measure of investor purchases, we scale this proxy by the ratio of quantity (constructed 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% under AUM but only 57% according to our market share definition.

[10] As a robustness check, we also calculate quantities as $Quantity_{jt} = Retail_j \times Trading\,Volume_{jt}$. This adjustment has no appreciable differences on our market share measure and consequently our empirical results.

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 as well as expense ratio. As discussed above, leveraged ETFs are smaller in AUM compared to trackers. Leveraged ETFs also charge substantially higher fees, and ETFs withmore leverage tend to be marginally more expensive. Figure 2b shows the trends in ETF fees. ETF fees are relatively stable over time, though the average fee for 1x trackers has been declining since 2013.

3 Empirical Framework

3.1 Demand for ETFs

We model an investor's investment 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). The parameter λ is the investor's coefficient of risk aversion, and can be interpreted to represent either constant absolute risk aversion or constant relative risk aversion. 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. We later extend the model to allow for heterogeneous risk aversion: $\lambda_i \sim G(\cdot)$. Though λ_i is defined 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. 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 model and the corresponding estimates, see Appendix B. 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.[11] 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 $k \neq j$. 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

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

[11]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.

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 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.[12] Therefore, an investor chooses leverage category j if and only if

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

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 $F(\cdot)$.

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. We discuss the merits of the assumptions with respect to our empirical implementation in Section 4. In estimation, we allow F to vary over time. Here, we provide conditions to identify the distribution that applies in each relevant period.

Identification is obtained by using two sources of variation. The first source is variation comes from the menu of choices facing investors. 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 chooses a 3x return. By observing the market shares of purchases in each leverage category, we can pin down features of the distribution of expected returns.

[12] 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.

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 a leverage $k \leq j$: $\sum_{k=-3}^{j} s_k$. We can add up the shares from equation (5) to obtain a system of equations satisfying

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

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 in each period. These six equations allow us to identify, in principle, a period-specific risk aversion parameter λ, as well as a period-specific distribution for F. The distribution of F can be estimated as a flexible distribution of up to five parameters, provided it is sufficiently well-behaved. 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 prices and volatility. Intuitively, if we observe how prices shift leverage choices, we can pin down the scale of risk aversion. We assume that prices p_j and the available leverage choices β_j are independent of investor expectations $F(\cdot)$. In the data, both prices and leverage choices are relatively fixed in the short run; this helps alleviate concerns that ETF issuers are endogenously changing fees and leverage choices, quarter-to-quarter, in response to changes in investor expectations.

Formally, suppose that there exists a realization of the data for which $\tilde{\sigma}^2 = \sigma^2$ and $\tilde{S}_k = S_j$ for $k \neq j$. Then it must be that $\frac{\lambda}{2}\left(2j+1\right)\sigma^2 + p_{j+1} - p_j = \frac{\lambda}{2}\left(2k+1\right)\sigma^2 + \tilde{p}_{k+1} - \tilde{p}_k$. Therefore, we have $\frac{\lambda}{2}\sigma^2\left(2j-2k\right) = \left(\tilde{p}_{k+1} - \tilde{p}_k\right) - \left(p_{j+1} - p_j\right)$, 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 $F^{-1}(S_j)$. Furthermore, we only have to identify λ once, so this single comparison provides identification at all quantiles $\{S_j\} \cup \{\tilde{S}_j\}$.

Likewise, exogenous variation 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 $\frac{\lambda}{2}\left(2j+1\right)\sigma^2 + p_{j+1} - p_j = \frac{\lambda}{2}\left(2j+1\right)\tilde{\sigma}^2 + \tilde{p}_{j+1} - \tilde{p}_j$. 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 $\frac{\partial S_j}{\partial p_j} = -f(S_j)$ and $\frac{\partial S_j}{\partial \sigma^2} = \frac{\lambda}{2}\left(2j+1\right)f(S_j)$, we can recover λ as

$\lambda = \frac{-2}{2j+1}\frac{\frac{\partial s_j}{\partial \sigma^2}}{\frac{\partial s_j}{\partial p_j}}$.

Our main empirical results use both sources of variation. We estimate the belief distribution at the quarterly level, allowing monthly variation in prices and volatility to assist in identification. To demonstrate the identifying power of the menu of choices only, we provide an alternative set of estimates in Appendix A. Using this alternative approach, we can allow the belief distribution to vary at the monthly level. These alternative estimates closely resemble our main results.

3.3 Extension: 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,[13] 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).[14] 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 realizations can be used to pin down properties of the distribution of risk aversion coefficients.

[13]See Chapter 8 of Greene and Hensher (2010) for a discussion of the literature.

[14]See Cunha et al. (2007) for further discussion of ordered choice models.

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.

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 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 scaled utility 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

$$\prod_{t \in T_s} \prod_{i \in N_t} \prod_{j \in J} \left(F_s(x_{jt}) - F_s(x_{(j-1)t})\right)^{\mathbf{1}[a_i = j]} \tag{6}$$

and the log-likelihood is

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

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

$$\sum_{t \in T_s} \sum_{j \in J} s_{jt} \ln\left(F_s(x_{jt}) - F_s(x_{(j-1)t})\right). \tag{8}$$

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, scale, skewness, and kurtosis; these are further described in Table 2.[15] 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, scale, 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, scale, 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

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

In our baseline specification, we hold λ constant over time. This has the interpretation that λ is a deep preference parameter, rather than a proxy for consumption and other factors.[16]

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 some alternative specifications, we allow λ_i to be heterogeneous across investors, and we hold the distribution of λ_i fixed over our sample.

[15] In estimation, we use the skewt package in R for calculating the skewed t distribution $\tilde{F}$ for $a = 0$ and $b = 1$. Thus, our routine parameterizes a and b as $\tilde{F}\left(\frac{x_{jt}-a}{b}\right)$.

[16] We think treating λ as a deep preference parameter is a reasonable assumption for ETF demand. In the data, leverage in ETFs increases during the crisis; models with time-varying risk aversion such as Campbell and Cochrane (1999) would suggest increased risk aversion and reduced leverage during this period. As a robustness check, we estimate an extension of the model where we allow risk aversion to vary annually, which we present in Appendix C. This extension generates a similar distribution of investor expectations to our baseline model. In contrast to Campbell and Cochrane (1999), our estimates suggest that risk aversion is slightly procyclical for ETF demand.

4.2 Baseline Results

Our estimates for investor expectations are plotted in Figure 3. 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, scale, 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 3a 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% 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%, which is similar to, albeit slightly smaller, than the median realized market risk premium in our sample (4.65%) and 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 scale 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 a 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 3a, 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 $c_t < 1$. 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.

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 1.262 implies fat tails that are roughly in line with the Cauchy distribution.[17] Our estimated risk aversion parameter of $\lambda = 0.982$ implies that investors are willing to pay an additional 40 basis points in fees for a one standard deviation reduction in volatility.[18] 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, λ and λ_i may not be directly interpretable as risk aversion 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 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 $\lambda_i \sim U[0.656, 1.000]$. Thus, the midpoint of 0.828 is slightly lower than the constant risk aversion parameter estimate of 0.982. The other parameters are only slightly affected by the change. Figure 4 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,

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

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

which is plotted in yellow. The vertical blue lines correspond to the cutpoints 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 cutpoints 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 $\mu_i = 2.8$ with risk aversion heterogeneity.

4.4 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 potentially 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 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.

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 5 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 5a 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 5b 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 5c 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 comparable.[19] 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.[20] The index is designed to capture a broad measure of U.S. investors' outlook on their finances and the economy based on their survey 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

$$M = \frac{\sum_{j=\{1,2,3\}} \hat{s}_j}{\sum_{j=\{1,2,3\}} \hat{s}_j + \sum_{k=\{-3,-2,-1\}} \hat{s}_j} \tag{9}$$

[19] 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.

[20] 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.

where $\hat{s}_j$ is the predicted share from our model.[21] 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 6 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 (9).

Figure 6 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 (bearish).[22] Because the percent bullish and percent bearish are highly correlated in the

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

[22] 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]

survey, we construct a single measure, $\frac{bullish}{bullish+bearish}$, which corresponds closely to the relative share positive versus negative from equation (9). Comparing each response separately to analogous measures from our estimates yields similar results.

Panel (c) in Figure 6 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.[23] 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 6 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). Although survey respondents are not explicitly asked about stock market, respondents do provide their annual forecast of real GDP growth. Panel (e) in Figure 6 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 5 and 6 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 estimates of investor beliefs help complement the survey data. While the survey data is

[23] 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]

representative of the population of interest, our belief measure comes from the actual investment decisions of investors.

4.5 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 scale 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.

We consider two alternative definitions of the outside option to test the sensitivity of our results to this measure. 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 three 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 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.

In the third set of results, we estimate the model while allowing investors to have time-varying risk aversion λ_t (Appendix C). On average, we estimate similar but slightly lower risk aversion compared to the baseline model (0.61 vs 0.98). More importantly, for the purposes of our analysis, the two models produce qualitatively similar distributions of investor expectations. The correlation of the mean expected return across investors in the time-varying risk aversion model and our baseline model is 0.96.

4.6 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 D, 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.[24]

Figure 7 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 (7e). Similarly, the negative effects of the European sovereign debt crisis on investor beliefs are immediately apparent in Figure 7c, as the average investor expected a decline in equity prices and there was an increase in disagreement across investors.

We estimate different risk aversion for each asset class, because the sample of investors trading each asset class may be different. For example, we estimate that investors in gold are slightly less risk averse than those in S&P 500 ($\lambda = 0.783$ vs. $\lambda = 0.982$). We estimate that investors in oil are much less risk averse, with a risk aversion parameter of 0.278. 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 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.

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 negative (-10%) return shock impacts investors' beliefs for up to two years into the future. Lastly, we compare our estimates of investor expectations with future returns and model-based expected 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

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

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.

Panel (a) of Figure 8 displays a binned scatter plot of our estimated mean expected excess return versus the previous year-over–year excess return of the stock market. Investor expectations are positively and significantly correlated with historical stock market returns (corr=0.67). We examine the relationship more systematically 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. Observations are at the quarterly level.

We report the corresponding estimates in column (1) of Table 4. 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% 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 co-vary with recent stock market returns. Panel (b) of Figure 8 displays a binned scatter plot of the standard deviation of expected returns across investors versus the previous year-over–year excess return of the stock market. The two series are negatively and significantly correlated (-0.58), indicating that investor beliefs become disperse following a downturn in stock market returns. Column (2) of Table 4 displays the corresponding regression estimates. 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.

Panel (c) of Figure 8 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. Column (3) of Table 4 displays the corresponding regression estimates. The results indicate that a one percentage-point increase in recent historical returns is correlated with a 0.27 percent increase in the skewness parameter. One potential explanation for the results is 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.

5.2 Persistence of Beliefs

Figure 3 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 a large increase in disagreement. As illustrated in the figure, these effects persist for up to two years.

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

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

Observations in eq. (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 the excess return of the stock market during the same quarter, $QuarterlyRet_q$. We report the corresponding estimates in Table 5. 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 9 displays the impulse response of how the expectation distribution evolves in response to positive/negative returns shocks in the stock market. Panel (a) displays how investor expectations respond to a 10% decrease in stock market returns occurring at time 1. As illustrated in the figure, the mean expectation across investors immediately falls and remains negative and below the steady state level for almost two years. The negative stock market return also 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. The effect is driven by changes to the scale and skew of the distribution. In response to the negative return shock, the 10th and 25th percentile of investors become dramatically more pessimistic. The expected return among investors in the 10th percentile falls by almost 15%.

Panel (b) of Figure 9 shows how the expectation distribution evolves in response to a 10% increase in stock market returns occurring at time 1. 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 disperse 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 such that 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 for the proceeding two years.

5.3 Future Returns and Model Returns

Finally, we explore whether investor expectations of returns can forecast future returns.

Figure 10a displays a binned scatter plot of our estimates of the mean expected excess returns versus future 12-month excess returns. Rather than predicting excess returns, the estimated mean expected returns instead have a weakly negative correlation with future returns. Figure 10b displays the relationship between future returns and the share of investors purchasing positive leverage minus the share of investors purchasing negative leverage. We again find little evidence suggesting that investor expectations predict future returns.

Our evidence is consistent with the findings in Greenwood and Shleifer (2014) that investor expectations do not forecast future returns. 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, and second, we use the consumption wealth ratio (cay) of Lettau and Ludvigson (2001) as a proxy for expected returns. Figure 11 displays binned scatter plots of the dividend-price ratio and cay versus 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. The correlation between our measure of expected returns and the dividend-price ratio is -0.83. This evidence is consistent with the findings from Greenwood and Shleifer (2014) that investor expected returns are negatively correlated with model-based measures of expected returns.

Our result is also related to the lower bound on equity premium that Martin (2017) constructs using prices on S&P 500 index options. Martin (2017) shows that this lower bound predicts future returns and is negatively correlated with survey measures. Figure 12a displays a binned scatter plot of the lower bound versus our estimate of the mean expected return. The correlation is also negative and significant, further confirming that our estimates are aligned with survey measures and do not predict future returns.

In addition, Martin and Papadimitriou (2019) highlight the distinction between the average belief and the belief of a representative investor who chooses to hold the market. This insight suggests another connection with Martin (2017). The lower bound on equity premium binds for an unconstrained rational investor with log utility who is 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 12b 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 on the belief of the representative investor.

5.4 Investor Expectations Across Asset Classes

An advantage in our setting is that we recover the distribution of investor beliefs across asset classes which allows us to 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. The panel setting allows us to exploit variation in asset returns over time, and a greater laboratory to study how investors respond to market crashes and booms.

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

$$\overline{E[R]}_{iq} = \beta AnnualRet_{iq} + \mu_i + \mu_q + \epsilon_{iq}. \tag{12}$$

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 6 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

$$FutureRet_{iq} = \overline{E[R]}_{iq} + \mu_i + \mu_t + \epsilon_{iq} \tag{13}$$

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 7. 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.

6 Value of Product Variety in ETF Choice

The wide dispersion of expectations about future stock market returns suggests that there are large ex ante welfare gains from product variety in the context of S&P 500 ETFs. Providing investors with a menu of leverage choices allow them to invest based on their idiosyncratic beliefs. For example, investors with a negative expected return of the S&P 500 would not choose to invest in an S&P 500 tracker, but they might invest in an inverse ETF. These investors comprise, on average, 10.5 percent of the market in our sample (Table 1). Thus, the availability of inverse ETFs provides a way for investors to express their view on the market when they have divergent beliefs.[25]

In this section, we quantify the welfare gains of product variety by comparing investor utility in our data to a counterfactual in which leveraged ETFs are eliminated. In our counterfactual, investors can only choose tracker ETFs (leverage = 1x) or the outside option. We consider ex ante expected utility in each scenario, i.e., the utility realized by investors if ex post returns matched each investor's ex ante expectation.

Our counterfactual is further motivated by Vanguard's recent ban on leveraged ETFs for users on their investment platform. In January 2019, Vanguard banned leveraged ETFs on their platform, limiting the ability of investors to pick an investment product matched to their individual expectations. Our estimated gains from variety correspond to the (ex ante) losses realized by investors on Vanguard's platform. Vanguard's stated motive for the ban was to reduce additional risk to investors who hold onto ETFs for a long period. For investors that hold the leveraged ETF for a longer period, the ex post leverage may differ from the nominal value. We find that the average holding period in our sample is less than one month, which suggests that most investors do not hold leveraged ETFs for a long period. Despite this, we analyze the potential impact of the ban, as well as the incremental profit Vanguard might realize from the ban, in Appendix E. We find that even investors who hold on to these products for two years benefit from additional product variety, despite the increased risk.

To measure the gains from the availability of leveraged ETFs, we calculate the welfare gains from these products relative to trackers and the outside option. Because our model generates a strict ordering of preference for leverage, eliminating leveraged ETFs will shift all investors in inverse ETFs to the outside option, and all investors in positively leveraged ETF to trackers. Using the recovered distribution of expected return μ_i and risk aversion, it is straightforward to compute the difference in utility measured in risk adjusted return from investors' original

[25]Inverse ETFs provide investors with a straightforward and simple way to short the market, relative to the other investment options available. For example, investors are not required to have a margin account to invest in inverse ETFs. For the purposes of the counterfactual, we assume that the set of investors that have access to more sophisticated instruments are not investing in S&P 500 ETFs.

choices to either the outside option or trackers.

As before, investor i's indirect (ex ante) utility from choosing leverage j is given by

$$u_{ij} = \beta_j \mu_i - p_{jt} - \frac{\lambda}{2} {\beta_j}^2 \sigma^2 \tag{14}$$

Denote the realized utility with the menu of choices in the data as $u_i^{(1)} = \max_j u_{ij},\ j \in \{-3, -2, -1, 0, 1, 2, 3\}$. Denote the counterfactual utility as $u_i^{(0)} = \max_j u_{ij},\ j \in \{0, 1\}$ with the restricted choice set. We calculate the gains from variety as $E\left[u_i^{(1)} - u_i^{(0)}\right]$ by assigning all investors that choose $\beta_j < 0$ to $\beta_j = 0$ and all investors with $\beta_j > 1$ to $\beta = 1$ and re-computing their utility.[26]

We calculate the gains separately for each period.[27] Figure 13 displays the quarterly average gain. On average, investors realize gains of 3.72 percentage points in ex ante excess return from the presence of leveraged ETFs. The gains are higher during the crisis period, averaging 7.43 percentage points in excess return from 2008 to 2011. Higher gains are the result of greater disagreement about the future performance of the stock market, which can be observed in the higher dispersion of expectations in Figure 3 before 2012. From 2012 on, the dispersion in expected return is much lower; the average gains from variety from 2012 through 2018 is 1.24 percentage points, which is lower but still economically meaningful.

One caveat to this exercise is that we take investor expectations as given when calculating the ex ante utility. If investors make systematic mistakes when forming expectations, then one might want to replace investor expectations with an alternative distribution, such as one based on rational expectations. A paternalistic utility function along these lines would imply a different value for product variety.

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

[26]We follow the standard for welfare calculations of assigning a utility of zero to the outside option. This assumption rules out substitution to assets that provide similar exposure to S&P 500 ETFs but that are not in our sample. These alternative assets are not in our model and are ruled out by construction.

[27]To calculate gains and losses, we make additional restrictions on the tails of the expected return distribution. Because we do not identify the tails in estimation, some investors in 3x and -3x leverage have extreme and unrealistic expected returns. Hence, we censor the distribution of μ_i at the lowest and highest level we can identify. Specifically, we censor $\mu_i \in [\underline{\mu}, \overline{\mu}]$, where $\underline{\mu}$ is the maximum μ_i that chooses the inverse ETF with highest leverage (e.g., -3x) and $\overline{\mu}$ is the minimum μ_i that chooses the ETF with highest positive leverage (e.g., 3x).

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 expected 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

We also use our framework to understand the welfare benefits of product variety in the ETF setting. Given that there is substantial heterogeneity in the distribution of investor beliefs, we find substantial welfare benefits to increasing the product variety (leverage choice) available to investors, even in light of the rebalancing concerns of leveraged ETFs.

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.

References

Amromin, G. and S. A. Sharpe (2013). From the horse's mouth: Economic conditions and investor expectations of risk and return. *Management Science 60*(4), 845–866.

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.

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

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

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*.

Campbell, J. Y. and J. H. Cochrane (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of political Economy 107*(2), 205–251.

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.

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

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.

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

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

Egan, M. (2019). Brokers vs. retail investors: Conflicting interests and dominated products. *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.

Egan, M., G. Matvos, and A. Seru (2019). The market for financial adviser misconduct. *Journal of Political Economy 127*(1), 000–000.

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.

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.

Graham, J. R. and C. R. Harvey (2008). The equity risk premium in 2008: evidence from the global cfo outlook survey. *Available at 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.

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

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.

Ivanov, I. and S. Lenkey (2014). Are concerns about leveraged etfs overblown?

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

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. *Available at SSRN 3383677*.

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

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*.

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.

Nagel, S. and Z. Xu (2019). Asset pricing with fading memory. Technical report, 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.

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

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.

Xiao, K. (2019). Monetary transmission through shadow banks. *Available at SSRN 3348424*.

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	Scale	Multiplicative; 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	
	Coef	SE	Coef	SE
Expected Return				
Location (Median)	2.856	(0.695)	2.484	(0.657)
Scale (Median)	1.081	(0.424)	0.992	(0.448)
Skewness (Median)	0.766	(0.343)	0.752	(0.259)
Kurtosis	1.262	(0.510)	1.340	(0.465)
Risk Aversion				
Mean	0.982	(0.020)	0.828	(0.069)
Dispersion			0.172	(0.077)
Implied Mean Expectation				
10 Pct	-2.545		-1.225	
25 Pct	0.106		0.391	
50 Pct	0.768		0.714	
75 Pct	1.140		1.131	
90 Pct	1.402		1.487	
Model Fit				
R^2	0.921		0.923	
Log Likelihood	-168.593		-168.468	

Notes: Table 3 shows estimation results with constant and heterogeneous risk aversion. The first panel displays parameters for the expected return distributions. Location, scale and skewness parameters are allowed to vary over time, and we estimate one set of coefficients for each quarter. We display the median location, scale, 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: Expected Returns versus Past 12-month Returns

	Mean	ln(SD)	ln(Skew)
	(1)	(2)	(3)
Annual Return	0.11***	-0.026***	0.0027***
	(0.030)	(0.0076)	(0.00092)
Observations	44	44	44
R-squared	0.582	0.310	0.189

Notes: Table 4 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 standard deviation parameter in logs, and in column (3) is the skew 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. Newey-West based standard errors are in parenthesis with four lags. *** p<0.01, ** p<0.05, * p<0.10.

Table 5: Evolution of the Parameters of the Expectation Distribution: Vector Autoregressions

	Location	Scale	Skewness
Lag Parameters	0.42***	0.52**	0.42***
	(0.14)	(0.085)	(0.15)
Quarterly Return	-0.27**	-0.24*	0.0023
	(0.12)	(0.12)	(0.0014)
Const	2.95***	1.69**	0.44***
	(0.84)	(0.65)	(0.11)
Observations	43	43	43
R-squared	0.638	0.675	0.247

Notes: Table 5 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 Quarterly Return is the excess return of S&P 500 during the same quarter measured in percentage points. Robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.10.

Table 6: Expected Returns versus Past 12-month Returns Across All Asset Classes

	Mean	ln(SD)	ln(Skew)
	(1)	(2)	(3)
Annual Return	0.054*	-0.012**	0.0014**
	(0.031)	(0.0052)	(0.00065)
Time Fixed Effects	X	X	X
Asset Fixed Effects	X	X	X
Observations	330	330	330
R-squared	0.394	0.667	0.654

Notes: Table 6 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. 12). Observations are at the asset class by quarter level over the period 2008-2018. See Appendix D 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 standard deviation parameter in logs, and in column (3) is the skew 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. Driscoll-Kraay based standard errors are in parenthesis with four lags and are grouped by asset class. *** p<0.01, ** p<0.05, * p<0.10.

Table 7: Future Annual Returns vs. Expected Returns Across all Asset Classes

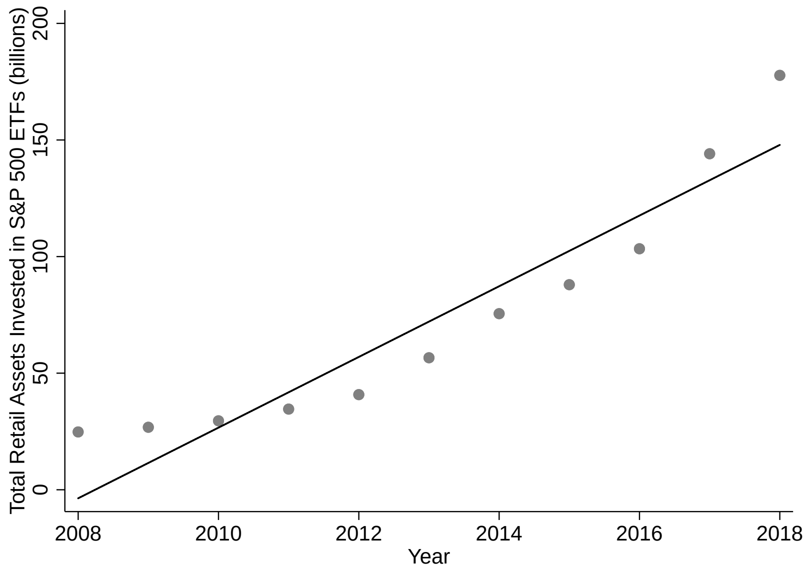
	(1)	(2)	(3)
$AverageExpectedReturn$	-0.42***	-0.29	0.10
	(0.14)	(0.24)	(0.22)
Asset Fixed Effects		X	X
Time Fixed Effects			X
Observations	330	330	330
R-squared	0.013	0.046	0.404

Notes: Table 7 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. 13). Observations are at the asset class by quarter level over the period 2008-2018. See Appendix D 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. Driscoll-Kraay based standard errors are in parenthesis with four lags and are grouped by asset class. *** p<0.01, ** p<0.05, * p<0.10.

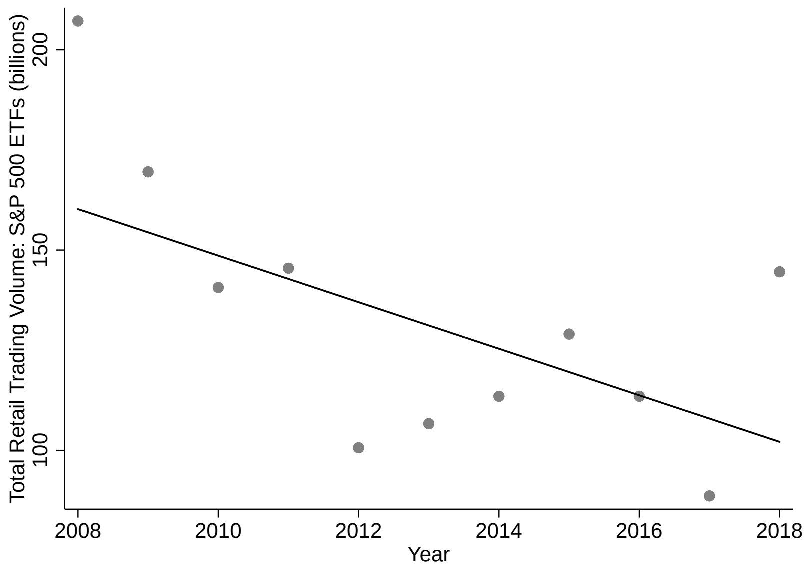
Figures

Figure 1: S&P 500 ETFs

(a) Assets Under Management (Retail Investors)



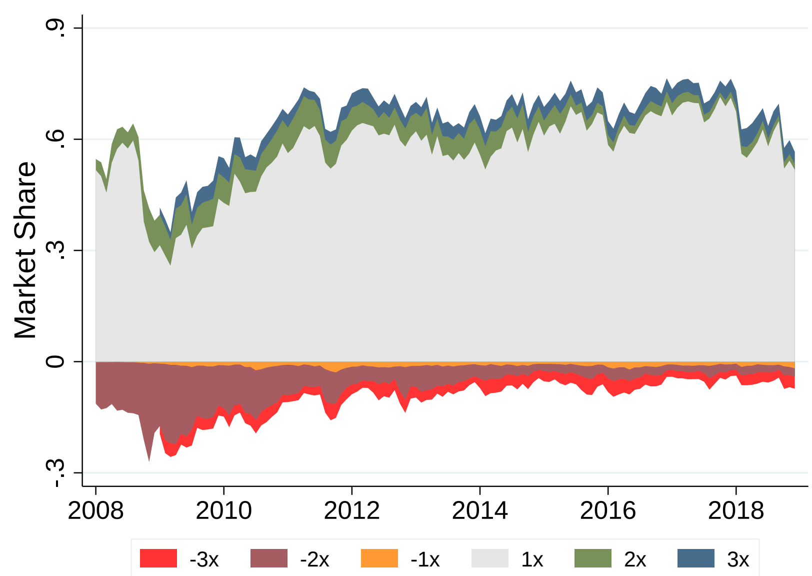
(b) Trading Volume (Retail Investors)



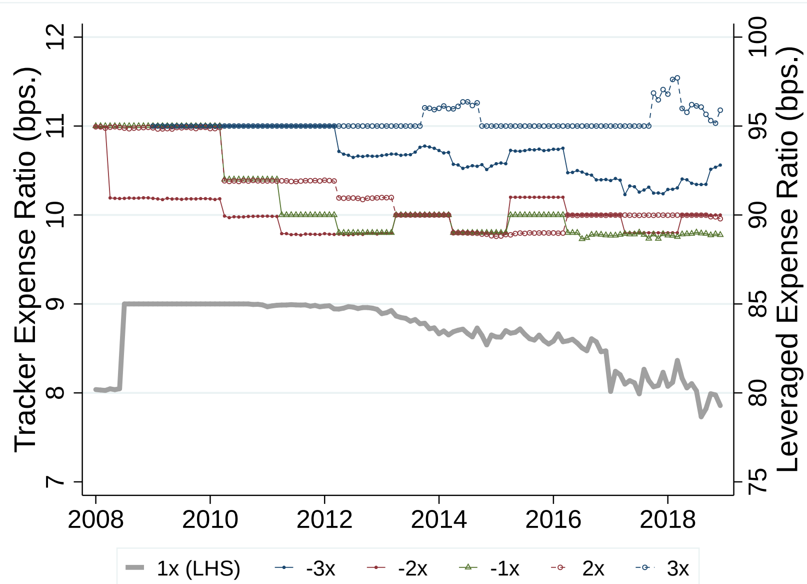
Notes: Figure 1 shows binned scatters at annual frequency along with the linear fitted lines for total retail AUM in panel (a) and total retail trading volume in panel (b) of ETFs that track S&P 500. Retail AUM is computed as $Retail_j \times AUM_{jt}$ and trading volume is computed as $Retail_j \times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

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

(a) Market Share



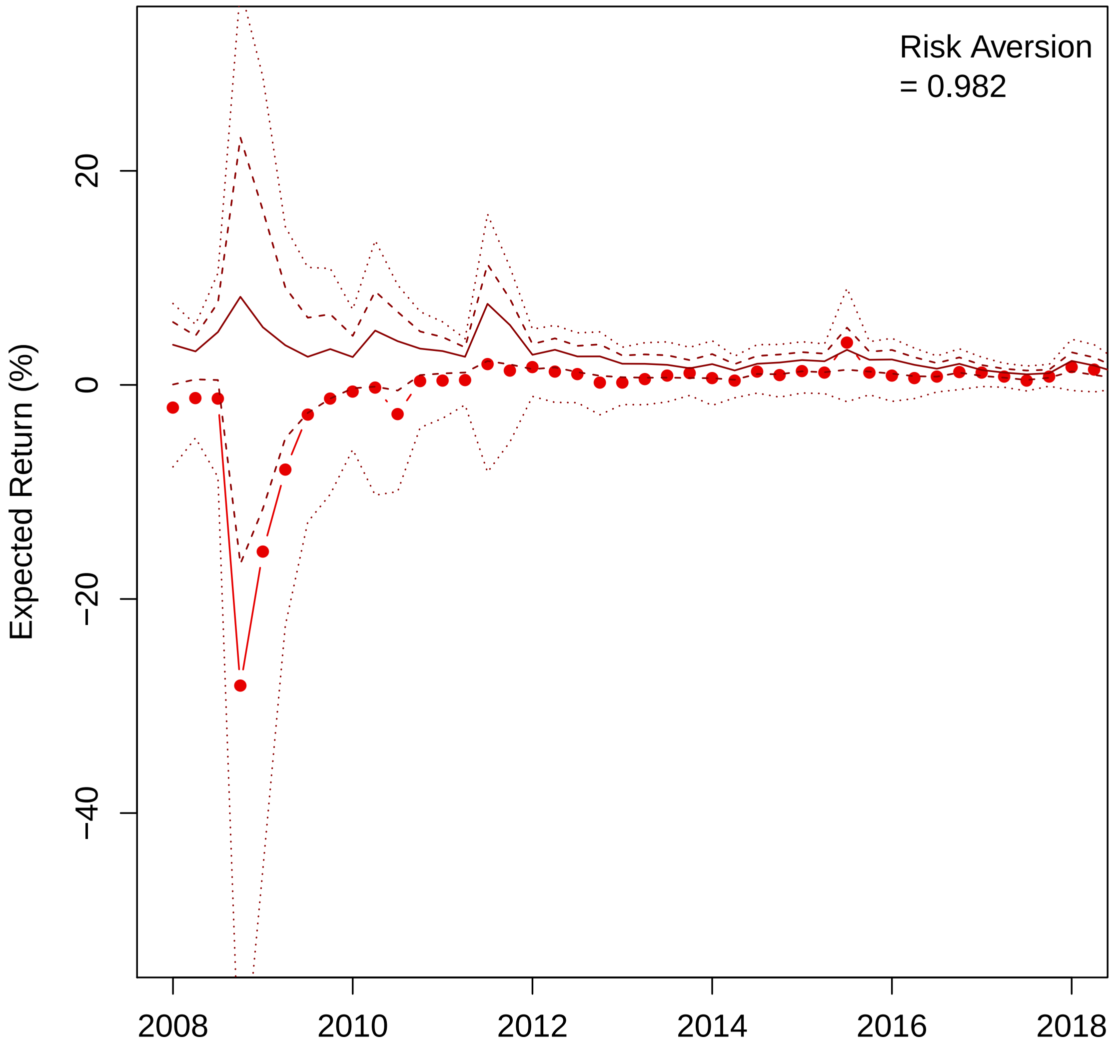
(b) Expense Ratio



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: Time-Varying Investor Expectations

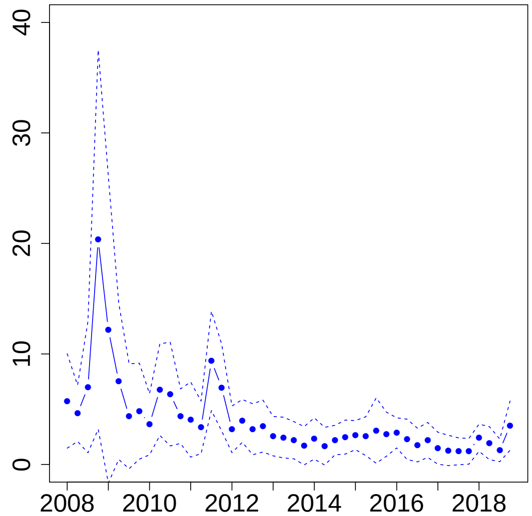
(a) Estimated Distribution



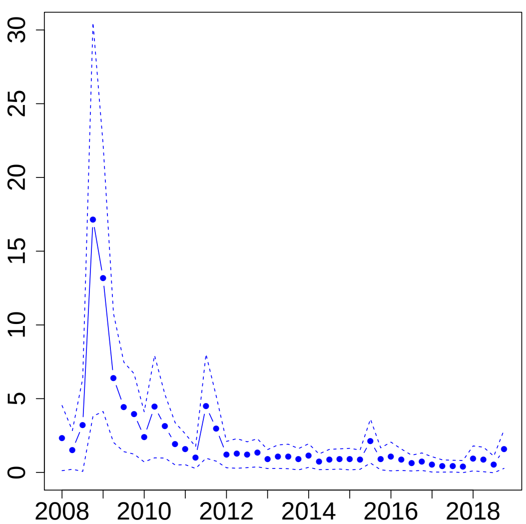
Notes: Figure 3 panel (a) plots the estimated distribution of investor expectations over time. Red dots represent mean expected return. Solid dark red lines indicate median, dashed lines indicate 25th and 75th percentiles, and dotted lines indicate 10th and 90th percentiles.

Figure 3: Time-Varying Investor Expectations (Cont.)

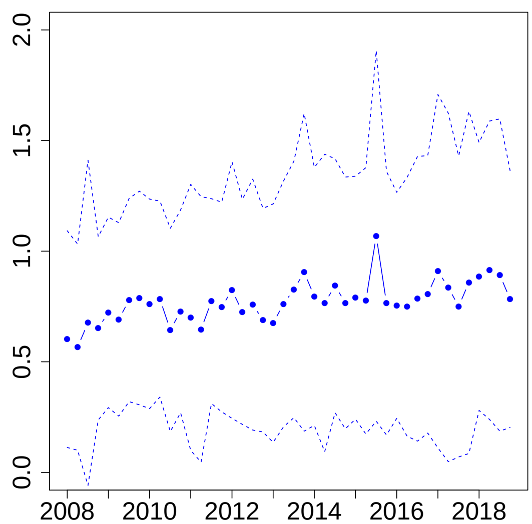
(b) Location Parameter (a_s)



(c) Scale Parameter (b_s)

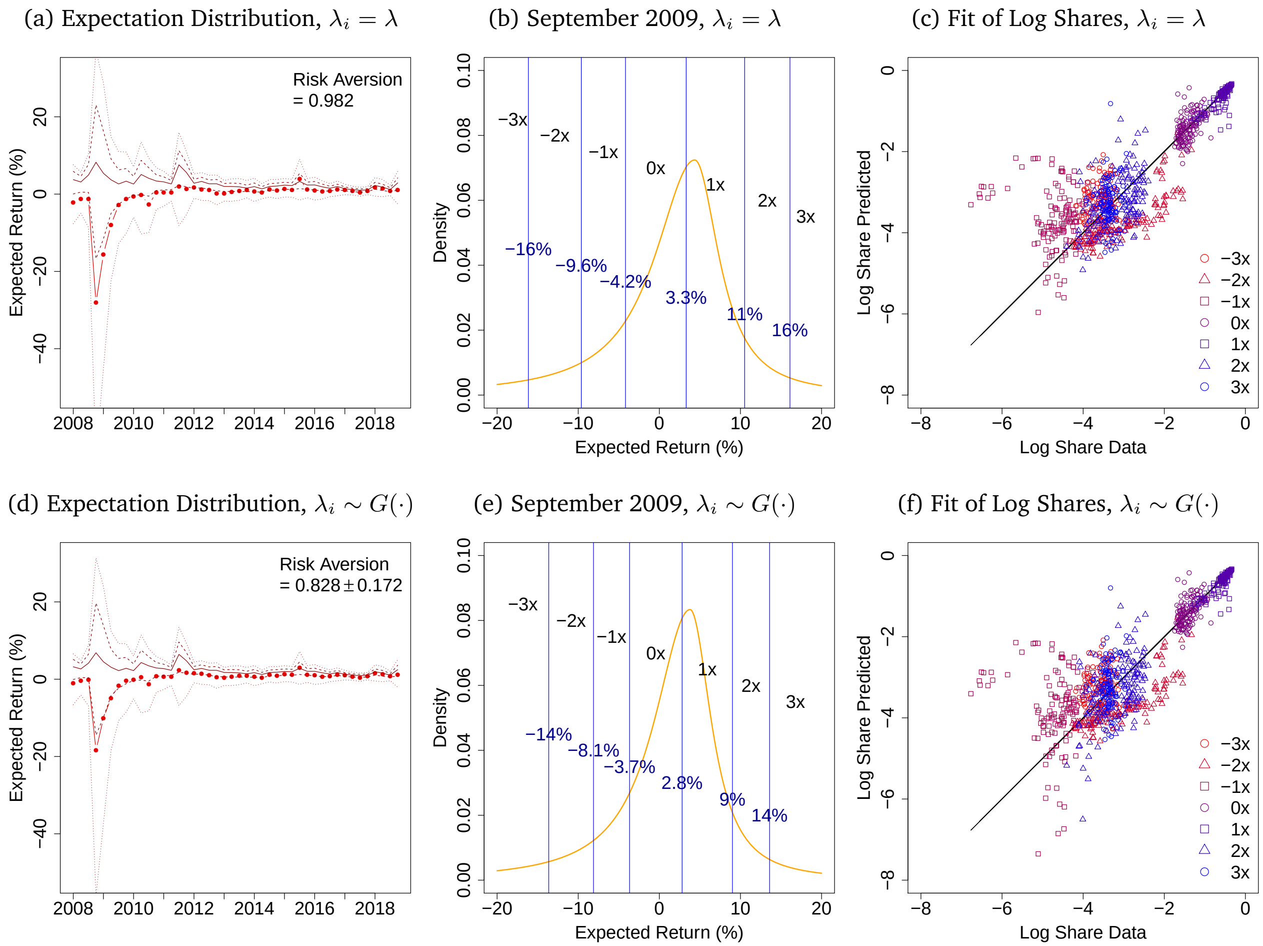


(d) Skewness Parameter (c_s)



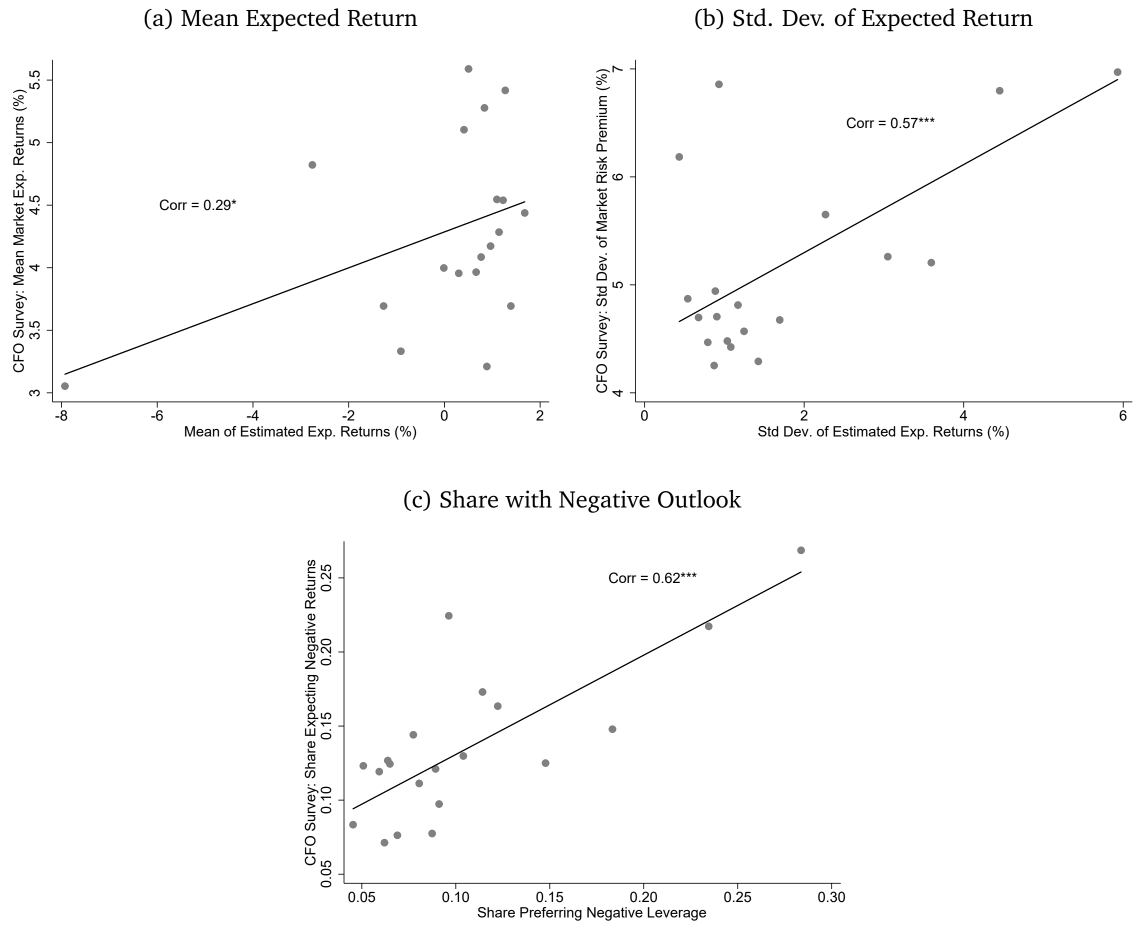
Notes: Figure 3 panels (b) to (d) show estimated time-varying location, scale, and skewness parameters for expectation distribution in blue dotted lines, and the 90 percent confidence intervals in blue dashed lines.

Figure 4: Expectations and Model Fit: Baseline and Heterogenous Risk Aversion (S&P 500)



Notes: Figure 4 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.

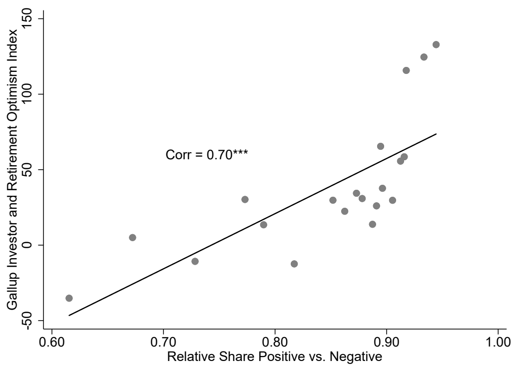
Figure 5: Comparison with Duke CFO Global Business Outlook Survey



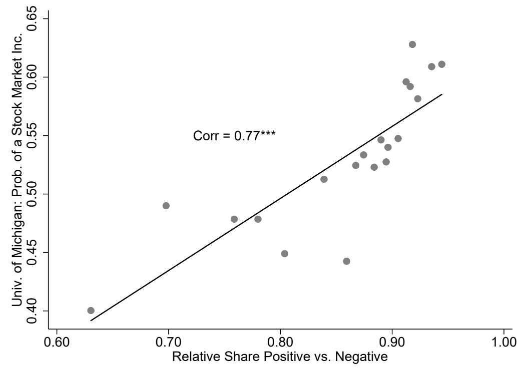
Notes: Figure 5 panels (a)-(c) display binned scatter plots of our estimated beliefs distribution versus results from the Duke CFO Global Business Outlook Survey. The estimated beliefs distribution corresponds to our baseline model estimates reported in column (1) of Table 3. Observations in each panel 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 6: Comparison with Surveys

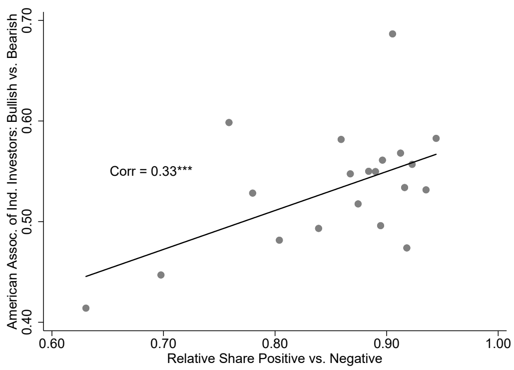
(a) Gallup



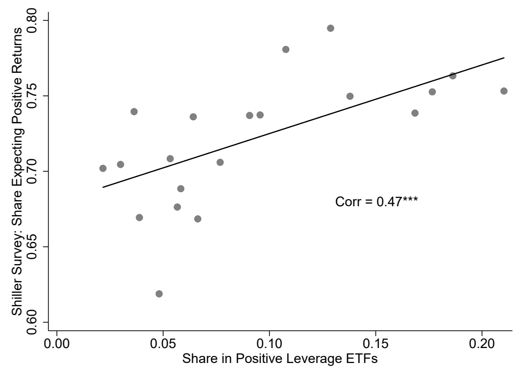
(b) University of Michigan



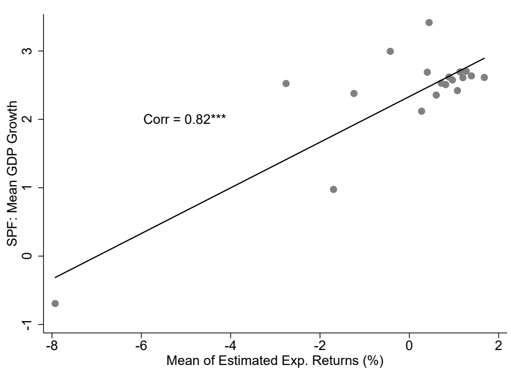
(c) AAII



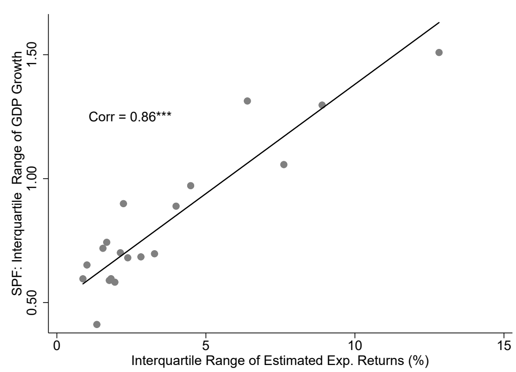
(d) Shiller Index



(e) SPF: Average GDP Forecasts



(f) SPF: Interquartile Range of GDP Forecasts



Notes: Figure 6 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 beliefs 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 4.4. 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. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Figure 7: Expectations: Other Asset Classes

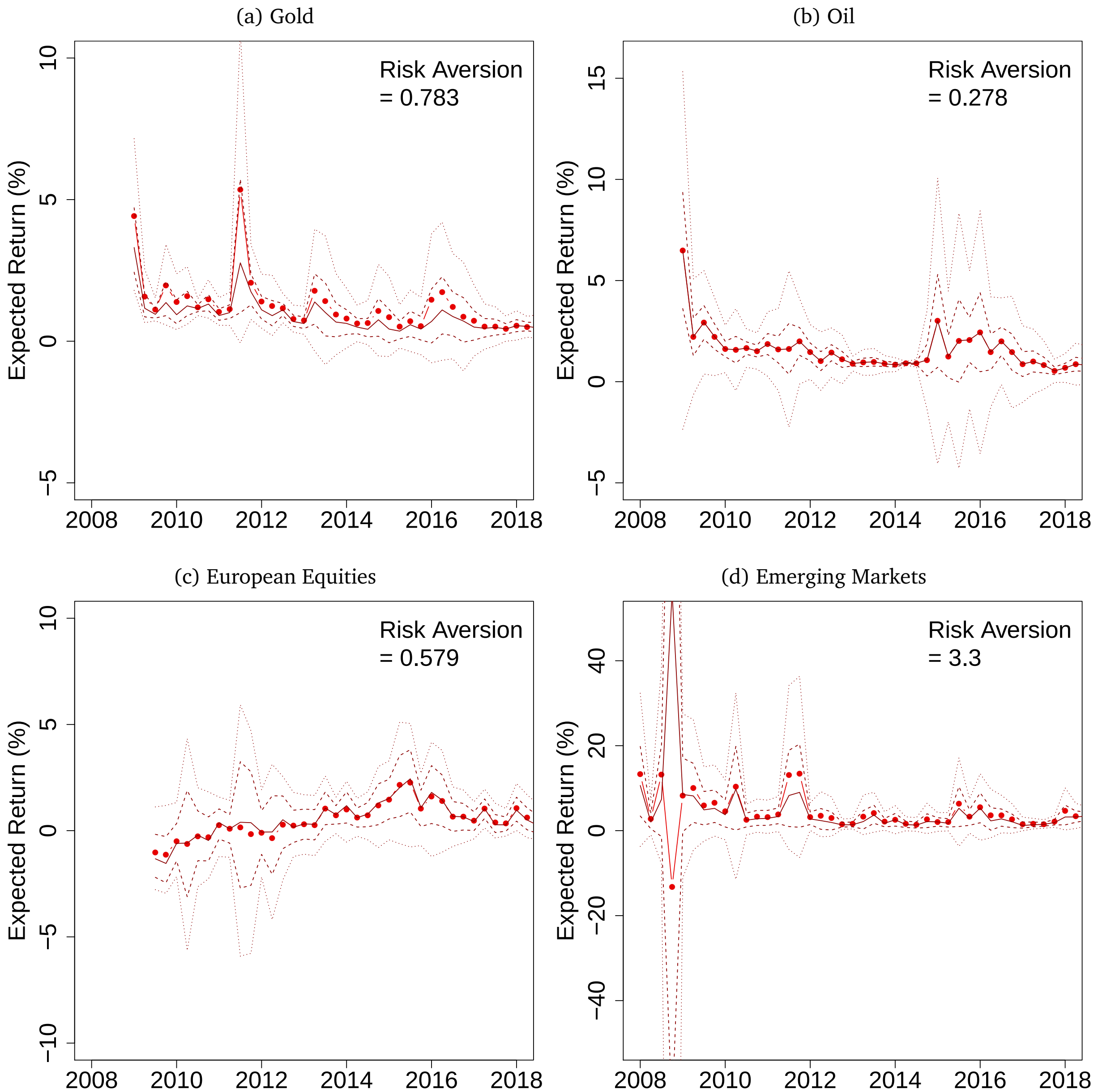
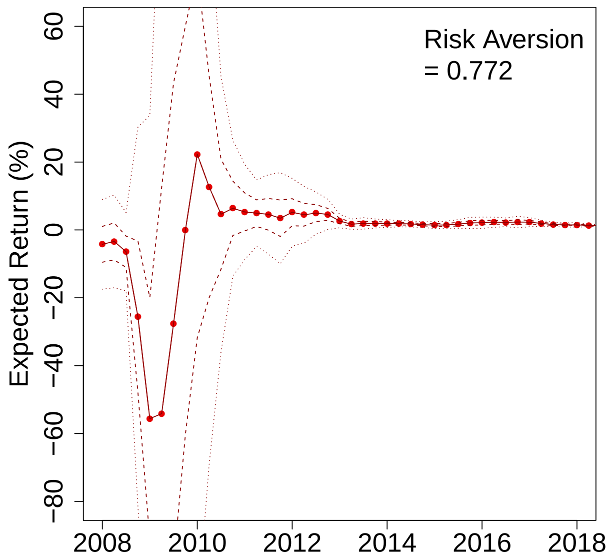
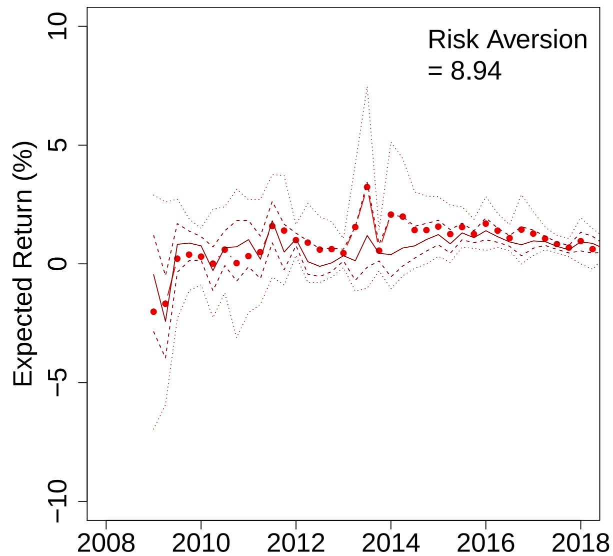


Figure 7: Expectations and Model Fit: Other Asset Classes (Cont.)

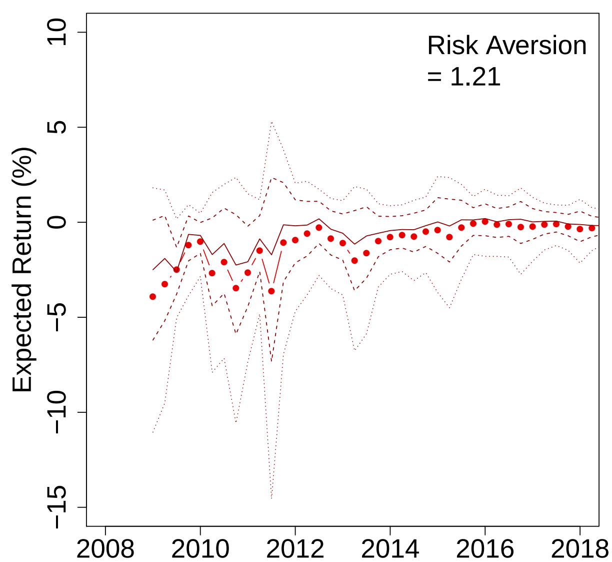
(e) US Real Estate



(f) Mid-Term Treasury



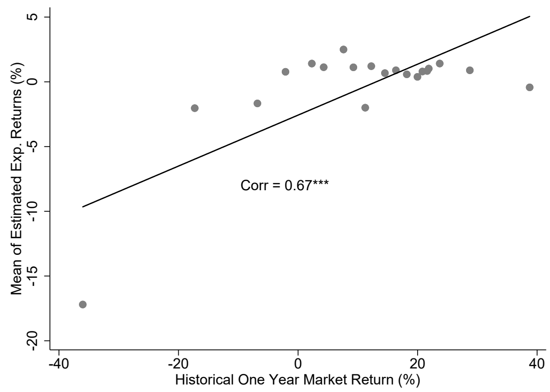
(g) Long-Term Treasury



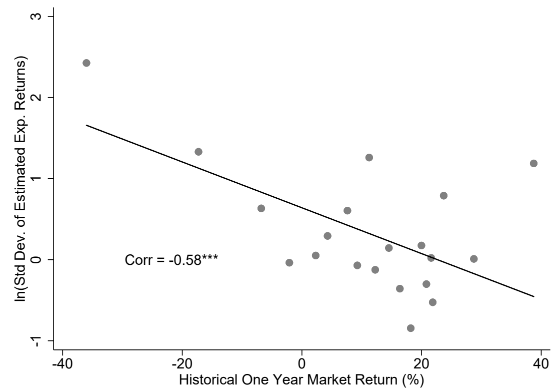
Notes: Figure 7 panels (a)-(g) displays the estimated expectation distribution corresponding to gold, oil, European equities, emerging market equities, US real estate, medium-term (7-10 year) Treasury, and long-term (20+ year) Treasury. Red dots represent mean expected return, solid dark red lines indicate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

Figure 8: Extrapolated Beliefs

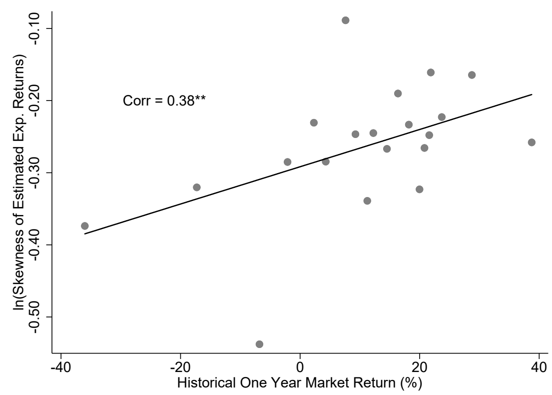
(a) Mean Expected Return vs. Prev 12m Return



(b) SD of Expected Return vs. Prev 12m Return



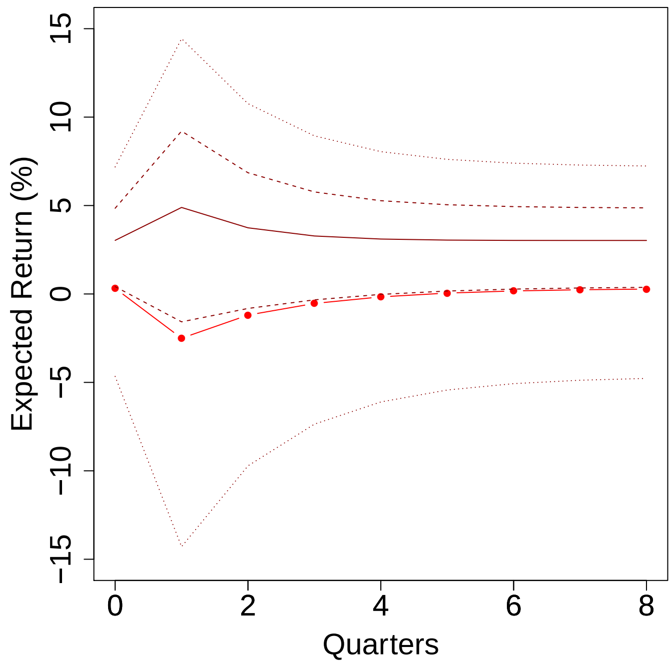
(c) Skewness of Expected Return vs. Prev 12m Ret



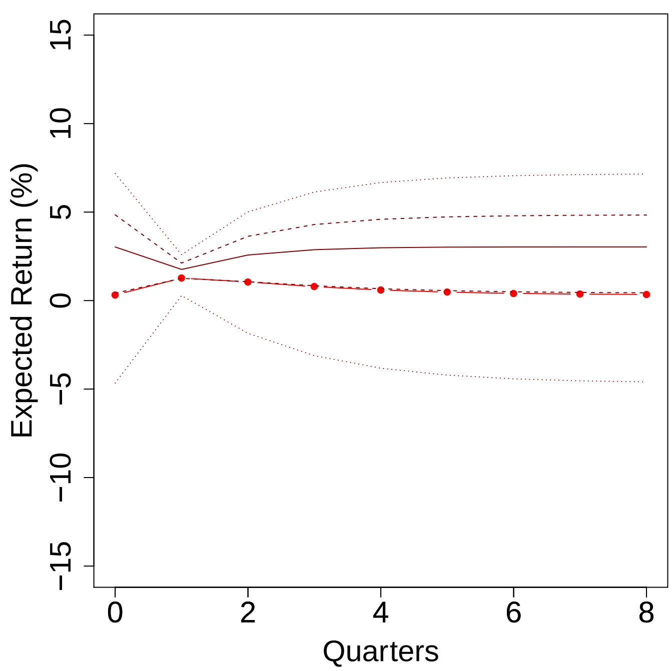
Notes: Figure 8 panels (a)-(c) display the relationship between the past twelve-month excess return of US stock market versus our estimated distribution of investor expected returns. The estimated expected return correspond to our baseline model estimates reported in column (1) of Table 3. Observations in each panel are at the quarterly level over the period 2008-2018. Figure 8a displays a binned scatter plot of the mean of the estimated distribution of expected returns versus the past twelve-month excess return of US stock market. Figure 8b displays a binned scatter plot of the standard deviation of the estimated distribution of expected returns versus the past twelve-month excess return of US stock market. Figure 8c displays a binned scatter plot of the skew of the estimated distribution of expected returns versus the past twelve-month excess return of US stock market. *** p<0.01, ** p<0.05, * p<0.10.

Figure 9: Impulse Response

(a) Impulse Response Following a 10% Decrease in Returns



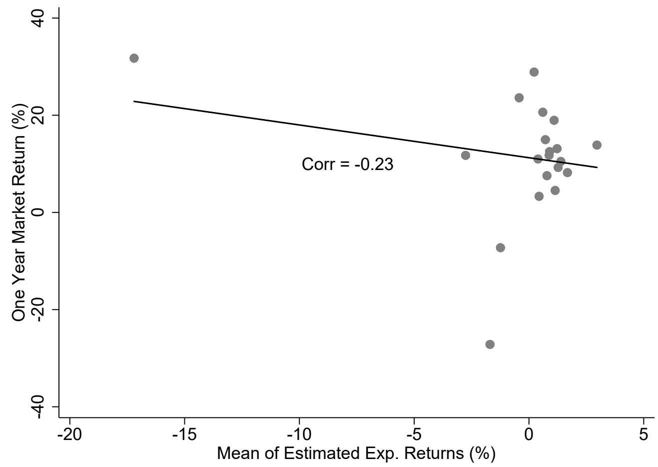
(b) Impulse Response Following a 10% Increase in Returns



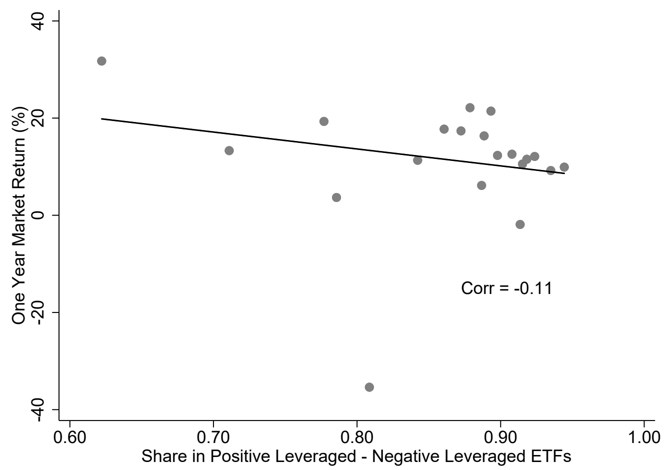
Notes: Figure 9 displays the impulse responses of a -10% S&P 500 return at $t = 1$ in the top panel and a 10% return at $t = 1$ in the bottom panels. In both panels, we assume that S&P 500 returns are 2.6% for $t = 0$ and $t > 1$. We predict each parameter separately using their lagged value in the previous quarter and the excess return of S&P 500 during the same quarter as reported in Table 5. The initial values are kept at steady state mean of each parameter. Red dots correspond to analytical mean. Solid dark red line shows median, and dashed dark red lines show 10, 25, 75, 90th percentiles.

Figure 10: Forecasting Returns

(a) Estimated Mean Expected Return vs. Fwd 12m Ret



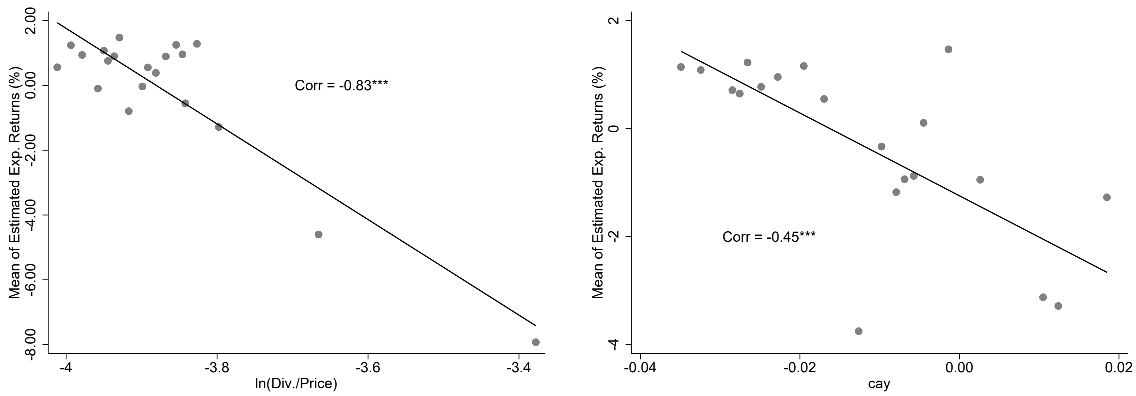
(b) Relative Share Positive vs. Fwd 12m Ret



Notes: Figure 10 displays the relationship between the estimated expected returns from our model and the future 12-month excess return of the U.S stock market. Observations in each panel are at the quarterly level over the period 2008-2018. The estimated expected returns correspond to our baseline model estimates reported in column (1) of Table 3. Panel (a) displays a binned scatter plot of the future 12-month excess return of the U.S stock market versus the mean estimated expected return from our model. We winsorize the mean of expected returns from our model at the 5% level to account for outliers during the financial crisis. Panel (b) displays a binned scatter plot of the future 12-month excess return of the U.S stock market versus the relative share of investors preferring positive to negative leverage, based on our estimated distribution of expectations. *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Figure 11: Comparision with Model Returns

(a) Mean Expected Return vs. ln(Div/Price) (b) Mean Expected Return vs. cay

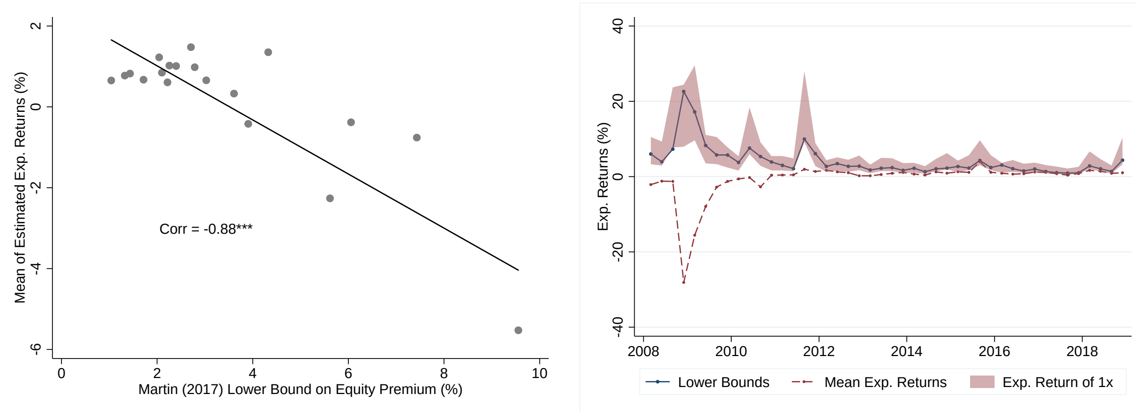


Notes: Figure 11 displays the relationship between the estimated expected returns from our model and model-based expected returns. The estimated expected returns correspond to our baseline model estimates reported in column (1) of Table 3. Observations in each panel are at the quarterly level over the period 2008-2018. Panel (a) displays a binned scatter plot of the mean estimated expected return from our model versus the log divided-price ratio. Panel (b) displays a binned scatter plot of the mean estimated expected return from our model versus cay from Lettau and Ludgivson (2001). In both panels, 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.

Figure 12: Comparision with Lower Bound on Equity Premium

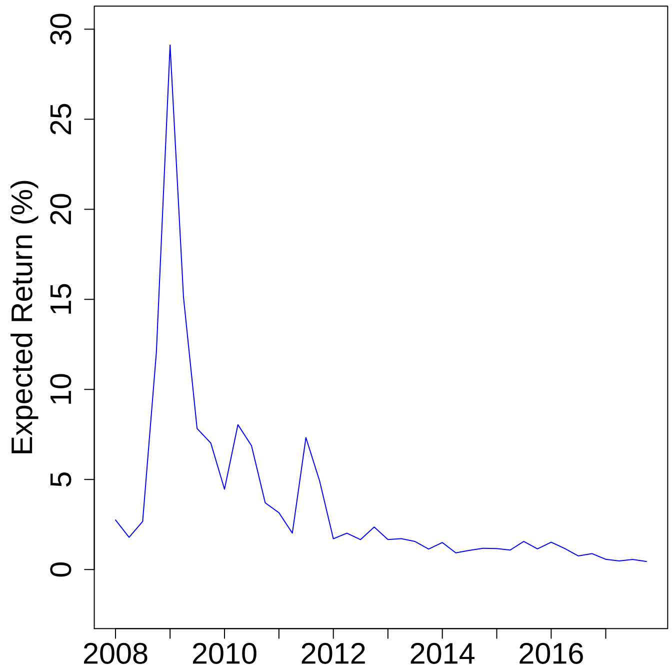
(a) Mean Expected Return vs. Lower Bound

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



Notes: Figure 12 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 13: Gains from Variety



Notes: Figure 13 displays quarterly average gains from variety, measured as the utility difference (in terms of expected return) between the full choice set in the data and a restricted choice set of only 1x trackers or the outside option.

Appendices

A Alternative Estimates

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 scale 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 scale 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 A2. 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 A2 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 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 sgn(β_j) $\neq$ sgn(β_{ω_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 directly identify 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 A4 displays estimates for investor expectations with hedging, which are close to our baseline results in Figure 3. The estimated mean expected returns in these two models are highly correlated with a correlation coefficient of 0.98. Our estimates suggest that portfolio demand is not meaningful in the context of S&P 500 ETFs. While we only recover $\beta_\omega/\delta = -0.813$ on average and 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 $\beta_\omega = -0.0813$. This suggests investors behave as if their wealth is nearly uncorrelated to 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 colinearity 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.

C Time-Varying Risk Aversion

In this appendix, we estimate an extension of our baseline specification where 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. Specifically, we estimate three parameters corresponding to the location, scale, and skewness of expectation distribution each quarter, risk aversion parameters each calendar year, and the kurtosis parameter kept constant over time. With 11 years of data, we estimate 144 parameters.

Figure A5a 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). We have also estimated a specification where we allow risk aversion to vary at the quarterly level, and we found similar time-varying patterns of both risk aversion and expectations.

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 A5b, we estimate lower risk aversion during those periods, and hence we need less dispersion in beliefs to rationalize leverage choices in the data. However, we find such cyclical risk aversion unintuitive. In typical models with time-varying risk aversion, such as Campbell and Cochrane (1999), investors are more risk averse during crises. Therefore, although this additional flexibility in time-varying risk aversion can also rationalize data, we prefer holding risk aversion constant in our baseline model.

D 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 7 and A6 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.[28]

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.

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

E Vanguard's Ban on Leveraged ETFs

In January 2019, Vanguard banned leveraged ETFs on their platform,[29] eliminating the product variety we analyze above. Vanguard's stated motive for the ban was investor protection. As we describe below, investors who hold on to the leveraged ETFs for a sufficiently long period may realize an ex post leverage that differs from the nominal leverage associated with the ETF. The difference between ex post and nominal leverages depends on stock market performance and volatility. In essence, Vanguard's stated motive is to product investors against additional risk.[30]

Accounting for Leverage Risk

Provided investors hold on to ETFs for a sufficiently short period, our gains from variety calculated above correspond to the losses for users of Vanguard's platform. Perhaps short-term holding is a reasonable assumption: ProShares and Direxion, two issuers of leveraged funds, say trading data show that most investors treat leveraged ETFs as short-term investments. In a letter to the SEC, Direxion estimated that its shareholders hold triple-leveraged funds for between one and four days.[31] In our data, the average holding period for ETFs is less than one month. Despite this, quantifying the welfare impacts for longer holds provides a valuable benchmark. To consider the impact of additional risk from the divergence between the nominal and ex post leverage, we first calculate the ex post leverage realized for hypothetical investors that hold on to each product for a period of 12 or 24 months.

To capture the ex post realization, we estimate ex post leverage by comparing the realized performance of the leverage category to the performance of S&P 500 for an investor who buys at month t and holds for 12 or 24 months. We construct a time-varying measure by running an OLS regression of the leverage category returns from t to $t + 12$ or $t + 24$ on S&P 500 returns over the same holding period, in a moving window of 7 months centered around t.

The ex post leverage for a 12-month holding period are displayed in Figure A7. As can be seen in the figure, the median ex post leverage is fairly close to nominal for a 12-month holding period across leverage categories. However, there are periods where the ex post leverage departs meaningfully from the nominal leverage. In July 2008, ETFs with negative nominal

[29]"Vanguard to stop accepting purchases in leveraged and inverse investments," Vanguard.com. https://investornews.vanguard/vanguard-to-stop-accepting-purchases-in-leveraged-and-inverse-investments. Accessed November 12, 2019.

[30]As Vanguard describes on their website, "On any given day, if you use a leveraged or inverse product, you can expect a return similar to the stated objective. However...extended holdings beyond one day or one month, depending on the investment objective, can lead to results different from a simple doubling, tripling, or inverse of the benchmark's average return over the same period. This difference in results can be magnified in volatile markets. As a result, these types of investments aren't generally designed for a buy-and-hold strategy... These funds are riskier than alternatives that don't use leverage." "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.

[31]https://www.wsj.com/articles/sec-moves-to-curb-leveraged-etfs-1465205401

leverage—i.e., a positive return during a downturn—generated a *positive* leverage for those that bought and held for a year. In 2011, increased volatility resulted in a negative shift for all ex post leverage. Around January 2015, the inverse products realized a positive shift, with the 3x leverage realizing a negative return. Likewise, the ex post leverage for a 24-month holding period can be higher or lower than the nominal leverage. Though our sample does not show a systematic bias one way or the other, the realized leverages can deviate by large factors for long holding periods, illustrating the additional risk of these products for buy-and-hold investors.

To calculate the investor protection benefit of a ban of these products, we simulate a counterfactual in which buy-and-hold investors have perfect foresight over the realized ex post leverage. We denote the counterfactual utility for a buy-and-hold investor as

$$\tilde{u}_{ij}(h) = \tilde{\beta}_j(h)\mu_i - p_{jt} - \frac{\lambda}{2}\tilde{\beta}_j(h)^2\sigma^2$$

where h is the holding period and $\tilde{\beta}_j(h)$ is the leverage for category j as a function of the holding period. First, we hold investors' choices fixed and adjust the utility based the ex post realization of leverage: $\tilde{u}_i^{(1)} = \tilde{u}_{ij}$, $j = \arg\max_j u_{ij}$. We then allow these investors to re-optimize and choose their preferred leverage based on the ex post leverage of the product, $\tilde{u}_i^{(2)} = \max_j \tilde{u}_{ij}$, For both $\tilde{u}_i^{(2)}$ and $\tilde{u}_i^{(1)}$, we hold fix investors' stock market expectation and allow them to choose from $j \in \{-3, -2, -1, 0, 1, 2, 3\}$.

Using these calculations, we compute two measures of regret from leverage risk. Our first measure is the fraction of investors with leverage regret, i.e., those investors that would change their product choice with foresight of the ex post leverage: $E_i\left[1\left\{\arg\max_j u_{ij} \neq \arg\max_j \tilde{u}_{ij}\right\}\right]$. Our second measure is expected loss, which we compute as the average difference between the utility from re-optimized choices based on ex post leverage and from the original choices: $E_i\left[\tilde{u}_i^{(2)} - \tilde{u}_i^{(1)}\right]$.[32]

Finally, we consider the investor protection gains from a ban on leveraged ETFs for buy-and-hold investors. We construct a third measure of utility based on ex post leverage when investors can only choose trackers or the outside option, $\tilde{u}_i^{(0)}$, but, as before, they make their choice based on the nominal leverage ($j = \arg\max_j u_{ij}, j \in \{0, 1\}$). Thus, $E_i\left[\tilde{u}_i^{(1)} - \tilde{u}_i^{(0)}\right]$ provides us with a measure of investor gains from product variety, taking into account the ex post leverage regret. Or, equivalently, the investor loss from the ban. Note that the welfare effects are composed of the gain from variety in product choice and the loss of protection from leverage risk. For the investor protection to be a net benefit, the losses from leverage risk must outweigh the gains from product variety.[33]

[32] In addition, when we allow investors to re-optimize after learning the ex post leverage, we restrict them to their original long or short directions. For example, suppose an investor buys a -2x ETF but learns that the ex post leverage of a 2x ETF is in fact -1.8x, which happens to be her most ideal leverage. In this case, we do not allow this investor to shift from -2x to 2x.

[33] Investors may realize significant ex post losses by making the "wrong" bet on the market. By looking at ex post leverage only, and not ex post return, we do not protect investors from ex post losses based on realized returns.

Figure A8 shows the respective welfare calculations for investors with a 12-month holding period. Panel (a) shows the fraction of investors with leverage regret. The fraction is highest during the crisis and declines during our sample. Panel (b) shows the expected loss in terms of excess return. The quarterly average expected loss is on average around 0.8 percentage points during the crisis and peaks at nearly 3 percentage points in the second half of 2008. Much of this is driven by the fact that the ex post leverage for inverse ETFs diverged significantly from the nominal beta (Figure A7) in this period. The expected loss drops to 0.1 percentage points on average after the crisis. Panel (c) shows the gains from variety, taking into account ex post leverage. Investors gain from the availability of leveraged ETFs in our sample. Investor gain is highest during the crisis, when investors have greater dispersion in expectation. This is despite the fact that the fraction with regret is also highest. Overall, the ability to trade on expected return μ_i more than offsets the loss from misunderstanding the product, though not all investors gain. Thus, we calculate that a ban would result in a net loss to investors.[34]

Our welfare results are summarized in Table A1. For comparison, we include gains for 12-month and 24-month holding periods. For 24-month holding periods, we find similar results for a 12-month holding period. Overall, investors gain from increased product variety. These gains are the largest during the crisis, despite a larger fraction of investors with leverage regret.

Our findings suggest that the benefits of protecting some investors do not offset the large losses from reduced product variety. Therefore, protecting the average investor does not seem to justify a ban on leveraged ETFs. Even if investors were naively buying and holding these products, the gains from product variety appear to dominate the leverage risk. Corroborating this finding, there is almost no record of investorcomplaints about these products. The Financial Industry Regulatory Authority (FINRA) requires that all investor complaints are reported through its BrokerCheck website. Using the BrokerCheck data, we parse through the universe of investor complaints reported on BrokerCheck (300k total complaints), and find fewer than one hundred related to leveraged ETFs.[35] One notable exception is a class-action lawsuit initiated by investors against ProShares and Direxion in 2009, which claimed that the companies misled investors by not adequately explain the risks of holding the products over time. A judge dismissed the suit against ProShares in 2012, while Direxion agreed in 2013 to pay $8 million to settle the lawsuit while denying wrongdoing.[36]

In addition to estimating the net loss to investors due to this ban, we also conduct a back-of-

[34]Note that the precise magnitude of the gains or losses from the ban, as well as other counterfactual outcomes, is sensitive to our assumption that investors in leveraged ETFs do not understand the rebalancing mechanism. If investors are sophisticated and realize the additional risk from rebalancing yet still choose high (positive or negative) nominal leverage, they must have more extreme μ_i than our estimates and would suffer even more from the ban. On the other hand, these sophisticated investors would not regret their purchase if nominal and ex post leverages are different, so we over-estimate the regret and expected loss. Therefore, although we are unable to measure what fraction of investors understand these products in our setting, our simplifying assumption that no investors understand leads to a lower bound of the overall net loss from the ban.

[35]See Egan et al. (2019) for further discussion of the data.

[36]https://www.wsj.com/articles/sec-moves-to-curb-leveraged-etfs-1465205401

the-envelope calculation of the potential profit impact to Vanguard. Given the vertical demand structure, all affected investors who would have chosen 2x and 3x ETFs will shift into trackers. Because we do not have a precise statistics of Vanguard's platform market share, we consider two different benchmarks for the fraction of affected investors. First, we assume the fraction of shifted investors that choose Vanguard's tracker is equal to the market share of its tracker (VOO) in terms of retail AUM at the end of 2018, which is around 24 percent. For our second measure, we use a rough upper bound estimate of 50 percent. This latter figure is motivated by the fact that, in addition to offering ETFs, Vanguard also provides brokerage accounts for retail investors. Compared to its main competitors: Fidelity, Charles Schwab, TD Ameritrade, and T. Rowe Price, Vanguard is the only platform that provides its own S&P 500 ETF, and it is generally recognized as a superior platform on which to trade ETFs. Our 50 percent benchmark captures the fact that investors may disproportionately trade in leveraged ETFs on Vanguard's platform.

We use the average retail AUM in 2x and 3x leverage categories over our sample as an estimate for the total amount of assets that would be affected by the ban. Vanguard would attract between 550 and 1,169 million dollars in assets, which would generate 0.16 to 0.35 million dollars of revenue based on its current expense ratio of 3 basis points. Using the same units as expected loss for investors, we compute Vanguard's expected gain across investors is 0.06 to 0.12 basis points. Comparing with the net loss due to the ban on leveraged ETF, the expected gain is an order of magnitude smaller. Our estimates suggest that Vanguard did not have a substantial profit motive for the ban.

Table A1: Gain from Variety

	Crisis 2008-2011	Post-Crisis 2012-2018	Full-Sample 2008-2018
Nominal Leverage			
Gain from Variety (pp.)	7.429	1.243	3.717
Ex Post Leverage: 12-Month Holding Period			
Fraction Regret	0.152	0.071	0.103
Expected Loss (pp.)	0.806	0.106	0.386
Gain from Variety Net of Leverage Risk (pp.)	7.252	1.139	3.584
Ex Post Leverage: 24-Month Holding Period			
Fraction Regret	0.161	0.094	0.124
Expected Loss (pp.)	1.329	0.181	0.691
Gain from Variety Net of Leverage Risk (pp.)	6.532	1.252	3.598

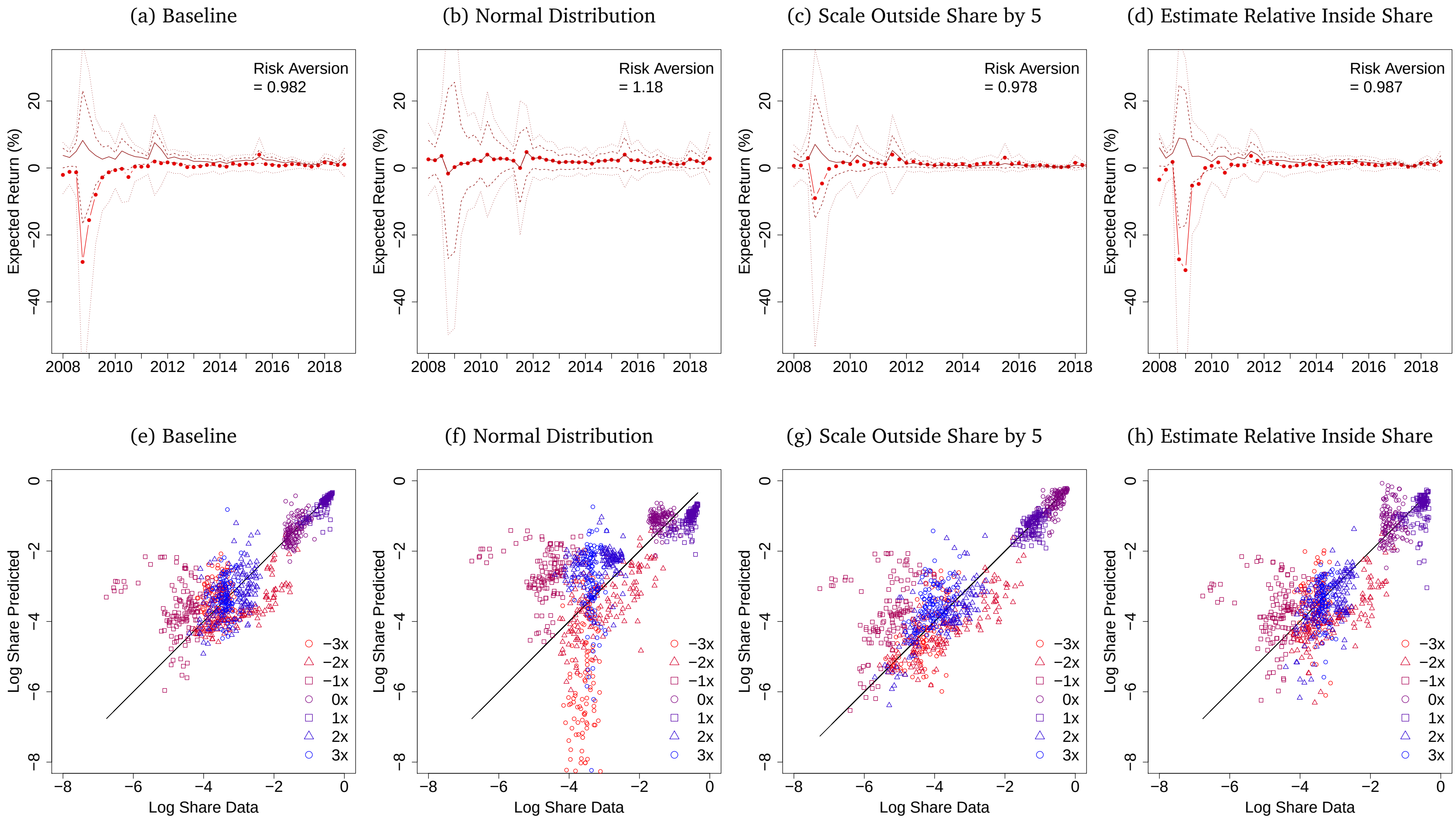
Notes: Table A1 summarizes counterfactual results. We display the average outcomes in the crisis period (2008-2011), post-crisis (2012-2018) and full sample (2008-2018). The first row corresponds to gains from variety, measured as the utility difference between the choice sets in the data based on nominal leverage and the restricted choice set of either outside option or trackers. The next three rows correspond to the welfare effect of leverage risk over time. Fraction regret measures the fraction of investors who would regret their leverage choices after learning ex post leverages. Expected loss is the difference between the utility from re-optimized choices based on ex post leverage and the utility from original choices. Gain from variety net of leverage risk measures the utility difference between the full choice set in the data and a restricted choice set of only 1x trackers or the outside option. Ex post leverage is computed assuming a 12-month holding period. The final three rows are the same as above, except that ex post leverage is computed assuming a 24-month holding period instead.

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	0.361***	-0.274**	1					
	(0.000)	(0.003)						
Real Estate	0.686***	-0.00523	-0.0690	1				
	(0.000)	(0.956)	(0.432)					
Treasury 7-10	0.751***	0.573***	-0.274**	0.617***	1			
	(0.000)	(0.000)	(0.002)	(0.000)				
Treasury 20+	0.595***	0.535***	-0.538***	0.467***	0.521***	1		
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)			
Gold	-0.471***	-0.306***	0.701***	-0.340***	-0.251**	-0.679***	1	
	(0.000)	(0.001)	(0.000)	(0.000)	(0.006)	(0.000)		
Oil	-0.761***	-0.128	0.421***	-0.653***	-0.517***	-0.478***	0.573***	1
	(0.000)	(0.173)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	
Observations	132							

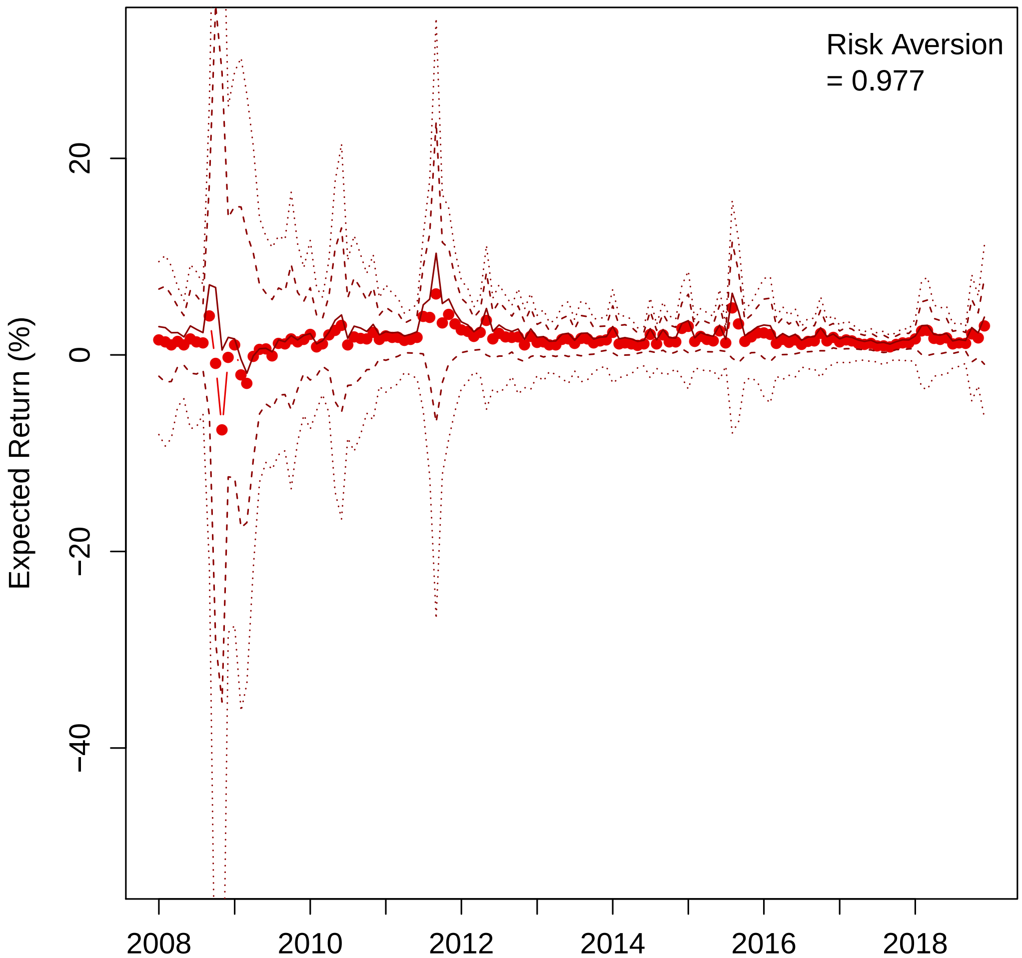
Notes: Table A2 displays the correlation of mean expectation each quarter across all assest classes. Standard errors are shown in paranthesis. *** p<0.01, ** p<0.05, * p<0.10.

Figure A1: Expectations and Model Fit: Robustness Checks



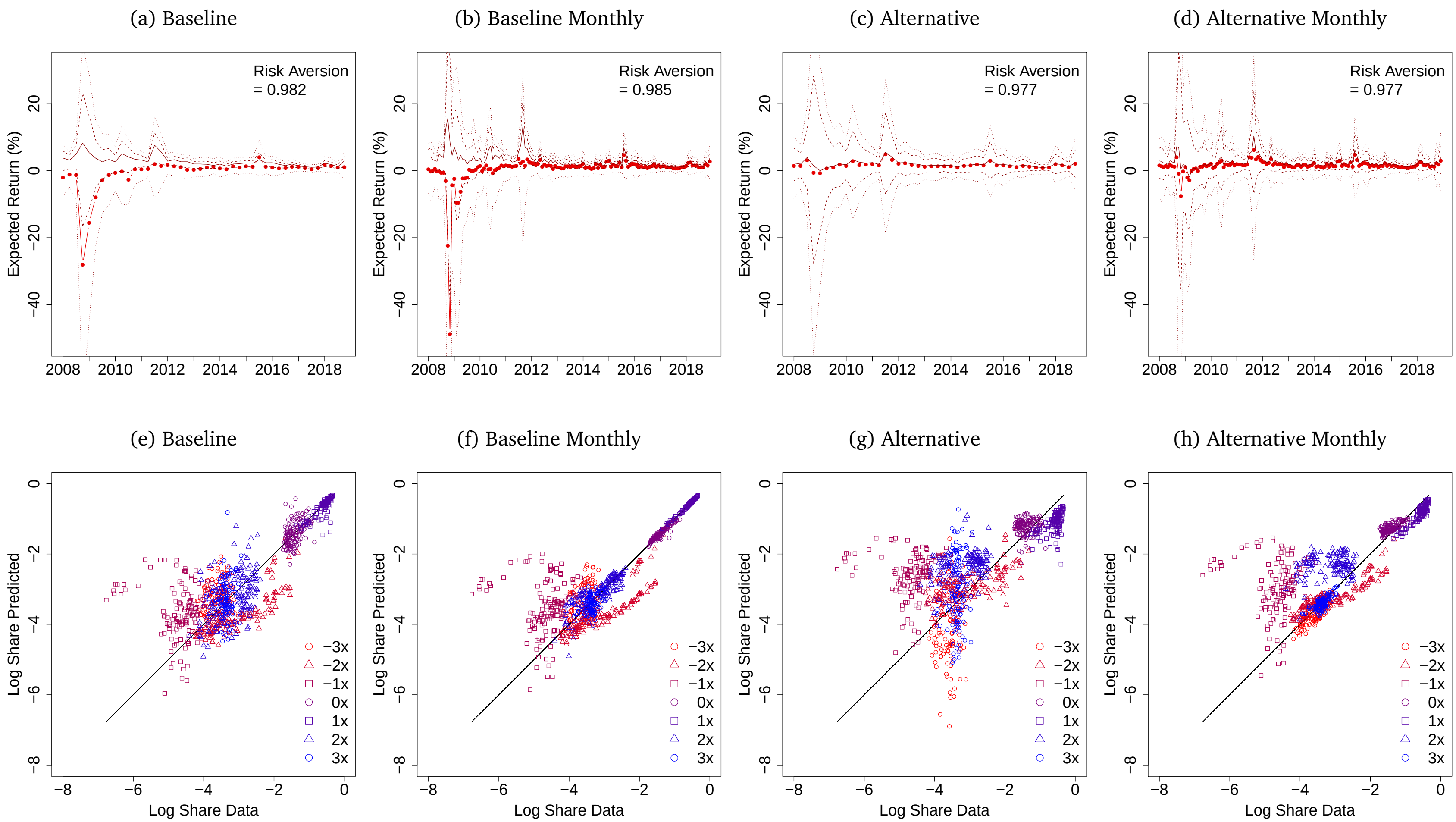
Notes: Figure A1 panel (a) and (e) correspond to the baseline estimates. In (b) and (f), we fit data assuming expectation follows normal distribution. In (c) and (d), we scale the outside share of our baseline definition by a factor of 5. In (d) and (h), we fit relative inside share only without using 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: Time-Varying Investor Expectations: Alternative Estimates



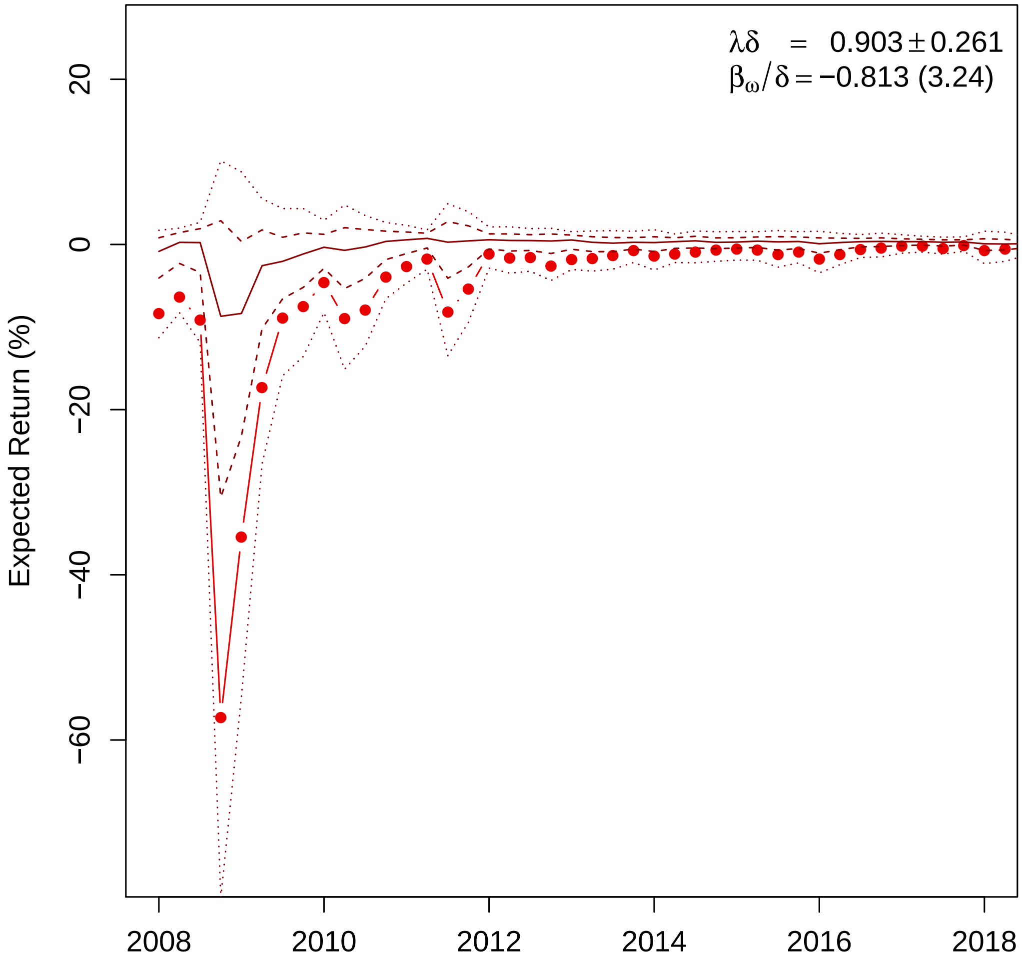
Notes: Figure A2 plots the estimated distribution of investor expectations over time in each month, using the alternative approach described in Appendix A. These estimates use only variation in the choices facing investors to recover the time-variation distribution of expecations. Red dots represent 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 A3: Expectations and Model Fit: Comparison of Baseline and Alternative Estimates



Notes: Figure A3 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.

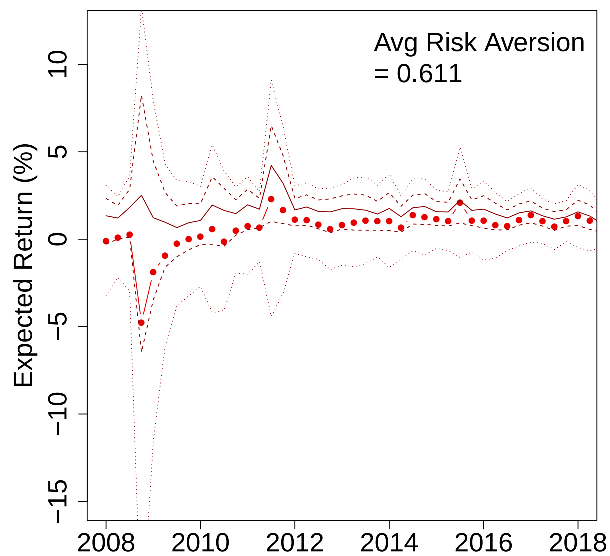
Figure A4: Time-Varying Investor Expectations: Hedging



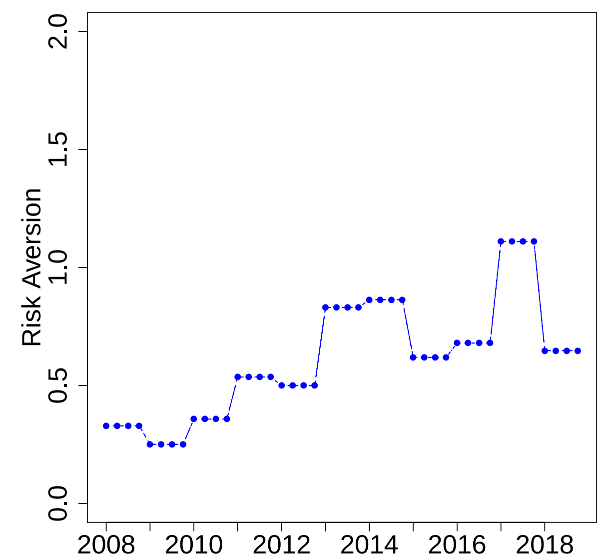
Notes: Figure A4 plots the estimated distribution of expectations over time for investors with hedging considerations. We display the risk aversion coefficient, which follows a uniform distribution, and the market risk beta of investors' wealth, which follows a normal distribution. These coefficients 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 A5: Time-Varying Risk Aversion

(a) Expectations



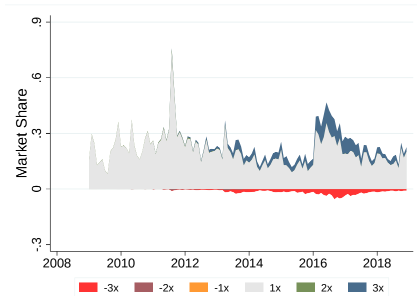
(b) Risk Aversion



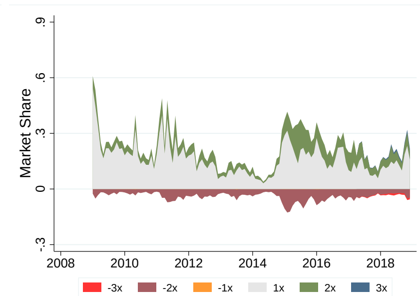
Notes: Figure A5 panel (a) plots the estimated distribution of expectations over time where we allow risk aversion to vary at annual level. Red dots represent mean expected return. Solid dark red lines indicate median, dashed lines indicate 25th and 75th percentiles, and dotted lines indicate 10th and 90th percentiles. Panel (b) plots the estimated risk aversion which varies at the annual level.

Figure A6: Market Shares: Other Asset Classes

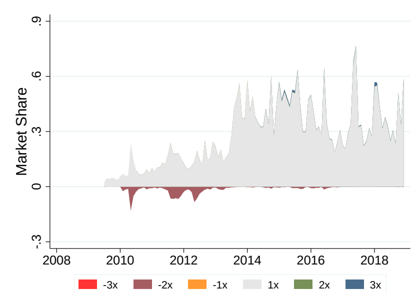
(a) Gold



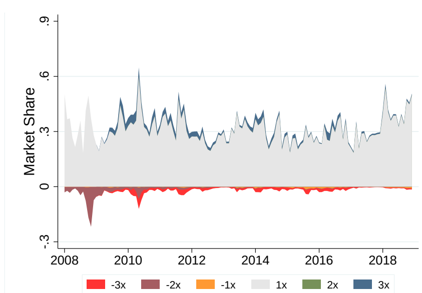
(b) Oil



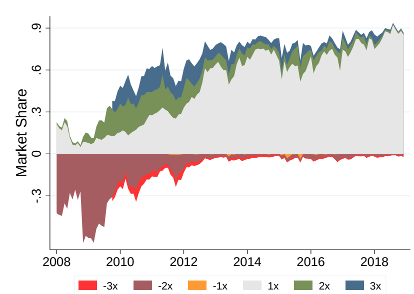
(c) European Equities



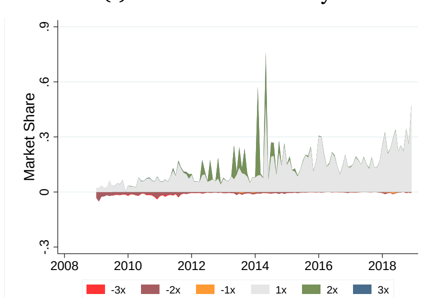
(d) Emerging Market



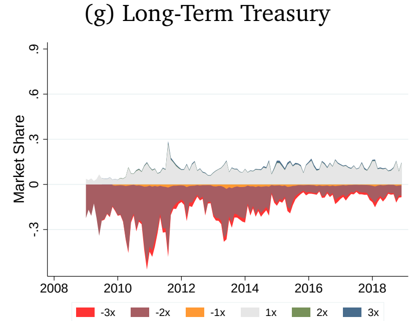
(e) US Real Estate



(f) Mid-Term Treasury

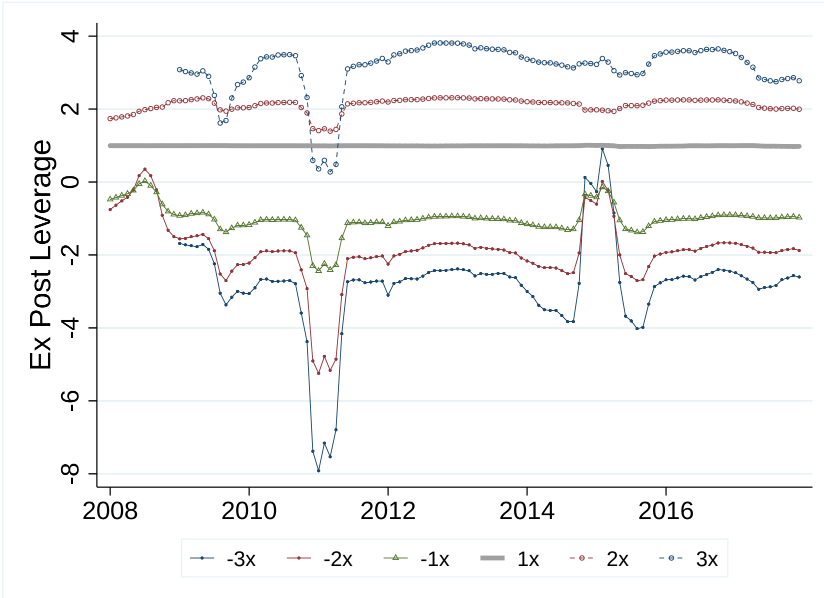


(g) Long-Term Treasury



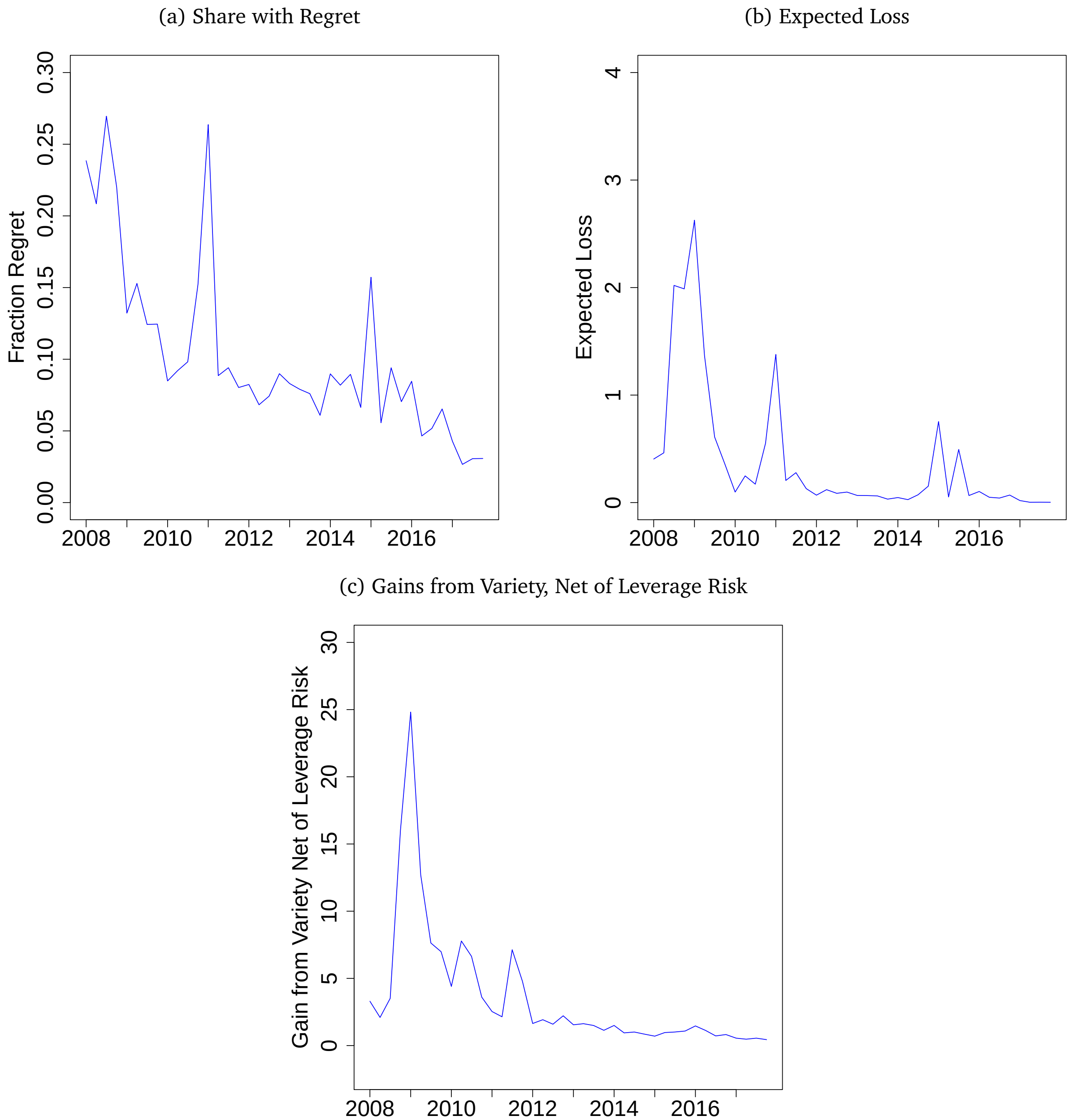
Notes: Figure A6 shows market share of each leverage for each asset class..

Figure A7: Ex Post Leverage: 12-Month Holding Period



Notes: Figure A7 plots ex post leverage for each leverage category over time. Ex post leverage is computed by running OLS regressions of leverage category returns over 12-month holding periods on S&P 500 returns over the same period, in a moving window of 7 months.

Figure A8: Welfare Effects with Leverage Risk: 12-Month Holding Period



Notes: Figure A8 shows the welfare effect of leverage risk over time. Panel (a) plots the fraction of investors who would regret their leverage category choices after learning ex post leverage. Panel (b) plots the expected loss as the difference between the utility from re-optimized choices based on ex post leverage and the utility from original choices. Panel (c) shows the gain from variety taking into account the leverage risk, measured as the utility difference between the full choice set in the data and a restricted choice set of only 1x trackers or the outside option.