# Demand-Based Expected Returns

November 2024


**Abstract**

This paper proposes a theoretical framework for recovering investors' subjective beliefs using holdings data and option prices under the assumption of no-arbitrage. We empirically document that the statistical properties of subjective expected returns and Sharpe ratios differ wildly across investor type and depend crucially on their portfolio composition. While expected returns estimated from price data alone suggest that expected returns are highly volatile and countercyclical, including holdings data can imply returns that are less volatile and procyclical. Using buy and sell orders on S&P500 options, we show that the expected returns inferred from retail and institutional investor beliefs increase in bad times when they become the net suppliers of crash insurance in option markets, mirroring price-based estimates. Market makers' expected returns decrease during bad times when they become the net buyers of crash protection when their constraints bind. Finally, we show that market makers' expected returns are highly correlated with survey measures of expected returns by sophisticated agents, while customers' expected returns are not.



*Keywords:* subjective expected returns, subjective risk, options, portfolio holdings, recovery
*JEL Classification:* G12, G40

# 1 Introduction

Canonical estimates of the expected return on the market inferred from asset prices suggest that they increase significantly during crises periods and are highly volatile. Estimates of expected returns from survey data, however, are less volatile and can be pro-, a-, or counter-cyclical depending on investor type and level of sophistication, see, e.g., Greenwood and Shleifer [2014], Nagel and Xu [2023], and Dahlquist and Ibert [2024], respectively. Price-based measures of expected returns ignore information on investors' holdings. Recent research, however, highlights the strong link between holdings data and investors' beliefs.[1] In this paper, we argue that including information about investors' portfolios is crucial to understanding the dynamics of subjective expected returns and risk.

To this end, we propose a theoretical framework for recovering beliefs of heterogeneous investors from prices and holdings data jointly under the assumption of no-arbitrage. More specifically, we theoretically show that investors' subjective expected return on the market as well as their perceived risk can be directly inferred from option prices and their corresponding option holdings in real-time at granular levels. We empirically document substantial heterogeneity in expected return estimates across investor types. Most importantly, we find that expected returns recovered from holdings data can deviate in interesting ways from price-based measures. Using transaction-level data on buy and sell orders on S&P500 index options, we show that the subjective expected return of financial intermediaries' can drop in times of distress, contrary to price-based measures, while customers' expected returns increase.

In normal times, financial intermediaries are net suppliers of deep-out-of-the-money puts to public investors, see, e.g., Gârleanu, Pedersen, and Poteshman [2008] and Chen, Joslin, and Ni [2019]. However, during crises times when financial intermediaries' constraints bind, public investors provide crash insurance to intermediaries. As a consequence, estimates of customers' subjective expected returns increase during bad times (due to their market exposure via the short puts), while intermediaries' expected returns decrease (due to their protection). We also find that our measures of expected returns correlate highly with survey measures of expected returns of professional investors while the correlation with households' expected returns is low. In line with a large literature that studies measures of expected returns inferred from survey data, we conclude that the dynamics of subjective measures of expected returns can vary greatly across (highly sophisticated) investors.

In arbitrage-free markets, prices are the expected value of future payoffs discounted by

[1] For example, Giglio et al. [2021] document a strong relationship between investors' expected returns and portfolio holdings using surveys. Beutel and Weber [2023] provide causal evidence for the link between beliefs and portfolio decisions using experiments. And Egan, MacKay, and Yang [2024] show that belief heterogeneity accounts for the majority of the variation in households' portfolio allocations.

some stochastic discount factor (SDF) $M$. The expectation is computed under the probability measure $\mathbb{P}$ supported by $M$. While $\mathbb{P}$ encodes the investor's subjective belief, the SDF encodes her risk preferences. Standard methods extract agents' beliefs from asset prices under some assumptions for $M$. These methods, however, ignore information about quantities (such as portfolio holdings, trading flows, or open interest) which, different from prices, are available on a granular level, that is, for each investor.

Our "demand-based" belief recovery extracts $\mathbb{P}_i$ for an investor $i$ by leveraging investor-level data on holdings together with option prices. More specifically, we assume that investors with potentially heterogeneous beliefs can hold wealth shares in the market index and a family of options written on the market index. Our main theoretical result posits that subjective expected returns under $\mathbb{P}$ can be directly inferred from investors' holdings and option prices. As is well-known, the risk-neutral pricing measure $\mathbb{Q}$ can be fully determined by observed option prices under no arbitrage (Breeden and Litzenberger [1978]). Since holdings are observable at the investor level and payoffs under $\mathbb{Q}$ can be recovered from option prices directly, we obtain not only measures of subjective expected returns in real-time for each investor type but it also allows us to recover measures of subjective risk.

With this methodology, we obtain SDFs that are joint functions of the index and the options returns. Therefore, the ensuing expected market returns and measures of risk may be less or more volatile depending on portfolio composition, with sign and cyclicality properties that depend on the contingent state of the economy. The shapes of the SDF projections also span a large variety of functional forms. For instance, we can recover loss averse investors with time-varying risk aversion, who expect a relatively stable market in the future and contribute to a low premium. These agents take short positions on out-of-the-money options. Similar intuition also allows us to recover SDF projections that are monotonically decreasing or monotonically increasing. Earlier literature ignores options because it is assumed they are in zero-net supply. Since in reality, options are non-redundant, we show that holdings in option portfolios are informative about investors' beliefs. For example, a larger investment in deep-out-of-the-money puts corresponds to conservative investors that are progressively more sensitive to higher-order risk factors and trade options to reallocate them profitably. Moreover, we show that not only the sign, but also the cyclicality of the market risk premium is endogenous to investor's belief. In order to illustrate our theoretical framework, we merge option price information with buy and sell orders of large investors in index option markets.

More specifically, we use our results to gain insights about the beliefs and subjective expected returns of two groups of option market participants: public investors (retail and institutional) and intermediaries. To this end, we leverage the CBOE Open-Close Database which records daily buy and sell orders per investor category for every option. Real-time holdings data allows us to recover each investor's beliefs such that the solution to our recovery problem is aligned with the observed portfolio positions.

We summarize our empirical findings as follows. First, we find that customers and market makers can have complementary patterns with regards to the shape of their SDFs. For example, during normal times, market makers hold large short positions in calls and puts which exposes them to changes in both the up- and downside. As a consequence, market makers' SDFs are $U$-shaped as a function of expected returns. Customers, on the other hand, who are net demander of these options, have inverted patterns. These regularities, however, changed dramatically during so-called crisis days when intermediaries' constraints start to bind. For example, we find that in November 2008, customers' SDF projections are monotonically decreasing, while market makers' SDFs are flat for negative returns and increasing for positive returns. The reason for this is that market makers become net demanders during this period for downside protection.

Second, the changing portfolio holdings and exposures to downside risk during crisis periods across the two investors has large effects on the time-series properties of expected returns. We find that the ensuing expected returns are pro-cyclical and very volatile for market makers. In fact, we observe that expected returns become negative during crises, up to -17% per year in Fall 2008. Intuitively again, this happens because of the large long put positions that they hold on their portfolios. Customers, on the other hand, have countercyclical expected returns because they make the market for crash insurance during bad times. As a consequence, we find the average correlation between the two expected return measures to be a mere 40%.

Third, we can also use our framework to recover subjective measures of risk, allowing us to measure subjective risk and return trade-offs. We find that measures of perceived risk are more highly correlated than their expected return measures while subjective Sharpe ratios are basically uncorrelated (mostly due to the low correlation in expected returns).

Finally, we also study the relation of our expected return measures with survey measures of expected market returns. We find that while the correlation with survey measures of expected returns for individuals or retail investors is low (or even negative), the correlation of expected return measures of professional investors with intermediaries' expected return is over 70%. Finally, we study determinants of our expected return measures. We find that standard predictors of realized returns such as the dividend-price ratio, do not have any statistically significant relation with expected returns. The only variable that loads significantly on expected returns are past returns where the slope coefficient is negative. This implies that when returns are low, both market makers and customers expected expected returns to be high. This is in contrast to survey evidence which has shown that retail investors' expected returns tend to be positively correlated with past returns. We also do not find any statistically significant relation between expected returns and measures of cyclicality.

Our measures of subjective expected returns and risk can be interpreted as the lower bounds on expected returns and perceived risk by heterogeneous investors, offering intuition for why some of the literature has documented different cyclicality patterns across various surveys. We emphasize that our approach does not recover the "true" beliefs of investors but provides a sensible benchmark for the potential beliefs of large players in the option and index market.

*Related Literature.* This paper is related to several strands of the literature. Starting from the seminal work of Ross [2015] a growing literature has proposed ways to recover investors' subjective beliefs, see, e.g., Borovička, Hansen, and Scheinkman [2016], Jensen, Lando, and Pedersen [2019], among others for recent refinements of the Ross [2015] recovery theorem. Our framework differs from these papers in at least two ways: First, we include demand-based data instead of just asset pricing data to extract investor-specific beliefs. Second, our framework allows us to recover conditional beliefs in real-time.

Chen, Hansen, and Hansen [2020], Ghosh and Roussellet [2023], and Korsaye [2024] use survey data in addition to price data to recover the representative agent's belief and study their properties relative to a rational expectations framework. As we show, holdings data allows us to recover beliefs on a much more granular level, that is at the investor level. More generally, our theoretical framework also allows for the inclusion of survey data. However, long time-series of granular survey data is hard to obtain.

Our paper is also related to the literature that studies the option demand of heterogeneous investors. For example, Chen, Joslin, and Ni [2019] document how variation in the net demand of deep OTM put options between intermediaries and public investors is driven by intermediaries' constraints. Almeida and Freire [2022] show how net option demand helps explain the pricing kernel puzzle. And Farago, Khapko, and Ornthanalai [2021] study a heterogeneous agent economy to explain index put trading volumes. We complement this literature by estimating intermediaries' and public investors' beliefs from observed option demand.

Our paper is most closely related to the literature that makes use of asset prices to recover measures of the expected return. Even tough these papers do not explicitly recover heterogeneous investors' beliefs, some of their results are nested in our framework. For example, Martin [2017], Martin and Wagner [2019], and Gao and Martin [2021] derive lower bounds on expected returns for stocks by assuming that the expected return of an asset can be inferred from the allocation of a growth-optimal portfolio that maximizes an investor's long-run growth. Expected returns are shown to be functions of risk-neutral variance. Chabi-Yo and Loudis [2020] use a Taylor series expansion of the inverse of the marginal utility to construct lower and upper bounds on the conditional expected excess market return that are functions of higher-order risk-neutral simple return moments. Gormsen and Jensen [2022] study physical (as opposed to risk-neutral) moments as perceived by a power utility investor. Gandhi, Gormsen, and Lazarus [2023] study the term structure of expected returns inferred from option prices and find that long-term expected returns are (excessively) countercyclical and volatile. Tetlock [2023] assumes that the SDF of a log investor is the reciprocal of a combination between the market return and the return of a portfolio of higher-order (risk-neutral) moments of $R$ whose weights come from regressing the variance premium on some risk-neutral moments to obtain point estimates of the expected return. Our findings show that the dynamics and statistical properties of measures of expected returns crucially depend on the weights allocated to the basis assets. While Martin [2017] assumes that investors choose to hold 100% of their wealth in the market (and none in the derivatives themselves), Tetlock [2023] allows for holdings in both the market and power contracts on the market. In our setting, we do not need to make any assumptions about the redundancy of option markets and optimal weights since our estimation framework incorporates information from actual weights as provided by transaction level data.

Our paper contributes to an empirical literature studying beliefs of heterogeneous investors using surveys. Dahlquist and Ibert [2024] document large heterogeneity in asset managers' beliefs, while Giglio et al. [2021] study the relationship between retail investors' beliefs and portfolio holdings. Meeuwis et al. [2022] document that political orientation determines households' beliefs and portfolio allocation into risky assets. Ghosh, Korteweg, and Xu [2022] recover heterogeneous beliefs from the cross-section of stock returns. Our paper is different from these papers since we recover beliefs from price and holdings data jointly, allowing us to measure beliefs for a long-time series at the daily frequency for large investors. We document, however, that intermediaries' subjective measure of returns is highly correlated with survey data of sophisticated investors, while households' expectations are not.

Finally, our paper contributes to the demand-based asset pricing literature starting with the seminal work of Koijen and Yogo [2019]. Similar to our approach, asset-demand systems impose constraints such that holdings data is matched and market clearing holds in equilibrium. While that literature is mainly interested in how heterogeneous investors affect movements in asset prices, our focus is on recovering subjective expected returns.

*Outline.* The rest of the paper is organized as follows. The key idea of our paper is that holdings data is informative about investors' risk perceptions. We illustrate this idea in an intuitive example in Section 2. Section 3 presents a general theoretical framework where we show how to infer subjective expected returns and risk from holdings and price data. Section 4 contains our main empirical results. All proofs and some additional mathematical details are provided in the Appendix. Additional results are gathered in an Online Appendix.

# 2 Illustrative Example

The key idea of our paper is that portfolio holdings are informative about risk perceptions of market participants. To provide some intuition, we start with an example to illustrate how portfolio holdings affect beliefs/SDFs. Assume heterogeneous investors who hold a growth-optimal portfolio, that is, investors maximize expected long-run wealth. As is well-known, in this case, the growth-optimal return is the reciprocal of the stochastic discount factor, i.e., $M^\star = 1/R^\star$, where $R^\star$ is the return of the optimal portfolio according to the investor's subjective view, see, e.g., Long [1990]. Even though investors have the same preferences and are subject to the same constraints, the optimal portfolio can vary across investors because they may have different beliefs.

To set a benchmark, assume there exists a specific constrained utility-maximization problem whose solution is a portfolio fully invested in the market. In that case, the SDF takes the following form: $M^0 := 1/R$, where $R$ is the return on the market index. This is the case studied in Martin [2017]. As we argue in our paper, a priori, there is no reason to exclude other traded assets (say, options) from the optimal portfolio. In fact, ample empirical evidence in the literature shows that options are non-redundant securities and demand for options can be in the order of trillions of dollars, especially following market crashes.[2] In that case, the corresponding optimal portfolio will have non-zero positions in the index options, and the return $R^\star$ will be different from $R$ (and in turn $M^\star \neq M^0$).

Let $M^*$ be the SDF supported by a portfolio being long some calls or puts. Figure 1 compares the value of $M^\star$ relative to $M^0$, as a function of the only state variable $R$ for calls (left panel) and puts (right panel) across different moneyness. For illustrative purposes, we assume that the investor holds a portfolio consisting of 85% in the index and 15% in an equally-weighted portfolio of calls or puts with the same maturity but different strikes.[3] As is immediately evident, even for a small investment in options (with respect to the investment in the underlying) and small changes in the market return, the ratio $M^0/M^\star$ can wildly differ from 1: $M^0/M^\star$ rises significantly for at-the-money (ATM) and especially out-of-the-money (OTM) calls and puts. For example, assuming the market excess return is +20% (-20%), the ratio increases to 7 (5.5) for OTM calls (puts).

Intuitively, if we would like to recover the probability measure $\mathbb{P}^\star$ (supported by SDF $M^\star$) from the probability measure $\mathbb{P}^0$ of an investor fully invested in the market, we can interpret the ratio $M^0/M^\star$ as the corresponding probability distortion. For instance, the belief of an investor who optimally chooses to be long in puts, is more left-skewed than the benchmark. Investing in puts, shifts the probability mass uniformly from the region of positive market return towards the region of negative returns proportional to the strikes and the moneyness.

[2] Theoretically, positive net demand for options can arise in settings with heterogeneous beliefs. For example, Buraschi, Trojani, and Vedolin [2014] show that agents with more pessimistic views about the future growth of the economy demand OTM puts from more optimistic agents. In a setting with frictions and market incompleteness Johnson, Liang, and Liu [2016] show that the primary reason for the high demand in index option is transfer of unspanned crash risk.

[3] In later sections, we will use transaction-level data from CBOE to track portfolio holdings in real-time.

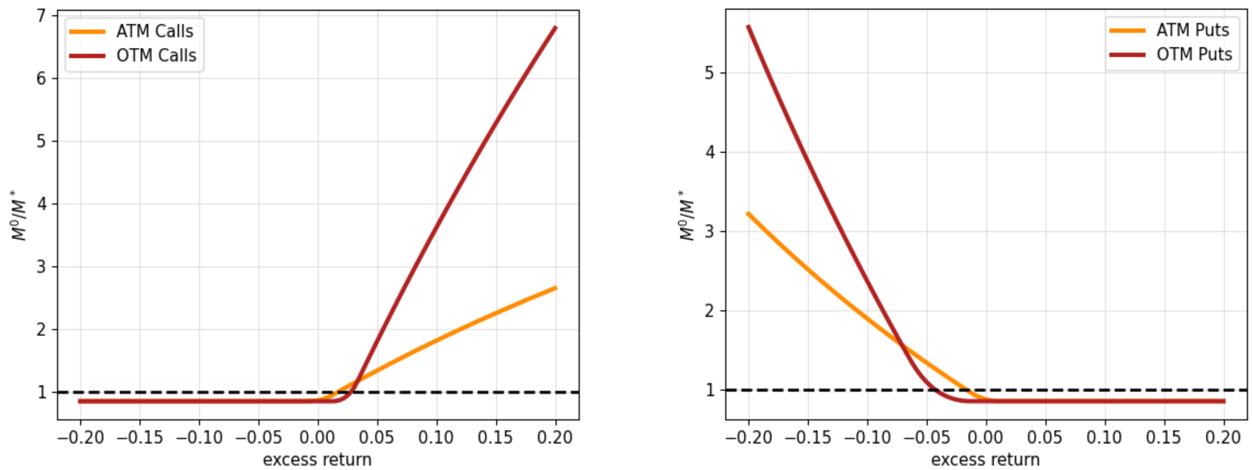



**Figure 1.** Ratio between Benchmark SDF $M^0$ and $M^\star$
*Notes:* This figure plots $M^0/M^\star$ as a function of the excess return on the market. $M^\star = 1/R^\star$, where $R^\star$ is the return of a portfolio investing 85% of the wealth in the underlying and 15% in an equally-weighted portfolio of calls (left plot) and puts (right plot) with different moneyness. ATM options have $|\Delta| \in (0.375, 0.625)$. OTM options have $|\Delta| \in [0.125, 0.375]$.

More generally, investors who assign higher weights to extreme events have higher demand for deep OTM puts. The reverse holds true when the investor is long calls.

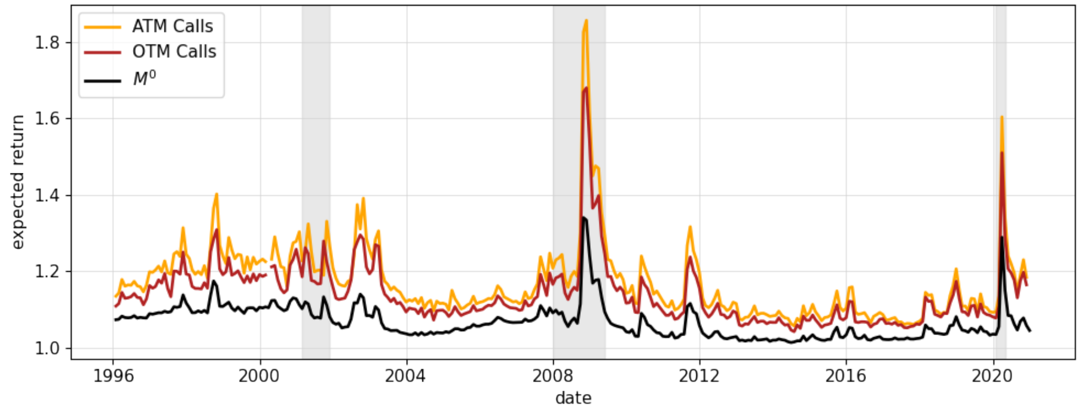
In a next step, we study the effect of investors' demand on the time-series properties of subjective expected returns. In Figure 2, we plot the time-series of the expected market return for different $M^\star$ implied by options (the same as in Figure 1) together with the benchmark log investor case who is 100% invested in the index ($M^0$).

Panel A plots the perceived expected returns for investors who hold call options in addition to the index. Not very surprisingly, the patterns mirror the benchmark case almost one-for-one. Expected returns increase in bad time, decrease in normal periods, and are highly volatile. Notice that expected returns are considerably higher even with a small investment in options. Intuitively, the size of the expected return increases relative to the benchmark since options represent a levered trade on the underlying itself. The portfolio with ATM calls has the highest premium, since ATM calls move one-for-one with the underlying market index.

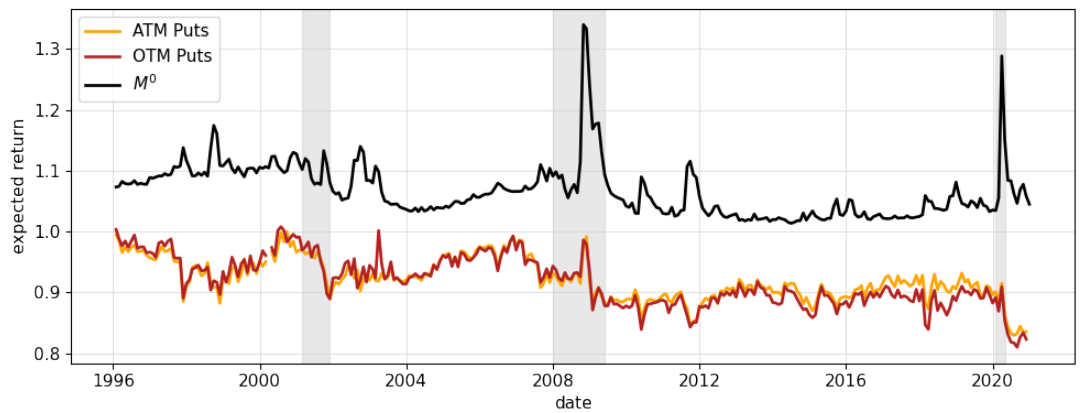
We can juxtapose this pattern with inferred expected returns from investors' who are long in puts. As can be seen from the middle panel, being long in puts decreases the exposure to market risk, and as a consequence, the corresponding expected return is lower. In fact, the expected return even becomes negative. As discussed before, holding puts reflects the view of investors holding more left-skewed beliefs, who expect higher (negative) market fluctuations. From their perspective, the risk-return ratio given by holding the index alone is not profitable - or equivalently, they find OTM puts to be under-priced. Buying puts leads to protection which reflects a pessimistic view and negative expected return. Accordingly, the volatility of the expected return is higher with respect to what we recover under $M^0$. Notice also that even small investments in puts can lead to a pro-cyclical pattern of the expected return. Given that, we conclude that the size and cyclicality properties of expected returns depend on the moneyness, and on the amount of wealth invested in the options.

While instructive, the above examples maybe too stylized. To study a more realistic setting, we now showcase two popular option strategies: collars and straddles. For example, some investors are known to hold the underlying and add protection via longing puts and shorting calls (collar); other investors bet on the underlying volatility by taking long positions in calls and puts (straddle). Agents are also known to typically delta-hedge their option positions. In Figure 2 Panel C, we plot the monthly time-series of expected returns recovered from hypothetical delta-hedged collar and straddle strategies with OTM options, compared to the benchmark case. Results align with our previous example: being long in calls/puts increases the distortion in the tails, the reverse is true for short positions. Thus the option component in a collar strategy significantly reduces the expected return.

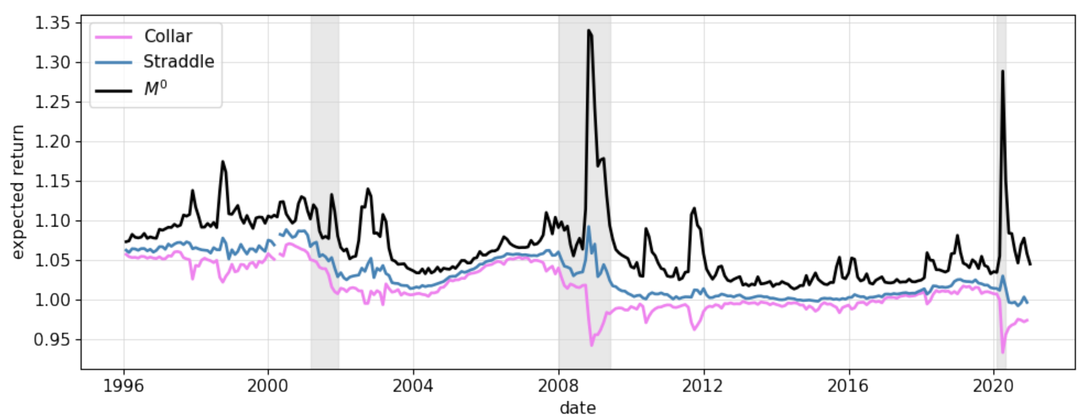
We conclude that not only the size, but also the sign, cyclicality, and volatility of the subjective expected return depend on asset demand. While these results are based on hypothetical portfolios that do not represent any specific investor, in our empirical section we will use transaction level data on buy and sell orders to track investors' beliefs over time.




**A.** Calls




**B.** Puts




**C.** Collar and Straddle

**Figure 2.** Monthly Time-Series of Expected Market Return

*Notes:* This figure plots the expected market return recovered from different stochastic discount factors. In each panel, we plot the expected return recovered by $M^0 = 1/R$, where $R$ is 100% invested in the index, as well as by $M^\star = 1/R^\star$, where $R^\star$ is the return of a portfolio investing 85% in the index and 15% in an equally-weighted portfolio of ATM (orange line) or OTM (red line) options for calls (Panel A) and puts (Panel B). The bottom panel shows a delta-hedged «collar» which is long in the index and in OTM puts, and short in OTM calls. The <delta-hedged <straddle» is long in OTM calls and OTM puts. Frequency and horizon are monthly, values are annualized. Grey areas indicate NBER recession periods.

# 3 Theoretical Framework

We now present a simple theoretical framework to explain how to recover subjective expected returns from options price and holdings data. Consider an investor, labeled $i$, with logarithmic preferences who has access to three types of assets: a risk-free asset with return $R_f$, a risky asset with forward return $R$, and an entire family of options written on the risky asset with a continuum of strike prices. Let $\mathbb{E}_i[\cdot]$ denote the subjective expectation of this investor over possible states of the world. The investor's subjective beliefs may or may not coincide with the true underlying data-generating process. In this paper, we assume that the risky asset is the S&P 500 index and the options are European calls and puts. Let $\boldsymbol{R}^e = \boldsymbol{R} - \mathbf{1}$ be the excess forward return of the index and options.

## 3.1 Subjective Expected Returns

Our goal is to recover the physical belief $\mathbb{P}_i$ for investor $i$ under the minimal assumptions stated above to infer the subjective expected return on the market, $\mathbb{E}_i[R]$. Since investors have logarithmic utility, it immediately implies that one can define an agent-specific SDF $M_i$ that prices all assets from the perspective of agent $i$ as follows:

$$M_i = \left(1 + \boldsymbol{\theta}_i' \boldsymbol{R}^e\right)^{-1}, \tag{1}$$

where $\boldsymbol{\theta}_i$ are the portfolio weights in the market index and the options by investor $i$, see, e.g., Long [1990]. SDF $M_i$ is the reciprocal of the return of the growth-optimal portfolio since it maximizes expected long-run growth of the investor $i$'s wealth. No arbitrage implies that

$$\mathbb{E}_i[M_i \boldsymbol{R}^e] = \mathbf{0}.$$

We can now define a change of measure, $\frac{d\mathbb{Q}}{d\mathbb{P}_i} = M_i$. The subjective expected return of the market for investor $i$ under the physical measure can hence be written as:

$$\mathbb{E}_i[R] = \mathbb{E}^{\mathbb{Q}}[M_i^{-1} R] = \mathbb{E}^{\mathbb{Q}}\left[\left(1 + \boldsymbol{\theta}_i' \boldsymbol{R}^e\right) R\right] \tag{2}$$

Equation (2) is the main identity studied in this paper. It relates investor-specific physical beliefs to risk-neutral expectations about prices of assets that can be traded and investor $i$'s holdings. Intuitively, we interpret equation (2) as the most conservative measure of the expected return of a log investor. Two remarks are in order. First, notice that we do not need to make any assumptions about whether investor $i$ is constrained or not since the constraints are reflected in the portfolio holdings. Second, as we will show later on, even if we do not believe that investors have log utility, identity (2) still provides us with a useful lower bound.

In the following, we will consider two different cases. First, the fact that portfolio holdings are observable in the data at high frequency, allows us to directly recover agent $i$'s expected return of the market as a function of holdings and option prices in real-time. Second, it is reasonable to assume that holdings data is measured with some error. The intuition for this is at least twofold. First, we only observe a subset of the "true" portfolio of investors. For example, while we observe the open and close orders on calls and puts for the S&P500 (SPX), we do not observe the holdings and neither the transactions on other derivatives with the same underlying, such as SPY options (that is on the ETF tracking the S&P500).[4] Since major market makers provide liquidity in both SPX and SPY option markets, we only observe a fraction of their true market exposure.[5] Second, our data is sampled at high frequency (every 30 minutes), however, we aggregate to the monthly frequency and across different types of customers (retail and institutional) to calculate expected returns. This aggregation will likely lead to further measurement error affecting our estimates. Given this, we consider a second case where we assume that portfolio holdings are observed with measurement error leading to lower and upper bounds on the subjective expected returns representing the most conservative and most aggresive value, respectively.

### 3.1.1 Log Utility Assumption

Before explaining technical details on how to recover subjective beliefs from the data, one might be worried that the log utility assumption used in equation (1) is unrealistic. Notice that our framework still provides insights into the subjective expected return in the form of a lower bound even if investors do not have log utility. To see this, suppose that the investor has some other utility function, and potentially unknown belief $\mathbb{P}^*$. We say that the negative covariance condition (NCC) holds if:

$$\text{Cov}^*(M^*\boldsymbol{\theta}_i'\boldsymbol{R}, R) \leq 0 \,, \tag{3}$$

where $M^* = d\mathbb{Q}/d\mathbb{P}^*$ is the true stochastic discount factor of the non-log investor and $\boldsymbol{\theta}_i$ is her portfolio, as above. In the case where the portfolio is fully invested in the index, we get the same NCC as in Martin [2017]. In particular, he shows that the NCC is likely to hold empirically

[4]While SPX options trade exclusively on the CBOE, SPY options trade across several exchanges. Institutional public investors mainly trade SPX options (due to larger contract sizes, tax treatment, etc.) while retail investors mostly trade in SPY options.

[5]Moussawi, Xu, and Zhou [2024] show that at least four market makers provide liquidity in both markets simultaneously: Susquehanna Securities, Citadel Securities, Wolverine Trading, and IMC Financial Markets.

and also holds in several leading macro-finance models. If the NCC holds for any portfolio $\boldsymbol{\theta}_i$ of agent $i$, we get the following lower bound:

$$\mathbb{E}^*[R] \geq \frac{\mathbb{E}^{\mathbb{Q}}[(1+\boldsymbol{\theta}_i'\boldsymbol{R}^e)R]}{\mathbb{E}^{\mathbb{Q}}[1+\boldsymbol{\theta}_i'\boldsymbol{R}^e]} = \mathbb{E}_i[R]\,, \tag{4}$$

where we formally recognize $(1+\boldsymbol{\theta}_i'\boldsymbol{R}^e)/\mathbb{E}^{\mathbb{Q}}[1+\boldsymbol{\theta}_i'\boldsymbol{R}^e]$ as the change of measure from $\mathbb{Q}$ to belief $\mathbb{P}_i$.[6] This change of measure has the functional form of a log-type SDF as in (1). The right-hand side of inequality (4) is the recovered expected return in (2). Therefore, the $\mathbb{E}_i[R]$ we extract under the log-utility assumption additionally provides a lower bound for the subjective moment of the agent holding the same portfolio but with non-log utility.

## 3.2 Recovering Subjective Expected Returns

We can now discuss how to recover subjective beliefs from options. As is well-known, an arbitrage-free cross-section of options suffices for the existence of a probability measure $\mathbb{Q}$, which determines the price of any payoff that is replicable by a delta-hedged option portfolio (see, e.g., Acciaio et al. [2016]). In our application, the pricing measure $\mathbb{Q}$ is a forward probability between times $t$ and $t+1$. This trivially implies that $\mathbb{E}^{\mathbb{Q}}[\boldsymbol{R}^e] = \mathbf{0}$.

Given this, we can directly compute equation (2) from the data, since we observe the portfolio holdings $\boldsymbol{\theta}_i$ and we can calculate the risk-neutral expectation of $(1+\boldsymbol{\theta}_i'\boldsymbol{R}^e)R$ using the Carr and Madan [2001] formula.

Before delving into details about implementation, several remarks are in order. Our set-up is similar to Martin [2017], who derives a lower bound on the expected return assuming an unconstrained rational investor who's risk aversion is at least one. The crucial difference to Martin [2017] is that he imposes that the optimal portfolio held by the investor consists of a 100% investment in the market return. A consequence of this assumption is that all options are redundant. As we argue in this paper, there are at least two reasons why this assumption seems too restrictive. First, several papers show that options are non-redundant assets since they allow to hedge crash risk and demand for OTM puts options is significant. Second, in the data, we find that the SDF defined as the inverse of the market return induces significant pricing errors when pricing options. For example, we find that for the 1996 to 2020 period, the average pricing error is: 30% for OTM puts, 20% for ATM puts, 23% for OTM calls, and 1% for ATM calls, respectively.

In our paper, we do not assume such redundancy. However, notice that our setting nests Martin [2017]'s case if one assumes a log investor and $\theta_i$ equals one for the market and zero

[6] But $\mathbb{E}^{\mathbb{Q}}[1+\boldsymbol{\theta}_i'\boldsymbol{R}^e] = 1$.

otherwise. In that case, the expected return of investor $i$ is equal to $\mathbb{E}^{\mathbb{Q}}[R^2]$ which is the risk-neutral variance measured under the forward-measure $\mathbb{Q}$.[7]

Another related paper is Tetlock [2023] who assumes that the investor can hold (integer) power-contracts written on the market index. The paper again makes an assumption about the redundancy of certain option contracts such as non-integer power contracts. In general, however, we do not observe holdings of power-contracts. To circumvent this issue, Tetlock [2023] estimates "hypothetical" portfolio holdings from expanding window regressions predicting risk-neutral (power) moments with physical counterparts. Our approach is different since we directly observe holdings on plain vanilla calls and puts.

## 3.3 Bounds on Subjective Expected Returns

Under the assumption that we observe holdings with no error, subjective expected returns can be recovered via the exact identity in equation (2). For the reasons discussed earlier, when holdings $\boldsymbol{\theta}_i$ are observed with error, we can provide lower and upper bounds on the subjective return on the market. Since in that case, we want to constrain the amount by which the optimal portfolio weights can deviate from the observed weights, we impose some constraint such that

$$d(\boldsymbol{\theta}_i, \boldsymbol{\theta}^\star) \leq \delta \, , \tag{5}$$

where $\boldsymbol{\theta}^\star$ are the growth-optimal portfolio weights and for some convex discrepancy function $d(\cdot,\cdot) \geq 0$ such that $d(\boldsymbol{x}, \boldsymbol{y}) = 0$ if and only if $\boldsymbol{x} = \boldsymbol{y}$. Intuitively, $\delta$ measures the amount that optimal weights can deviate from the observed weights. To be concrete, in the following, we assume a $L^2$-norm.[8]

In that case, $d(\boldsymbol{\theta}_i, \boldsymbol{\theta}^\star) = \frac{1}{2}||\boldsymbol{\theta}^\star - \boldsymbol{\theta}_i||_2^2$. To solve for the subjective expected return, we now have to solve the following optimization problem:

$$\inf_{\boldsymbol{\theta}^\star \in \mathbb{R}^N} \left\{ \mathbb{E}^{\mathbb{Q}}[(1 + \boldsymbol{\theta}^{\star\prime} \boldsymbol{R}^e) R] + \lambda \left( \frac{1}{2}||\boldsymbol{\theta}^\star - \boldsymbol{\theta}_i||_2^2 - \delta \right) \right\} \, . \tag{6}$$

This leads us to our second main result which are closed-form solutions on the bounds of subjective expected returns.

[7] If the investor has log utility, Martin [2017]'s lower bound becomes an exact identity since $\text{Cov}(M_i R, R) = 0$ in that case.

[8] The main reason we use an $L^2$-norm (as opposed to other norms) is for tractability, as in this case, we get closed-form solutions.

**Proposition 1** (Bounds on Subjective Expected Returns)**.** *Assume that portfolio weights* $\boldsymbol{\theta}_i$ *are observed with error, in that case, the lower bound for the subjective return on the market for investor* $i$ *is:*

$$\mathbb{E}_i[R] \geq \mathbb{E}^{\mathbb{Q}}[(1+\boldsymbol{\theta}_i'\boldsymbol{R}^e)R] - \sqrt{2\delta}||\mathbb{E}^{\mathbb{Q}}[\boldsymbol{R}^e R]||_2\,. \tag{7}$$

*The upper bound for the subjective return on the market for investor* $i$ *is given by:*

$$\mathbb{E}_i[R] \leq \mathbb{E}^{\mathbb{Q}}[(1+\boldsymbol{\theta}_i'\boldsymbol{R}^e)R] + \sqrt{2\delta}||\mathbb{E}^{\mathbb{Q}}[\boldsymbol{R}^e R]||_2 \tag{8}$$

It is obvious that if the measurement error is assumed to be zero, i.e., $\delta = 0$, that equations (7) and (8) coincide with equation (2). Intuitively, we can interpret the lower bound as the most conservative assessment of the subjective expected return for any investor who's portfolios align in a neighborhood around the observed portfolios, $\theta$.

## 3.4 Subjective Risk

Our theoretical framework also allows us to study subjective expectations of higher-order moments beyond returns. To this end, we next study the subjective risk-return trade-off. While the relationship between measures of realized risk and returns is normally weak, recent literature generally reports a strong positive relationship using survey based measures, see, e.g., Couts, Goncalves, and Loudis [2023] for the aggregate stock and bond markets and Jensen [2024] for individual stocks. Our theoretical framework can be trivially extended to study measures of subjective risk from prices and holdings using equation (2). It follows immediately, that subjective risk for investor $i$ is given by:

$$\mathbb{E}_i[R^2] = \mathbb{E}^{\mathbb{Q}}[M_i^{-1}R^2] = \mathbb{E}^{\mathbb{Q}}\left[\left(1+\boldsymbol{\theta}_i'\boldsymbol{R}^e\right)R^2\right]\,, \tag{9}$$

and the subjective volatility reads as $\mathbb{E}_i[R^2] - \mathbb{E}_i[R]^2$. In the following section, we explain how we recover subjective expected returns and risk using equations (2) and (9).

# 4 Empirical Analysis

This section describes the data used and how we empirically implement our main theoretical results.

## 4.1 Data

To empirically implement our theory, we make use of the CBOE Open-Close dataset that provides daily buy and sell volumes of SPX options since 1996 separately for type of position (opening/closing) and origin: (i) customer; (ii) brokers-dealer; (iii) firm; and (iv) market maker. As is common practice, we aggregate these daily volumes to cumulative positions for the last three categories[9] and label them "market makers". Customers include retail and institutional investors. Our data starts in January 1996 and ends in December 2020. The label "broker-dealer" is available only from 2011 and accounts for less than 3% of the trades.

The volume data comes without pricing information. To this end, we obtain end-of-day bid-ask prices from the OptionMetrics database and use best closing bid- and ask-prices to compute mid-point prices. We merge the CBOE Open-Close database with price data and apply standard filters from Bakshi, Cao, and Chen [1997]. That is, we remove option contracts: with a price less than \$3/8; with implied volatility smaller than 0.1% or greater than 1; with bid price exceeding ask price; with relative bid-ask spread larger than 1/2; that are traded for less than 5 units. We also filter out events where the sum of the transactions across investors is not zero, which happens for less than 2% of the times. No-arbitrage filters apply as well.

Throughout our empirical analysis, we use monthly frequency and horizons. We compute every quantity at daily frequency, then we aggregate to monthly averages. On each day $t$, separately for calls and puts, we linearly interpolate options volatility, options delta, and investors' holdings, on a grid of strike prices, for the required maturity (30 days if not explicitly stated otherwise).[10] The grid consists of $n$ uniformly distributed values between the smallest and the largest available strike in $t$, where $n$ is the number of calls/puts actively traded in $t$.

Table 1 reports summary statistics on the option sample for customers' net demand for calls (top panel) and puts (lower panel). Net demand is defined as total buy minus total sell orders aggregated over a day. On average, customers are net sellers for call options and net buyers for put options with an average maturity of 60 days. Traded options on average are out-of-the-money. In our empirical analysis, we focus on OTM options given that trading is mostly concentrated in these options.[11]

Portfolios investing in the risk-free asset, S&P 500, and OTM options with the same maturity are assembled at date $t$ with options having the same expiration (30 days if not explicitly stated). Since we only observe option holdings but not holdings in the underlying, we follow the literature and assume that investors build delta-hedged strategies in which the investment

[9]The CBOE Regulatory Circular defines firms as «OCC clearing member firm proprietary accounts». Thus we aggregate firms and brokers-dealers with market makers because they mainly trade against public customers, although they are not designated as intermediaries.

[10]When extrapolation is required, we look for the nearest value outside the convex hull in strikes and maturities.

[11]The results do not change qualitatively if we include all options.

**Table 1.** Summary statistics options data.

| | $K/S_t$ | | Net demand | | Maturity (days) | |
|---|---|---|---|---|---|---|
| **Calls** | All | OTM | All | OTM | All | OTM |
| mean | 1.02 | 1.06 | -2,127 | -625 | 59.9 | 59.6 |
| std | 0.06 | 0.05 | 11,127 | 10,980 | 76.3 | 71.7 |
| min | 0.08 | 1.00 | -112,908 | -123,146 | 2 | 2 |
| max | 2.03 | 2.03 | 61,835 | 71,017 | 473 | 473 |
| **Puts** | | | | | | |
| mean | 0.92 | 0.89 | 2,093 | 795 | 63.9 | 64.0 |
| std | 0.11 | 0.10 | 14,689 | 15,208 | 77.8 | 76.5 |
| min | 0.10 | 0.10 | -140,275 | -140,639 | 2 | 2 |
| max | 2.72 | 1.02 | 92,694 | 92,408 | 473 | 473 |

*Notes*: This table reports summary statistics for the options data. The relative moneyness and the maturity are computed over single option contracts that are traded on every date. Customers' net demand is instead aggregated over all the options traded on a single day; net demand is defined as the total opening/closing buy orders minus sell orders. For each variable, the first column refers to the full dataset, while the second to OTM options only. Data runs from January 1996 to December 2020.

in the index is the negative of the delta of the portfolio at time $t$, see, e.g., Gayda, Grünthaler, and Harren [2023] and Baltussen, Jerstegge, and Whelan [2024].

## 4.2 Implementation

In order to implement the expression in equation (2), we apply the Carr and Madan [2001] formula to approximate $\mathbb{E}^{\mathbb{Q}}[(1+\boldsymbol{\theta}'\boldsymbol{R}^e)R]$. More specifically, let $X(K)$ be the payoff of an option with strike $K$, and let's define $f(R) := (1+\boldsymbol{\theta}'\boldsymbol{R}^e)R$, then the subjective expected return is given by:

$$\begin{aligned} \mathbb{E}_i[R] &= \mathbb{E}^{\mathbb{Q}}[f(R)], \\ &\approx \mathbb{E}^{\mathbb{Q}}\left[f(1) + f'(1)(R-1) + \sum_k f''(k)\, x(k)\, \Delta k\right], \end{aligned}$$

where $x(k) := \frac{X(K)}{F}$ with $F$ being the forward price of the market index, and $k = \frac{K}{F}$ the relative moneyness.

### 4.2.1 Constructing Option Holdings

Subjective expected returns depend on portfolio holdings, $\theta_i$, where $\theta_i$ are the total portfolio holdings of investor $i$ as opposed to her portfolio flows. This distinguishes our approach from the more standard demand-based literature that exploits opening positions at time $t$, that is, e.g., daily buy and sell orders on new contracts. To get a measure of total portfolio holdings, notice that our database reports the total daily opening and closing positions on every option contract. Opening orders represent shocks to demand (flows), while aggregate opening and closing positions represent changes to the holdings. In order to get a measure of investors' holdings, we aggregate opening and closing positions from issuance of a contract until date $t$.[12]

Figure 3 and Figure 4 compare our monthly holdings proxy with flows for customers for aggregated OTM options.[13] On aggregate, customers' are long OTM put options but short OTM call options echoing the summary statistics in Table 1. In the run-up to the 2008 crisis, aggregate holdings for OTM put options spike in September 2007, then rapidly drop and then spike up again in August 2008.

Notice that holdings and flows exhibit different patterns. For example, holdings tend to remain positive on aggregate for puts and negative for calls, while trading flows are more erratic, oscillating between positive and negative values both for calls and puts. During periods of crisis such as in 2008, customers' option demand drops and becomes negative. This property lines up with the intuition of constrained arbitrageurs during the financial crisis when customers become net sellers, see, e.g., Chen, Joslin, and Ni [2019]. However, this pattern is much more muted when looking at aggregate portfolios, because investors buy or sell options to close previous contracts in their portfolios. Intuitively that explains why holdings and shocks not always co-move together: the correlation between the two monthly time-series is 28% for calls and 55% for puts.[14]

[12]The CBOE Open-Close Database explicitly assigns a unique identification number to every option contract which can be used to track it day-by-day from issuance to expiration.

[13]Since options are in zero net supply, market makers' aggregated holdings are just the mirror image.

[14]Notice that our measure of option holdings may differ also in sign with trading flows because they are not always large enough to offset the overall investors' positions: Figure C.1 in the Appendix shows that it is quite common to have negative changes in the option demand while holdings remain positive. The total traded volume is increasing in the recent years (see Figure C.2), while net holdings are decreasing as investors' portfolios reflect a mix of positive and negative positions.

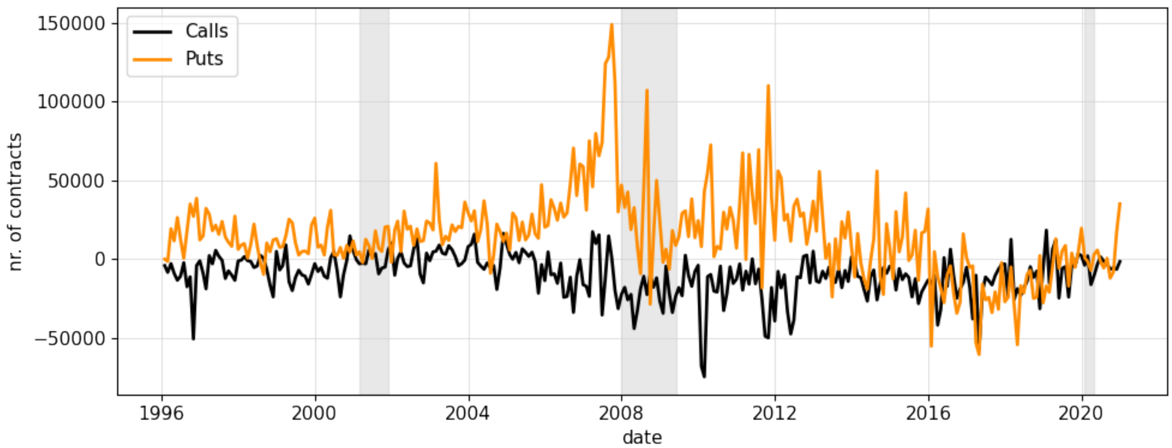



**Figure 3.** Customers' holdings of monthly OTM options
*Notes:* This figure plots the time-series of our proxy for customers' portfolio holdings of OTM calls and puts, expiring in 30 days, from 1996 to 2020. Options holdings are the sum of opening and closing positions on the same contract that customers enter from issuance. Daily data are averaged on a monthly basis. Gray bars indicate NBER recessions.


**Figure 4.** Customers' demand shocks of monthly OTM options
*Notes:* This figure plots the time-series of our proxy for customers' demand shocks of OTM calls and puts, expiring in 30 days, from 1996 to 2020. Options demand shocks are the sum of opening and closing positions on the same contract traded every day. Daily data are averaged on a monthly basis. Gray bars indicate NBER recessions.

## 4.3 Subjective Expected Returns

We now have all the ingredients to calculate subjective expected returns for the two investor types. We again stress that our theoretical framework does not allow us to recover the "true" beliefs of a particular investor, say market makers. Instead, we interpret our measure as a conservative lower bound on subjective expectations of heterogeneous agents. We start from our main identity (2): the variation of expected returns over time as perceived by different agents, assuming that the portfolio weights are observed without error.

Although we do not know the true real-time investment in the index, empirical evidence suggests that people hold the underlying and adjust their positions with delta-hedged portfolios of OTM options. Thus, we build agents' portfolios with the observed option positions from the data and we perfectly delta-hedge them. In addition, we assume that each investor allocates 50% of their wealth in the underlying. The rest of the wealth is invested in the risk-free asset. We find that customers typically take large and positive position in the index, mainly stemming from hedging their long positions in OTM puts and short positions in OTM calls (see Figure C.3).

Figure 5 plots expected returns for customers and market makers over time, in comparison with the benchmark case.[15] Table 2 lists their summary statistics. There are several interesting observations. First, we notice strikingly different dynamics of the expected return measures across investors and in comparison to the benchmark, both in size and comovement. For example, in terms of size, the expected excess market return under $M^0$ is 6.6% per year, while it is 4.1% for customers, and 2.9% for market makers. Investors' expected returns are systematically smaller than the case of Martin [2017] and often negative. Also, we notice that they are more similar in size and trends before the 2008 Great Financial Crisis; after that, they tend to be more erratic and differ more starkly from the benchmark. Intuitively, this results from the larger trading volume in the options and as a consequence, larger option leverage with respect to the underlying. Customers reach their highest expectation during Covid (excess return of 26% in March 2020); conversely, this is the lowest expectation for market makers (-15%). Between 2002 and 2008, market makers' expected returns are higher than customers; the reverse holds after the crisis almost everywhere. The correlation between customers' and market makers' time-series is 59%, while it is 86% between customers and the benchmark (60% for market makers). In the post-2008 period, correlations decrease to 33% for market makers and 48% for customers, respectively.

Our evidence can be explained by the holdings effects. We have already discussed the typical portfolio compositions of the two investor categories. Customers are long in the index and they buy protection, potentially selling OTM calls to finance their long (and costly)

[15] Notice that our recovery can lead to violations of the no-arbitrage condition in which case we drop the corresponding observation.

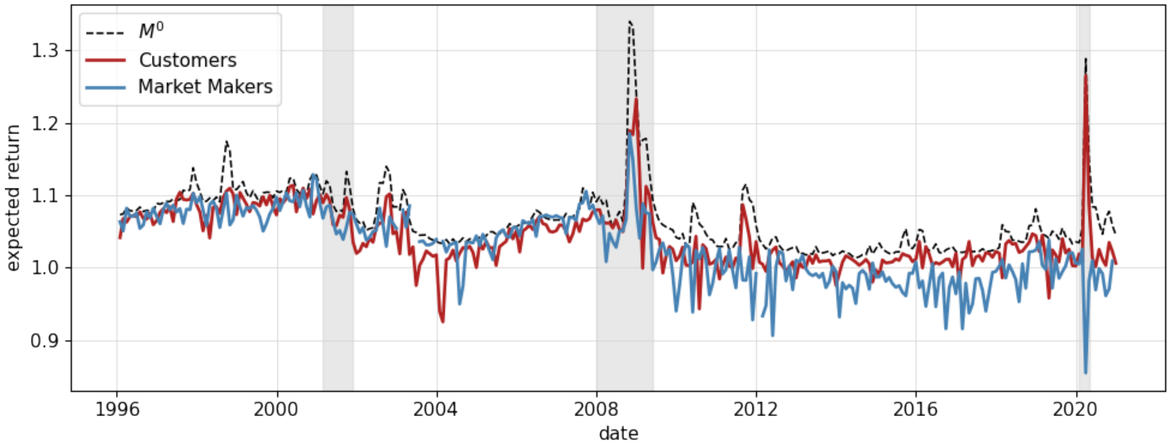



**Figure 5.** Lower Bound Subjective Expected Market Return
*Notes:* This figure plots the time-series of the expected market return implied by customers' and market markers' holdings. The black dashed line displays the monthly expected return recovered by $M^0 = 1/R$. Frequency and horizon are monthly, values are annualized. Data runs from January 1996 to December 2020. Gray bars indicate NBER recessions.

**Table 2.** Summary statistics of $\mathbb{E}_i[R]$.

| | mean | std | min | median | max | counts | AR(1) |
|---|---|---|---|---|---|---|---|
| Customers | 1.041 | 0.041 | 0.925 | 1.032 | 1.265 | 300 | 0.72 |
| Market Makers | 1.029 | 0.048 | 0.854 | 1.033 | 1.185 | 294 | 0.78 |
| Benchmark | 1.066 | 0.045 | 1.013 | 1.057 | 1.340 | 300 | 0.84 |

*Notes*: This table reports summary statistics for the monthly averaged time-series of annualized expected return recovered by the benchmark $M^0$, and by the observed positions of customers and market makers (as depicted in Figure 5). Data is monthly and runs from January 1996 to December 2020.

positions in OTM puts. Market makers take opposite positions, resulting in a long call/short put portfolio mimicking the underlying; but because of the negative delta-hedging, this reduces the exposition to the index. As a consequence, both investors' exposures to market risk are more mitigated, coherently with having more pessimistic belief (than the guy who is fully invested in the market) and formulating lower expectations on future market returns.

Notice also that the time-series recovered by $M^0$ is by construction counter-cyclical only. Although the investors' expected returns have high correlations with that, they do not perfectly move together. This suggests that the cyclicality of subjective expected returns is not perfectly aligned with $M^0$, being contingent on the state of the economy and depending on the option demand.

Subjective expected returns are less persistent than those from $M^0$, as the last column in Table 2 shows. Our measure of $\mathbb{E}_i[R]$ is noisy because it absorbs shocks both to option prices and to portfolio rebalancing (i.e. option demand). This second effect is absent in $\mathbb{E}_0[R]$.

## 4.4 Bounds on Subjective Expected Returns

We now turn to the case when portfolio holdings contain some measurement error. We solve the linear optimization problem in equation (6) to determine the most conservative and the maximum expected market return compatible with investors' observed positions. We set $\delta$ equal to half of the average bid-ask spread in the options cross-section at every date $t$. As before, we average daily data to get monthly averages.

Figures 6 and 7 plot the resulting time-series of subjective expected returns, in comparison with the benchmark case. Summary statistics are reported in Table 3. More specifically, the upper plots show the most conservative subjective expected return for each investor, while the lower plots show all the admissible values that lie between the minimum and the maximum. These act as lower and upper bounds for the subjective expected return perceived by all the possible investors whose portfolios are aligned (to some degree) with the observed positions of customers and market makers. Intuitively, the lower bound represents the expectation of the "most pessimistic" investor in the group - or, equivalently, it represents the "worst-case" expectation that customers and market makers may formulate.

Both lower bounds are mostly positive before 2008 and negative afterwards, when the option leverage gets higher and the optimized bounds become wider. Between 2008 and 2020, customers' lower bound is mostly pro-cyclical, with a huge negative peak during Covid (-24%). Conversely, market makers' lower bound is more aligned to Figure 5 in sign and cyclicality; it is just smaller in size, reaching a minimum of -39% in Covid (vs. -15% in the case with no measurement error).

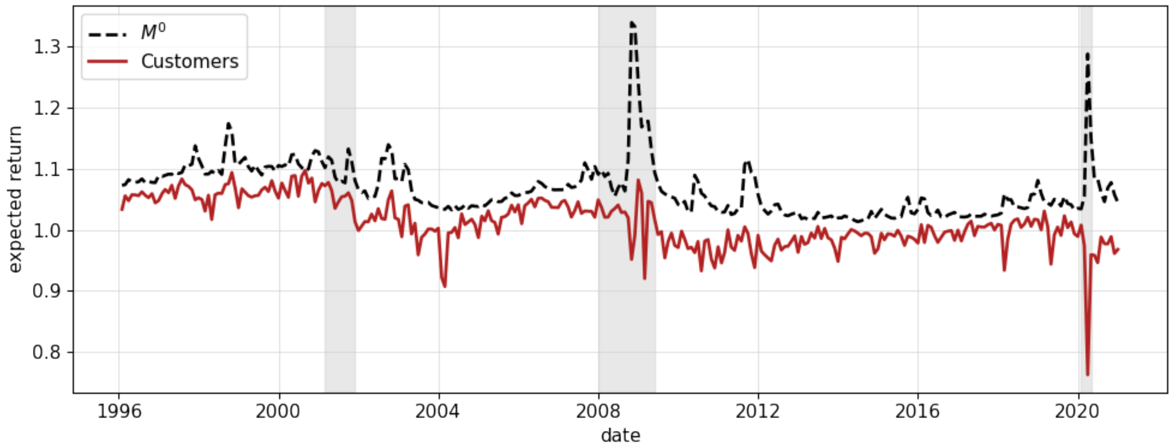
From our illustrative example discussed before, we expect that the lower bound is attained with a portfolio that hedges volatility risk with long positions in calls and puts. The upper bound, however, is most likely supported by portfolios which tend to have short positions in calls and puts. Therefore they will be more sensitive to periods of high volatility. Indeed the degree of belief heterogeneity in the market results larger during times of increasing volatility. This explains why the lower bound is largely pro-cyclical (after 2008, correlation with the $M^0$ time-series is -12% for customers and -31% for market makers) and the upper bound is mostly counter-cyclical (correlations 70% for customers and 83% for market makers).

By construction, optimized portfolio weights tend to assign more mass to the option components supporting the lowest and the largest expected returns on each date (compatible with the delta-hedging and the $\delta$-constraint). That's why the lower bounds are more correlated across investors than in the case without measurement errors (70% vs. 59%).

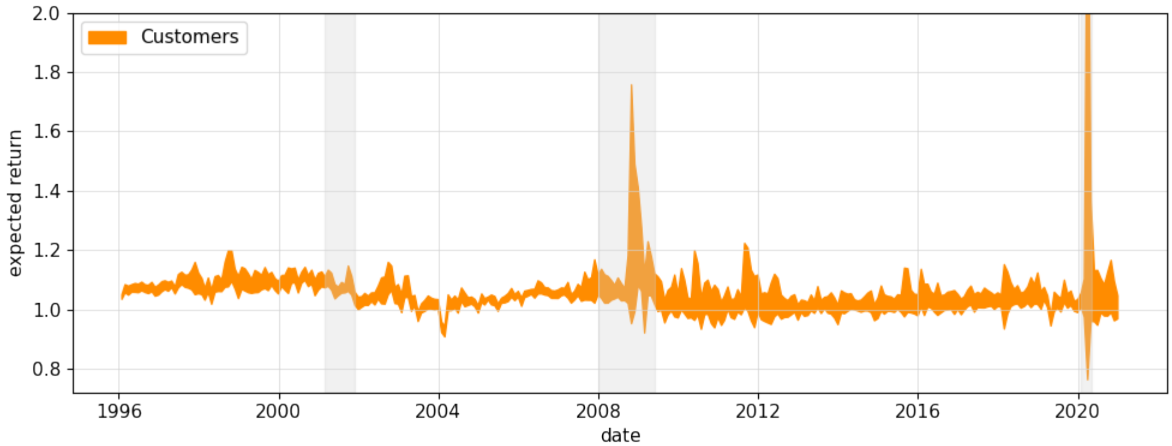
**Table 3.** Descriptive statistics of $\mathbb{E}_i[R]$ with measurement errors.

| | mean | std | min | median | max | counts |
|---|---|---|---|---|---|---|
| **Lower Bound** | | | | | | |
| Customers | 1.013 | 0.040 | 0.762 | 1.009 | 1.097 | 300 |
| Market Makers | 1.005 | 0.053 | 0.605 | 1.010 | 1.109 | 300 |
| **Upper Bound** | | | | | | |
| Customers | 1.091 | 0.117 | 0.943 | 1.074 | 2.735 | 300 |
| Market Makers | 1.079 | 0.076 | 0.950 | 1.073 | 1.919 | 300 |

*Notes*: This table reports summary statistics for the time-series of lower bound and upper bound on $\mathbb{E}_i[R]$, as recovered by solving the optimization problem on the customers' and market makers' positions. Data is monthly and runs from January 1996 to December 2020.

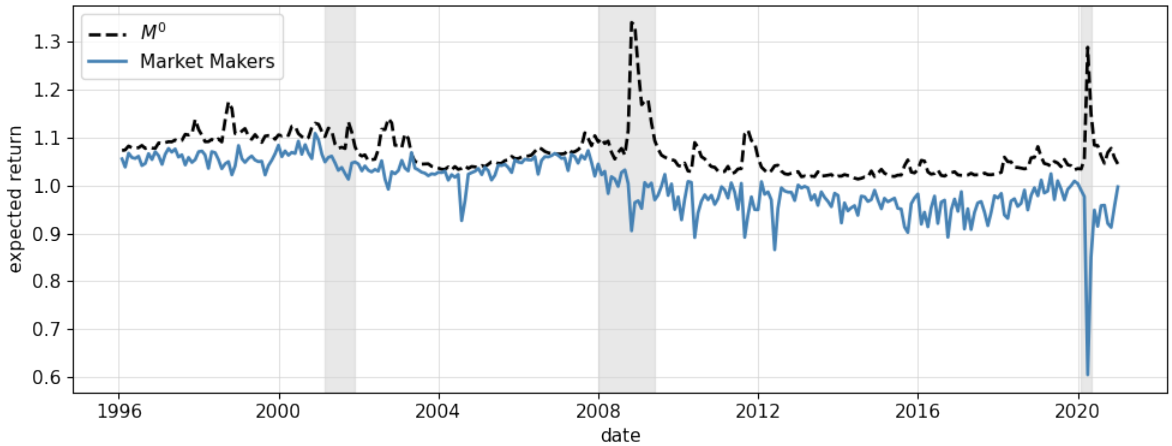



**A.** Lower Bound on Subjective Market Expected Return

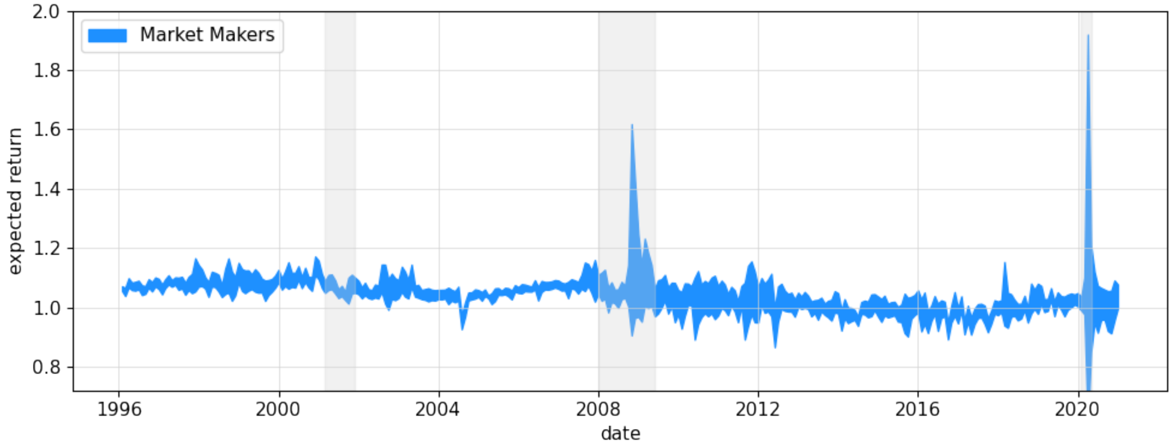



**B.** Admissible Values for Subjective Market Expected Return

**Figure 6.** Subjective Expected Market Return with measurement errors
*Notes:* This figure plots the time-series of the lower bound on the subjective expected market return (panel A) and all the possible values for the expected market return (panel B) as recovered by SDFs compatible with customers' positions in delta-hedged options. Frequency and horizon are monthly, values are annualized. Data runs from January 1996 to December 2020. Gray bars indicate NBER recessions.




**A.** Lower Bound on Subjective Market Expected Return




**B.** Admissible Values for Subjective Market Expected Return

**Figure 7.** Subjective Expected Market Return with measurement errors
*Notes:* This figure plots the time-series of the lower bound on the subjective expected market return (panel A) and all the possible values for the expected market return (panel B) as recovered by SDFs compatible with market makers' positions in delta-hedged options. Frequency and horizon are monthly, values are annualized. Data runs from January 1996 to December 2020. Gray bars indicate NBER recessions.

## 4.5 Survey Data and Time-Series Properties

We now want to investigate the relation of our expected return measures with survey measures and study their cyclicality properties in more detail. To make our measures comparable with survey measures, we construct excess returns and subtract the one-month Treasury Bill rate.[16] We use the following three surveys about expected market returns: The individual investor series by Nagel and Xu [2023], the Graham and Harvey CFO survey [Ben-David, Graham, and Harvey 2013], and the Livingston Survey available from the Federal Reserve Bank of Philadelphia. Nagel and Xu [2023] extend the UBS/Gallup survey backward and forward using other surveys such as the Conference Board survey, and the Michigan Survey of Consumers and is available at the quarterly frequency.[17] The Graham and Harvey survey polls financial officers about the one-year expected return on the S&P500 and is also available at the quarterly frequency. Finally, the Livingston Survey is released in June and December every year and polls economists at financial, non-financial, and academic institutions, as well as labor organizations, government, and insurance companies.

Figure 8 plots expected return measures for market makers and customers together with three survey measures. As with regards to the overall co-movement, we find that the correlation between market makers' expected (excess) return and the Livingston Survey is -22% (-8% for customers' expected return), -35% with the CFO survey (same for customers' expected returns), and -60% with the Nagel and Xu [2023] measure of expected returns (-40% for customers' expected return). Our findings echo Dahlquist and Ibert [2024] who study surveys from asset managers and find that their expectations behave quite different from those of retail investors. Similarly, market makers and their customers in the SPX market are highly sophisticated hedge funds and banks and their expectations about the market's return are highly negatively correlated.

We now turn to studying market makers' and customers' expected return determinants. A large literature has studied drivers of investors' expectations by estimating time-series regressions from expected returns on standard predictors of realized returns. Interestingly, most survey measures do not load significantly on standard predictors, see, e.g., Nagel and Xu [2023]. We follow this literature and regress the expected return measures on the S&P500 P/D ratio, the consumption wealth ratio (CAY) of Lettau and Ludvigson [2001], as well as net equity expansion (NTIS) from Welch and Goyal [2008], calculated as the ratio of twelve-month moving sums of net issues by NYSE-listed stocks divided by the total market capitalization of NYSE stocks at the end of the twelve-month window. In all of our regressions, we also include past realized returns as in Greenwood and Shleifer [2014]. Panel A of Table 4 reports the results.

[16] Notice that most surveys measure expected stock returns over a one-year horizon. We do not observe many options with maturity of one year. We therefore use a one-month horizon (as before).

[17] We download data from Zhengyang Xu's webpage.

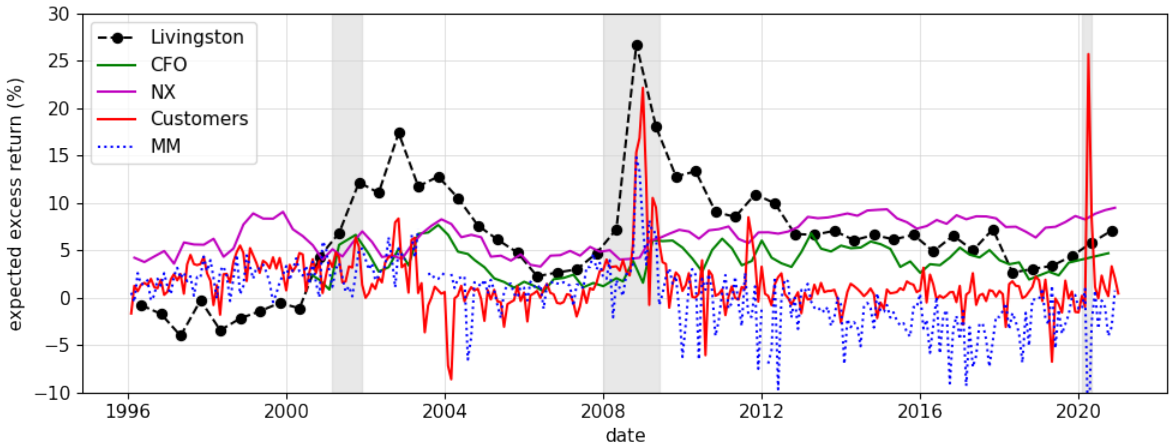



**Figure 8.** Subjective Expected Market Return from Survey Data, Customers and Market Makers
*Notes:* This figure plots the time-series of the expected market premium (p.a. in %) for market makers and customers together with survey data from Nagel and Xu [2023] (NX), the Graham and Harvey survey (CFO), and the Livingston Survey. Data runs from January 1996 to December 2020. Gray bars indicate NBER recessions.

None of the standard realized return predictor coefficients are statistically significant. This result is inline with Nagel and Xu [2023] who also find that subjective risk premia do not load on standard predictors. Past realized returns, however, load highly statistically significantly on subjective returns of market makers and customers with coefficients of similar size.

The negative slope coefficient is interesting, because most survey measures load positively on past realized returns. For example, Greenwood and Shleifer [2014] show that when past realized returns are high, investors expect higher returns going forward. The authors interpret this finding as evidence for extrapolation. Using subjective expected returns of market makers and customers, we find that they actually expect lower expected returns.

As discussed earlier, the cyclicality properties of expected returns of market makers' and customers may depend on the composition of their portfolios. To study this relation more formally, we run time-series regressions of expected returns on measures of cyclicality. To proxy for cyclicality, we take industrial production growth (IP), the 10-year minus 3-month Treasury term spread (TERM), the default spread defined as the difference between Moody's BAA and AAA corporate bond yields (DEFAULT), and the real factor of Ludvigson and Ng [2009] (F1). Moving to Panel B in Table 4, we find that with the exception of the term spread, none of the regression coefficients is statistically significant when controlling for past realized returns. We therefore conclude that the cyclicality properties of subjective expected returns of market makers and customers is muted.

**Table 4.** Determinants and Cyclicality Expected Returns

| | Panel A: Determinants | | | | | | Panel B: Cyclicality | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CAY | | DP | | NTIS | | IP | | TERM | | DEFAULT | | F1 | |
| **Expected Returns: Market Makers** | | | | | | | | | | | | | | |
| Const | 0.03 | 0.03 | 0.02 | 0.02 | 0.04 | 0.05 | 0.03 | 0.04 | 0.04 | 0.04 | 0.03 | 0.03 | 0.03 | 0.03 |
| ($p$-value) | (0.00) | (0.00) | (0.18) | (0.05) | (0.00) | (0.00) | (0.01) | (0.00) | (0.00) | (0.00) | (0.01) | (0.00) | (0.01) | (0.00) |
| Coeff | 1.77 | 1.76 | -2.05 | -2.51 | 1.51 | 2.26 | -1.63 | -3.87 | -1.44 | -1.52 | 0.34 | -1.22 | 0.48 | -0.32 |
| ($p$-value) | (0.11) | (0.12) | (0.08) | (0.01) | (0.30) | (0.07) | (0.37) | (0.02) | (0.09) | (0.05) | (0.81) | (0.39) | (0.49) | (0.59) |
| $R^e_{past}$ | | -1.15 | | -1.93 | | -1.90 | | -2.38 | | -1.32 | | -1.78 | | -1.39 |
| ($p$-value) | | (0.04) | | (0.04) | | (0.01) | | (0.00) | | (0.11) | | (0.00) | | (0.07) |
| Adj. $R^2$ | 0.21 | 0.25 | 0.16 | 0.26 | 0.05 | 0.15 | 0.02 | 0.14 | 0.08 | 0.13 | -0.00 | 0.05 | 0.00 | 0.04 |
| **Expected Returns: Customers** | | | | | | | | | | | | | | |
| Const | 0.04 | 0.05 | 0.04 | 0.04 | 0.04 | 0.06 | 0.04 | 0.05 | 0.05 | 0.06 | 0.04 | 0.05 | 0.04 | 0.05 |
| ($p$-value) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) | (0.00) |
| Coeff | 0.98 | 0.95 | -0.80 | -1.32 | 0.41 | 1.24 | -0.27 | -2.70 | -1.38 | -1.51 | 1.48 | -0.21 | 0.98 | -0.11 |
| ($p$-value) | (0.15) | (0.18) | (0.47) | (0.11) | (0.74) | (0.20) | (0.85) | (0.03) | (0.04) | (0.01) | (0.25) | (0.89) | (0.17) | (0.85) |
| $R^e_{past}$ | | -1.77 | | -2.19 | | -2.18 | | -2.61 | | -1.94 | | -1.91 | | -1.87 |
| ($p$-value) | | (0.00) | | (0.01) | | (0.00) | | (0.00) | | (0.01) | | (0.00) | | (0.00) |
| Adj. $R^2$ | 0.08 | 0.21 | 0.03 | 0.21 | 0.00 | 0.17 | -0.00 | 0.20 | 0.10 | 0.25 | 0.05 | 0.13 | 0.04 | 0.13 |

*Notes*: This table reports estimated coefficients from regressing market makers' and customers' expected returns on determinants and measures of cyclicality. Data runs from January 1996 to December 2020.

# 5 Subjective Measures of Risk

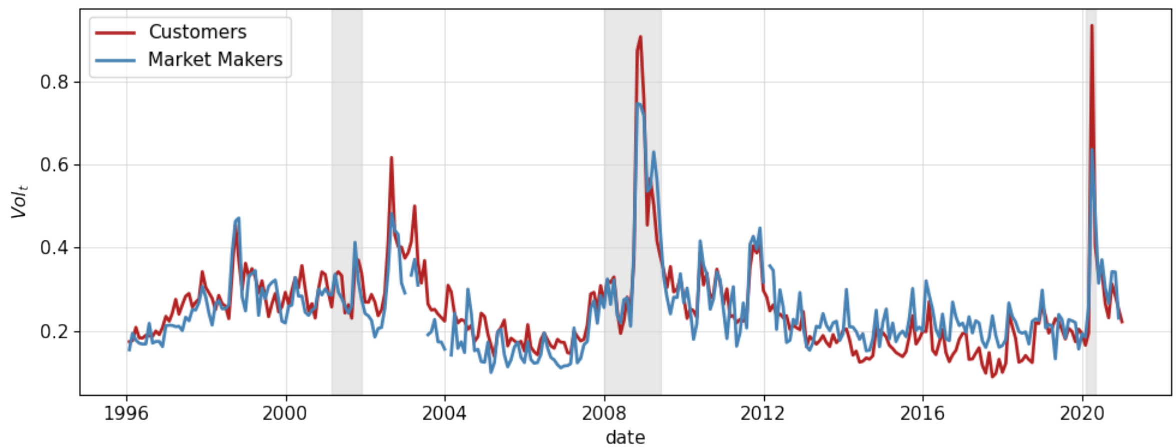
We can now study subjective risk as outlined in equation (9). Table 5 presents summary statistics for subjective risk and the subjective Sharpe ratio, that is the subjective expected excess return divided by the subjective volatility. The average volatility for both customers and market makers is around 25%. However, in terms of Sharpe ratios, customers exhibit slightly larger value (5.9% vs. 3.7%), because after 2008 the subjective return of market makers turns negative systematically. In general, both low Sharpe ratios reflect the investors' conservative view about the market.

The time-series of subjective volatility and Sharpe ratios for customers and market makers are plotted in Figure 9. The subjective risks of both customers and market makers are highly similar and clearly countercyclical, high in bad times and low in good times. Significant differences arise only in the huge peaks during crisis, where volatility perceived by customers gets much higher than the other.

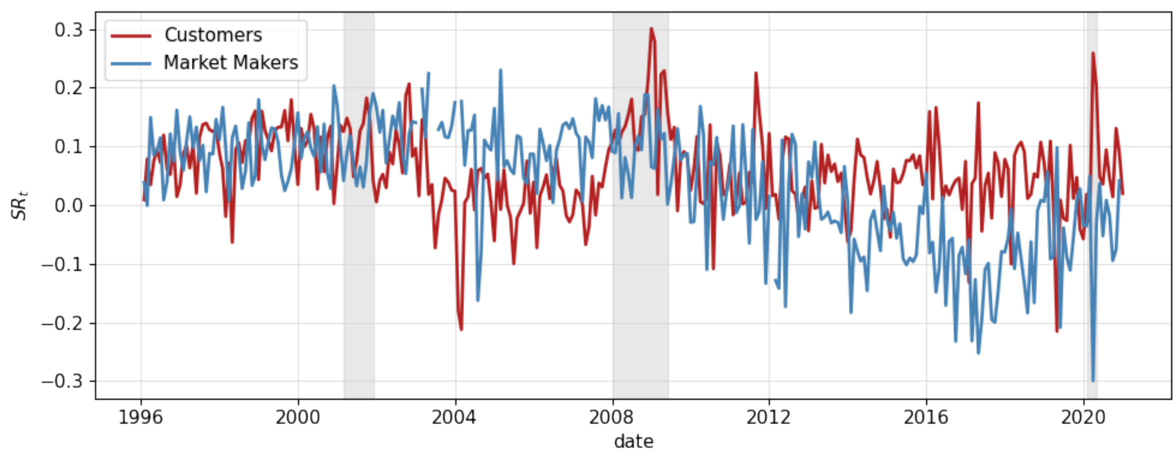
**Table 5.** Summary statistics of subjective volatility and Sharpe Ratio

| | mean | std. | min | median | max | counts |
|---|---|---|---|---|---|---|
| **Volatility** | | | | | | |
| Customers | 0.251 | 0.110 | 0.089 | 0.232 | 0.935 | 300 |
| Market Makers | 0.248 | 0.100 | 0.099 | 0.226 | 0.747 | 294 |
| **Sharpe Ratio** | | | | | | |
| Customers | 0.059 | 0.071 | -0.215 | 0.054 | 0.301 | 300 |
| Market Makers | 0.037 | 0.100 | -0.300 | 0.058 | 0.230 | 294 |

*Notes*: This table reports summary statistics for the time-series of monthly averaged subjective volatility and Sharpe Ratio recovered from the observed positions of customers and market makers (showed in Figure 9). Values are annualized. Data runs from January 1996 to December 2020.




**A.** Subjective Volatility




**B.** Subjective Sharpe Ratio

**Figure 9.** Subjective Volatility and Sharpe Ratio with no measurement errors
*Notes:* This figure plots the time-series of the subjective volatility (panel A) and the subjective Sharpe Ratio (panel B) as recovered through the SDF supported by the observed positions of customers and market makers, as described in the main text. Frequency and horizon are monthly, values are annualized. Gray bars indicate NBER recessions.

# 6 Demand-Based SDFs

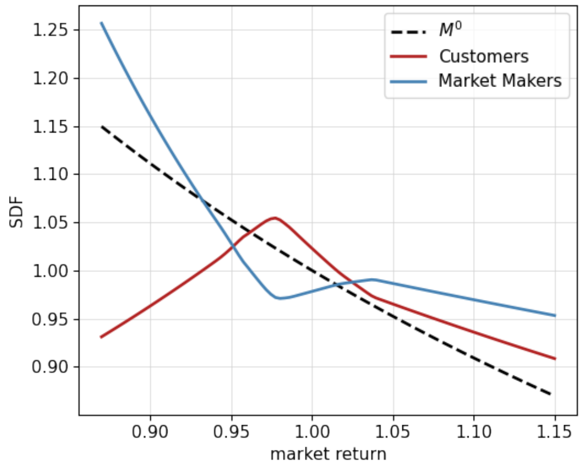
As our final exercise, we now recover the SDFs of our two investors groups and we show how they depend on their portfolio compositions. To this end, we show the SDF shapes and the customers' portfolio weights[18] on options in two complementary situations, summarized in Figure 10 and Figure 11.

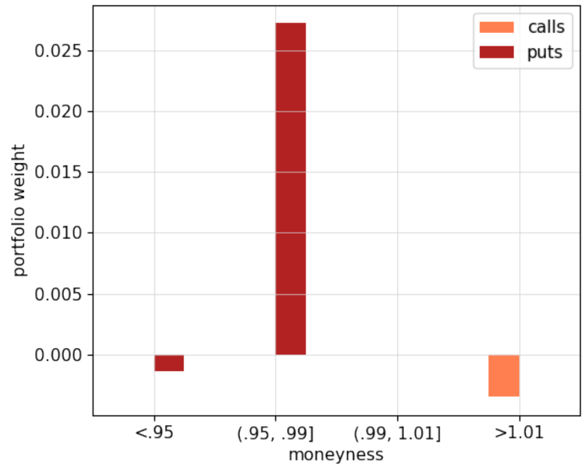
On both dates, we find that the investment in puts is larger than in calls, but the type of investment is opposite. Figure 10 depicts a day in October 2012 where customers are long in puts and short in calls. We already know that on average this is the typical customers' portfolio of monthly OTM options. Then, their usual SDF is not monotonically decreasing (as $M^0$). In the downside region, it is increasing because the risk has been hedged by buying protection through the OTM puts. The market makers' SDF has a complementary form, being almost $U$-shaped as they have taken larger exposition to the index volatility.

Despite this is the most frequent observation in our data, many interesting deviations are allowed depending on the contingent portfolio structure. Figure 11 is a prime example of the case where customers have taken overall negative positions in both calls and puts. The picture refers to March 10, 2020 during Covid crisis. This echoes earlier findings in Chen, Joslin, and Ni [2019] who argue that while market makers are net suppliers of insurance in normal times, they become net demanders in bad times when their financial constraints bind. Although this effect is not evident on average on a monthly scale, it can be registered when looking at the daily level. The regime is switched from Figure 10: customers' SDF is now increasing in the upside and the downside risk, exhibiting a clear $U$-shaped form. Market makers' SDF is conversely peaked in the middle, and decreasing towards extreme values in both directions.

The different compositions in the portfolios have significant effects on the ensuing expected returns. When investors increase their exposure to the index volatility, they perceive larger risk premia: for instance, for customers it is 1% in October 22, 2012 and 21% in March 10, 2020.

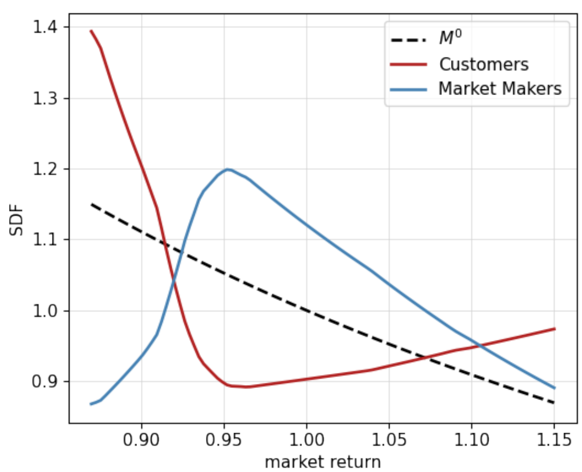
[18] Since options are in zero net supply, market makers' aggregated holdings are just the mirror image.

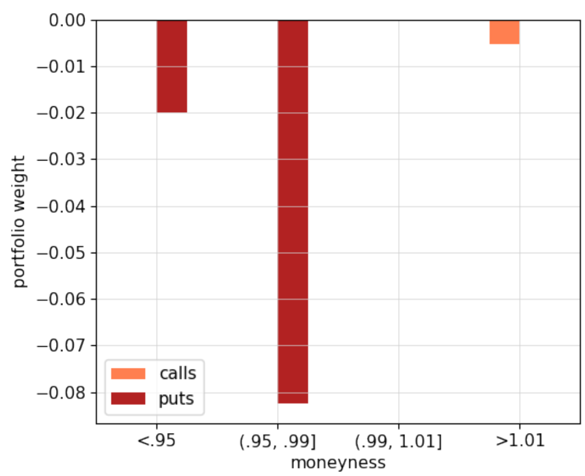



**A.** SDF **B.** Customers' Option Investment

**Figure 10.** Recovered SDF and Option Portfolio Weights
*Notes:* This Figure plots the stochastic discount factor recovered from customers and market makers' observed positions in 30-days expiring options on October 22, 2012 (left panel), and the corresponding distribution of the portfolio weights for OTM calls and puts across different levels of moneyness (right panel). Portfolio weights are summed in every bin.




**A.** SDF **B.** Customers' Option Investment

**Figure 11.** Recovered SDF and Option Portfolio Weights
*Notes:* This Figure plots the stochastic discount factor recovered from customers and market makers' observed positions in 30-days expiring options on March 10, 2020 (left panel), and the corresponding distribution of the portfolio weights for OTM calls and puts across different levels of moneyness (right panel). Portfolio weights are summed in every bin.

# 7 Conclusion

A large literature has documented different time-series properties of survey measures of expected returns depending on the level of sophistication. Our findings show that even across investors with very high levels of sophistication such as hedge funds (who act both as market makers and customers), expected returns can differ wildly. Measures of expected returns that uniquely rely on pricing information are ill-suited to explain such large heterogeneity.

In this paper, we propose a theoretical framework for recovering investors' beliefs using demand-based data. Information about investors' holdings allows us to recover possible beliefs of individual investors when observing a cross-section of option prices. Our main empirical result is that the size, dynamics, and cyclicality properties of belief-implied expected returns and subjective Sharpe ratios vary significantly across investor types. Using granular transaction data on buy and sell orders of financial intermediaries and public investors, we show that beliefs are heterogeneous and the implied expected returns may vary considerably across the two investors as they depend on the structure of the underlying portfolio and on the state of the economy.

While market makers' expected returns are highly correlated with survey measures of sophisticated agents, they are not correlated with households' expectations. The opposite holds true for customers. Finally, we find that the cyclicality properties of subjective expected returns are muted.

Our setting abstracts from frictions such as trading costs. The sensitivity of portfolio holdings to expectations depends, however, on such costs. For example, Giglio et al. [2021] show that such sensitivity increases as investors face lower costs. In future work, we plan to jointly model survey data, holdings, and prices in the presence of frictions.

# A Proofs and Derivations

**Proposition 2** (Upper and lower bounds on expected payoffs of a log investor)**.** *Suppose $\boldsymbol{\theta} \in \Theta$, where $\Theta$ is some closed convex set, indexes a log investor holding an optimal portfolio $\boldsymbol{\theta}$, with return $R$, and having belief $\mathbb{P}$. Further let $f(R)$ be some payoff depending on $R$. Then, the following upper and lower bounds hold:*

$$\mathcal{U}(f) := \sup_{\boldsymbol{\theta} \in \Theta} \mathbb{E}^{\mathbb{Q}}[Rf(R)] \geq \mathbb{E}_i[f(R)] \,, \tag{10}$$

*and*

$$\mathcal{L}(f) := \inf_{\boldsymbol{\theta} \in \Theta} \mathbb{E}^{\mathbb{Q}}[Rf(R)] \leq \mathbb{E}_i[f(R)] \,. \tag{11}$$

*Proof.* Since portfolio $\boldsymbol{\theta}$ is optimal for the log utility investor and the associated belief $\mathbb{P}$, we get:

$$\mathbb{E}_i[f(R)] = \mathbb{E}^{\mathbb{Q}}[Rf(R)] \,, \tag{12}$$

where by construction $\mathbb{E}^{\mathbb{Q}}[R] = 1$. Given convex set $\Theta$ of admissible portfolio weights, the worst case expected payoff across admissible maximum growth portfolios is:

$$\mathcal{L}(f) := \inf_{\boldsymbol{\theta} \in \Theta} \mathbb{E}_i[f(R)] = \inf_{\boldsymbol{\theta} \in \Theta} \mathbb{E}^{\mathbb{Q}}[Rf(R)] \,. \tag{13}$$

Analogously, the best case expected payoff across admissible maximum growth portfolios is:

$$\mathcal{U}(f) := \sup_{\boldsymbol{\theta} \in \Theta} \mathbb{E}_i[f(R)] = \sup_{\boldsymbol{\theta} \in \Theta} \mathbb{E}^{\mathbb{Q}}[Rf(R)] \,. \tag{14}$$

This concludes the proof. ■

**Corollary 1** (Upper and lower bounds on expected payoffs from observed investor's holding)**.** *In the context of Proposition 2, suppose that $\boldsymbol{\theta}_0^\star$ is the optimal portfolio of a log investor with belief $\mathbb{P}$, which is however not observable. Assume further that there exists an observable portfolio $\boldsymbol{\theta}_0$ such that*

$$d(\boldsymbol{\theta}_0, \boldsymbol{\theta}_0^\star) \leq \delta \,, \tag{15}$$

*for some convex discrepancy function $d(\cdot,\cdot) \geq 0$ such that $d(x,y) = 0$ if and only if $x = y$. Then, the upper and lower bounds in Proposition 2 are such that:*

$$\mathcal{L}(f) = \inf_{d(\boldsymbol{\theta},\boldsymbol{\theta}_0)\leq\delta} \mathbb{E}^{\mathbb{Q}}[Rf(R) \leq E_i[Rf(R)] \leq \mathcal{U}(f) = \sup_{d(\boldsymbol{\theta},\boldsymbol{\theta}_0)\leq\delta} \mathbb{E}^{\mathbb{Q}}[R(\boldsymbol{\theta})f(R)] \,. \tag{16}$$

*In the case where $\delta = 0$, i.e., there is no portfolio measurement error, then*

$$\mathcal{L}(f) = \mathbb{E}_i[f(R)] = \mathcal{U}(f) \,.$$

**Example 1.** If $d(\boldsymbol{\theta},\boldsymbol{\theta}_0) = \frac{1}{2}||\boldsymbol{\theta}-\boldsymbol{\theta}_0||_2^2$, then:

$$g_{\mathcal{L}(f)}(\lambda) = \inf_{\boldsymbol{\theta}\in\mathbb{R}^N} \left\{ \mathbb{E}^{\mathbb{Q}}[R(\boldsymbol{\theta})f(R)] + \lambda\left(\frac{1}{2}||\boldsymbol{\theta}-\boldsymbol{\theta}_0||_2^2 - \delta\right)\right\} \,. \tag{17}$$

This gives the optimality condition:

$$\mathbf{0} = \nabla g_{\mathcal{L}(f)}(\lambda) = \mathbb{E}^{\mathbb{Q}}[\boldsymbol{R}^e f(R)] + \lambda(\boldsymbol{\theta}-\boldsymbol{\theta}_0) \,, \tag{18}$$

and, whenever the constraint is binding:

$$\frac{1}{2}||\mathbb{E}^{\mathbb{Q}}[\boldsymbol{R}^e f(R)]||_2^2 = \frac{1}{2}\lambda^2||\boldsymbol{\theta}-\boldsymbol{\theta}_0||_2^2 = \lambda^2\delta \,, \tag{19}$$

i.e., an optimal Lagrange multiplier given by:

$$\lambda^\star = \frac{1}{\sqrt{2\delta}}||\mathbb{E}^{\mathbb{Q}}[\boldsymbol{R}^e f(R)]||_2 \,. \tag{20}$$

Therefore, the optimal portfolio supporting the lower bound is such that:

$$\boldsymbol{\theta}^\star = \boldsymbol{\theta}_0 - \frac{1}{\lambda^\star}\mathbb{E}^{\mathbb{Q}}[\boldsymbol{R}^e f(R)] = \boldsymbol{\theta}_0 - \frac{\sqrt{2\delta}}{||\mathbb{E}^{\mathbb{Q}}[\boldsymbol{R}^e f(R)]||_2}\mathbb{E}^{\mathbb{Q}}[\boldsymbol{R}^e f(R)] \,. \tag{21}$$

This gives the closed-form lower bound:

$$\mathcal{L}(f) = g_{\mathcal{L}(f)}(\lambda^\star) = \mathbb{E}^{\mathbb{Q}}[R(\boldsymbol{\theta}^\star)f(R)] = \mathbb{E}^{\mathbb{Q}}[f(R)] + \mathbb{E}^{\mathbb{Q}}[\boldsymbol{\theta}_0'\boldsymbol{R}^e f(R)] - \sqrt{2\delta}||\mathbb{E}^{\mathbb{Q}}[\boldsymbol{R}^e f(R)]||_2 \,. \tag{22}$$

In an analogous vein, we obtain:

$$\mathcal{U}(f) = g_{\mathcal{U}(f)}(\lambda^\star) = \mathbb{E}^{\mathbb{Q}}[f(R)] + \mathbb{E}^{\mathbb{Q}}[\boldsymbol{\theta}_0'\boldsymbol{R}^e f(R)] + \sqrt{2\delta}||\mathbb{E}^{\mathbb{Q}}[\boldsymbol{R}^e f(R)]||_2 \ . \tag{23}$$

**Proposition 3** (Lower bound on expected log return of optimally invested wealth)**.** *Suppose $\boldsymbol{\theta} \in \Theta$ indexes a log investor holding an optimal portfolio $\boldsymbol{\theta}$, with return $R$, and having belief $\mathbb{P}$. Then, the following lower bound holds:*

$$\mathcal{L} := \inf_{\boldsymbol{\theta}\in\Theta} \ \mathbb{E}^{\mathbb{Q}}[R\log R] \leq \mathbb{E}_i[\log R] \ . \tag{24}$$

*Proof.* Since portfolio $\boldsymbol{\theta}$ is optimal for the log utility investor and the associated belief $\mathbb{P}$, we get:

$$\mathbb{E}_i[\log R] = \mathbb{E}^{\mathbb{Q}}[R\log R] \ , \tag{25}$$

where by construction $\mathbb{E}^{\mathbb{Q}}[R] = 1$. Given convex set $\Theta$ of admissible portfolio weights, the worst case expected log utility over maximum growth portfolios is:

$$\mathcal{L} := \inf_{\boldsymbol{\theta}\in\Theta} \ \mathbb{E}_i[\log R(\boldsymbol{\theta})] = \inf_{\boldsymbol{\theta}\in\Theta} \ \mathbb{E}^{\mathbb{Q}}[R(\boldsymbol{\theta})\log R(\boldsymbol{\theta})] \ . \tag{26}$$

This problem is convex, with solution obtained using standard duality methods. This concludes the proof. ■

**Corollary 2** (Lower bound extracted from observed investor's holdings)**.** *In the context of Proposition 3, suppose that $\boldsymbol{\theta}_0^\star$ is the optimal portfolio of a log investor with belief $\mathbb{P}$, which is however not observable. Assume further that there exists and observable portfolio $\boldsymbol{\theta}_0$ such that*

$$d(\boldsymbol{\theta}_0, \boldsymbol{\theta}_0^\star) \leq \delta \ , \tag{27}$$

*for some convex discrepancy function $d(\cdot,\cdot) \geq 0$ such that $d(x,y) = 0$ if and only if $x = y$. Then, the lower bound in Proposition 3 becomes:*

$$\mathcal{L} = \inf_{\boldsymbol{\theta}} \ \mathbb{E}^{\mathbb{Q}}[R\log R(\boldsymbol{\theta})] \quad \textit{s.t.} \quad d(\boldsymbol{\theta}, \boldsymbol{\theta}_0) \leq \delta \ . \tag{28}$$

*In the case where $\delta = 0$, i.e., there is no portfolio measurement error, then*

$$\mathcal{L} = \mathbb{E}_i[\log R(\boldsymbol{\theta}_0^\star)] \ .$$

*Proof.* The lower bound follows from Proposition 3 once we define $\Theta := \{\boldsymbol{\theta} \mid d(\boldsymbol{\theta}, \boldsymbol{\theta}_0) \leq \delta\}$. In the case where there is no measurement error, $\delta = 0$ and $\Theta = \{\boldsymbol{\theta}_0^\star\}$, i.e.:

$$\mathcal{L} = \mathbb{E}^{\mathbb{Q}}[R \log R] = \mathbb{E}_i[\log R] \, . \tag{29}$$

This concludes the proof. ■

**Corollary 3** (Dual formulation)**.** *In the context of Proposition 3, for any $\lambda \geq 0$ it follows:*

$$\mathcal{L} \geq g(\lambda) := \inf_{\boldsymbol{\theta} \in \mathbb{R}^N} \left\{ \mathbb{E}^{\mathbb{Q}}[R(\boldsymbol{\theta}) \log R(\boldsymbol{\theta})] + \lambda(d(\boldsymbol{\theta}, \boldsymbol{\theta}_0) - \delta) \right\} \, . \tag{30}$$

*Therefore, $\mathcal{L} \geq \sup_{\lambda \geq 0} g(\lambda)$. Moreover, when suitable Constraints Qualification conditions hold, then $\mathcal{L} = \sup_{\lambda \geq 0} g(\lambda)$. In particular, if there exists $0 < \delta' < \delta$ such that $\mathbb{E}^{\mathbb{Q}}[R(\boldsymbol{\theta}) \log R(\boldsymbol{\theta})] < \infty$ for all $\boldsymbol{\theta}$ such that $d(\boldsymbol{\theta}, \boldsymbol{\theta}_0) < \delta'$ then Slater's Constraint Qualification conditions hold.*

*Proof.* The proof follows with standard Lagrangian duality arguments. ■

**Example 2.** If $d(\boldsymbol{\theta}, \boldsymbol{\theta}_0) = \frac{1}{2}||\boldsymbol{\theta} - \boldsymbol{\theta}_0||_2^2$, then:

$$g(\lambda) = \inf_{\boldsymbol{\theta} \in \mathbb{R}^N} \left\{ \mathbb{E}^{\mathbb{Q}}[R(\boldsymbol{\theta}) \log R(\boldsymbol{\theta})] + \lambda \left( \frac{1}{2}||\boldsymbol{\theta} - \boldsymbol{\theta}_0||_2^2 \right) \right\} \, . \tag{31}$$

# B Demand-based expected returns in economies with heterogenous beliefs

We explore expected returns and their relation to option-implied moments in a simple economy with $n$ investors, who may allocate their wealth to the aggregate market index (assumed in positive net supply) and a set of options (assumed in zero net supply).

We denote by $R_i$ the (forward) return on $i$-th investor's wealth over horizon $[0, T]$, and by $R$ the (forward) return on the market index. Investor's $i$ wealth at time $t$ is denoted by $W_{it}$. Given the market index value $I_t$ at time $t$, the market clearing condition yields:

$$I_t = W_t = \sum_i W_{it} , \tag{32}$$

and

$$R = \frac{I_T}{I_0} = \frac{W_T}{W_0} = \sum_i \frac{W_{i0}}{W_0} R_i =: \sum_i w_i R_i , \tag{33}$$

with $i-$th investor's share $w_i$ of the aggregate wealth. Let further $\mathbb{P}_i$ denote $i-$th investor's subjective probability belief and $\mathbb{E}_i[\cdot]$ expectations under this belief. The consensus market expected return among all investors is then defined by:

$$\bar{\mathbb{E}}[R] := \sum_i w_i \mathbb{E}_i[R] . \tag{34}$$

The next proposition gives a first characterization of the consensus belief in an economy where all investors are optimally investing in the index and option markets.

**Proposition 4.** *Let investor $j$ be a log investor optimally investing in the index market and the (complete) option market. It then follows:*

$$\mathbb{E}_j[R] = 1 + \mathbb{C}ov(R_j, R) = 1 + w_j \mathbb{V}ar^{\mathbb{Q}}[R_j] + \sum_{i \neq j} w_i \mathbb{C}ov^{\mathbb{Q}}[R_j, R_i] . \tag{35}$$

*If all investors are log investors optimally investing in the index market and the (complete) option market, then:*

$$\bar{\mathbb{E}}[R] = 1 + \mathbb{V}ar^{\mathbb{Q}}(R) \tag{36}$$

*If all index weights $w_i$ are positive, there exists a consensus belief $\bar{\mathbb{P}} := \sum_i w_i \mathbb{P}_i$ such that:*

$$\bar{\mathbb{E}}[R] = \mathbb{E}^{\bar{\mathbb{P}}}[R] \,. \tag{37}$$

*Proof.* We first obtain, using investor $j$ optimality conditions for investment in the index and the (complete) option markets:

$$\mathbb{E}_j[R] = \mathbb{E}^{\mathbb{Q}}[R_j R] = \mathbb{E}^{\mathbb{Q}}[R_j \sum_i w_i R_i] = \sum_i w_i \mathbb{E}^{\mathbb{Q}}[R_j R_i] \,.$$

The market clearing condition in the option market further implies $\sum_i w_i = 1$, which gives:

$$\mathbb{E}_j[R-1] = \sum_i w_i \mathbb{E}^{\mathbb{Q}}[R_j R_i - 1] = \sum_i w_i \mathbb{C}ov^{\mathbb{Q}}(R_j, R_i) = 1 + \mathbb{C}ov(R_j, R) \,.$$

Furthermore,

$$\bar{\mathbb{E}}[R] - 1 = \sum_j w_j \mathbb{E}_j[R-1] = \sum_j w_j \mathbb{C}ov^{\mathbb{Q}}(R_j, R) = \mathbb{V}ar^{\mathbb{Q}}\left(\sum_i w_i R_i\right) = \mathbb{V}ar^{\mathbb{Q}}(R) \,.$$

Finally, if all index weights are positive, then:

$$\bar{\mathbb{E}}[R] = \sum_i w_i \mathbb{E}_i[R] = \mathbb{E}^{\bar{\mathbb{P}}}[R] \,.$$

This concludes the proof. ■

One obvious special case of Proposition 4 arises when all investors have identical beliefs, in which case $\bar{\mathbb{P}}$ is the common belief of each investor in the economy. A key requirement for the validity of formula (36) in Proposition 4 – which is based on a consensus belief aggregated across all investors – is that all investors in the economy are unconstrained in their investment to the index and the option market. This ensures that all their associated returns on optimally invested wealth are the inverse of a stochastic discount factor for jointly pricing the index and all options. Therefore, whenever an investor exists who (i) invests in the market and (ii) is constrained in her investment in some options, a consensus belief cannot be identified via formula (36). This situation trivially arises, e.g., when an investor exists who can only invest in the index.

While formula (36) can be expected to fail in general, it is still possible to identify a consensus belief aggregated only among all investors who can optimally invest in both the index and the option market. To this end, let $\mathcal{I}$ be the index set indexing the subset of such investors having a non zero optimal wealth allocation to the index and option markets. For each investor $i \in \mathcal{I}$ it then follows:

$$\mathbb{E}_i[R] = 1 + \mathbb{C}ov(R_i, R) \; . \tag{38}$$

The associated consensus belief across such investors is analogously defined as:

$$\bar{\mathbb{E}}_{\mathcal{J}}[R] := \sum_{i \in \mathcal{I}} w_i^{\mathcal{I}} \mathbb{E}_i[R] \; , \tag{39}$$

with the wealth shares

$$w_i^{\mathcal{I}} = \frac{W_{i0}}{\sum_{i \in \mathcal{I}} W_{i0}} \tag{40}$$

The following corollary then yields the corresponding consensus belief characterization.

**Corollary 4.** *In the setting of Proposition 4, if there is a subset $\mathcal{I}$ of investors optimally investing in the index market and the (complete) option market, then:*

$$\bar{\mathbb{E}}_{\mathcal{I}}[R] = 1 + \mathbb{C}ov^{\mathbb{Q}}(R_{\mathcal{I}}, R) \; , \tag{41}$$

*where*

$$R_{\mathcal{I}} := \sum_{i \in \mathcal{I}} w_i^{\mathcal{I}} R_i \; . \tag{42}$$

*If all weights $w_i^{\mathcal{I}}$ are positive, then:*

$$\bar{\mathbb{E}}_{\mathcal{I}}[R] = \mathbb{E}^{\bar{\mathbb{P}}_{\mathcal{J}}}[R] \; , \tag{43}$$

*for a consensus belief $\bar{\mathbb{P}}_{\mathcal{J}} := \sum_{i \in \mathcal{I}} \mathbb{P}_i$.*

It follows form Corollary 4 that the consensus equity premium among investors optimally trading in the index and option markets is positive if and only if return $R_{\mathcal{J}}$ correlates positively with the market index return under pricing measure $\mathbb{Q}$. This feature is more likely to emerge, e.g., when the average index holding across these investors is positive. For instance, whenever the net aggregate wealth allocated to options by investors $i \in \mathcal{I}$ is nil, then:[19]

$$\bar{E}_{\mathcal{I}}[R] = \bar{E}[R] = 1 + \mathbb{V}ar^{\mathbb{Q}}(R) > 1 \ .$$

Conversely, when these investors additionally have on aggregate a zero exposure to market risk, then $R_{\mathcal{I}} = 0$ and $\bar{E}_{\mathcal{I}}[R] = 1$.

Given consensus belief $\bar{E}_{\mathcal{I}}[R]$, the belief dispersion about index returns of investors optimally investing in the index and options markets is conveniently defined by:

$$\mathbb{D}_{\mathcal{I}}[R] := \sum_{i \in \mathcal{I}} w_i |\mathbb{E}_i[R] - \bar{\mathbb{E}}_{\mathcal{I}}[R]| \ . \tag{44}$$

The next corollary gives the corresponding characterization of the belief dispersion.

**Corollary 5.** *In the setting of Proposition 4, if there is a subset $\mathcal{I}$ of investors optimally investing in the index market and the (complete) option market, then:*

$$\bar{\mathbb{D}}_{\mathcal{I}}[R] = \sum_{i \in \mathcal{I}} w_i |\mathbb{C}ov^{\mathbb{Q}}(R_i - R_{\mathcal{I}}, R)| \ . \tag{45}$$

*Proof.* Given equation (38), we have:

$$\begin{aligned} \mathbb{E}_i[R] &= 1 + \mathbb{C}ov^{\mathbb{Q}}(R_i, R) \ , \\ \bar{\mathbb{E}}_{\mathcal{I}}[R] &= 1 + \mathbb{C}ov^{\mathbb{Q}}(R_{\mathcal{I}}, R) \ . \end{aligned}$$

Therefore,

$$\mathbb{D}_{\mathcal{I}}[R] = \sum_{i \in \mathcal{I}} w_i |\mathbb{C}ov^{\mathbb{Q}}(R_i, R) - \mathbb{C}ov^{\mathbb{Q}}(R_{\mathcal{J}}, R)| = \sum_{i \in \mathcal{I}} w_i |\mathbb{C}ov^{\mathbb{Q}}(R_i - R_{\mathcal{J}}, R)| \ .$$

This concludes the proof. ■

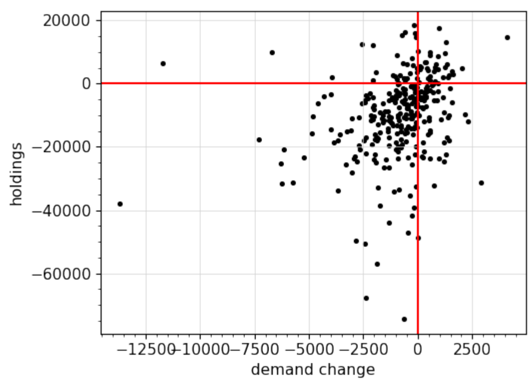
According to Corollary 5, belief dispersion is large when on average individual excess return $R_i - R_{\mathcal{I}}$ has a large forward-neutral covariance with index returns. Therefore, the belief dispersion increases, ceteris paribus, with the forward variance of the index, the average forward variance of excess return $R_i - R_{\mathcal{I}}$, and the average absolute forward correlation between $R_i - R_{\mathcal{I}}$ and index returns. In particular, following scaled definition of belief dispersion

[19] Because of the option market clearing condition, this identity holds, e.g., when set $\mathcal{J}$ contains all option investors in the conomy.

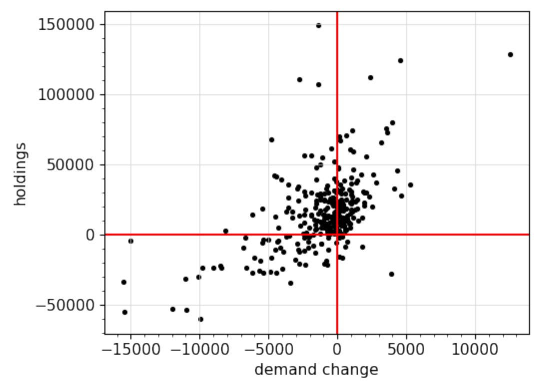
isolates the belief heterogeneity implications of the latter two effects from the mechanical effect of the index forward variance:

$$\mathbb{S}\bar{\mathbb{D}}_{\mathfrak{I}}[R] := \frac{\bar{\mathbb{D}}_{\mathfrak{I}}[R]}{\sqrt{\mathbb{V}ar^{\mathbb{Q}}(R)}} = \sum_{i\in\mathfrak{I}} w_i |\mathbb{C}orr^{\mathbb{Q}}(R_i - R_{\mathfrak{I}}, R)| \sqrt{\mathbb{V}ar^{\mathbb{Q}}(R_i - R_{\mathfrak{I}})}\,. \tag{46}$$

# C Additional Figures




**A.** Calls




**B.** Puts

**Figure C.1.** Customers' demand changes vs. holdings on monthly OTM options
*Notes:* This figure plots the holdings and the demand shocks in the customers' portfolios for OTM calls (left panel) and puts (right panel) expiring in 30 days. Data are computed daily from 1996 to 2020, then aggregated to monthly frequency.

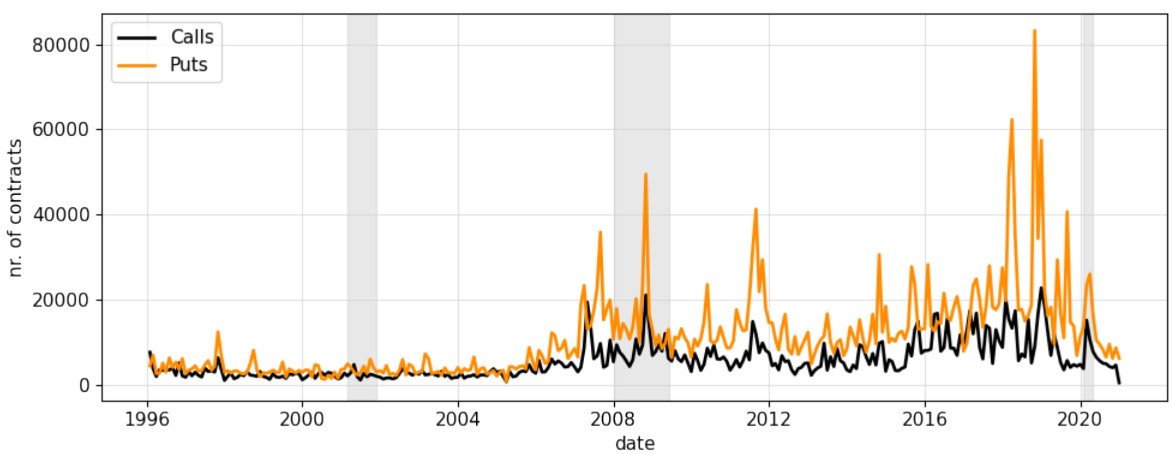



**Figure C.2.** Trading Volume in OTM options
*Notes:* This figure plots the time-series of the trading volume in OTM calls and puts, expiring in 30 days, from 1996 to 2020. Trading volume is the sum of the absolute value of opening and closing positions recorded every day. Data are monthly averages. Gray bars indicate NBER recessions.

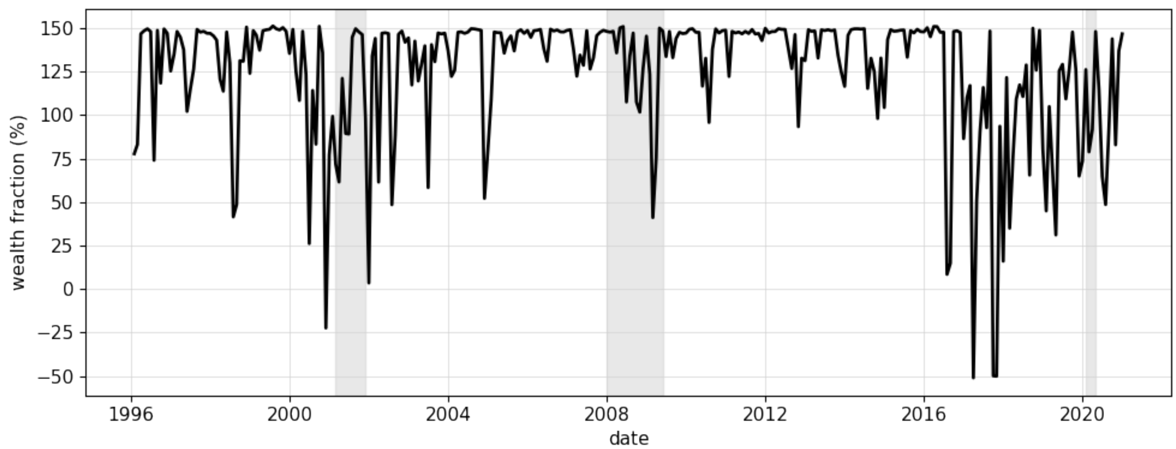



**Figure C.3.** Investment in the Index
*Notes:* This figure plots the time-series of the wealth fraction (in percentage) invested in the S&P500 index by customers. Data are monthly averages. Gray bars indicate NBER recessions.

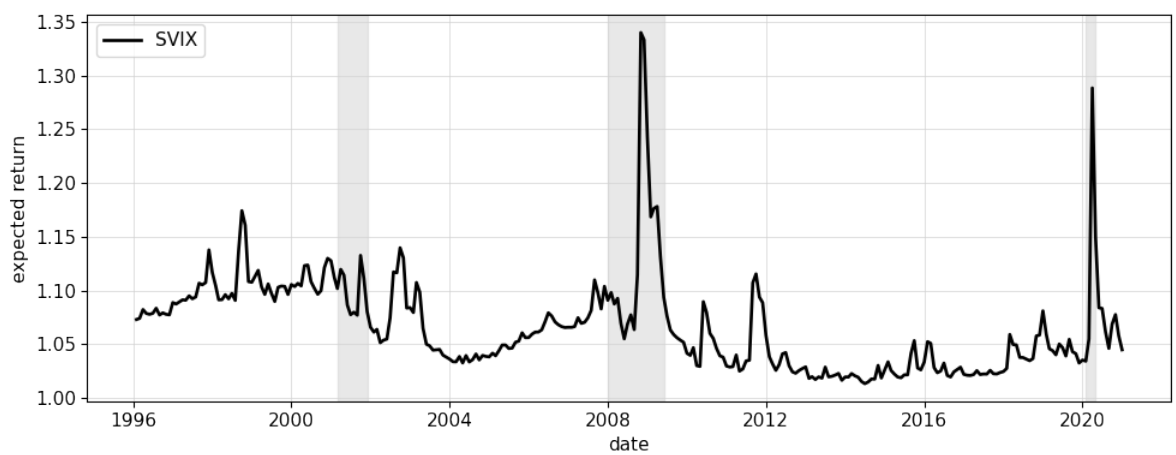



**Figure C.4.** SVIX
*Notes:* This figure plots the time-series of $\mathrm{E}^0[R]$, i.e. the expected return recovered by $M = 1/R$. The graph shows the monthly average of daily recovered expected returns, over the horizon of one month. Gray bars indicate NBER recessions.

# References


Acciaio, B., M. Beiglböck, F. Penkner, and W. Schachermayer. 2016. "A Model-Free Version of the Fundamental Theorem of Asset Pricing and the Super-Replication Theorem." *Mathematical Finance* 26 (2): 233–251.

Almeida, Caio, and Gustavo Freire. 2022. "Demand in the Option Market and the Pricing Kernel." *Working Paper, Princeton University.*

Bakshi, Gurdip, Charles Cao, and Zhiwu Chen. 1997. "Empirical Performance of Alternative Option Pricing Models." *Journal of Finance* 52 (5): 2003–2049.

Baltussen, Guido, Julian Jerstegge, and Paul Whelan. 2024. "The Derivative Payoff Bias." *Working Paper, Erasmus University.*

Ben-David, Itzhak, John R. Graham, and Campbell R. Harvey. 2013. "Managerial Miscalibration." *Quarterly Journal of Economics* 128: 1547–1584.

Beutel, Johannes, and Michael Weber. 2023. "Beliefs and Portfolios: Causal Evidence." *Working Paper, Deutsche Bundesbank.*

Borovička, Jaroslav, Lars Peter Hansen, and José A Scheinkman. 2016. "Misspecified Recovery." *Journal of Finance* 71 (6): 2493–2544.

Breeden, Douglas T, and Robert H Litzenberger. 1978. "Prices of State-Contingent Claims Implicit in Option Prices." *Journal of Business*: 621–651.

Buraschi, Andrea, Fabio Trojani, and Andrea Vedolin. 2014. "When Uncertainty Blows in the Orchard: Comovement and Equilibrium Volatility Risk Premia." *Journal of Finance* 101 (1): 101–137.

Carr, Peter, and Dilip Madan. 2001. "Optimal Positioning in Derivative Securities." *Quantitative Finance* 1: 19–37.

Chabi-Yo, Fousseni, and Johnathon Loudis. 2020. "The Conditional Expected Market Return." *Journal of Financial Economics* 137 (3): 752–786.

Chen, Hui, Scott Joslin, and Sophie Xiaoyan Ni. 2019. "Demand for Crash Insurance, Intermediary Constraints, and Risk Premia in Financial Markets." *Review of Financial Studies* 32 (1): 228–265.

Chen, Xiaohong, Lars Peter Hansen, and Peter G Hansen. 2020. "Robust Identification of Investor Beliefs." *Proceedings of the National Academy of Sciences* 117 (52): 33130–33140.

Couts, Spencer J., Andrei S. Goncalves, and Johnathan Loudis. 2023. "The Subjective Risk and Return Expectations of Institutional Investors." *Working Paper, Ohio State University.*

Dahlquist, Magnus, and Markus Ibert. 2024. "Equity Return Expectations and Portfolios: Evidence from Large Asset Managers." *Review of Financial Studies* 37: 1887–1929.

Egan, Mark, Alexander MacKay, and Hanbin Yang. 2024. "What Drives Variation in Investor Portfolios? Estimating the Roles of Beliefs and Risk Preferences." *Working Paper, Harvard Business School.*

Farago, Adam, Mariana Khapko, and Chayawat Ornthanalai. 2021. "Asymmetries and the market for put options." *Swedish House of Finance Research Paper* (21-12).

Gandhi, Mihir, Niels Joachim Gormsen, and Eben Lazarus. 2023. "Forward Return Expectations." *Working Paper, University of Chicago.*

Gao, Can, and Ian Martin. 2021. "Volatility, valuation ratios, and bubbles: An empirical measure of market sentiment." *Journal of Finance* 76 (6): 3211–3254.

Gârleanu, Nicolae, Lasse Heje Pedersen, and Allen M Poteshman. 2008. "Demand-based Option Pricing." *Review of Financial Studies* 22 (10): 4259–4299.

Gayda, Leander, Thomas Grünthaler, and Jan Harren. 2023. "Option Liquidity and Gamma Imbalances." *Working Paper, University of Münster*.

Ghosh, Anisha, Arthur Korteweg, and Qing Xu. 2022. "Recovering Heterogeneous Beliefs and Preferences from Asset Prices." *Working Paper, McGill University*.

Ghosh, Anisha, and Guillaume Roussellet. 2023. "Identifying Beliefs from Asset Prices." *Working Paper, McGill University*.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus. 2021. "Five Facts About Beliefs and Portfolios." *American Economic Review* 111 (5): 1481–1521.

Gormsen, Niels Joachim, and Christian Skov Jensen. 2022. "Higher-Moment Risk." *Working Paper, University of Chicago*.

Greenwood, Robin, and Andrei Shleifer. 2014. "Expectations of Returns and Expected Returns." *Review of Financial Studies* 27 (3): 714–746.

Jensen, Christian Skov, David Lando, and Lasse Heje Pedersen. 2019. "Generalized Recovery." *Journal of Financial Economics* 133 (1): 154–174.

Jensen, Theis Ingerslev. 2024. "Subjective Risk and Return." *Working Paper, Yale University*.

Johnson, Timothy, Mo Liang, and Yan Liu. 2016. "What Drives Index Options Exposures?" *Review of Finance*: 1–33.

Koijen, Ralph S. J., and Motohiro Yogo. 2019. "A Demand System Approach to Asset Pricing." *Journal of Political Economy* 127 (4): 1475–1515.

Korsaye, Sofonias Alemu. 2024. "Investor Beliefs and Market Frictions." *Working Paper, Johns Hopkins University*.

Lettau, Martin, and Sydney Ludvigson. 2001. "Consumption, Aggregate Wealth, and Expected Stock Returns." *Journal of Finance* 56: 815–849.

Long, John B. 1990. "The Numeraire Portfolio." *Journal of Financial Economics* 26 (1): 29–69.

Ludvigson, Sydney, and Serena Ng. 2009. "Macro Factors in Bond Risk Premia." *Review of Financial Studies* 22: 5027–5067.

Martin, Ian. 2017. "What is the Expected Return on the Market?" *Quarterly Journal of Economics* 132 (1): 367–433.

Martin, Ian, and Christian Wagner. 2019. "What is the Expected Return on a Stock?" *Journal of Finance* 74 (4): 1887–1929.

Meeuwis, Maarten, Jonathan A. Parker, Antoinette Schoar, and Simester Duncan. 2022. "Belief Disagreement and Portfolio Choice." *Journal of Finance* 77 (6): 3191–3247.

Moussawi, Rahib, Lai Xu, and Zhaoque Zhou. 2024. "A Market Maker of Two Markets: The Role of Options in ETF Arbitrage." *Working Paper, Villanova University*.

Nagel, Stefan, and Zhengyang Xu. 2023. "Dynamics of Subjective Risk Premia." *Journal of Financial Economics* 150 (2): 1–24.

Ross, Steve. 2015. "The Recovery Theorem." *Journal of Finance* 70 (2): 615–648.

Tetlock, Paul C. 2023. "The Implied Equity Premium." *Working Paper, Columbia Business*

*School.*
Welch, Ivo, and Amit Goyal. 2008. "A Comprehensive Look at The Empirical Performance of Equity Premium Prediction." *Review of Financial Studies* 21: 1455–1508.