Investor Beliefs and Trading Frictions

Sofonias Alemu Korsaye*†

January 18, 2024

Abstract

I develop a theoretical framework to identify investors' subjective beliefs consistent with survey expectations and asset prices in markets with trading frictions. I introduce a metric to quantify the deviation of these beliefs from Rational Expectations (RE), interpretable as a bound on the difference between the maximum Sharpe Ratios under investors' beliefs and RE. Empirically, I show that a significant share of the deviation from RE, assessed assuming frictionless markets, can be attributed to small trading costs. This deviation and the impact of trading costs differ across investor characteristics, with sophisticated investors' expectations more closely aligning with asset prices.

Keywords: Subjective Beliefs, Rational Expectations, Survey Expectations, Market Frictions

JEL Classification: G10, G40

*Carey Business School, Johns Hopkins University. Email: sofonias.korsaye@jhu.edu

†I thank Tony Berrada, Lars P. Hansen, Peter G. Hansen, Ralph Koijen, Yueran Ma, Stefan Nagel, Alberto Quaini, Alireza Tahbaz-Salehi, Julian Terstegge, Fabio Trojani, Andrea Vedolin, and seminar participants at the Chicago Booth, University of Geneva, Johns Hopkins Carey Business School, Olin Business School, Goizueta Business School, UNC Kenan-Flagler Business School, University of Zurich, Rotman School of Management, Imperial College Business School, INSEAD, The Desautels Faculty of Management, Kellogg School of Management, Carlson School of Management, Questrom School of Business, and Western Finance Association 2023. All errors are mine.

1 Introduction

The Rational Expectation (RE) hypothesis posits an alignment between investors' subjective beliefs and the objective belief, the latter reflecting the true data-generating process governing asset prices. Investor beliefs about future cash flows and prices can be inferred from two sources: asset prices and survey data. One common finding in the literature is that subjective beliefs recovered from either surveys or asset prices are at odds with RE. For instance, Chen, Hansen, and Hansen [2020] find that expectations and probabilities recovered from asset prices deviate from those under the objective belief.

This disconnect between investor beliefs and RE is further echoed in the broader literature, where investor expectations from surveys are difficult to reconcile with some predictions of standard RE asset pricing models (Adam and Nagel [2022] and references therein). Moreover, deviations from predictions of RE vary with investor and asset characteristics (Ghosh, Korteweg, and Xu [2020]) and depend on market frictions, confidence in one's own belief, and investor sophistication (Giglio, Maggiori, Stroebel, and Utkus [2021]). While the relation between deviations from RE and market frictions is largely documented, there is limited evidence in the literature that quantifies the magnitude of these deviations that can be attributed to frictions. This paper fills this gap.

To what extent do investor beliefs actually deviate from RE if trading frictions faced by investors are taken into account? To answer this question, I develop a theoretical framework to identify investor beliefs using survey data jointly with asset prices while allowing for a broad class of trading frictions. In particular, I quantify investor belief distortion as the deviation from RE that is necessary to jointly accommodate asset prices and survey expectations. The proposed belief distortion metric can be economically interpreted as an upper bound on the difference between the maximum Sharpe Ratio according to investors and the one inferred by the econometrician. The empirical results of this paper underscore the significant influence of market frictions on our understanding of the magnitude of deviations from RE. More specifically, I document that a statistically and economically significant component of the investor belief distortion, initially estimated assuming frictionless markets, can be attributed to empirically observable transaction costs.

Intuitively, market frictions and trading costs, in particular, can introduce a wedge between investor beliefs about assets' future returns and the objective asset return distribution. For example, leverage constraints may limit the actions of optimistic investors while short selling costs constrain the actions of pessimistic investors. As a result, future returns may not fully reflect investor beliefs recovered by the econometrician who assumes frictionless markets, resulting in significant belief distortion. However, investor beliefs identified by taking into account trading frictions deviate from RE to a lesser extent. To formally quantify the impact of such frictions on the distortion of investor beliefs, I use the proposed framework to measure the deviation from RE for various levels of trading costs.

The proposed framework identifies investor beliefs that are consistent with both observed survey information and no-arbitrage pricing relations. To this end, I jointly use two sets of moment conditions: First, investor expectations obtained from various survey data and second, asset pricing relations characterized by the absence of arbitrage opportunities. These no-arbitrage pricing relations are given by the standard Euler equations in frictionless markets and by inequalities in the presence of market frictions. Therefore, a belief is consistent with asset prices if there exists some investor preference model for which the no-arbitrage pricing relations hold.

In the empirical analysis, I quantify deviations from the objective belief necessary to simultaneously (i) match the one-year ahead market return expectations of individual investors, CFOs, and professional forecasters, and (ii) satisfy asset pricing relations for annual returns of equity portfolios formed on various characteristics. My findings can be summarized as follows. In a standard CAPM setting, investor belief distortions are significant when markets are assumed to be frictionless. The maximum Sharpe Ratio according to the investor can deviate from the one inferred by the econometrician by almost 90%. Incorporating transaction costs remarkably lowers these distortions. A 0.5% (1%) annual trading cost results in a 30% (45%) reduction of the deviation from RE necessary to accommodate asset prices and survey expectations that is initially estimated assuming frictionless markets.[1]

The framework further allows to study belief distortions across different levels of investor sophistication. To this end, I compare the market return expectations of professional forecasters, individual investors, and CFOs in terms of consistency with prices of portfolios formed on various characteristics. I find that, once asset pricing relations are satisfied, the additional deviation from RE necessary to match survey expectations is an order of magnitude smaller for professional forecasters' expectations compared to those of individual investors and CFOs. This evidence indicates that the expectations of professional forecasters are more in line with asset prices than those of individual investors and CFOs. Moreover, CFOs' expectations about the aggregate market are not more closely aligned with asset prices than those of individual investors.

The empirical results further document substantial heterogeneity across investors in different segments of the equity market in terms of the magnitude of investor belief distortion that can be attributed to trading costs. 60% of the belief distortion for investors in small stocks estimated assuming frictionless markets can be explained by a 0.5% annual trading cost. The same degree of trading cost explains 36% of the belief distortion when large stocks are considered, 38% for growth and value stocks, 20% for winner stocks, and 14% for loser stocks.[2]

[1]For instance, the trading cost faced by investors exclusively due to bid-ask spreads on US equities from 1993 to 2020, estimated as the effective half bid-ask spread using the New York Stock Exchange's Trade and Quote database, stands at 0.75%, see Appendix A.

[2]Ghosh, Korteweg, and Xu [2020] report heterogeneity across segments of the equity market in terms of absolute levels of belief distortion, when investor beliefs are inferred solely from asset prices.

When comparing costs relative to long and short portfolio positions, I find that the costs associated with the short legs of trades are significantly influential in accounting for the deviations from RE obtained assuming frictionless markets. This effect is particularly pronounced in the context of small stocks, where over two-thirds of what was initially perceived as a reduction in belief distortion due to symmetric trading costs impacting both legs of trades is, in reality, attributable to the costs incurred on the short positions.

Finally, I use the recovered investor beliefs to scrutinize behavioral biases suggested by the literature, especially those related to market anomalies such as the value-growth anomaly. In analyzing the probability distributions of annual returns for value and growth stock portfolios under both objective and subjective beliefs, a distinct pattern emerges. Investors display an excessive pessimism towards value stocks, lending support to the mispricing interpretation commonly associated with the value-growth anomaly, see Lakonishok, Shleifer, and Vishny [1994], Porta, Lakonishok, Shleifer, and Vishny [1997], among others. Specifically, the probability of experiencing a negative net return for value stock portfolio stands at approximately 29% under the objective distribution, compared to a heightened 39% under the identified investor belief. Contrary to common assertion in the literature, however, the findings do not confirm the notion of excessive investor optimism toward growth stocks.

This paper contributes to a large literature investigating the joint role of investor beliefs and preferences in the determination of asset prices. A growing strand of literature exploits survey expectations data to characterize potential inconsistencies between observed asset prices and RE, see Adam and Nagel [2022]. Two main inconsistencies have been documented in this literature. First, most individual investor expectations are procyclical, see Greenwood and Shleifer [2014], whereas RE asset pricing models predict countercyclical expected returns. Second, the sensitivity of investor portfolio share in equity to changes in investor expectations is much lower compared to the predictions of standard RE asset pricing models, see Vissing-Jorgensen [2003] and Giglio, Maggiori, Stroebel, and Utkus [2021], among others. In contrast, Dahlquist and Ibert [2021] find results in line with RE predictions when focusing on expectations of large sophisticated investors and professional forecasters. In particular, they report countercyclicality of market return expectations of large investors and a higher sensitivity of portfolios of such investors to their expectations. Moreover, Giglio, Maggiori, Stroebel, and Utkus [2021] show that the sensitivity of portfolio holdings to expectations increases as investors face lower costs, implying that trading frictions are important in quantitatively explaining deviations from RE.

Overall, the empirical evidence from the literature suggests a strong relation between deviations of reported expectations from RE and market frictions that investors face when forming their portfolios. I contribute to this literature, by providing a systematic framework for identifying the deviation from the objective belief necessary to accommodate asset pricing relations in the presence of market frictions and to match evidence on investors' expectations from survey data. The framework directly quantifies the impact of trading costs on the distortion of investors' subjective beliefs. Empirically, I show that deviations from RE are substantially smaller when accounting for trading costs.

From a methodological point of view, this paper is related to recent papers proposing different approaches to recover the properties of investor beliefs consistently with observed asset prices, see Chen, Hansen, and Hansen [2020], Chen, Hansen, and Hansen [2021], and Ghosh and Roussellet [2023]. Similarly, Ghosh, Julliard, and Taylor [2017] use an information theoretical framework to recover a multiplicative missing component of consumption-based SDF. This paper differs from this literature in two main aspects. First, I obtain asset pricing relations used to identify investor beliefs in the presence of market frictions with the objective of quantifying the impact of trading costs on the deviation of investor beliefs from RE. Second, I identify investor beliefs using both asset prices and survey expectations.

Finally, while the main focus of this paper is quantifying the impact of trading costs on the distortion of investor beliefs, the proposed framework can be used to study examples of economic settings with frictions originating from other sources. Different types of frictions have been proposed in the literature as potential explanation for the discrepancy between the objective distribution of future outcomes and expectations of economic agents reported in surveys. Such examples include economic settings with information rigidity (Mankiw and Reis [2002] and Coibion and Gorodnichenko [2015]), rational inattention caused by informational constraints (Sims [2003]), bounded rationality (Sargent [2001] and Gabaix [2014], among others), and investors' uncertainty about their own beliefs (Enke and Graeber [2019]). Whenever the no-arbitrage asset pricing implications of such frictional economies can be characterized by constrained mispricings, the framework provided in this paper can be used to study the properties of investor beliefs consistent with survey data and asset prices.

2 Investor Beliefs, Preferences and Trading Frictions

This section outlines the fundamental relations between asset prices, risk preferences, and investor beliefs that hold in a market free of arbitrage but inclusive of trading costs. These interrelations, along with investor expectations obtained from survey data, identify investor beliefs that are consistent with both asset prices and survey-derived expectations.

Let us consider a two-period economy with a set of n basis assets. These assets have a price vector $\boldsymbol{P}_t$ at time t and a payoff vector $\boldsymbol{X}_{t+1}$ at time $t+1$. The probability framework is represented as $(\Omega, \mathcal{F}, \mathbb{P}_0)$, where Ω is the set of possible outcomes, $\mathcal{F}$ is the collection of events, and $\mathbb{P}_0$ is the objective probability belief governing the random payoff vector $\boldsymbol{X}_{t+1}$. A marginal investor assigns probabilities to events in $\mathcal{F}$ based on her subjective belief, denoted $\mathbb{P}_I$. This subjective belief may diverge from the objective belief $\mathbb{P}_0$, deviating from Rational Expectations.

In constructing a portfolio at time t, an investor faces trading costs. These market frictions are modeled through a positive transaction cost function $h(\boldsymbol{w})$, with $\boldsymbol{w} \in \mathbb{R}^n$ representing the portfolio weight. The transaction cost functions obey two properties: $h(\alpha \boldsymbol{w}) = \alpha h(\boldsymbol{w})$ for any $\alpha \geq 0$ and $h(\boldsymbol{w}_1 + \boldsymbol{w}_2) \leq h(\boldsymbol{w}_1) + h(\boldsymbol{w}_2)$. This implies that splitting a given portfolio into multiple orders might not reduce costs, and in some cases, it could even increase them. This approach enables the modeling of trading costs that optimal execution strategies cannot eliminate.[3] Examples of such transaction costs include no-short selling constraints, costly short selling, bid-ask spreads, and proportional transaction costs, among others.

2.1 Trading Frictions and Absence of Arbitrage

Given the set of basis assets, the total cost of a tradeable payoff $x := \boldsymbol{X}'_{t+1}\boldsymbol{w}$ is defined as

$$\pi(x) := \boldsymbol{P}'_t \boldsymbol{w} + h(\boldsymbol{w}) \ , \tag{1}$$

where $\boldsymbol{P}'_t \boldsymbol{w}$ is the price before costs of the payoff x. Therefore, $\pi(x)$ represents the aggregate cost, encompassing the transaction costs for trading the portfolio $\boldsymbol{w}$.

In an ideal market devoid of frictions, i.e., $h = 0$, there exists a relationship between asset prices, investor beliefs, and preferences, given by the Euler equation:

$$E_{\mathbb{P}_I}[M_{t+1}\boldsymbol{X}_{t+1}] = \boldsymbol{P}_t \ , \tag{2}$$

with M_{t+1} denoting the investor's stochastic discount factor (SDF). This implies that prices are discounted cash flows under the investor's subjective belief. The existence of a strictly positive SDF satisfying equation (2) is characterized by the absence of arbitrage opportunities or by equilibrium conditions under the subjective belief of a utility-maximizing investor. However, when market frictions are present, the relationship between asset prices, risk preferences, and investor beliefs is characterized by inequality conditions, as outlined in the works

[3]The transaction cost functions addressed in this paper do not consider price impact, where one could potentially reduce costs by splitting portfolios. However, they are adequate for examining the trading costs empirically investigated here, such as bid-ask spreads and short selling costs.

of Luttmer [1996], Korsaye, Quaini, and Trojani [2019] in the arbitrage pricing framework, and He and Modest [1995] in the context of consumption-based pricing.

In markets with frictions, investors use the total cost function π and assess probabilities under their subjective belief $\mathbb{P}_I$ to rule out arbitrage opportunities. Arbitrages considered here are defined as free-lunch opportunities, which get arbitrarily close to a positive payoff with a non-positive price.[4] The following proposition characterizes the absence of arbitrage in markets with frictions under the investor subjective belief.

Proposition 1. The absence of arbitrage under investor belief $\mathbb{P}_I$ is equivalent to the existence of a strictly positive SDF M_{t+1}, ensuring that for any portfolio $\boldsymbol{w} \in \mathbb{R}^n$: [5]

$$E_{\mathbb{P}_I}[M_{t+1}\boldsymbol{X}'_{t+1}\boldsymbol{w}] - \boldsymbol{P}'_t\boldsymbol{w} \leq h(\boldsymbol{w}). \tag{3}$$

While equation (2) establishes the interrelation between investor belief, preference, and asset prices in frictionless markets, Proposition 1 suggests this interrelation in markets with trading costs is governed by (3). Alternatively, it posits that mispricing can emerge in economies where transaction costs are non-negligible. Nonetheless, any strategy exploiting such mispricings is rendered unprofitable due to the associated transaction costs.

The "Limits to Arbitrage" literature elucidates why asset prices may not always mirror their fundamental values given by discounted cash flows. It posits that various frictions, such as institutional constraints and costs of implementation faced by arbitrageurs, inhibit their ability to correct market anomalies caused by less sophisticated investors, see Gromb and Vayanos [2010] for a review of the literature. In a similar vein, Proposition 1 illustrates that in a simple economic setting, the transaction costs borne by investors can result in

[4] A free-lunch under belief $\mathbb{P}_I$ entails payoff sequences U_k and V_k with $U_k \leq V_k$, $\liminf \pi(V_k) < 0$, and $U_k \to U \geq 0$, where inequalities are to be considered under belief $\mathbb{P}_I$.

[5] Proposition 1 is an application of Theorem 2.1 of Jouini and Kallal [1995] on the set of the basis assets.

systematic mispricings that remain unexploited, as any potential gains are dominated by the transaction costs.

2.2 Investor Beliefs Consistent with Asset Prices and Survey Expectations

Asset prices reveal information about both preferences and beliefs of investors. However, as stated by equation (2) in absence of market frictions and by inequality (3) in presence of frictions, what asset prices reveal is the interaction between investors' preferences, characterized by the SDF M_{t+1}, and investors' belief, $\mathbb{P}_I$. The intertwining nature of these two components – investors' beliefs and preferences – makes it challenging to discern one without making additional assumptions about the other.

Typically, asset pricing literature assumes that investor belief aligns with the objective belief, as per the Rational Expectations hypothesis, and focuses on specifying investor preferences to explain observed asset prices. This approach involves using a parametric family of SDFs as described by:

$$M_{t+1} := M(\boldsymbol{\theta}, \boldsymbol{Z}_{t+1}) \ , \tag{4}$$

where $\boldsymbol{Z}_{t+1}$ is a vector of variables underlying investor risk preferences, and $\boldsymbol{\theta} \in \mathbb{R}^d$ is the parameter vector for the SDF.[6] In my analysis, I retain the parametric SDF family as in equation (4), but I allow for a divergence between investor beliefs and RE, yet still consistent with the absence of arbitrage in frictional markets as per Proposition 1. Accordingly, I define the set of admissible investor beliefs implied by a given SDF model that satisfies pricing condition (3) in frictional markets as consistent with asset prices in the absence of arbitrage

[6]For instance, a standard power utility consumption-based SDF is given by $M(\boldsymbol{\theta}, Z_{t+1}) = \beta Z_{t+1}^{-\gamma}$, where $Z_{t+1} = C_{t+1}/C_t$ indicates consumption growth and $\boldsymbol{\theta} = (\beta, \gamma)$ encapsulates the SDF parameters, including the time discounting parameter β and the risk aversion parameter γ.

opportunities.

Definition 1 (Investor beliefs consistent with asset prices)**.** Given an SDF $M(\cdot, \boldsymbol{Z}_{t+1})$ and transaction cost h, the set of beliefs consistent with asset prices is defined as

$$\mathcal{P}_h := \{\mathbb{P} : \quad \exists \boldsymbol{\theta} \in \mathbb{R}^d : \quad E_{\mathbb{P}}[M(\boldsymbol{\theta}, \boldsymbol{Z}_{t+1})\boldsymbol{X}'_{t+1}\boldsymbol{w}] - \boldsymbol{P}'_t\boldsymbol{w} \leq h(\boldsymbol{w}) \quad \text{for all } \boldsymbol{w} \in \mathbb{R}^n\} \ . \tag{5}$$

Under this definition, investor beliefs are probability distributions for which there exists at least one suitable SDF within a parametric family that satisfies the pricing relations outlined by the absence of arbitrage opportunities in markets with frictions. In particular, in a frictionless market, investor beliefs are probabilities for which the Euler equation (1) holds for a given parametric SDF, as expressed by:

$$\mathcal{P}_{h=0} := \{\mathbb{P} : \quad \exists \boldsymbol{\theta} \in \mathbb{R}^d : \quad E_{\mathbb{P}}[M(\boldsymbol{\theta}, \boldsymbol{Z}_{t+1})\boldsymbol{X}_{t+1}] = \boldsymbol{P}_t\} \ . \tag{6}$$

Survey data regarding investor expectations offer additional insights into investor beliefs, extending beyond asset prices. These surveys typically query participants at time t about their expectations for a variable $\boldsymbol{Y}_{t+1}$, observable at time $t+1$. These expectations can include forecasts of asset returns, volatility, or the likelihood of returns falling below a certain threshold. Given the survey expectation, denoted as $\mathcal{E}[\boldsymbol{Y}_{t+1}]$, I define the set of beliefs consistent with this expectation:

Definition 2 (Investor beliefs consistent with survey expectations)**.** The set of beliefs consistent with survey expectation $\mathcal{E}[\boldsymbol{Y}_{t+1}]$ is

$$\mathcal{P}_{SE} := \{\mathbb{P} : \quad E_{\mathbb{P}}[\boldsymbol{Y}_{t+1}] = \mathcal{E}[\boldsymbol{Y}_{t+1}]\} \ . \tag{7}$$

The investor beliefs outlined in equation (7) match survey expectations precisely, though these are often aggregate measures and can be prone to measurement errors. To mitigate this, one approach is to allow $E_{\mathbb{P}}[\boldsymbol{Y}_{t+1}]$ to fall within a certain range of $\mathcal{E}[\boldsymbol{Y}_{t+1}]$, such as a confidence region for the average survey expectation, i.e., $E_{\mathbb{P}}[\boldsymbol{Y}_{t+1}] \in CI(\mathcal{E}[\boldsymbol{Y}_{t+1}])$, where CI represents a confidence region with a specified significance level.

The additional information provided by surveys is helpful to further narrow the set of admissible investor beliefs identified from asset prices alone, simply by imposing admissible investor beliefs to be consistent with observable information generated by both asset prices and survey expectations.

Definition 3 (Investor beliefs consistent with asset prices and survey expectations)**.** The set of beliefs consistent with both asset prices and survey expectations is defined as

$$\mathcal{P}_I := \mathcal{P}_h \cap \mathcal{P}_{SE}\ . \tag{8}$$

Subsequently, I analyze the characteristics of investor beliefs within $\mathcal{P}_I$, focusing particularly on the degree to which investor beliefs must diverge from the objective belief to align with both asset prices and survey expectations.

3 Investor Belief Distortion

In this section, I present a theoretical framework designed to assess the deviation of investor beliefs from Rational Expectations. This belief distortion is conceptualized as the discrepancy between investors' subjective beliefs and the objective belief, particularly in presence of market frictions. For this purpose, I extend the framework proposed by Chen, Hansen, and Hansen [2020] to account for market frictions. In absence of market frictions, Chen, Hansen, and Hansen [2020] previously quantified investor belief distortion as the extent of divergence from the objective belief needed to satisfy the pricing condition (2).

Consider a belief $\mathbb{P}$, which is absolutely continuous with respect to $\mathbb{P}_0$. Let N denote the change of belief from $\mathbb{P}_0$ to $\mathbb{P}$, represented as $N = \frac{d\mathbb{P}}{d\mathbb{P}_0}$. The deviation of belief $\mathbb{P}$ from $\mathbb{P}_0$ can be quantified using what is termed as $\phi-$divergence, defined as $E_{\mathbb{P}_0}[\phi(N)]$. Here, ϕ denotes a convex function, satisfying $\phi(1) = 0$, implying that the divergence is zero when both beliefs coincide. Examples of ϕ-divergences include χ^2, total variation, and entropy divergences, among others.

However, I will focus on the χ^2-divergence due to its economic interpretability as a measure of belief distortion. The χ^2-divergence of belief $\mathbb{P}$ from $\mathbb{P}_0$ is given by

$$\chi^2(\mathbb{P}, \mathbb{P}_0) := E_{\mathbb{P}_0}[(N-1)^2]. \tag{9}$$

The following proposition posits that if beliefs $\mathbb{P}$ and $\mathbb{P}_0$ are aligned regarding the risk profile of portfolios, the square root χ^2-divergence between the two beliefs is an upper bound for the difference between the maximum Sharpe Ratios under the two beliefs.

Proposition 2. If the basis assets have the same variance under beliefs $\mathbb{P}$ and $\mathbb{P}_0$, then

$$|SR_{max}(\mathbb{P}) - SR_{max}(\mathbb{P}_0)| \leq \sqrt{\chi^2(\mathbb{P}, \mathbb{P}_0)}, \tag{10}$$

where $SR_{max}(\mathbb{P})$ and $SR_{max}(\mathbb{P}_0)$ are the maximum Sharpe Ratios under the respective beliefs.

Inequality (10) is closely related to the well-known Hansen-Jagannathan bound on the maximum Sharpe Ratio, Hansen and Jagannathan [1991]. When the two beliefs in question are given by the investor belief and the risk-neutral belief, the change of belief, N, is the forward SDF and the square root of the χ^2-divergence is its volatility. This and the fact that the Sharpe Ratio under the risk-neutral belief is zero result in the Hansen-Jagannathan bound.

Inequality (10) gains in informative value as the basis assets span a more extensive set of contingent claims. When the $\chi^2(\mathbb{P}, \mathbb{P}_0)$ value is large, it implies a significant potential disparity between the Sharpe Ratio as perceived by the investor and the Sharpe Ratio as retrospectively inferred by the econometrician.

As discussed in Section 2.2, investor beliefs can be identified through moment conditions involving both asset prices and survey expectations. Subsequently, I define investor belief distortion as the minimal divergence from the objective belief necessary to align with both survey expectations and asset prices in markets characterized by frictions.

Definition 4 (Investor Belief Distortion)**.** Investor belief distortion is the smallest deviation from the objective belief required for investor beliefs to be consistent with asset prices and survey expectations. Formally,

$$\delta := \inf_{\mathbb{P}} \chi^2(\mathbb{P}, \mathbb{P}_0) \qquad \text{s.t.} \qquad \mathbb{P} \in \mathcal{P}_h \cap \mathcal{P}_{SE}. \tag{11}$$

Here, δ represents the divergence of the set of investor beliefs from the objective belief. An infinite δ implies the nonexistence of any belief consistent with both asset prices and survey expectations. Conversely, a δ of zero indicates that the objective belief aligns with both asset pricing moment conditions and survey expectations, thereby upholding the Rational Expectations hypothesis. Incorporating more transaction costs broadens the set of investor beliefs that accord with asset prices, thus diminishing the necessary deviation from Rational Expectations to align with both asset prices and survey expectations. In the empirical section, the analysis will focus on quantifying the extent to which the required deviation from Rational Expectations diminishes as empirically observable levels of trading costs are incorporated.

The following proposition characterizes the metric δ of investor belief distortion through an equivalent dual optimization problem. This formulation is more tractable compared to problem (11) and instrumental in deriving the estimator of the belief distortion δ and its large-sample properties.

Proposition 3 (Dual Characterization)**.** Assuming the regularity condition that for every $\boldsymbol{\theta} \in \mathbb{R}^d$, there exists a corresponding $\tilde{\mathbb{P}} \sim \mathbb{P}_0$ such that $E_{\tilde{\mathbb{P}}}[M(\boldsymbol{\theta}, \boldsymbol{Z}_{t+1})\boldsymbol{X}_{t+1}'\boldsymbol{w}] - \boldsymbol{P}_t'\boldsymbol{w} \leq h(\boldsymbol{w})$

for all $\boldsymbol{w} \in \mathbb{R}^n$, with strict inequality when $h(\boldsymbol{w}) > 0$, then we can state[7,8]

$$\delta = - \sup_{\boldsymbol{\theta} \in \mathbb{R}^d} \inf_{\gamma \in \mathbb{R}, \boldsymbol{w} \in \mathbb{R}^n, \boldsymbol{\alpha} \in \mathbb{R}^k} E_{\mathbb{P}_0}[(1 - M(\boldsymbol{\theta}, \boldsymbol{Z}_{t+1})\boldsymbol{X}'_{t+1}\boldsymbol{w} - \mathbf{Y}'_{t+1}\boldsymbol{\alpha} - \gamma)_+^2/4] \\ + \boldsymbol{P}'_{t+1}\boldsymbol{w} + h(\boldsymbol{w}) + \mathcal{E}[\mathbf{Y}_{t+1}]'\boldsymbol{\alpha} + \gamma. \tag{12}$$

Moreover, the least distorted investor belief, $\mathbb{P}^*$, as a solution to problem (11), is characterized by the change of belief $N^* = \frac{d\mathbb{P}^*}{d\mathbb{P}_0}$, with

$$N^* = (1 - M(\boldsymbol{\theta}^*, \boldsymbol{Z}_{t+1})\boldsymbol{X}'_{t+1}\boldsymbol{w}(\boldsymbol{\theta}^*) - \mathbf{Y}'_{t+1}\boldsymbol{\alpha}(\boldsymbol{\theta}^*) - \gamma(\boldsymbol{\theta}^*))_+/2, \tag{13}$$

where $\boldsymbol{\theta}^*$ solves problem (12).

This proposition characterizes investor belief distortion as the result of a finite-dimensional optimization problem, wherein the transaction cost function regularizes the dual parameter $\boldsymbol{w}$ related to assets subject to trading costs. The regularity condition assumption in Proposition 12 is in line with standard interior-point requirements commonly invoked in the characterization of optimality conditions under inequality constraints.

Problem (12) can be used to estimate the minimum deviation from RE necessary to accommodate only asset prices by setting the dual parameter $\boldsymbol{\alpha}$ to zero, effectively disregarding the survey expectations moment condition. Conversely, setting the dual parameters $\boldsymbol{w}$ to zero enables the evaluation of the minimal deviation from RE needed to conform exclusively to survey expectations.

[7] $\tilde{\mathbb{P}}$ is equivalent to $\mathbb{P}_0$, $\tilde{\mathbb{P}} \sim \mathbb{P}_0$, if both beliefs concur on zero-probability events.

[8] More general results are detailed in Appendices B and C, where, instead of exactly matching survey expectations aggregated across participants, investor expectations are constrained within a neighborhood of $\mathcal{E}[\boldsymbol{Y}_{t+1}]$.

4 Empirical Analysis

This section applies the theoretical framework developed in Section 2 to achieve three objectives: 1) to identify investor beliefs by jointly using asset prices and survey expectations, 2) to evaluate the deviation of these beliefs from Rational Expectations (RE), and 3) to determine the proportion of these deviations attributable to trading frictions. The ensuing paragraphs outline the data utilized and the methodology employed for estimating distortions of investor beliefs.

4.1 Data

Investor beliefs and preferences are identified by using moment conditions related to both asset prices and survey expectations. The initial set of moment conditions is based on asset pricing relations, employing portfolios sorted by characteristics from the Kenneth French data library. Specifically, six portfolios sorted by size and book-to-market are utilized to represent a broader range of assets. Additionally, portfolios formed based on size, book-to-market, and momentum are used to represent different categories of assets - namely, large and small assets, growth and value assets, and winners and losers - to examine the variability in investor belief distortion across equity market segments.

The second set of moment conditions for identifying investor beliefs is derived from survey expectations of the one-year ahead market return. The primary source of these expectations is the Livingston survey data from the Federal Reserve Bank of Philadelphia. Here, professional forecasters semiannually provide their expectations of the one-year ahead S&P500 index price. The expected market return is defined as $\mathcal{E}[R_{M,t+1}] = \mathcal{E}\left[\frac{P_{M,t+1}}{P_{M,t}}\right] + \frac{D_{M,t}}{P_{M,t}}\mathcal{E}\left[\frac{D_{M,t+1}}{D_{M,t}}\right]$, where the expected price growth, $\mathcal{E}\left[\frac{P_{M,t+1}}{P_{M,t}}\right]$, from 1992 onwards is determined by the ratio of twelve-month to zero-month mean level forecasts and from 1952 to 1992 by the annualized ratio of twelve-month to six-month forecasts. The expected dividend growth, $\mathcal{E}\left[\frac{D_{M,t+1}}{D_{M,t}}\right]$, in line with Nagel and Xu [2022], is set at 1.064, the average S&P500 annual dividend growth from 1946 to 2020.

The second source of survey expectations is from Nagel and Xu [2022], where the authors aggregate investor expectations of the market return from various sources, including the UBS/Gallup Survey, the Conference Board survey, and the Michigan survey of consumers. This dataset includes quarterly observations of one-year ahead investor expectations from 1990 to 2019. A third source of survey expectations, used in the empirical analysis, is the Graham-Harvey CFO survey. In this survey, participants quarterly provide their expectations of the one-year ahead market return, covering the period from 2000 to 2020. Figure 1 presents the time series of investor expectations of market return from these three sources.

4.2 Methodology

The empirical analysis employs mainly the CAPM SDF, defined as $M_{t+1} = a + bR_{M,t+1}$ with a and b being real numbers, and $R_{M,t+1}$ representing the market return. The choice of CAPM is driven by two primary reasons. Firstly, its widespread adoption as a benchmark model in Asset Pricing, owing to its versatility across diverse economic scenarios and distributional assumptions. Secondly, and more crucially, in the context where the investor SDF is given by the CAPM and survey expectations are of the market return, both the SDF variable $\boldsymbol{Z}_{t+1}$ and the expectation variable Y_{t+1} converge on the market return $R_{M,t+1}$. This implies that survey expectations reveal information not only about investor beliefs but also about investor preferences. We may thus gain a more informative assessment of the investor belief distortion by combining survey expectations of the market return in addition to investor SDF provided by the CAPM. The investigation extends to encompass power and recursive utility-based SDFs related to the market return, which is further elaborated in Section 4.7.

Proposition 3 provides a dual framework to estimate investor belief distortion across various market frictions. The estimator is formulated as

$$\hat{\delta} := -\sup_{\boldsymbol{\theta}} \inf_{\gamma,\boldsymbol{w},\alpha} \frac{1}{T}\sum_{t}\Big((1 - M(\boldsymbol{\theta}, R_{M,t+1})\boldsymbol{R}'_{t+1}\boldsymbol{w} - \alpha R_{M,t+1} - \gamma)_+^2/4 + \pi(\boldsymbol{R}'_{t+1}\boldsymbol{w}) + \alpha\mathcal{E}[R_{M,t+1}]\Big) + \gamma, \tag{14}$$

where $\boldsymbol{R}_{t+1} = \{\boldsymbol{X}_{i,t+1}/\boldsymbol{P}_{i,t}\}_{i=1,\ldots,n}$ is the pre-cost return vector of the basis assets. Here, both the SDF and survey expectation variables are given by the market return. The large sample properties of $\hat{\delta}$ are detailed in Appendix B. Given that asset pricing moment relations (3) are based on inequalities and transaction cost functions lack differentiability, the estimator's asymptotic distribution is not standard, necessitating the implementation of a bootstrap approach.

A related metric that is used throughout this empirical analysis, alongside δ, is given by the following expression:

$$\delta^{SR} := \sqrt{\delta}/SR_{max}(\mathbb{P}_0)\ , \tag{15}$$

where δ denotes the minimum χ^2-divergence, as specified in equation (11). Here, $SR_{max}(\mathbb{P}_0)$ is the maximum Sharpe Ratio under the objective belief that is retrospectively inferred by the econometrician. In scenarios where investor and objective beliefs concur on the variance of basis payoffs, δ^{SR} sets an upper bound for the deviation of the maximum Sharpe Ratio according to the investor relative to the maximum Sharpe Ratio according to the econometrician, i.e.,

$$\frac{|SR_{max}(\mathbb{P}^*) - SR_{max}(\mathbb{P}_0)|}{SR_{max}(\mathbb{P}_0)} \leq \delta^{SR}\ ,$$

with $\mathbb{P}^*$ representing the least distorted belief consistent with asset prices and survey expectations. A lower value of δ^{SR} implies the existence of an investor belief yielding a maximum Sharpe Ratio close to that under the objective belief.

The empirical analysis encompasses three market scenarios: frictionless markets, markets with bid-ask spreads, and markets bearing short selling costs.[9] The transaction function h

[9]While the focus of this paper is on trading frictions such as bid-ask spreads and short selling costs, the

is plotted in figure 2 for the different market settings considered.

Frictionless Market: In a frictionless market framework, investors incur no transaction costs when trading, which is equivalent to a zero transaction cost function h. This assumption leads to the utilization of standard Euler Equations, complemented by survey expectations, for identifying investor beliefs, as represented by

$$\mathcal{P}_I = \left\{ \mathbb{P} : \ \exists \boldsymbol{\theta} \in \mathbb{R}^d : \quad E_{\mathbb{P}}[M(\boldsymbol{\theta}, R_{M,t+1}) R_{t+1,i}] - 1 = 0 \quad \forall i, \quad E_{\mathbb{P}}[R_{M,t+1}] = \mathcal{E}[R_{M,t+1}] \right\} \ .$$

Bid-Ask Spreads: In this context, investors bear transaction costs in the form of bid-ask spreads when trading assets. The (half) bid-ask spread is given by $\epsilon_{bas} = (P_{ask} - P_{bid})/2P$, where P_{ask}, P_{bid} and P are ask, bid, and mid prices, respectively. To simplify the terminology, I will use 'bid-ask spread' to refer to what is technically the 'half bid-ask spread'. This spread represents the cost incurred by an investor when purchasing a share of an asset at the ask price or selling at the bid price, expressed as a fraction of the asset's mid price. Thus, the transaction cost for a portfolio $\boldsymbol{w}$ is given by

$$h(\boldsymbol{w}) = \sum_{w_i \geq 0} \epsilon_{bas} P_i w_i - \sum_{w_i \leq 0} \epsilon_{bas} P_i w_i \ .$$

When translating to returns, the bid-ask spread cost results as $\epsilon_{bas} \|\boldsymbol{w}\|_1$, introducing a LASSO penalty (Tibshirani [1996]) in the optimization equation (12). The investor belief set in the presence of bid-ask spreads is consequently

$$\mathcal{P}_I = \left\{ \mathbb{P} : \ \exists \boldsymbol{\theta} \in \mathbb{R}^d : \quad |E_{\mathbb{P}}[M(\boldsymbol{\theta}, R_{M,t+1}) R_{t+1,i}] - 1| \leq \epsilon_{bas} \quad \forall i, \quad E_{\mathbb{P}}[R_{M,t+1}] = \mathcal{E}[R_{M,t+1}] \right\} \ .$$

Short Selling Costs: Short selling costs include direct costs, such as asset loan fees, or indirect costs, such as limited lending supply of assets and asset recall risk. When borrowing

reported values of trading costs can be viewed as aggregated quantities across various types of transaction costs.

assets, investors need to provide the lender with collateral on which the latter pays a rebate rate. The difference between the risk-free rate and the rebate rate is a direct measure of the loan fee.

In this setting, I consider only direct costs faced by investors when short selling assets and the shorting cost is proportionate to the price of the asset, i.e.,

$$h(\boldsymbol{w}) = -\sum_{w_i \leq 0} \epsilon_{ss} P_i w_i \ , \quad \text{with} \quad \epsilon_{ss} > 0 \ .$$

In presence of shorting costs, the set of investor beliefs consistent with asset prices and survey expectations is given by

$$\mathcal{P}_I = \Big\{ \mathbb{P} : \ \exists \boldsymbol{\theta} \in \mathbb{R}^d : \quad -\epsilon_{ss} \leq E_{\mathbb{P}}[M(\boldsymbol{\theta}, R_{M,t+1}) R_{t+1,i}] - 1 \leq 0 \quad \forall i, \quad E_{\mathbb{P}}[R_{M,t+1}] = \mathcal{E}[R_{M,t+1}] \Big\} \ .$$

The scenario where ϵ_{ss} is set to infinity ($\epsilon_{ss} = \infty$) equates to a market condition where investors are not able to short sell.

The following empirical analysis employs conservative estimates for trading costs of 0.5% and 1%. For context, Appendix A illustrates that the average (half) bid-ask spread alone from 1993 to 2020, as per TAQ data and following the method by Holden and Jacobsen [2014], is 0.75%. When addressing short selling costs and constraints, the literature reports varied figures for asset loan fees. For instance, Cohen et al. [2007] note an average fee as high as 2.60% for small stocks. Moreover, the next empirical analysis does not consider other significant limitations when short selling, such as the limited availability of stocks for borrowing and the risk of stock recalls. As a result, the findings presented in the following sections are based on conservative estimates of trading costs.

4.3 Evaluating Investor Belief Distortion in Presense of Trading Costs

This section focuses on the extent to which investor beliefs diverge from Rational Expectations under varying scenarios of bid-ask spreads. In essence, the analysis seeks to understand how the assumed levels and types of transaction costs impact our understanding of the deviation of investor beliefs from Rational Expectations.

Let us consider investor beliefs identified by pricing portfolios formed on size and book-to-market ratios, and matching one-year ahead market return expectations of professional forecasters from the Livingston survey data. Figure 3 displays the magnitude of investor belief distortion, initially determined using asset prices alone and later combined with survey expectations. Without market frictions, the minimal belief distortion necessary to be consistent with only asset prices, as measured by δ^{SR}, stands at 70.6%, whereas accommodating both asset prices and survey expectations raises this distortion to 88.8%. This suggests a significant discrepancy, up to 88.8%, from the maximum Sharpe Ratio under the objective belief when investor beliefs are identified using both asset prices and survey expectations.

Incorporating bid-ask spreads of 0.5% and 1% for annual holdings markedly lowers the necessary belief distortion to align with only asset prices, to 52.6% and 41.1% respectively. This represents a notable reduction of 25% and 42% from the frictionless market baseline, where the minimum deviation from Rational Expectations (RE) was recorded at 70.6%. Moreover, when these levels of bid-ask spreads are factored in, the distortion needed for consistency with both asset prices and survey expectations is further reduced to 61.8% and 48.2%. This shift signifies a substantial decline from the 88.8% distortion observed when markets are assumed to be frictionless. Panel (b) of Figure 3 illustrates these changes, highlighting a 30% and 46% decrease in distortion when annual costs of 0.5% and 1% are assumed, respectively.

The analysis, as illustrated in figure 3, maps out a nuanced landscape of transaction cost and belief distortion. Three distinct regions emerge: 1) A white area indicating levels of transaction cost and belief distortion where no belief aligns with asset prices. 2) A light gray area showing levels where there are beliefs consistent with either asset prices or survey expectations, but not both, signaling a need for additional adjustment in beliefs. 3) A dark gray area encompassing levels where there are beliefs consistent with both asset prices and survey expectations.

The discrepancy between asset prices and survey expectations is quantifiable by examining the extra distortion required to meet survey expectations in addition to existing asset pricing conditions. This additional distortion is represented by the light gray area in Figure 3. In a frictionless market, this discrepancy amounts to 18.3%, which notably decreases to 9.2% and 7.1% with the introduction of bid-ask spreads of 0.5% and 1%, respectively.

When extremely high transaction costs are assumed, exceeding a 4% annual bid-ask spread, the asset pricing relation (3) becomes uninformative, as it identifies a very large set of admissible investor beliefs. Hence, the estimated belief distortion, derived from reconciling both asset prices and survey expectations, reduces to the minimal level of distortion required solely for aligning with survey expectations. This minimum level of belief distortion is identified to be 14.5%. This means that there exists a belief consistent with survey expectations and has a maximum Sharpe Ratio that differs by no more than 14.5% from the one computed under the objective belief.

4.4 Investor Belief Distortion Across Different Investor Sophistication Levels

The next analysis focuses on three distinct sources of survey expectations, differentiated by the sophistication level of the respondents.

Firstly, the Livingston Survey data, which encompasses one-year ahead market price expectations from professional forecasters, is used as a representation of sophisticated investors. In contrast, the survey data from Nagel and Xu [2022] provides insights into the one-year ahead market return expectations of individual investors. Additionally, the CFO survey data offers perspectives on the expectations of chief financial officers and others with financial decision-making roles. The core of this analysis hinges on comparing the expectations across these investor categories, primarily in terms of their deviation from Rational Expectations and their alignment, or lack thereof, with asset prices.

Consider, initially, a frictionless market scenario. Within this construct, the minimal deviation from Rational Expectations necessary to pricing portfolios formed on size and book-to-market ratios, while accommodating the expectations of sophisticated investors, is 88%. In comparison, this deviation escalates to 146.8% and 150.8% for individual investors and CFOs, respectively. These figures suggest two key points: firstly, the expectations of professional forecasters tend to be more closely aligned with Rational Expectations than those of individual investors and CFOs; secondly, the expectations of CFOs are not better than those of individual investors in terms of proximity to Rational Expectations.

Incorporating a 0.5% (1%) for annual holdings, the belief distortion reduces to 61.8% (48.2%), 92.5% (52.4%), and 87.5% (32.4%) for sophisticated investors, individual investors and CFOs, respectively. Compared to the frictionless market assumption, a 0.5% bid-ask spread assumption alone leads to a drop of the investor belief distortion by 30%, 37%, and 41% when expectations of sophisticated, individual investors and CFOs are considered.

How do aggregate market expectations differ among investors of varying sophistication, particularly in relation to alignment with portfolio returns? As discussed in the preceding section, the additional belief distortion necessary to accommodate survey expectations once asset pricing relations are satisfied is an indication of the discrepancy between asset prices and expectations of the aggregate market. Under the assumption of frictionless markets, this additional distortion is notably higher for individual investors and CFOs, at 88.2% and 106.7% respectively. These values are substantially larger – approximately five to six times –

compared to the additional distortion required to align with the expectations of sophisticated investors. This finding reinforces the notion that CFOs' expectations about the aggregate market are not more closely aligned with portfolio returns than those of individual investors.

The comprehensive comparison of the three sets of survey data, as depicted in figures 3 and 4, indicates a significant pattern. Among the surveys examined, the Livingston survey data, which represents aggregate market expectations of sophisticated investors, shows a closer alignment with actual asset prices. This observation is derived from the fact that investor beliefs, which are consistent with asset pricing moments, require less adjustment to be in line with the expectations outlined in the Livingston survey.

4.5 Investor Belief Distortion Across Different Equity Market Segments

In this section the analysis focuses on the variation of investor beliefs across different segments of the equity market. In particular, it examines belief distortions among investors trading in various categories of stocks, including large, small, value, growth, winner, and loser stocks. Following Ghosh, Korteweg, and Xu [2020], portfolios based on size, book-to-market ratios, and momentum are employed for this purpose. In what follows, the survey-driven expectations of the aggregate market is given by sophisticated investors.

The empirical analysis is driven by two main questions: Firstly, to what degree investors' deviations from Rational Expectations vary based on the market segments they engage in? And secondly, To what extent these belief distortions depend on the assumed trading costs. A key part of the analysis is to explore the effects of two types of trading costs on belief distortions: bid-ask spreads, and short selling costs. This approach will help in discerning the relationship between trading costs and investor beliefs across different segments of the equity market.

4.5.1 Bid-Ask Spreads

Let us consider investors who face transaction costs, specifically in the form of bid-ask spreads. The ensuing analysis, reported in Table 1, is based on: firstly, by pricing annual returns from 1960 to 2019 across a spectrum of assets – large, small, growth, value, winners, and losers; and secondly, by matching survey expectations, drawing from the Livingston survey data. The table reports the results encompassing both the maximum Sharpe Ratio deviation bound, δ^{SR}, of equation (15) and the minimum χ^2-divergence, δ, of equation (11). The corresponding bootstrapped 95% confidence intervals are reported alongside the belief distortion metrics. The analysis is conducted under two scenarios: a frictionless market setting and a market scenario factoring in an annual 0.5% bid-ask spread.

The analysis of investor beliefs reveals a diverse picture of distortions, when comparing investors in different market segments. Notably, beliefs of investors in small stocks exhibit more distortion when contrasted with those trading in larger assets. In a similar vein, the beliefs of those investing in growth (or loser) stocks are markedly more distorted compared to their counterparts in value (or winner) stocks. Focusing on the extent of these distortions, a pronounced deviation in the beliefs of loser stock investors is observed, with the maximum Sharpe Ratio deviation bound at 100%. Investors in small stocks follow, with a deviation bound of 60%.

When factoring in a 0.5% annual bid-ask spread, the observed deviations in investor beliefs significantly change compared to a frictionless market scenario. For investors in large assets, the Sharpe Ratio deviation bound declines from 49% to 31%. For small assets, this deviation drops from 60% to 23%. Distortion for investors in growth assets has a reduction from 65% to 40%, while that in value assets decreases from 54% to 34%. For winner asset investors, the deviation bound lessens from 57% to 45%, and for those investing in loser assets, it falls from 100% to 86%.

Attributable to the 0.5% annual bid-ask spread, this adjustment in belief distortion explains

36% of the distortion for large asset investors, 60% for small asset investors, 38% for both growth and value asset investors, 20% for winner asset investors, and 14% for loser asset investors. On average, across these market segments, the inclusion of a 0.5% annual bid-ask spread leads to a 34% reduction in belief distortion. These findings clearly indicate that deviations from Rational Expectations, initially estimated based on a frictionless market assumption, are indeed partly due to trading costs. The impact of these trading costs is particularly significant in the context of small assets. For value and growth stocks, the effect of trading costs is not markedly different, whereas their influence is comparatively less for momentum stocks.

4.5.2 Short Selling Costs

This section probes whether divergences from Rational Expectations in investor beliefs are differently impacted by the costs associated with the short and long positions in trading.

Let us consider investors who incur transaction costs only when short selling assets. Figure 5 delineates the divergence from Rational Expectations among investors, categorized by the type of assets they trade – encompassing large, small, growth, value, winning, and losing stocks. The analysis spans four market conditions: a frictionless environment, markets with short selling costs of 0.5% and 1% per annum, and a market setting where short selling is entirely precluded.

An annual short selling cost of 0.5% (1%) accounts for 42% (60%) of the belief distortion, estimated assuming frictionless markets, in investors of small stocks. For those trading in large and value stocks, these figures are 21% (38%), and for growth stocks, they are 23% (36%). In the categories of winner and loser stocks, the distortions are quantified at 12% (23%) and 6% (11%) respectively. Notably, entirely eliminating short selling aligns the belief distortion closely with that derived from survey expectations alone, as depicted by the horizontal blue line in Figure 5. However, across all these market scenarios, the Rational

Expectation Hypothesis is consistently rejected at a 95% confidence level.

These results underscores the critical impact of transaction costs on the deviation from Rational Expectations, particularly highlighting the distinct effects of short selling costs and bid-ask spreads. Short selling costs exclusively affect the short side of trades, while bid-ask spreads impact both sides. Let us consider two scenarios: one with a 0.5% bid-ask spread, imposing a 0.5% cost on both long and short legs of each trade, and another with only a 0.5% short selling cost on the short legs.

The results, presented in Table 1 and Figure 6, demonstrate that short selling costs are more influential in explaining deviations from Rational Expectations - estimated assuming frictionless market assumptions - than bid-ask spreads. This effect is most pronounced with small stocks, where 80% of the reduction in belief distortion, attributed to bid-ask spreads, is actually due to the costs on the short leg of trades. Across different market segments, this impact consistently exceeds 60%.

In summary, this analysis demonstrates that transaction costs, particularly those associated with short selling, play a pivotal role in explaining the deviations from Rational Expectations initially estimated assuming frictionless markets. This is most evident in the case of small stocks, where a significant portion of belief distortion, attributed to bid-ask spreads, is in fact primarily due to the costs on the short leg of trades.

4.6 Least Distorted Investor Belief

This section focuses on the characteristics of a specific investor belief that is 1) consistent with asset prices, 2) aligned with investor expectations about the aggregate market, and 3) minimally deviating from the objective belief. This least distorted belief serves as a lens to examine biases proposed by behavioral finance literature as explanations for market anomalies, particularly the value-growth anomaly.

The value-growth anomaly, the observation that over long periods value stocks tend to outperform growth stocks, has always received attention in finance literature that provides both risk- or mispricing-based explanations. A predominant mispricing explanation posits that investors display systematic optimism for growth stocks and pessimism for value stocks, often extrapolating future growth from historical patterns, see Lakonishok, Shleifer, and Vishny [1994], Porta, Lakonishok, Shleifer, and Vishny [1997] and Piotroski and So [2012], among others.

The identified investor belief in this study is suitable for probing whether investors indeed exhibit excessive pessimism towards value stocks and optimism towards growth stocks. This analysis also explores if these biases diminish when trading costs are considered.

To this end, the probability distribution of returns for specific portfolios is investigated: large and growth stocks, small and growth stocks, large and value stocks, and small and value stocks. The study involves three probability distributions: 1) distribution under the objective belief, estimated as the empirical distribution, 2) distribution under the investor belief characterized by equation (13) under frictionless market assumption, and 3) distribution under the investor belief with a 1% annual bid-ask spread assumption.[10] Figure 7 displays the probability distributions of annual returns for these portfolios under both objective and subjective beliefs. These beliefs are aligned with one-year ahead market return expectations of professional forecasters and the pricing of the assets in question.

A key finding is that a difference in the book-to-market ratio notably separates the subjective and objective return distributions, more so than a difference in size. The subjective probability distribution for value stocks (both small and large) leans left compared to the objective distribution, indicating a pronounced pessimistic bias towards value stocks. This

[10]The probability distribution under the investor belief is a transformation of the empirical distribution under the objective belief weighted by the change of belief given by equation (13).

aligns with the mispricing explanation of the value/growth effect. Notably, the probability of a negative net return for value assets is about 39% under the least distorted investor belief, while this probability is 29% under the objective distribution. This tilt in the subjective probability distortion persists across different levels of trading costs, suggesting that trading costs are not the primary drivers of the disparities between objective and subjective value stock return distributions.

In contrast, Figure 7 reveals no substantial bias in the probability distribution of growth stock returns. The probability of a negative net return is roughly 35% for growth stocks under both subjective and objective beliefs, challenging the notion of investor over-optimism towards growth stocks. Additionally, as trading costs are factored in, the distributions under investor and objective beliefs converge.

In conclusion, investor beliefs, extracted from both asset prices and survey expectations, support the idea of undue pessimism towards value stocks but do not corroborate the literature's assertion of excessive investor optimism towards growth stocks.

4.7 Robustness

In this section, I revisit some of the results in the previous sections, under alternative settings. So far, investor beliefs are identified using pricing relation (3) in which the asset pricing model is given by the CAPM. In what follows, I analyze investor belief distortion when beliefs are implied by two other asset pricing models, namely, power utility-based SDF and a recursive utility-based SDF.[11]

Figure 8 reports the fraction of investor belief distortion, estimated assuming frictionless

[11] The power utility-based SDF is $M_{t+1} = R_{M,t+1}^{-\lambda}$, where λ is the price of market risk. The recursive utility-based SDF reads: $M_{t+1} = \beta \left(\frac{C_{t+1}}{C_t}\right)^{-\rho} R_{M,t+1}^{\rho-\lambda}$, where C_{t+1}/C_t is consumption growth, β is time discount parameter, $1/\rho$ is the elasticity of intertemporal substitution and λ is the risk aversion parameter.

markets, that can be explained by a 0.5% bid-ask spread. The results are based on investor beliefs that are identified by matching survey expectations from the Livingston survey and asset pricing moment conditions for portfolios of large, small, growth, value, winner, and loser stocks.

When investor beliefs are identified using the CAPM SDF or the recursive utility-based SDF, a 0.5% bid-ask spread explains a similar fraction of the belief distortion computed assuming frictionless markets: 35% for the CAPM SDF and 36% for the recursive utility-based SDF. However, when the power utility-based SDF is used to identify investor beliefs, the component explained by a 0.5% bid-ask spread is 26%. Moreover, in all the settings, bid-ask spreads explain the largest fraction of the belief distortion for small stock investors.

As Figure 8 illustrates, the extent to which transaction costs explain belief distortions in different market segments varies with the SDF model used. Nonetheless, in all scenarios, the component of the belief distortion estimated assuming frictionless markets that is explained by a 0.5% bid-ask spread is economically important.

Figure 9 compares results using χ^2 and Entropy divergences to quantify the deviation of investor beliefs from the objective belief. The Entropy divergence between beliefs $\mathbb{P}$ and $\mathbb{P}_0$ is defined as $entr(\mathbb{P}, \mathbb{P}_0) := E_{\mathbb{P}_0}[N \log(N)]$, where $N = \frac{d\mathbb{P}}{d\mathbb{P}_0}$.

As Figure 9 reveals, whether using χ^2 or Entropy divergence, the findings are virtually the same regarding the component of belief distortion, computed in frictionless markets, that can be attributed to transaction costs.

5 Conclusions

This paper quantifies the extent to which investor beliefs must diverge from Rational Expectations (RE) to align with asset prices and survey data in a market characterized by frictions. A key innovation of the paper is to identify investor beliefs using jointly asset prices and expectations about the aggregate market from survey data. Moreover, I provided a metric of deviation from RE that is interpretable as an upper bound for the difference between the maximum Sharpe Ratios according to the investor and the econometrician. This approach facilitates an economic assessment of how varying trading costs influence the distortion of investor subjective beliefs.

The measurement of the deviation from RE is based on satisfying the asset pricing conditions for annual returns of portfolios with diverse characteristics and matching the one-year ahead market return expectations of individual investors, CFOs, and professional forecasters. The results reveal that for CAPM-driven investor beliefs, about 30% (46%) of the belief distortion assumed in frictionless markets can be attributed to a 0.5% (1%) annual trading cost.

I document significant heterogeneity in terms of the belief distortion magnitude that can be explained by market frictions, across marginal investors of different segments of the equity market. In particular, I show that 60% of the belief distortion for investors in small stocks computed assuming frictionless markets can be explained by a 0.5% annual trading cost. In the cross-section of survey market return expectations, I find that, while deviations from RE are always significant, they are an order of magnitude smaller for the beliefs of professional forecasters compared to the beliefs of individual investors and CFOs. Moreover, CFOs' expectations about the aggregate market are not more closely aligned with asset prices than those of individual investors.

When comparing costs relative to long and short legs of trades, it emerges that the costs linked to short selling, are critical factors in explaining the deviations from RE that are initially assumed in models of frictionless markets. The effect is most striking for small stocks, where a substantial 80% of the reduction in belief distortion, initially attributed to bid-ask spreads, is in fact attributable to costs incurred on the short side of trades.

Finally, comparing the probability distributions of returns across portfolios with varying characteristics under investor beliefs, I find strong evidence for overly pessimistic subjective expectations of value stock returns. However, the results do not corroborate the commonly held view in the literature of excessive investor optimism towards growth stocks.

Tables

	No Transaction Cost		Costly Short Selling (0.5%)		Bid-Ask Spread (0.5%)	
	δ	δ^{SR}	δ	δ^{SR}	δ	δ^{SR}
Large	5.36 [3.14,7.94]	49.08	3.33 [1.06,5.48]	38.70	2.15 [0,50,3.91]	31.09
Small	10.63 [7.74,14.73]	59.74	3.58 [1.23,5.99]	34.71	1.62 [0,3.33]	23.30
Growth	10.62 [7.96,14.38]	64.69	6.22 [2.79,9.11]	49.51	4.07 [2.02,5.94]	40.07
Value	12.05 [8.58,15.99]	54.38	7.37 [3.95,9.99]	42.52	4.63 [0,7.02]	33.70
Winners	12.28 [7.53,16.42]	56.83	9.42 [5.59,13.01]	49.76	7.78 [4.32,10.84]	45.23
Losers	53.90 [42.39,63.24]	100.20	47.49 [37.27,56.63]	94.06	39.42 [31.89,47.05]	85.69

Table 1: **Investor Belief Distortion and Trading Costs.**
This table reports the deviation from Rational Expectations necessary to pricing various segments of the equity market under CAPM, and to align expectations about the aggregate market of professional forecasters. Assets are given by portfolios of securities reported in rows and market return expectations are from Livingston survey data. Columns denominated with δ report the χ^2-divergence between the objective belief and investor beliefs along with the corresponding bootstrapped 95% confidence intervals. Columns denominated with δ^{SR} report the upper bound for the difference between maximum Sharpe Ratios under the investor and objective beliefs as a fraction of the Sharpe Ratio under the objective belief. Investor beliefs are identified assuming three different market settings: frictionless markets, markets with a 0.5% short selling costs, and markets in presence of a 0.5% per annum (half) bid-ask-spread, i.e., imposing a 0.5% cost on both long and short legs of each trade. CAPM is used as investor SDF. Bootstrapped 95% confidence intervals are reported in brackets. The data runs from 1960 to 2019.

Figures

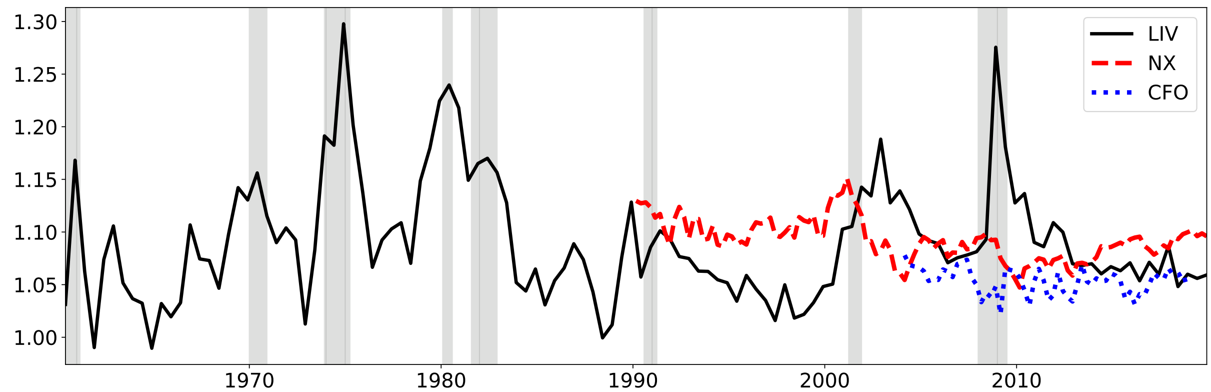
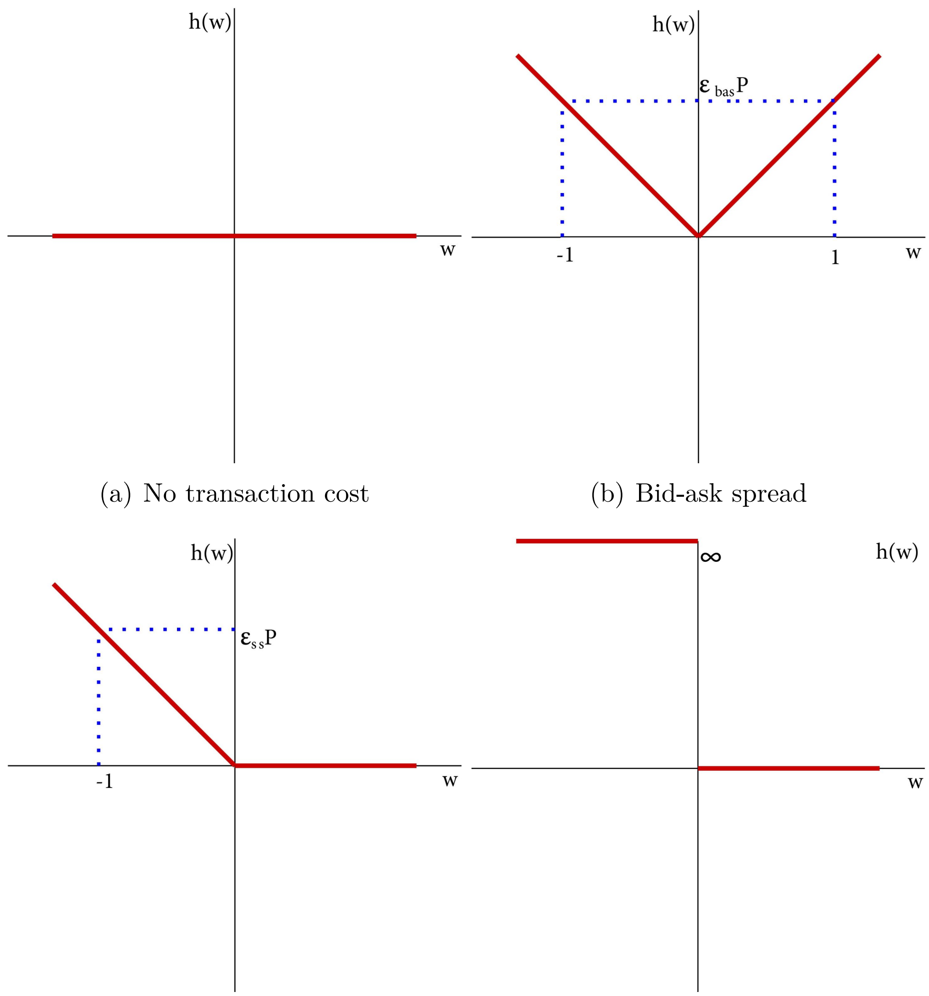


Figure 1: **Survey Expectations.**

This figure plots the time series of the one-year ahead market return expectations reported in three survey datasets. The black solid line corresponds to semiannual observations of expectations, averaged across survey participants, in Livingston survey data from 1960 to 2019. The red dashed line refers to quarterly observations of investor expectations from 1990 to 2019 from Nagel and Xu [2022] constructed by combining surveys from different sources, including UBS/Gallup survey, the Conference Board survey, and the Michigan Survey of Consumers. Quarterly observations of Graham-Harvey CFO survey expectations are given by the blue dotted line from 2005 to 2019. Shaded areas are NBER recession periods.

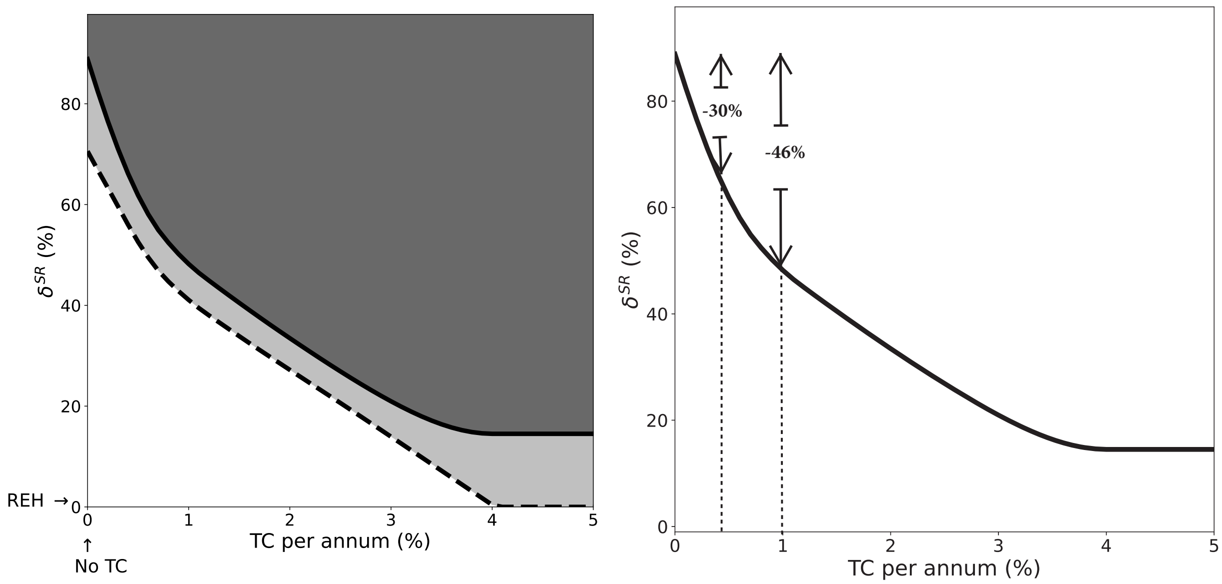


(a) No transaction cost (b) Bid-ask spread

(c) Costly short selling (d) No short selling

Figure 2: **Transaction Cost Functions.**

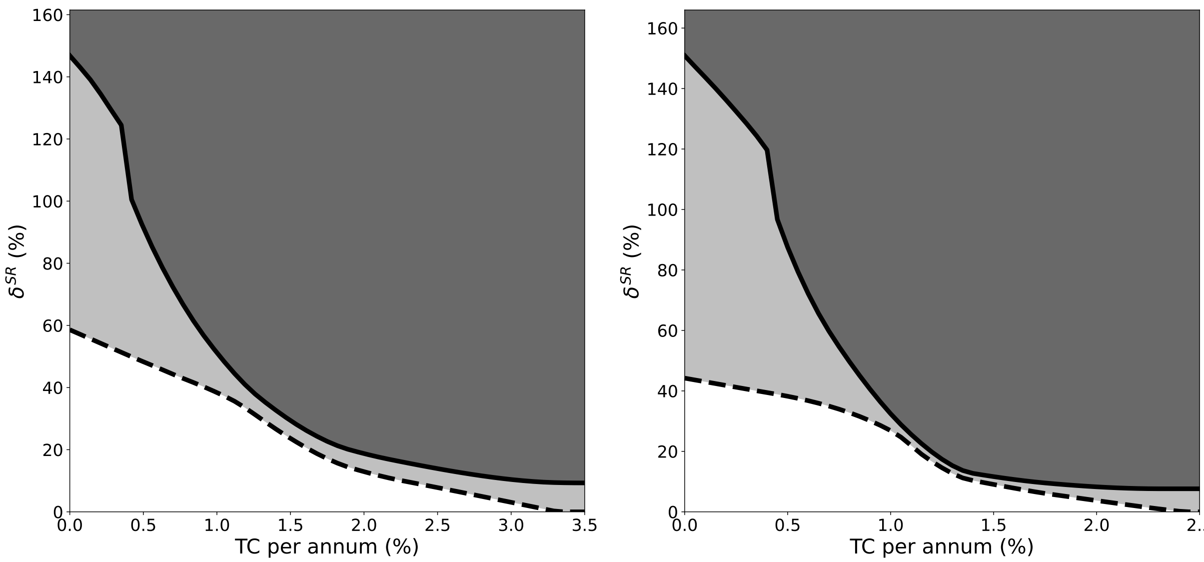
This figure plots specifications of transaction cost function $h(w)$. Panel (a) reports the no-transaction cost case, $h(w) = 0 \; \forall w$, panel (b) reports transaction cost given by bid-ask spread, in which both long and short positions entail trading cost given by fraction ϵ_{bas} of the price wP. Panel (c) corresponds to the case in which only short positions have trading costs equal to a fraction ϵ_{ss} of the price wP. Panel (d) corresponds to the case in which short selling is not possible.



(a) Belief Distortion

(b) Belief Distortion Reduction

Figure 3: **Investor Belief Distortion and Transaction Costs.**
This figure plots the minimum belief distortion necessary to accommodate asset prices (dashed line) and both asset prices and survey expectations (solid line) as functions of bid-ask spread. The y-axis reports the upper bound for the difference between maximum Sharpe Ratios under the investor and objective beliefs as fraction of the Sharpe Ratio under the objective belief. The two curves in Panel (a) delineate three regions, 1) white area where there are no beliefs consistent with asset prices, 2) light gray area where there exist beliefs consistent with asset prices or survey expectations but not both and, 3) dark gray area where there exist beliefs consistent with both asset prices and survey expectations. Panel (b) reports the reduction in belief distortion when 0.5% and 1% annual transaction is assumed. Annual returns of portfolios formed on size and book-to-market, and market return expectations from Livingston survey data are used to identify investor beliefs. CAPM is used as investor SDF. The data runs from 1960 to 2019.



(a) Individual Investors

(b) CFOs

Figure 4: **Investor Belief Distortion of Individual Investors and CFOs.** This figure plots the minimum belief distortion necessary to accommodate asset prices (dashed line) and both asset prices and survey expectations (solid line) as functions of bid-ask spread. The y-axis reports the upper bound for the difference between maximum Sharpe Ratios under the investor and objective beliefs as fraction of the Sharpe Ratio under the objective belief. The two curves delineate three regions, 1) white area where there are no beliefs consistent with asset prices, 2) light gray area where there exist beliefs consistent with asset prices but not with survey expectations, and 3) dark gray area where there exist beliefs consistent with both asset prices and survey expectations. Yearly returns of portfolios formed on size and book-to-market and market return expectations from Nagel and Xu [2022] data in panel (a) and CFO data in panel (b) are used to identify investor beliefs. CAPM is used as investor SDF. The data runs from 1990 to 2019 in panel (a) and from 2005 to 2019 in panel (b).

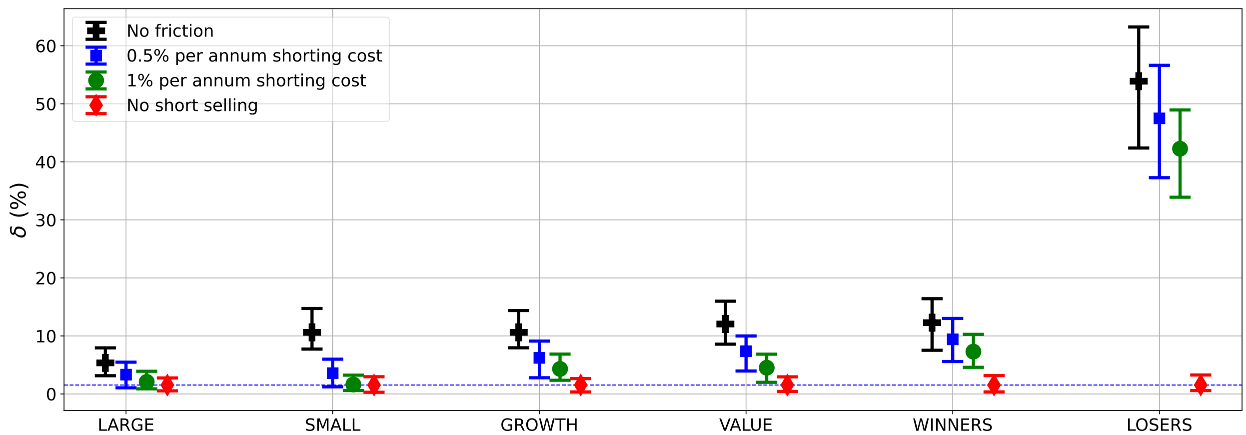


Figure 5: **Investor Belief Distortion and Short Selling Costs.**
This figure plots investor belief distortion, measured by the minimum χ^2-divergence necessary to accommodate annual returns of assets listed on the x-axis and survey expectations from Livingston survey data. Investor beliefs are identified assuming frictionless markets (black cross), assuming a 0.5% per annum short selling cost (blue square), a 1% per annum short selling cost (green circle), and by eliminating short selling (red diamond). The vertical intervals represent bootstrapped 95% confidence intervals. The blue horizontal dashed line corresponds to the distortion of investor beliefs consistent with only survey expectations. CAPM is used as investor SDF. The data runs from 1960 to 2019.

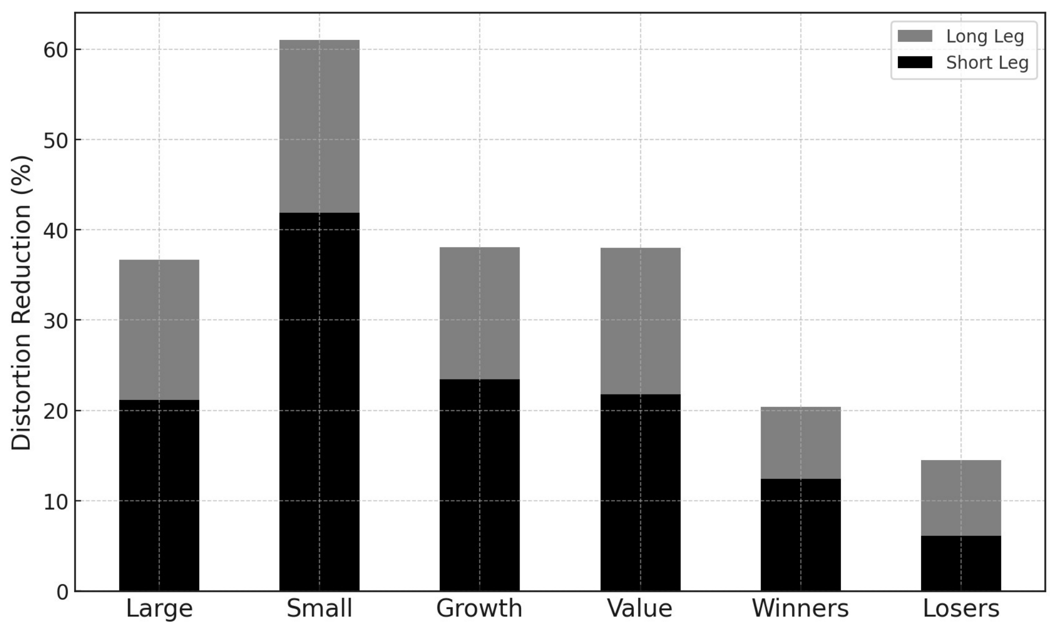
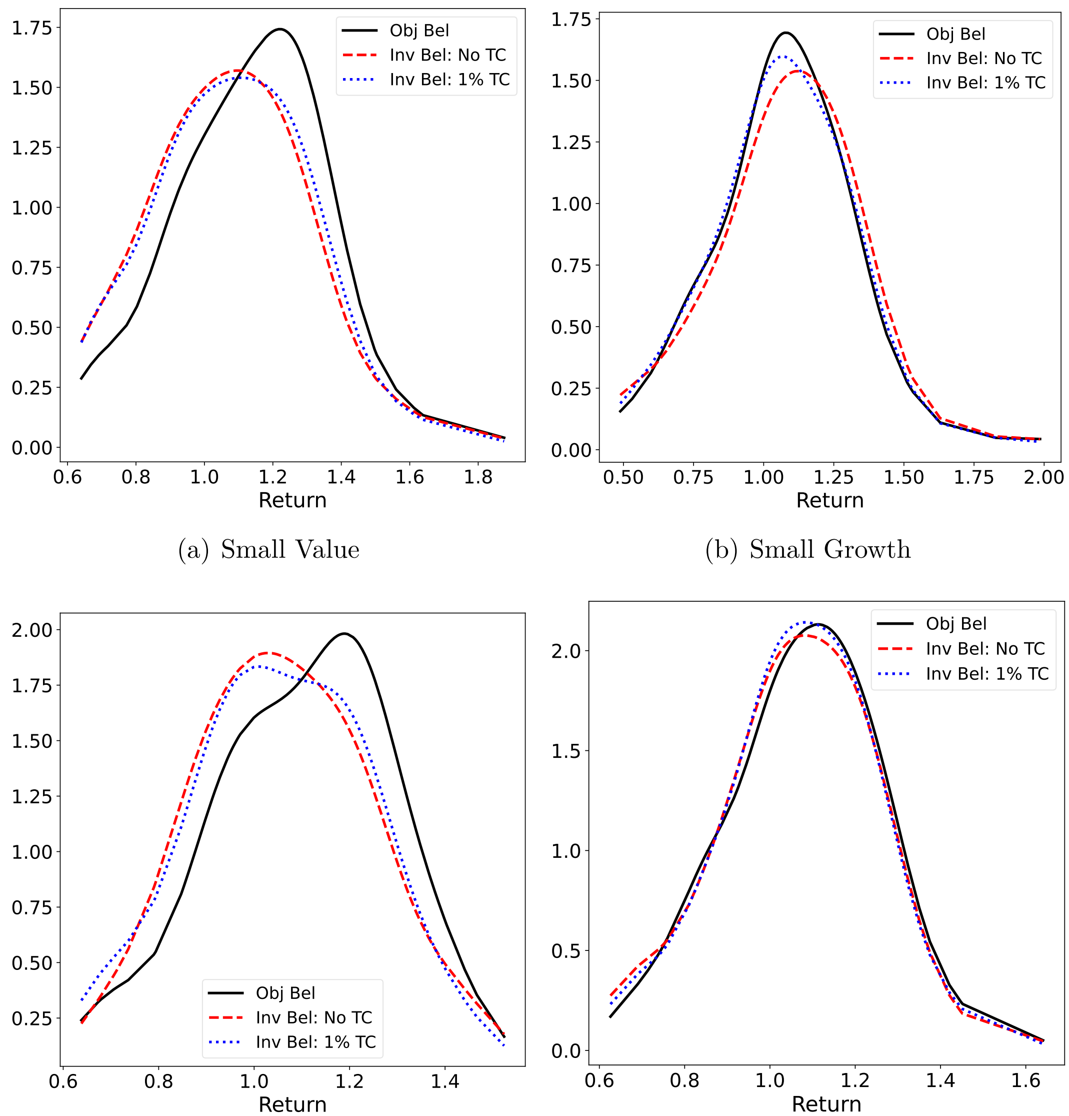


Figure 6: **Investor Belief Distortion and Costs Related to Long and Short Legs of Portfolios.**
This figure demonstrates the reduction in the deviation from Rational Expectations when incorporating a 0.5% short selling cost (Black) and an additional 0.5% cost for the long leg of trades (gray). The distortion is measured using the maximum Sharpe Ratio deviation bound, necessary to accommodate asset prices and survey expectations. The y-axis reports the reduction in belief distortion when accounting for trading costs, i.e., $(\delta_0^{SR} - \delta^{SR})/\delta_0^{SR}$, where δ_0^{SR} is investor belief distortion measured assuming frictionless markets, δ^{SR} is instead investor belief distortion assuming the presence of short selling cost or bid-ask spreads. Market return expectations are from Livingston survey data. CAPM is used as investor SDF. The data runs from 1960 to 2019.



(a) Small Value

(b) Small Growth

(c) Large Value

(d) Large Growth

Figure 7: **Probability Distribution of Portfolio Returns.**
This figure plots the probability distribution of the annual return of a portfolio of small value assets in panel (a), small growth assets in panel (b), large value assets in panel (c), and large growth assets in panel (d). Least distorted investor beliefs are identified by satisfying prices of portfolios formed on size and book-to-market and matching market return expectations from the Livingston survey data. Red-dashed lines represent settings where frictionless markets are assumed and blue-dotted lines assuming an annual 1% trading cost. Black solid line represents the probability distributions under the objective belief. Market return expectations are from Livingston survey data. The data runs from 1960 to 2019.

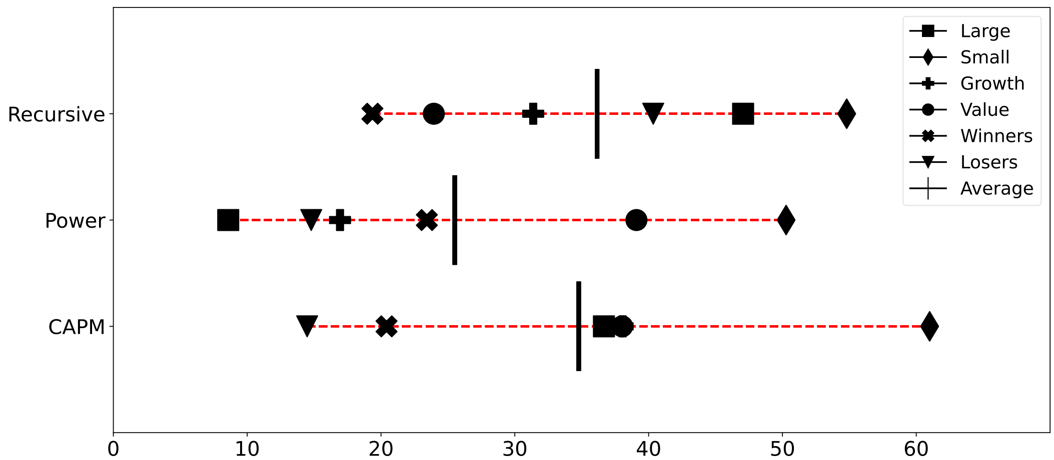


Figure 8: **Asset Pricing Model Comparison**

This figure compares three asset pricing models used to identify investor beliefs which are functions of the market return $R_{M,t+1}$, 1) CAPM: $M_{t+1} = a + bR_{M,t+1}$, 2) Power utility-based SDF: $M_{t+1} = R_{M,t+1}^{-\lambda}$, where λ is the price of market risk, and 3) Recursive utility-based SDF: $M_{t+1} = \beta \left(\frac{C_{t+1}}{C_t}\right)^{-\rho} R_{M,t+1}^{\rho-\lambda}$, where C_{t+1}/C_t is consumption growth, β is time discounting parameter, $1/\rho$ is elasticity of intertemporal substitution and λ is the risk aversion parameter. The x-axis reports the reduction in belief distortion when accounting for trading costs, i.e., $(\delta_0^{SR} - \delta^{SR})/\delta_0^{SR}$, where δ_0^{SR} is investor belief distortion measured assuming frictionless markets, δ^{SR} is instead investor belief distortion assuming a bid-ask spread of 0.5%. Market return expectations are from Livingston survey data. The data runs from 1960 to 2019.

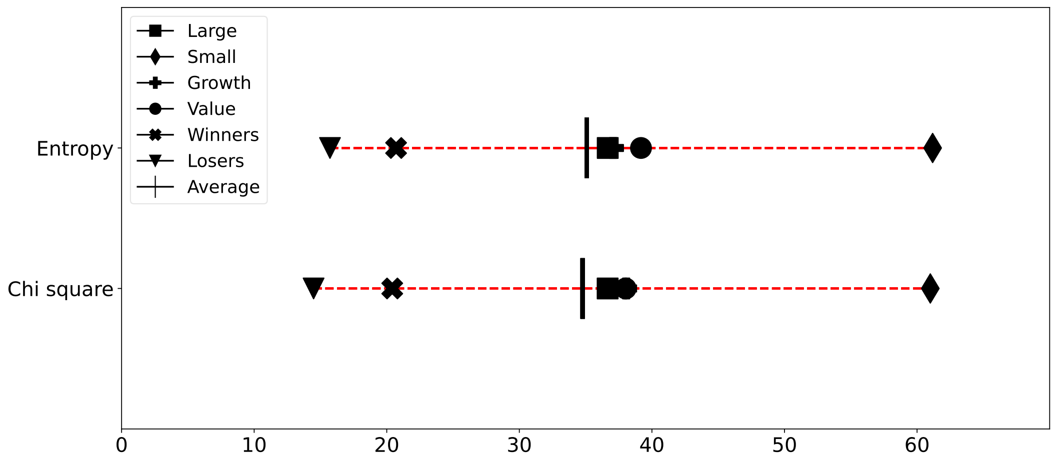


Figure 9: **Distortion Metric Comparison**

This figure compares the χ^2 and the Entropy divergence metrics used to measure investor belief distortion. The x-axis reports the reduction in belief distortion when accounting for trading costs, i.e., $(\delta_0^{SR} - \delta^{SR})/\delta_0^{SR}$, where δ_0^{SR} is investor belief distortion measured assuming frictionless markets, δ^{SR} is instead investor belief distortion assuming the presence of bid-ask spreads of 0.5%. Market return expectations are from Livingston survey data. The data runs from 1960 to 2019.

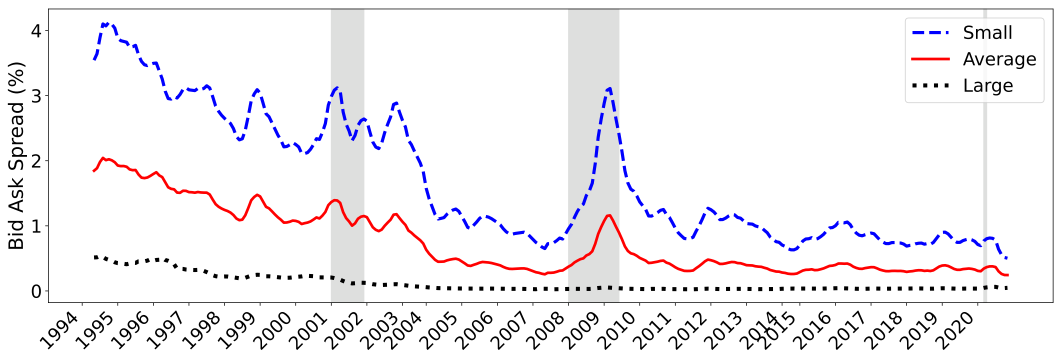
Appendix A Trading Cost Evidence

The literature has proposed both high-frequency measurements of the bid-ask spread using Trade and Quote (TAQ) data from the New York Stock Exchange (NYSE) and low-frequency approximations using end-of-day quotes from the Center for Research in Security Prices (CRSP). Figure 10 plots the half effective bid-ask spread of firms included in both Trade and Quote (TAQ) and Center for Research in Security Prices (CRSP) databases. The effective spreads are estimated following Holden and Jacobsen [2014].[12] The pooled average bid-ask spread for stocks in both databases is of 0.75% from 1994 to 2020. However, there is a large variation both over time and across assets as reported in Figure 10. Panel (a) shows the difference in terms of the bid-ask spread between large and small stocks, given by the top and bottom tertiles. The average bid-ask spread is 0.1% for large assets and 1.7% for small assets, respectively. Value stocks have wider bid-ask spreads relative to Growth stocks as shown in panel (b). The average bid-ask spread for value stocks is 1.04% and 0.58% for growth stocks, respectively. As reported in panel (c), the difference between bid-ask spreads for loser and winner stocks is less pronounced. The average bid-ask spread for loser stocks is 0.94% and for winner stocks is 0.72%.

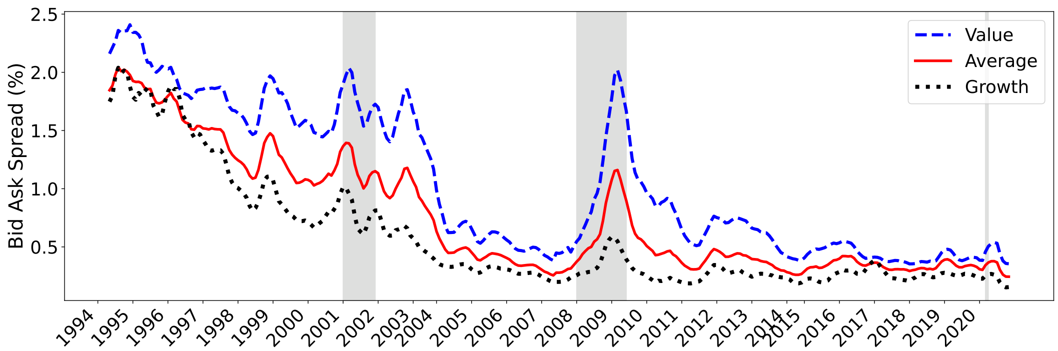
One of the costs investors face when short selling is the loan fee. Saffi and Sigurdsson [2011] document that the average loan fee is close to 0.68% per annum with a standard deviation of 1.61%, by analyzing lending activities between 2005 and 2008 of large firms in the US security lending industry. Cohen, Diether, and Malloy [2007] show that between 1999 and 2003 the average loan fee is 2.60% per annum based on lending data of a large market maker in small stock lending market. D'avolio [2002] finds that the value-weighted loan fee is 0.25% per annum using information on the loan supply of a large financial institution from 2000 to 2001. Moreover, loan fees vary across assets based on various characteristics. Cohen, Diether, and Malloy [2007] find that stocks above (below) the NYSE median market capitalization have an average lending fee of 0.39% (3.94%) per annum. Geczy, Musto, and Reed [2002] show that the average fee for growth (loser) stocks is higher than value (winner) stocks based on data from a US custody bank between 1998 and 1999. Saffi and Sigurdsson [2011] similarly report that firms with higher book-to-market ratios have lower loan fees.

An important limitation to short selling is the limited lending supply of stocks. D'avolio [2002] documents that 16% of the stocks in the CRSP database, mostly small, are impossible to short. Saffi and Sigurdsson [2011] find that the average lending supply amounts only to 23.5% of total shares outstanding and that less than 2% of market capitalization is available to borrow for about 25% of firms. These empirical findings imply that limited lending supply is the main obstacle investors face when short selling.

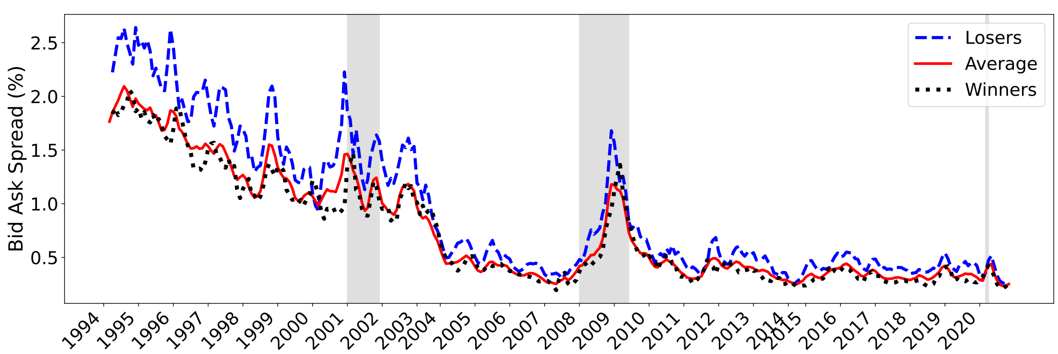
[12] The authors provide the SAS codes on their Web page.



(a) Size



(b) Book to Market



(c) Momentum

Figure 10: **Effective Half Bid-Ask Spreads.** This figure plots the time series of the effective half bid-spreads computed following Holden and Jacobsen [2014], using the Daily and Monthly Trade and Quote database. Bid-ask spreads are averaged across all assets in both TAQ and CRSP databases across large and small assets separately in panel (a), across growth and value assets in panel (b), and across winner and loser assets in panel (c). The solid red line represents the pooled average half bid-ask spread. Gray bars indicate recessions according to NBER. Time series are six-month moving averages calculated from daily data running from 1994 to 202.

Appendix B Inference

In this section, I provide an estimator of the investor belief distortion, δ, and derive its asymptotic distribution by using the dual characterization in equation (12). I provide the framework for the general setting in which investors' expectation $E_{\mathbb{P}}[\boldsymbol{Y}_{t+1}]$ belongs to a convex set $D \subset \mathbb{R}^k$. This allows to account for measurement error by setting, for example, by setting D equal to the 95% confidence interval of the survey expectation $\mathcal{E}[\boldsymbol{Y}_{t+1}]$.

Let us collect all random variables in one vector as $\boldsymbol{v}_{t+1}(\boldsymbol{\theta}) = (M(\boldsymbol{\theta}, \boldsymbol{Z}_{t+1})\boldsymbol{X}'_{t+1}, \boldsymbol{Y}_{t+1})'$ and all the dual parameters as $\boldsymbol{u} = (\boldsymbol{w}', \boldsymbol{\alpha}', \gamma)'$. Then by Proposition 3 we can write the belief distortion as:

$$\begin{aligned}&\delta = \inf_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \mathcal{J}(\boldsymbol{\theta}) - 1, \qquad \text{where} \quad \mathcal{J}(\boldsymbol{\theta}) := \sup_{\boldsymbol{u}} F(\boldsymbol{v}_{t+1}(\boldsymbol{\theta}), \boldsymbol{u}) \ , \quad \text{with} \\ &F(\boldsymbol{v}_{t+1}(\boldsymbol{\theta}), \boldsymbol{u}) := E_{\mathbb{P}_0}[-(-M(\boldsymbol{\theta}, \boldsymbol{Z}_{t+1})\boldsymbol{X}'_{t+1}\boldsymbol{w} - \boldsymbol{Y}'_{t+1}\boldsymbol{\alpha} - \gamma)^2_+/4] - \pi(\boldsymbol{X}'_{t+1}\boldsymbol{w}) - \sigma_D(\boldsymbol{\alpha}) - \gamma\end{aligned} \tag{B1}$$

and $\sigma_D(\boldsymbol{\alpha}) := \sup_{\boldsymbol{z}\in D} \boldsymbol{z}'\boldsymbol{\alpha}$. The estimator of δ, which is denoted by $\hat{\delta}$, is defined as the empirical counterpart given by

$$\begin{aligned}&\hat{\delta} := \inf_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \mathcal{J}_T(\boldsymbol{\theta}) - 1, \qquad \text{where} \quad \mathcal{J}_T(\boldsymbol{\theta}) := \sup_{\boldsymbol{u}} F_T(\boldsymbol{v}_{t+1}(\boldsymbol{\theta}), \boldsymbol{u}) \ , \quad \text{with} \\ &F_T(\boldsymbol{v}_{t+1}(\boldsymbol{\theta}), \boldsymbol{u}) := \frac{1}{T}\sum_{t=1}^{T} - \left(-M(\boldsymbol{\theta}, \boldsymbol{Z}_{t+1})\boldsymbol{X}'_{t+1}\boldsymbol{w} - \boldsymbol{Y}'_{t+1}\boldsymbol{\alpha} - \gamma\right)^2_+ - \boldsymbol{P}'_t\boldsymbol{w} - h(\boldsymbol{w}) - \sigma_D(\boldsymbol{\alpha}) - \gamma \ .\end{aligned} \tag{B2}$$

Let us now consider the following assumptions which allow us to derive the asymptotic properties of the estimator δ_T.

Assumption 1. The parameter space Θ for $\boldsymbol{\theta}$ is compact with non empty interior.

Assumption 2. The parametric SDF $M(\boldsymbol{\theta}, \boldsymbol{Z}_{t+1})$ is continuous in $\boldsymbol{\theta}$ for almost every $\boldsymbol{Z}_{t+1}$.

Assumptions 1 and 2 restrict the family of parametric SDF family and its parameter space and will be useful to use the asymptotic properties of $\mathcal{J}_T(\boldsymbol{\theta})$ to derive those of $\hat{\delta}$. The family of parametric SDFs considered in this paper satisfies such assumptions.

Assumption 3. $\{\boldsymbol{v}_{t+1}\}_t$ is i.i.d..

While the DGP restriction in Assumption 3 considers only i.i.d. process, following Chen et al. [2021] it can be relaxed to accommodate stationary and $\beta-$mixing data.

Assumption 4. Given the variable $S_{t+1}(\boldsymbol{\theta}) := (-M(\boldsymbol{\theta}, \boldsymbol{Z}_{t+1})\boldsymbol{X}'_{t+1}\boldsymbol{w}^*(\boldsymbol{\theta}) - \boldsymbol{Y}'_{t+1}\boldsymbol{\alpha}^*(\boldsymbol{\theta}) -$

$\gamma^*(\boldsymbol{\theta}))_+/2$, where $\boldsymbol{u}^*(\boldsymbol{\theta}) := (\boldsymbol{w}^{*'}(\boldsymbol{\theta}), \boldsymbol{\alpha}^{*'}(\boldsymbol{\theta}), \gamma^*(\boldsymbol{\theta}))'$ maximizes $F(\boldsymbol{v}_{t+1}(\boldsymbol{\theta}), \boldsymbol{u})$, the following three processes 1) $\sqrt{T}\left(\frac{1}{T}\sum_t S_{t+1}(\boldsymbol{\theta})M(\boldsymbol{\theta}, \boldsymbol{Z}_{t+1})\boldsymbol{X}_{t+1}\right) - E_{\mathbb{P}_0}[S_{t+1}(\boldsymbol{\theta})M(\boldsymbol{\theta}, \boldsymbol{Z}_{t+1})\boldsymbol{X}_{t+1}]$, 2) $\sqrt{T}\left(\frac{1}{T}\sum_t S_{t+1}(\boldsymbol{\theta})\mathbf{Y}_{t+1} - E_{\mathbb{P}_0}[S_{t+1}(\boldsymbol{\theta})\mathbf{Y}_{t+1}]\right)$, and 3) $\sqrt{T}\left(\frac{1}{T}\sum_t S_{t+1}(\boldsymbol{\theta}) - E_{\mathbb{P}_0}[S_{t+1}(\boldsymbol{\theta})]\right)$ are bounded in probability for any $\boldsymbol{\theta}$.

In the case of differentiable transaction cost function h, Assumption 4 can be substituted with standard equicontinuity condition. Furthermore, let us consider the next high-level Central Limit assumption for the function F_T.

Assumption 5. $\sqrt{T}\left(\frac{1}{T}\sum_t F(\boldsymbol{v}_{t+1}(\boldsymbol{\theta}), \boldsymbol{u}^*(\boldsymbol{\theta})) - E[F(\boldsymbol{v}_{t+1}(\boldsymbol{\theta}), \boldsymbol{u}^*(\boldsymbol{\theta}))]\right)$ uniformly converges weakly to a Gaussian process $\mathcal{G}(\boldsymbol{\theta})$.

The next proposition characterizes the asymptotic distributions of the estimator of investor belief distortion under the above-mentioned assumptions.

Proposition 4. Under Assumptions 1-5, $\sqrt{T}(\hat{\delta}-\delta)$ converges in distribution to $\min_{\boldsymbol{\theta}\in\Theta^*}\mathcal{G}(\boldsymbol{\theta})$, where $\Theta^* := \arg\min \mathcal{J}(\boldsymbol{\theta})$. If the set Θ^* is a singleton, then the estimator converges to a normally distributed variable.

Bootstrap

I now provide, following Chen, Hansen, and Hansen [2020], a suitable bootstrap method for estimating the limit distribution of the belief distortion metric. The following proposition is an application of the numerical delta method, see Hong and Li [2018][Th. 3.1].

Proposition 5. If $\sqrt{T}(\mathcal{J}_T^B(\boldsymbol{\theta}) - \mathcal{J}_T(\boldsymbol{\theta})) \rightsquigarrow_{\mathbb{P}_0} \mathcal{G}(\boldsymbol{\theta})$, where $\mathcal{J}_T^B(\boldsymbol{\theta})$ is a bootstrap version of $\mathcal{J}_T(\boldsymbol{\theta})$, then for some $\epsilon_T \to 0$ with $\epsilon_T\sqrt{T} \to \infty$ we have [13]

$$\Delta_T := \frac{\inf_{\boldsymbol{\theta}}\{\mathcal{J}_T(\boldsymbol{\theta}) + \epsilon_T\sqrt{T}(\mathcal{J}_T^B(\boldsymbol{\theta}) - \mathcal{J}_T(\boldsymbol{\theta}))\} - \inf_{\boldsymbol{\theta}}\mathcal{J}_T(\boldsymbol{\theta})}{\epsilon_T} \rightsquigarrow_{\mathbb{P}_0} \inf_{\boldsymbol{\theta}\in\Theta^*}\mathcal{G}(\boldsymbol{\theta}) \ . \tag{B3}$$

Proposition 5 states that given a consistent bootstrap scheme for $\mathcal{J}_T(\boldsymbol{\theta})$, such consistency

[13] Weakly convergence in probability of process H_T to H ($H_T \rightsquigarrow_{\mathbb{P}_0} H$) means that $\sup_{f\in BL_1}\left|E_{\mathbb{P}_0}[f(H_T)|\{H_t\}_{t=1}^T] - E[f(H)]\right| = o_p(1)$, where BL is the set of bounded Lipschitz functions.

is inherited by $\inf_{\boldsymbol{\theta}} \mathcal{J}_T(\boldsymbol{\theta})$, and the numerical derivative Δ_T converges weakly in probability to the limit law of the belief distortion metric, which can be used to construct a confidence interval for δ. In particular, a $1-\tau$ two-sided confidence interval for δ is given by $[\Delta_T - \frac{1}{\sqrt{T}} c_{1-\tau/2}, \Delta_T + \frac{1}{\sqrt{T}} c_{1-\tau/2}]$, where c_a is the $a-$empirical percentile of Δ_T.

Appendix C Proofs

Proof of Proposition 1

Proof. Transaction cost function h is such that $h(\alpha \boldsymbol{w}) = \alpha h(\boldsymbol{w})$ and $h(\boldsymbol{w}_1 + \boldsymbol{w}_2) \leq h(\boldsymbol{w}_1) + h(\boldsymbol{w}_2)$, which implies that the cost functional, π, is sublinear and that the set of tradeable payoffs, $\mathcal{X} := \{\boldsymbol{X}'_{t+1}\boldsymbol{w} : h(\boldsymbol{w}) < \infty\}$, is a convex cone. Hence, by Chen [2001, Thm. 1, 5 and Cor. 1], if $\boldsymbol{X}_{t+1} \in \mathcal{L}^p(\mathbb{P}_I)$, the absence of $\mathbb{P}_I$-free lunches is equivalent to the existence of ($\mathbb{P}_I$-almost surely) strictly positive SDF, $M_{t+1} \in \mathcal{L}^q(\mathbb{P}_I)$ with $1/q + 1/p = 1$, such that $\mathbb{E}_{\mathbb{P}_I}[M_{t+1}x] \leq \pi(x)$ for any $x \in \mathcal{X}$. By definition of cost functional π and payoff space $\mathcal{X}$, it equivalently follows for any $\boldsymbol{w} \in \mathbb{R}^n$ that $\mathbb{E}_{\mathbb{P}_I}[M_{t+1}\boldsymbol{X}'_{t+1}\boldsymbol{w}] - \boldsymbol{P}'_t\boldsymbol{w} \leq h(\boldsymbol{w})$. $\square$

Proof of Proposition 2

Proof. The variational representation of χ^2-divergence reads:

$$\chi^2(\mathbb{P}, \mathbb{P}_0) = \sup_{R \in \mathcal{R}} \frac{(E_{\mathbb{P}}[R] - E_{\mathbb{P}_0}[R])^2}{Var_{\mathbb{P}_0}(R)} ,$$

where $\mathcal{R} = \mathcal{L}^2(\mathbb{P}_0) \cap \mathcal{L}^1(\mathbb{P})$. Taking the square root gives us

$$\sqrt{\chi^2(\mathbb{P}, \mathbb{P}_0)} = \sup_{R \in \mathcal{R}} \left| \frac{E_{\mathbb{P}}[R]}{\sigma_{\mathbb{P}_0}(R)} - \frac{E_{\mathbb{P}_0}[R]}{\sigma_{\mathbb{P}_0}(R)} \right| .$$

If $\boldsymbol{X}_{t+1} \in \mathcal{L}^2(\mathbb{P})$, the assumption that $Var_{\mathbb{P}}(\boldsymbol{X}_{t+1}) = Var_{\mathbb{P}_0}(\boldsymbol{X}_{t+1})$, implies that excess returns generated by the vector of excess returns of basis assets $\boldsymbol{R}_{t+1} = [X_{t+1,0}/P_{t,0} - R^f_{t+1}, \ldots, X_{t+1,n}/P_{t,n} - R^f_{t+1}]$, where R^f_{t+1} is the risk-free return, are square-integrable under both $\mathbb{P}_0$ and $\mathbb{P}$, thus form a subset of $\mathcal{R}$.

By restricting the feasible set $\mathcal{R}$ to the set generated by the basis returns, we obtain the following inequality

$$\sqrt{\chi^2(\mathbb{P}, \mathbb{P}_0)} \geq \sup_{\boldsymbol{w}} \left| SR_{\mathbb{P}}(\boldsymbol{R}'_{t+1}\boldsymbol{w}) - SR_{\mathbb{P}_0}(\boldsymbol{R}'_{t+1}\boldsymbol{w}) \right| \geq \left| \sup_{\boldsymbol{w}} SR_{\mathbb{P}}(\boldsymbol{R}'_{t+1}\boldsymbol{w}) - \sup_{\boldsymbol{w}} SR_{\mathbb{P}_0}(\boldsymbol{R}'_{t+1}\boldsymbol{w}) \right| ,$$

where $SR_{\mathbb{P}}(\boldsymbol{R}'_{t+1}\boldsymbol{w}) = E_{\mathbb{P}}[\boldsymbol{R}'_{t+1}\boldsymbol{w}]/\sigma_{\mathbb{P}}(\boldsymbol{R}'_{t+1}\boldsymbol{w})$ is the Sharpe Ratio of portfolio $\boldsymbol{w}$ under belief $\mathbb{P}$ and $SR_{\mathbb{P}_0}(\boldsymbol{R}'_{t+1}\boldsymbol{w})$ the Sharpe Ratio under belief $\mathbb{P}_0$.

□

Proof of Proposition 3

In what follows I prove Proposition 3 for the general setting where investor expectation $E_{\mathbb{P}}[\boldsymbol{Y}_{t+1}]$ belongs to a convex set $D \subset \mathbb{R}^k$. This allows to account for measurement error by setting, for example, D equal to the 95% confidence interval of the aggregated survey expectation $\mathcal{E}[\boldsymbol{Y}_{t+1}]$. The proof follows mainly from Korsaye, Quaini, and Trojani [2019].

Proof. Since the beliefs in $\mathcal{P} = \mathcal{P}_h \cap \mathcal{P}_{SE}$ are absolutely continuous with respect to $\mathbb{P}_0$, for each $\mathbb{P} \in \mathcal{P}$ there exists a random variable N such that $N = d\mathbb{P}/d\mathbb{P}_0$. Assuming that $M(\boldsymbol{\theta})\boldsymbol{X}_{t+1} \in \mathcal{L}^p(\mathbb{P}_0)$ for some $p > 1$, and using the fact that $E[(N-1)^2] = E[N^2] - 1$, we can write belief distortion as $\delta = \inf_{\boldsymbol{\theta}} \delta(\boldsymbol{\theta}) - 1$, with

$$\delta(\boldsymbol{\theta}) := \inf_{N \in \mathcal{L}^q} E[\phi_+(N)] \quad \text{s.t. } E[N] = 1, \quad E[NM(\boldsymbol{\theta})\boldsymbol{X}_{t+1}] - \boldsymbol{P}_t \in C_h, \quad E[N\boldsymbol{Y}_{t+1}] \in D, \tag{C4}$$

where $C_h := \{\boldsymbol{\eta} \in \mathbb{R}^n : \ \boldsymbol{\eta}'\boldsymbol{w} \leq h(\boldsymbol{w})\}$, $1/q + 1/p = 1$, all expectations and $\mathcal{L}^p$-spaces are under the objective belief $\mathbb{P}_0$, and $\phi_+(N)$ is defined as N^2 if N is positive and infinity otherwise. For simplicity, I suppress the SDF variable $\boldsymbol{Z}_{t+1}$ from the notation.

Let us rewrite (C4) as follows:

$$\inf_{N \in \mathcal{L}^p} \underbrace{E[\phi_+(N)]}_{s(N)} + \underbrace{\zeta_{\{1\}}(E[N]) + \zeta_{\boldsymbol{P}_t + C_h}(E[NM(\boldsymbol{\theta})\boldsymbol{X}_{t+1}]) + \zeta_D(E[N\boldsymbol{Y}_{t+1}])}_{g(L_{\boldsymbol{\theta}}N)}, \tag{C5}$$

where $\zeta_A(\cdot)$ is indicator function of set A which is equal to zero if the argument belongs to set A and infinity otherwise, $s : \mathcal{L}^q \to \mathbb{R}$ with $s(N) = E[\phi_+(N)]$, $L_{\boldsymbol{\theta}} : \mathcal{L}^q \to \mathbb{R}^{n+1+k}$ with $L_{\boldsymbol{\theta}}(N) = (E[NM(\boldsymbol{\theta})\boldsymbol{X}_{t+1}], E[N], E[N\boldsymbol{Y}_{t+1}])$ and $g : \mathbb{R}^{n+1+k} \to \mathbb{R}$ with $g(\boldsymbol{u}_1, \xi, \boldsymbol{u}_2) = \zeta_{\boldsymbol{P}_t + C_h}(\boldsymbol{u}_1) + \zeta_{\{1\}}(\xi) + \zeta_D(\boldsymbol{u}_2)$.

By assumption, for a given $\boldsymbol{\theta}$ there exists a belief $\tilde{\mathbb{P}}$ equivalent to $\mathbb{P}_0$ such that for any $\boldsymbol{w} \in \mathbb{R}^n$ $E_{\tilde{\mathbb{P}}}[M(\boldsymbol{\theta})\boldsymbol{X}'_{t+1}\boldsymbol{w}] - \boldsymbol{P}'_t\boldsymbol{w} < h(\boldsymbol{w})$, which with further assumption that $E_{\tilde{\mathbb{P}}}[\boldsymbol{Y}_{t+1}] \in ri(D)$, implies that the following set is not empty:

$$\mathcal{N}(\boldsymbol{\theta}) := \{N \in \mathcal{L}^q_{++}, \quad E[N] = 1, \quad E[NM(\boldsymbol{\theta})\boldsymbol{X}_{t+1}] - \boldsymbol{P}_t \in ri(C_h), \quad E[N\boldsymbol{Y}_{t+1}] \in ri(D)\},$$

where $\mathcal{L}^q_{++}$ consists of $\mathbb{P}_0$-a.s. strictly positive elements and ri is the relative interior.[14] This will be used now to prove that

$$L_{\boldsymbol{\theta}}(qri(\text{dom } s)) \cap ri(\text{dom } g) \neq \emptyset \ .^{15} \tag{C6}$$

In particular given some $\tilde{N} \in \mathcal{N}(\boldsymbol{\theta})$, we will show that $L_{\boldsymbol{\theta}}(\tilde{N}) \in L_{\boldsymbol{\theta}}(qri(\text{dom}\, s)) \cap ri(\text{dom}\, g)$. The domain of the function g is given by $(\boldsymbol{P} + C_h) \times \{1\} \times D$, and its relative interior is given by $(\boldsymbol{P} + ri(C_h)) \times \{1\} \times ri(D)$. Since $E[\tilde{N}] = 1$, $E[\tilde{N}M(\boldsymbol{\theta}\boldsymbol{X}_{t+1})] - \boldsymbol{P}_t \in ri(C_h)$ and $E[\tilde{N}f(Y_{t+1})] \in ri(D)$, we have that $L_{\boldsymbol{\theta}}(\tilde{N}) \in ri(\text{dom}\, g)$. Moreover, the fact that $(0, \infty) \subset \text{dom}\, \phi$ by [Borwein and Lewis, 1991, Lemma 2.3] implies that $\text{dom}\, E[\phi(\cdot)] \cap qri(\mathcal{L}^q_+) \subset qri(E[\phi_+(\cdot)])$ and since $\tilde{N} \in \mathcal{L}^q_{++}$ and $E[\phi(\tilde{N})] < \infty$, we obtain that $\tilde{N} \in qri(\text{dom}\, s)$ and thus $L_{\boldsymbol{\theta}}(\tilde{N}) \in L_{\boldsymbol{\theta}}(qri(\text{dom}\, s))$, thus showing that the set in (C6) is not empty.

Since s and g are both convex, closed and lower semi-continuous, by [Borwein and Lewis, 1992, Corollary 4.3], we obtain

$$\delta(\boldsymbol{\theta}) = \max_{\boldsymbol{w}\in\mathbb{R}^n, \boldsymbol{\alpha}\in\mathbb{R}^k, \gamma\in\mathbb{R}} -s^*(L^T_{\boldsymbol{\theta}}(-\boldsymbol{w}, -\boldsymbol{\alpha}, -\gamma)) - g^*(\boldsymbol{w}, \boldsymbol{\alpha}, \gamma) \ ,$$

where s^* and g^* are legendre transforms of s and g, respectively, and $L^T_{\boldsymbol{\theta}}$ is the adjoint operator of the linear operator $L_{\boldsymbol{\theta}}$.[16] The adjoint map corresponding to $L_{\boldsymbol{\theta}}$ is given by $L^T_{\boldsymbol{\theta}}(-\boldsymbol{w}, -\boldsymbol{\alpha}, -\gamma) = -M(\boldsymbol{\theta})\boldsymbol{X}'_{t+1}\boldsymbol{w} - \boldsymbol{Y}'_{t+1}\boldsymbol{\alpha} - \gamma$. Let us now compute s^* and g^*. By [Rockafellar, 1968, Th. 2] $s^* = E[\phi^*_+(\cdot)]$. Moreover, we have $g^*(\boldsymbol{w}, \gamma, \boldsymbol{\alpha}) = \boldsymbol{P}'_t\boldsymbol{w} + h(\boldsymbol{w}) + \sigma_D(\boldsymbol{\alpha}) + \gamma$, where $\sigma_D(\cdot)$ is the support function of set D. Thus, we obtain

$$\delta(\boldsymbol{\theta}) = \max_{\boldsymbol{w}\in\mathbb{R}^n, \boldsymbol{\alpha}\in\mathbb{R}^k, \gamma\in\mathbb{R}} E[-\phi^*_+(-M(\boldsymbol{\theta})\boldsymbol{X}'_{t+1}\boldsymbol{w} - \boldsymbol{Y}'_{t+1}\boldsymbol{\alpha} - \gamma)] - \boldsymbol{P}'_t\boldsymbol{w} - h(\boldsymbol{w}) - \sigma_D(\boldsymbol{\alpha}) - \gamma \ ,$$

therefore

$$\delta = -\sup_{\boldsymbol{\theta}} \inf_{\boldsymbol{w}\in\mathbb{R}^n, \boldsymbol{\alpha}\in\mathbb{R}^k, \gamma\in\mathbb{R}} E[\phi^*_+(-M(\boldsymbol{\theta})\boldsymbol{X}'_{t+1}\boldsymbol{w} - \boldsymbol{Y}'_{t+1}\boldsymbol{\alpha} - \gamma)] + \boldsymbol{P}'_t\boldsymbol{w} + h(\boldsymbol{w}) + \sigma_D(\boldsymbol{\alpha}) + \gamma - 1 \ .$$

Similarly to Korsaye, Quaini, and Trojani [2019], the solution to problem (C4) is given by $N^*(\boldsymbol{\theta}) = \phi^{*'}(-M(\boldsymbol{\theta})\boldsymbol{X}'_{t+1}\boldsymbol{w}(\boldsymbol{\theta}) - \boldsymbol{Y}'_{t+1}\boldsymbol{\alpha}(\boldsymbol{\theta}) - \gamma(\boldsymbol{\theta}))$.

[14] The relative interior of a set C is its interior within the affine hull of C.

[15] $qri(A)$ is the quasi relative interior of set A. $qri(A) := \{x \in A : c\bar{o}ne(A - x)\}$, where $c\bar{o}ne$ is closure of conic hull.

[16] For $L : \mathcal{H} \mapsto \mathcal{K}$, we have that $L^T : \mathcal{K}^* \mapsto \mathcal{H}^*$, such that $\forall X \in \mathcal{K}$ and $\forall Y \in \mathcal{H}$ we have $\langle X, L(Y)\rangle_{\mathcal{K},\mathcal{K}^*} = \langle Y, L^T(X)\rangle_{\mathcal{H},\mathcal{H}^*}$, where H^* and K^* are the dual spaces of $\mathcal{H}$ and K, respectively, with duality pairing $\langle, \rangle_{\mathcal{K},\mathcal{K}^*}$ and $\langle, \rangle_{\mathcal{H},\mathcal{H}^*}$ the corresponding duality pairings.

Moreover, since $\phi_+^*(\boldsymbol{z}) = \boldsymbol{z}_+^2/4$, we have

$$\delta = -\sup_{\boldsymbol{\theta}} \inf_{\boldsymbol{w}\in\mathbb{R}^n, \boldsymbol{\alpha}\in\mathbb{R}^k, \gamma\in\mathbb{R}} E[(-M(\boldsymbol{\theta})\boldsymbol{X}_{t+1}'\boldsymbol{w} - \boldsymbol{Y}_{t+1}'\boldsymbol{\alpha} - \gamma)_+^2]/4 + \boldsymbol{P}_t'\boldsymbol{w} + h(\boldsymbol{w}) + \sigma_D(\boldsymbol{\alpha}) + \gamma - 1\ .$$

In the particular case where the survey expectation is exactly matched, i.e., $E_{\mathbb{P}}[\boldsymbol{Y}_{t+1}] = \mathcal{E}[\boldsymbol{Y}_{t+1}]$, we have $\sigma_D(\boldsymbol{\alpha}) = \mathcal{E}[\boldsymbol{Y}_{t+1}]'\boldsymbol{\alpha}$, and thus we obtain the claims in equations (12) and (13). □

Proof of Proposition 4

Proof. By Assumption 3 F_T converges pointwise to F $\mathbb{P}_0$-almost surely. Moreover, since F_T and F are concave functions in $\boldsymbol{u}$, by [Rockafellar, 1970, Th. 10.8], F_T converges uniformly to F on any compact set in Θ. Hence, by Assumption 1 $\boldsymbol{u}_T(\boldsymbol{\theta}) - \boldsymbol{u}^*(\boldsymbol{\theta}) = o_p(1)$.

Now, given $\boldsymbol{\theta}$, let us write the difference between empirical and population optimized values with respect to the dual parameter $\boldsymbol{u}$ as follows

$$\begin{aligned}\sqrt{T}\left(\mathcal{J}_T(\boldsymbol{\theta}) - \mathcal{J}(\boldsymbol{\theta})\right) = &\underbrace{\sqrt{T}\left(F_T(\boldsymbol{v}_{t+1}(\boldsymbol{\theta}), \boldsymbol{u}_T(\boldsymbol{\theta})) - F_T(\boldsymbol{v}_{t+1}(\boldsymbol{\theta}), \boldsymbol{u}^*(\boldsymbol{\theta}))\right)}_{G_T(\boldsymbol{\theta})} \\ &+ \sqrt{T}\left(F_T(\boldsymbol{v}_{t+1}(\boldsymbol{\theta}), \boldsymbol{u}^*(\boldsymbol{\theta})) - F(\boldsymbol{v}_{t+1}(\boldsymbol{\theta}), \boldsymbol{u}^*(\boldsymbol{\theta}))\right)\end{aligned} \tag{C7}$$

Since F_T is concave in $\boldsymbol{u}$, the first component of the right hand side of equation (C7), $G_T(\boldsymbol{\theta})$, satisfies the following relation

$$G_T(\boldsymbol{\theta}) \leq (\boldsymbol{u}_T(\boldsymbol{\theta}) - \boldsymbol{u}^*(\boldsymbol{\theta}))'\boldsymbol{R}_T(\boldsymbol{\theta}), \qquad \text{for any } \boldsymbol{R}_T(\boldsymbol{\theta}) \in \partial\sqrt{T}F_T(\boldsymbol{v}_{t+1}(\boldsymbol{\theta}), \boldsymbol{u}^*(\boldsymbol{\theta}))\ , \tag{C8}$$

where $\partial\sqrt{T}F_T(\boldsymbol{v}_{t+1}(\boldsymbol{\theta}), \boldsymbol{u}^*(\boldsymbol{\theta}))$ is the subdifferential of $\sqrt{T}F_T$ in $\boldsymbol{u}^*(\boldsymbol{\theta})$, which means that for any $\boldsymbol{\beta}_1(\boldsymbol{\theta}) \in \partial h(\boldsymbol{w}^*(\boldsymbol{\theta}))$ and $\boldsymbol{\beta}_2(\boldsymbol{\theta}) \in \partial\sigma_D(\boldsymbol{\alpha}^*(\boldsymbol{\theta}))$ we have

$$\begin{aligned}G_T(\boldsymbol{\theta}) \leq &\sqrt{T}(\boldsymbol{u}_T(\boldsymbol{\theta}) - \boldsymbol{u}^*(\boldsymbol{\theta}))' \\ &\left(\frac{1}{T}\sum_t S_{t+1}M(\boldsymbol{\theta}, \boldsymbol{Z}_{t+1})\boldsymbol{X}_{t+1} - \boldsymbol{P}_t - \boldsymbol{\beta}_1(\boldsymbol{\theta}), \frac{1}{T}\sum_t S_{t+1}\boldsymbol{Y}_{t+1} - \boldsymbol{\beta}_2(\boldsymbol{\theta}), \frac{1}{T}\sum_t S_{t+1} - 1\right)\ .\end{aligned} \tag{C9}$$

However, since $\boldsymbol{u}^*(\boldsymbol{\theta})$ maximizes $F(\boldsymbol{v}_{t+1}(\boldsymbol{\theta}), \boldsymbol{u})$, we have $\boldsymbol{0} \in \partial F(\boldsymbol{v}_{t+1}(\boldsymbol{\theta}), \boldsymbol{u}^*(\boldsymbol{\theta}))$. Hence, there exist $\boldsymbol{\beta}_1^*(\boldsymbol{\theta}) \in \partial h(\boldsymbol{w}^*(\boldsymbol{\theta}))$ and $\boldsymbol{\beta}_2^*(\boldsymbol{\theta}) \in \partial \sigma_D(\boldsymbol{\alpha}^*(\boldsymbol{\theta}))$ such that

$$\begin{cases} E\left[S_{t+1} M(\boldsymbol{\theta}, \boldsymbol{Z}_{t+1}) \boldsymbol{X}_{t+1} - \boldsymbol{P}_t\right] - \boldsymbol{\beta}_1^*(\boldsymbol{\theta}) = 0 \\ E\left[S_{t+1} \mathbf{Y}_{t+1}\right] - \boldsymbol{\beta}_2^*(\boldsymbol{\theta}) = 0 \\ E\left[S_{t+1}\right] - 1 = 0 \ . \end{cases} \tag{C10}$$

Since inequality (C9) is true for any $\boldsymbol{\beta}_1(\boldsymbol{\theta}) \in \partial h(\boldsymbol{w}^*(\boldsymbol{\theta}))$ and $\boldsymbol{\beta}_2(\boldsymbol{\theta}) \in \partial \sigma_D(\boldsymbol{\alpha}^*(\boldsymbol{\theta}))$, it holds also for $\boldsymbol{\beta}_1^*(\boldsymbol{\theta})$ and $\boldsymbol{\beta}_2^*(\boldsymbol{\theta})$. Hence, by substituiting $\boldsymbol{\beta}_1(\boldsymbol{\theta})$ and $\boldsymbol{\beta}_2(\boldsymbol{\theta})$ in (C9) with $\boldsymbol{\beta}_1^*(\boldsymbol{\theta})$ and $\boldsymbol{\beta}_2^*(\boldsymbol{\theta})$ from (C10), we obtain

$$G_T(\boldsymbol{\theta}) \leq (\boldsymbol{u}_T(\boldsymbol{\theta}) - \boldsymbol{u}^*(\boldsymbol{\theta}))' \boldsymbol{R}_T^*(\boldsymbol{\theta}) \ , \tag{C11}$$

where

$$\begin{aligned} \boldsymbol{R}_T^*(\boldsymbol{\theta}) := \sqrt{T} \Bigg(& \frac{1}{T} \sum_t S_{t+1} M(\boldsymbol{\theta}, \boldsymbol{Z}_{t+1}) \boldsymbol{X}_{t+1} - E_{\mathbb{P}_0}[S_{t+1} M(\boldsymbol{\theta}, \boldsymbol{Z}_{t+1}) \boldsymbol{X}_{t+1}] \ , \\ & \frac{1}{T} \sum_t S_{t+1} \mathbf{Y}_{t+1} - E_{\mathbb{P}_0}[S_{t+1} \mathbf{Y}_{t+1}], \\ & \frac{1}{T} \sum_t S_{t+1} - E_{\mathbb{P}_0}[S_{t+1}] \Bigg) \ , \end{aligned} \tag{C12}$$

which by Assumption 4 is bounded in probability. The fact that $\boldsymbol{R}_T^*(\boldsymbol{\theta})$ is bounded in probability and that $\boldsymbol{u}_T(\boldsymbol{\theta}) - \boldsymbol{u}^*(\boldsymbol{\theta}) = o_p(1)$ implies that $G_T(\boldsymbol{\theta}) = o_p(1)$ and thus

$$\sqrt{T}(\mathcal{J}_T(\boldsymbol{\theta}) - \mathcal{J}(\boldsymbol{\theta})) = \sqrt{T} \left(F_T(\boldsymbol{v}_{t+1}(\boldsymbol{\theta}), \boldsymbol{u}^*(\boldsymbol{\theta})) - F(\boldsymbol{v}_{t+1}(\boldsymbol{\theta}), \boldsymbol{u}^*(\boldsymbol{\theta})) \right) + o_p(1) \ ,$$

which by Assumption 5 converges uniformly to $\mathcal{G}(\boldsymbol{\theta})$. This and Assumption 1 and 2, by [Shapiro, 1991, Th. 3.2], imply that

$$\sqrt{T}(\hat{\delta} - \delta) \longrightarrow \min_{\boldsymbol{\theta} \in \boldsymbol{\Theta}^*} \mathcal{G}(\boldsymbol{\theta}) \ .$$

□

References

Klaus Adam and Stefan Nagel. Expectations data in asset pricing. Technical report, National Bureau of Economic Research, 2022.

Jonathan M Borwein and Adrian S Lewis. Duality relationships for entropy-like minimization problems. *SIAM Journal on Control and Optimization*, 29(2):325–338, 1991.

Jonathan M Borwein and Adrian S Lewis. Partially finite convex programming, part i: Quasi relative interiors and duality theory. *Mathematical Programming*, 57(1):15–48, 1992.

Xiaohong Chen, Lars Peter Hansen, and Peter G Hansen. Robust identification of investor beliefs. *Proceedings of the National Academy of Sciences*, 117(52):33130–33140, 2020.

Xiaohong Chen, Lars Peter Hansen, and Peter G Hansen. Robust inference for moment condition models without rational expectations. *Journal of Econometrics, forthcoming*, 2021.

Zhiwu Chen. Viable costs and equilibrium prices in frictional securities markets. *Annals of Economics and Finance*, 2(2):297–323, 2001.

Lauren Cohen, Karl B Diether, and Christopher J Malloy. Supply and demand shifts in the shorting market. *The Journal of Finance*, 62(5):2061–2096, 2007.

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

Magnus Dahlquist and Markus Ibert. Return expectations and portfolios: Evidence from large asset managers. *Swedish House of Finance Research Paper*, (21-1), 2021.

Gene D'avolio. The market for borrowing stock. *Journal of financial economics*, 66(2-3): 271–306, 2002.

Benjamin Enke and Thomas Graeber. Cognitive uncertainty. Technical report, National Bureau of Economic Research, 2019.

Xavier Gabaix. A sparsity-based model of bounded rationality. *The Quarterly Journal of Economics*, 129(4):1661–1710, 2014.

Christopher C Geczy, David K Musto, and Adam V Reed. Stocks are special too: An analysis of the equity lending market. *Journal of Financial Economics*, 66(2-3):241–269, 2002.

Anisha Ghosh and Guillaume Roussellet. Identifying beliefs from asset prices. In *Proceedings of Paris December 2019 Finance Meeting EUROFIDAI-ESSEC*, 2023.

Anisha Ghosh, Christian Julliard, and Alex P Taylor. What is the consumption-capm missing? an information-theoretic framework for the analysis of asset pricing models. *The Review of Financial Studies*, 30(2):442–504, 2017.

Anisha Ghosh, Arthur G Korteweg, and Qing Xu. Recovering heterogeneous beliefs and preferences from asset prices. *Available at SSRN 3650199*, 2020.

Stefano Giglio, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus. Five facts about beliefs and portfolios. *American Economic Review*, 111(5):1481–1522, 2021.

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

Denis Gromb and Dimitri Vayanos. Limits of arbitrage. *Annu. Rev. Financ. Econ.*, 2(1): 251–275, 2010.

Lars Peter Hansen and Ravi Jagannathan. Implications of security market data for models of dynamic economies. *Journal of political economy*, 99(2):225–262, 1991.

Hua He and David M Modest. Market frictions and consumption-based asset pricing. *Journal of Political Economy*, 103(1):94–117, 1995.

Craig W Holden and Stacey Jacobsen. Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions. *The Journal of Finance*, 69(4):1747–1785, 2014.

Han Hong and Jessie Li. The numerical delta method. *Journal of Econometrics*, 206(2): 379–394, 2018.

Elyes Jouini and Hédi Kallal. Martingales and arbitrage in securities markets with transaction costs. *Journal of Economic Theory*, 66(1):178–197, 1995.

Sofonias A Korsaye, Alberto Quaini, and Fabio Trojani. Smart sdfs. In *Proceedings of Paris December 2019 Finance Meeting EUROFIDAI-ESSEC*, 2019.

Josef Lakonishok, Andrei Shleifer, and Robert W Vishny. Contrarian investment, extrapolation, and risk. *The journal of finance*, 49(5):1541–1578, 1994.

Erzo GJ Luttmer. Asset pricing in economies with frictions. *Econometrica: Journal of the Econometric Society*, pages 1439–1467, 1996.

N Gregory Mankiw and Ricardo Reis. Sticky information versus sticky prices: a proposal to replace the new keynesian phillips curve. *The Quarterly Journal of Economics*, 117(4): 1295–1328, 2002.

Stefan Nagel and Zhengyang Xu. Asset pricing with fading memory. *The Review of Financial Studies*, 35(5):2190–2245, 2022.

Joseph D Piotroski and Eric C So. Identifying expectation errors in value/glamour strategies: A fundamental analysis approach. *The Review of Financial Studies*, 25(9):2841–2875, 2012.

Rafael La Porta, Josef Lakonishok, Andrei Shleifer, and Robert Vishny. Good news for value stocks: Further evidence on market efficiency. *the Journal of Finance*, 52(2):859–874, 1997.

Ralph Rockafellar. Integrals which are convex functionals. *Pacific journal of mathematics*, 24(3):525–539, 1968.

Ralph Tyrell Rockafellar. *Convex analysis*. Princeton university press, 1970.

Pedro AC Saffi and Kari Sigurdsson. Price efficiency and short selling. *The Review of Financial Studies*, 24(3):821–852, 2011.

Thomas J Sargent. *The conquest of American inflation*. Princeton University Press, 2001.

Alexander Shapiro. Asymptotic analysis of stochastic programs. *Annals of Operations Research*, 30(1):169–186, 1991.

Christopher A Sims. Implications of rational inattention. *Journal of monetary Economics*, 50(3):665–690, 2003.

Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, 1996.

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