Learning Time Varying Risk Preferences from Investment Portfolios using Inverse Optimization with Applications on Mutual Funds

Shi Yu

The Vanguard Group, Malvern, PA, USA, shi.yu@hotmail.com

Yuxin Chen

The Vanguard Group, Malvern, PA, USA

Chaosheng Dong*

Amazon, Seattle, WA, USA

The fundamental principle in Modern Portfolio Theory (MPT) is based on the quantification of the portfolio's risk related to performance. Although MPT has made huge impacts on the investment world and prompted the success and prevalence of passive investing, it still has shortcomings in real-world applications. One of the main challenges is that the level of risk an investor can endure, known as *risk-preference*, is a subjective choice that is tightly related to psychology and behavioral science in decision making. This paper presents a novel approach of measuring risk preference from existing portfolios using inverse optimization on the mean-variance portfolio allocation framework. Our approach allows the learner to continuously estimate real-time risk preferences using concurrent observed portfolios and market price data. We demonstrate our methods on real market data that consists of 20 years of asset pricing and 10 years of mutual fund portfolio holdings. Moreover, the quantified risk preference parameters are validated with two well-known risk measurements currently applied in the field. The proposed methods could lead to practical and fruitful innovations in automated/personalized portfolio management, such as Robo-advising, to augment financial advisors' decision intelligence in a long-term investment horizon.

1. Introduction

Risk preference (risk tolerance or risk aversion) is a fundamental concept modeling individual preference for certainty under uncertainty. In portfolio allocation, one primary goal has been to reconcile empirical information about securities prices with theoretical models of asset pricing under conditions of inter-temporal uncertainty (Cohn 1975). The notion of risk preference has been an essential assumption underlying almost all such models (Cohn 1975). However, measurement of risk preference has been treated in separate paths. In finance domain, quantification of risk can be roughly summarized as ratio comparisons between wealth and risky assets under market

* Work done prior to joining Amazon

volatility. However, the underlying biological, behavioral, and social factors behind risk appetite are commonly studied in many other disciplines, i.e., social science (Payne et al. 2017, Guiso and Paiella 2008), behavior science (Brennan and Lo 2011), mathematics (von Neumann and Morgenstern 1947), psychology (Sokol-Hessner et al. 2009, Mcgraw et al. 2010), and genetics (Linnér 2019). For more than half of a century, many measures of risk preference have been developed in various fields, including curvature measures of utility functions (Arrow 1971, Pratt 1964), human subject experiments and surveys (Rabin and Thaler 2001, Holt and Laury 2002), portfolio choice for financial investors (Guiso and Paiella 2008), labor-supply behavior (Chetty 2006), deductible choices in insurance contracts (Cohen and Einav 2007, Szpiro 1986), contestant behavior on game shows (Post et al. 2008), option prices (Aït-Sahalia and Lo 2000) and auction behavior (Lu and Perrigne 2008). Nowadays, investor-consumers' risk preferences are mainly investigated through one or combination of three ways. The first one is assessing actual behavior. Two examples are inferring households' risk attitudes using regression analysis on historical financial data (Schooley and Worden 1996) and inferring investors' risk preferences from their trading decision using reinforcement learning (Alsabah et al. 2020, Wang et al. 2020). The second one is assessing responses to hypothetical scenarios about investment choices (see Barsky et al. (1997) and (Hey 1999)). In practice, online questionnaires are widely adopted by Robo-advising firms to evaluate investors' risk profiles (Alsabah et al. 2020). The third one is subjective questions (see Hanna et al. (1998) for a survey of these different techniques).

Despite its profound importance in economics, there remain some limitations with respect to learning the risk preference using classic approaches. First, many existing methods assume inherent small deviations of the risk preference, which is "*valid only for potential losses that are relatively unthreatening to the individuals' wealth*" (Thomas 2016). Second, most practices in place are often insufficient to deal with scenarios where investment decisions are managed by machine learning processes (Robo-advising), and risk preferences are expected as input parameters that can generate new decisions directly (auto-rebalancing). Currently, Robo-advisors first communicate and categorize clients' risk preferences based on human interpretation, and later map them to the nearest values of a finite set of representative risk preference levels (Capponi et al. 2019). Those limitations make existing theories and approaches challenging to deal with prominent situations in which risk preference changes dramatically in an inter-temporal dimension, such as savings, investment, consumption problems, dynamic labor supply decisions, and health decisions (O'Donoghue and Somerville 2018). Real-world risk preference is clearly not as straightforward as many theories have assumed, and perhaps individuals even do not exhibit risk appetite consistently in their behaviors across domains (O'Donoghue and Somerville 2018). Namely, risk preference is time varying in reality. Nowadays, due to the technological advances, we already have overwhelming behavioral data

across all domains, providing a myriad of additional sources to help us decipher the perplexity of risk preference from different angles. Although these new approaches might not ultimately prove to be the best models for studying risk preference, they can be used in conjunction with traditional methods with additional, more nuanced implications that are borne out by data (O'Donoghue and Somerville 2018).

To tackle aforementioned limitations, we present a novel inverse optimization approach to measure risk preference directly from market signals and portfolios. In a volatile market, the same portfolio created by *buy and hold* investors could reflect very different risk levels when market conditions change, regardless most of the time, investor's subjective risk preference remains unchanged. Our approach's primary motivation is to complement traditional approaches through continuous monitoring and evaluation of embedded portfolio risks and to facilitate the automated decision process of portfolio adjustment when necessary to ensure that real portfolio risk is aligned with investor's true risk preference. Our method is developed based on two fundamental methodologies: convex optimization based Modern Portfolio Theory (MPT) and learning decision-making scheme through inverse optimization. We assume investors are rational and their portfolio decisions are near-optimal. Consequently, their decisions are affected by the risk preference factors through portfolio allocation model. We propose an inverse optimization framework to infer the risk preference factor that must have been in place for those decisions to have been made. Furthermore, we assume risk preference stays constant at the point of decision, but varies across multiple decisions over time, and can be inferred from joint observations of time-series market price and asset allocations.

Our inverse optimization approach represents an integral component of a general view of using machine learning to learn individual investor's risk preference, which embodies data and models from three aspects. The first aspect is leveraging demographic features, such as education, financial status, gender, age, to learn risk preference. Such type of information is collected once, and its impact on risk preference can be learned by applying statistical models on a large amount of client data. The second aspect reflects influences on risk preference from lifetime events, such as mortgage payment, disaster loss, accidental medical conditions and expense. These significant events typically have huge impacts on investment goals or contribution plans, leading to subsequent changes in risk preferences. The third aspect is extracting insights from consumer financial behavior, which involves understanding how investor-consumers make financial decisions, and how these decisions are reflected in the interactions of financial products. The data that captures these behaviors is probably the best source to investigate the underlying risk preferences behind decisions. In an automated financial advice system, risk evaluation coming from the first and second aspects are often sporadic, whereas the estimation from the third aspect is real-time insights based on continuous monitoring and measurement. Therefore, a successful investment management system shall always

compare and validate results from all three components and ensure their consistency throughout the entire investment horizon. Unfortunately, the third aspect was considered intractable until the recent advances of big data and machine learning offer new approaches to solve those challenging problems. Our approach offers a tractable way to monitor and measure investors' risk preferences by observing their portfolios, since portfolios are considered as outcomes of an investor's investment decisions.

We demonstrate our approach using 20 years of market data and 10 years of mutual fund quarterly holding data. The reasons of using mutual fund portfolios are threefold. First, mutual fund holdings are freely accessible public data and it is easier to discuss and interpret results using public data than clients' private data. Second, mutual funds are usually constructed by tracking industry capitalization indices, or managed by fund managers. Thus, they can be considered as optimal portfolios constructed through rational decision making. Moreover, risk measurements of mutual funds are common knowledge reported by several well-known metrics and estimations with our methods can be directly validated by these metrics. Third, mutual funds usually are diversified on large numbers of assets and they fit the underlying MPT used in our model. Moreover, high-dimensional portfolios are challenging to optimize in the inverse optimization problem and they really test the efficiency of our approach. Numerical experiments show that our approach can directly tackle learning tasks involving hundreds of assets. For tasks involving more than one thousand assets, we propose two different approaches, sector aggregation and factor projection, to transform the problems into lower dimensional space.

Our contributions We summarize the major contributions of our paper as follows:

• To the best of our knowledge, we propose the first inverse optimization approach for learning time varying risk preference parameters of the mean-variance portfolio allocation model, based on a set of observed mutual fund portfolios and underlying asset price data. The flexibility of our approach also enables us to move beyond mean-variance and adopt more general risk metrics.

• The proposed method provides an effective solution in Robo-advising where risk-return trade-off needs to be dynamically updated based on the risk profile communicated by the client. Risk preference values learned from our approach can be used directly as input parameters, or as references of market risk preference, in Robo-advising portfolio construction.

• Our inverse optimization approach is able to handle learning task that consists of hundreds of assets in portfolio, and efficiently learn from long sequences of time-series data composed by thousands of observations. For portfolios composed by more than one thousand assets, we propose Sector-based and Factor-base aggregation to improve the computational efficiency. In particular, to our knowledge, it is the first time factor analysis is introduced in inverse optimization approach on portfolio allocation.

• We collect and process 10 years of mutual fund portfolio holding and 20 years of market price data to demonstrate the proposed algorithms. Our data collection and engineering process is scalable to gather all available mutual fund historical holdings. To the best of our knowledge, historical fund portfolio holding data has been very difficult to find, and we aim to share this data to the public to facilitate related researches.

2. Background

The fundamental mean-variance portfolio optimization model developed by Markowitz (1952) and its variants assume that investors estimate the risk of the portfolio according to the variability of the expected return. Moreover, Markowitz (1952) assumes that investors make decisions solely based on the preferences of two objectives: the expected return and the risk. The trade-off between the two objectives is typically denoted by a positive coefficient and referred to as *risk tolerance* (or *risk aversion*). Later, Black and Litterman (1992) extends the framework in Markowitz (1952) by blending investors' private expectations, known as Black-Litterman (BL) model. A Bayesian statistical interpretation of BL model is proposed in He and Litterman (2002) and an inverse optimization perspective is derived in Bertsimas et al. (2012a). Most of these mean-variance based approaches assume an investor's risk preference is known. For example, in Bertsimas et al. (2012a), the risk-award trade-off δ is denoted by the ratio between expected profit and variance. As they mention: *Even though there are various proposals in the literature, there is no consensus on how to fit δ.* (Bertsimas et al. 2012a) Later, in their experiments, δ is exogenously set to 1.25 based on the suggestion of He and Litterman (2002). Suggested empirical risk aversion values also vary from expert to expert. Ang (2014) suggests a range from 1 to 10 for retail investors and believes it is rare to have risk aversion greater than 10 (Ang 2014). Fabozzi et al. (2007), however, states that risk aversion value should be somewhere between 2 and 4.

In expected utility theory, risk aversion relies on the choice of (usually nonlinear) a von Neumann–Morgenstern utility function (von Neumann and Morgenstern 1947) $u(c)$, where c represents value change in wealth. Absolute risk aversion $-u''(c)/u'(c)$ and relative risk aversion $-u''(c)c/u'(c)$ (Arrow 1971, Pratt 1964) are used to measure how much utility an investor gains (or loses) as the increase (or decrease) of wealth, and how risk aversions are compared across different individuals. Unfortunately, selection of such utility functions to fit representative investors is very challenging because the exact form and parameters of utility functions are generally unknown, and their selections depend on the objectives and preferences of the investor (Warren 2018). This is essentially a subjective process, and the literature has reached no consensus over which utility function provides the best description of individual behavior (Starmer 2000).

In CAPM (Captial Asset Pricing Model) (Treynor 1961, Sharpe 1964, Lintner 1969, Mossin 1966), the relationship between risk and expected return is modeled through *beta*, an indicator

measures relative volatility of the target security/portfolio comparing to market. Market benchmark portfolio has a *beta* of 1.0, and larger *beta* value usually means the security/portfolio can potentially outperform the average market return in a larger margin, and thus allow investors to gauge whether the cost (price) is consistent with such a likely return.

2.1. Related work

Our work is related to the inverse optimization problem (IOP), in which the learner seeks to infer the missing information of the underlying decision model from observed data, assuming that the decision maker is rationally making decision (Ahuja and Orlin 2001). IOP has been extensively investigated in the operations research community from various aspects (Ahuja and Orlin 2001, Iyengar and Kang 2005, Schaefer 2009, Wang 2009, Keshavarz et al. 2011, Chan et al. 2014, Bertsimas et al. 2015, Aswani et al. 2018, Esfahani et al. 2018, Birge et al. 2017, Bärmann et al. 2017, Chan and Lee 2018, Dong and Zeng 2018, Dong et al. 2018, Chan et al. 2019, Chan and Kaw 2020, Dong and Zeng 2020a,b). Due to the time varying nature of risk preferences, our work particularly takes the online learning framework in Dong et al. (2018), which develops an online learning algorithm to infer the utility function or constraints of a decision making problem from sequentially arrived observations of (signal, decision) pairs. This approach makes few assumptions and is generally applicable to learn the underlying decision making problem that has a convex structure. Dong et al. (2018) provides the regret bound and shows the statistical consistency of the algorithm, which guarantees that the algorithm will asymptotically achieves the best prediction error permitted by the underlying inverse model.

Also related to our work is Bertsimas et al. (2012b), in which creates a novel reformulation of the BL framework by using techniques from inverse optimization. The major advantage of this approach is the increased flexibility for specifying views and the ability to consider more general notions of risk than the traditional mean-variance approach. Computational evidence suggests that this inverse optimization approach provide certain benefits over the traditional BL model, especially in scenarios where views are not known precisely. There are two main differences between Bertsimas et al. (2012b) and our paper. First, the problems we study are essentially different. Bertsimas et al. (2012b) seeks to reformulate the BL model while we focus on learning specifically the investor's risk preferences. Second, Bertsimas et al. (2012b) considers a deterministic setting in which the parameters of the BL model are fixed and uses only one portfolio as the input. In contrast, we believe that investor's risk preferences are time varying and propose an inverse optimization approach to learn them in an online setting with as many historical portfolios as possible. Such a data-driven approach enables the learner to better capture the time-varying nature of risk preferences and better leverage the power and opportunities offered by 'Big data'. Similar

to Bertsimas et al. (2012b), Utz et al. (2014) also presents a static inverse optimization framework that aims to learn non-time varying risk tolerances in a multiobjective Markowitz portfolio model.

Recently, we note that researchers in reinforcement learning (RL) propose an exploration-exploitation algorithm to learn the investor's risk appetite over time by observing her portfolio choices in different market environments (Alsabah et al. 2020). The method is explicitly designed for Robo-advisors. In each period, the Robo-advisor places an investor's capital into one of several pre-constructed portfolios, where each portfolio decision reflects the Robo-advisor's belief of the investor's risk preference. The investor interacts with the Robo-advisor by portfolio selection choices, and such interactions are used to update the Robo-advisor's estimations about the investor's risk profile.

We believe RL has tremendous potential in investment Robo-advising. However, Alsabah et al. (2020) has assumptions that are fundamentally different from our approach:

• The investor's decision function in Alsabah et al. (2020) is a simple voting model among different pre-constructed candidate portfolios, and the portfolio allocation process is separate from that decision model. In contrast, we use the portfolio allocation model directly as the decision making model, and IOP allows us to infer risk preference directly from portfolios.

• The learning process of Alsabah et al. (2020) relies on interactions between the Robot and investors, and thus essentially belongs to interaction-based models (see Capponi et al. (2019)), making it useful in scenarios where training data is limited while persistent interaction channels are available to update the model. From the machine learning perspective, the settings of our approach are much different and the approach in Alsabah et al. (2020) is not directly applicable to our settings. We are interested in extracting decision insights from a wealthy amount of existing data, and we do not rely on interactions from the decision-makers. Therefore, our approach is ideal for analyzing a large amount of real-time or batch portfolio data from fund managers or retail investors. Moreover, our method can be used to reveal the inter-temporal change of risk preferences of the same product or the same individual. It can also compare the risk preferences of different products/individuals using their portfolios observed under the same market conditions.

• RL methods suffer from being sample inefficiency (Deisenroth and Rasmussen 2011) and prone to hyper-parameter tuning (Henderson et al. 2018). Learning risk preferences from high dimensional portfolios such as mutual funds would require lots of interactions, leading to high costs in the real world. In contrast, our inverse optimization-based methods are demonstrated to be sample efficient and have very limited number of hyper-parameters to tune (at most three).

Despite these differences, in Robo-advising, the approach in Alsabah et al. (2020) and ours can be combined. For example, an investor's risk appetites can be estimated from complete reference portfolios (mutual funds, market index, or historical investment portfolios) using our methods.

Subsequently, the outputs of our estimations can be considered as endogenous reference states for Alsabah et al. (2020) to personalize the investor's risk preference. Another potential advantage of our approach is computational efficiency and interpretation: According to our experiments on learning tasks composed of hundreds to thousands of assets, our method proves to be very efficient because of the convex setting; The learned risk parameters from real mutual fund portfolios are intuitively consistent with their standard market risk evaluations.

3.   Problem Setting

Before going further into different sections, we give the main notations that will be used throughout the paper. The detailed descriptions of the notations could be found in Table 1.

3.1.   Portfolio Optimization Problem

We mention that our method is generally applicable to convex portfolio optimization models. Besides the mean-variance model adopted in this paper, our method is also applicable to other extended formulations proposed in literature, such as portfolio selection model with transaction cost (Lobo et al. 2007). Without loss of generality, we consider the Markowitz mean-variance portfolio optimization problem (Markowitz 1952):

$$\min_{\mathbf{x}\in\mathbb{R}^n} \tfrac{1}{2}\mathbf{x}^T Q\mathbf{x} - r\mathbf{c}^T\mathbf{x} \qquad\qquad \text{PO}$$
$$s.t. \ \ A\mathbf{x} \geq \mathbf{b},$$

where $Q \in \mathbb{R}^{n\times n} \succeq 0$ is the positive semi-definite covariance matrix, $\mathbf{x}$ is portfolio allocation where each element x_i is the holding weight of asset i in portfolio, $\mathbf{c} \in \mathbb{R}^n$ is the expected asset level portfolio return, $r > 0$ is the *risk-tolerance* factor, $A \in \mathbb{R}^{m\times n}(n \leq m)$ is the structured constraint matrix, and $\mathbf{b} \in \mathbb{R}^m$ is the corresponding right-hand side in the constraints.

In this paper, we have $x_i \geq 0$ for each $i \in [n]$ as we do not consider shorting position. In general, $\mathbf{x}$ represents the portfolio optimized for n assets. Thus, n is around 500 for a S&P 500 portfolio. We further assume each element of $\mathbf{x}$ takes finite value.

In PO, the coefficient r is assigned to the linear term, and thus it represents *risk-tolerance* and larger r indicates more preferable to profit. In literature, some formulations assign r to the quadratic risk term in the objective as *risk-aversion* coefficient, in which larger value leads to more conservative portfolio. Throughout the remainder of this paper, we assign r to the linear term, and the learning task is to estimate *risk-tolerance*. In fact, the propensity for accepting risk regarding investments can be considered on a continuum (Roszkowski et al. 1993), thus *risk-aversion* and *risk-tolerance* are antonyms.

In PO, if variables $Q, r, \mathbf{c}, A, \mathbf{b}$ are given, the optimal solution $\mathbf{x}^*$ can be obtained efficiently via convex optimization. In financial investment, the process is known as finding the optimal portfolio allocations, and we call it the *Forward Problem.*

Table 1   Notation

Indices:	
T	number of learning periods
t	time-varying observation index ($t \in [T]$)
n	number of assets in portfolio
m	number of linear constraints
m_s	number of linear constraints at sector level
K	number of factors
Sets:	
$\mathcal{S}$	the set of sectors
$\mathcal{S}_i$	the set of assets in i-th sector
Θ	parameter space for θ
Data parameters:	
$\mathbf{p}_t$	the vector of monthly profit ratio, where each element corresponds to one asset at month t
w	window size
Y_t	observed portfolio data till time t
P_t	asset price data till time t
Model parameters:	
$\mathbf{x} \in \mathbb{R}^n$	general form of portfolio allocation vector
$\mathbf{x}_s \in \mathbb{R}^{|\mathcal{S}|}$	sector-level portfolio allocation vector
$\mathbf{x}_f \in \mathbb{R}^K$	factor-level portfolio allocation vector
$Q \in \mathbb{S}_+^{n \times n}$	general form of covariance matrix of portfolio
$Q_s \in \mathbb{S}_+^{|\mathcal{S}| \times |\mathcal{S}|}$	covariance matrix of sector-level portfolio
$\Sigma \in \mathbb{S}_+^{K \times K}$	covariance matrix of factor-level portfolio
$F \in \mathbb{R}^{n \times K}$	matrix whose columns are eigenvectors corresponding to the K largest eigenvalues of Q
$\mathbf{c} \in \mathbb{R}^n$	general form of expected profit vector
$\mathbf{c}_s \in \mathbb{R}^{|\mathcal{S}|}$	sector-level expected profit vector
$\mathbf{c}_f \in \mathbb{R}^K$	factor-level expected profit vector
$r \in \mathbb{R}$	general form of risk tolerance value
$r_s \in \mathbb{R}$	risk tolerance value at sector-level
$r_f \in \mathbb{R}$	risk tolerance value at factor-level
$A \in \mathbb{R}^{m \times n}$	structural linear constraint matrix
$A_s \in \mathbb{R}^{m_s \times |\mathcal{S}|}$	structural linear constraint matrix at sector level
$\mathbf{b} \in \mathbb{R}^m$	right-hand side in constraints
$\mathbf{b}_s \in \mathbb{R}^{m_s}$	right-hand side in constraints at sector level
$\mathbf{b}_f \in \mathbb{R}^m$	right-hand side in constraints at factor level
θ	parameter for the decision making problem
θ_t	parameter for the decision making problem learned at time t
u	noisy signal for the decision making problem
$\mathbf{y} \in \mathbb{R}^n$	observed noisy decision (investment portfolio)
$\mathbf{y}_s \in \mathbb{R}^{|\mathcal{S}|}$	observed noisy decision (investment portfolio) at sector level
$\mathbf{y}_f \in \mathbb{R}^K$	observed noisy decision (investment portfolio) at factor level
$\mathbf{u} \in \mathbb{R}^m$	dual variables
$\mathbf{u}_s \in \mathbb{R}^{m_s}$	sector level dual variables
$\mathbf{u}_f \in \mathbb{R}^K$	factor level dual variables
$\mathbf{z} \in \{0,1\}^m$	binary variables
$\mathbf{z}_s \in \{0,1\}^{m_s}$	sector level binary variables
$\mathbf{z}_f \in \{0,1\}^K$	factor level binary variables

3.2. Inverse Optimization through Online Learning

Now we consider a reverse scenario, assuming we can somehow observe a sequence of optimized portfolios while some parameters in PO such as r are unknown. The problem then becomes learning the hidden decision variables that control the portfolio optimization process. Such type of problem has been systematically investigated in Dong et al. (2018) in an online setting. Formally, consider the family of parameterized decision making problem

$$
\begin{aligned}
\min_{\mathbf{x} \in \mathbb{R}^n} \; & f(\mathbf{x}, u; \theta) \\
s.t. \; & \mathbf{g}(\mathbf{x}, u; \theta) \leq \mathbf{0}.
\end{aligned} \qquad \text{DMP}
$$

We denote $X(u; \theta) = \{x \in \mathbb{R}^n : \mathbf{g}(\mathbf{x}, u; \theta) \leq \mathbf{0}\}$ the feasible region of DMP. We let

$$
S(u, \theta) = \arg\min\{f(\mathbf{x}, u, \theta) : \mathbf{x} \in X(u, \theta)\}
$$

be the optimal solution set of DMP.

Consider a learner who monitors the signal $u \in \mathcal{U}$ and the decision maker' decision $\mathbf{x} \in X(u, \theta)$ in response to u. Assume that the learner does not know the decision maker's utility function or constraints in DMP. Since the observed decision might carry measurement error or is generated with a bounded rationality of the decision maker, we denote $\mathbf{y}$ the observed noisy decision for $u \in \mathcal{U}$.

In the inverse optimization problem, the learner aims to learn the parameter θ of DMP from (signal, noisy decision) pairs. In online setting, the (signal, noisy decision) pair, i.e., $(u, \mathbf{y})$, becomes available to the learner one by one. Hence, the learning algorithm produces a sequence of hypotheses $(\theta_0, \ldots, \theta_T)$. Here, T is the total number of rounds, θ_0 is an initial hypothesis, and θ_t for $t \geq 1$ is the hypothesis chosen after observing the t-th (signal,noisy decision) pair.

Given a *(signal, noisy decision)* pair $(u, \mathbf{y})$ and a hypothesis θ, the loss function is set as the minimum distance between $\mathbf{y}$ and the optimal solution set $S(u, \theta)$ in the following form:

$$
l(\mathbf{y}, u; \theta) = \min_{\mathbf{x} \in S(u; \theta)} \|\mathbf{y} - \mathbf{x}\|_2^2. \tag{1}
$$

Once receiving the t-th (signal, noisy decision) pair $(u_t, \mathbf{y}_t)$, θ would be updated by solving the following optimization problem:

$$
\theta_t = \arg\min_{\theta \in \Theta} \frac{1}{2} \|\theta - \theta_{t-1}\|_2^2 + \eta_t l(\mathbf{y}_t, u_t; \theta), \tag{2}
$$

where η_t is the learning rate in round t, and $l(\mathbf{y}, u; \theta)$ is the loss function defined in (1).

3.3. Learning Time Varying Risk Preferences

In the portfolio optimization problem we consider, the (*signal, noisy decision*) pair corresponds to (*market price, observed portfolio*). A learner aims to learn the investor's risk preference from (*price, portfolio*) pairs. More precisely, the goal of the learner is to estimate r in PO. In online setting, the (*price, portfolio*) pairs become available to the learner one by one. Hence, the learning algorithm produces a sequence of hypotheses $(r_0, ..., r_T)$. Here, T is the total number of rounds, r_0 is an arbitrary initial hypothesis, and r_t for $t \geq 1$ is the hypothesis chosen after observing the t)-th (*price, portfolio*) pair.

Given a (*price, portfolio*) pair $(u, \mathbf{y})$ and a hypothesis r, similar to (1), we set the loss function as follows:

$$l(\mathbf{y}, u; r) = \min_{\mathbf{x} \in S(Q, \mathbf{c}, A, \mathbf{b}; r)} \|\mathbf{y} - \mathbf{x}\|_2^2. \tag{3}$$

where $S(Q, \mathbf{c}, A, \mathbf{b}; r)$ is the optimal solution set of PO.

PROPOSITION 1. *Consider PO, and assume we are given a candidate portfolio* $\mathbf{x}$. *Then,* $\mathbf{x} \in S(Q, \mathbf{c}, A, \mathbf{b}; r)$ *if and only if there exists a* $\mathbf{u}$ *such that:*

$$S(Q, \mathbf{c}, A, \mathbf{b}; r) = \{ \mathbf{x} : A\mathbf{x} \geq \mathbf{b}, \mathbf{u} \in \mathbb{R}_+^m, \mathbf{u}^T(A\mathbf{x} - \mathbf{b}) = 0, Q\mathbf{x} - r\mathbf{c} - A^T\mathbf{u} = 0. \} \tag{4}$$

Proof. If we seek to learn r, the optimal solution set for PO can be characterized by KKT conditions. For any fixed values of the data with $Q \in \mathbb{R}^{n \times n} \succeq 0$ and $\mathbf{c}$, the forward problem is convex and satisfies a Slater Condition. Thus, it is necessary and sufficient that any optimal solution $\mathbf{x}$ satisfy the Karush-Kuhn-Tucker (KKT) conditions. The KKT conditions are precisely the equations (4). Here, $\mathbf{u}$ can be interpreted as the dual variable for the constraints in PO. $\square$

PROPOSITION 2. *The equation* $\mathbf{u}^T(A\mathbf{x} - \mathbf{b}) = 0$ *in (4) is equivalent to that there exist* $M > 0$ *and* $\mathbf{z} \in \{0, 1\}^m$ *such that*

$$\mathbf{u} \leq M\mathbf{z}, \ A\mathbf{x} - \mathbf{b} \leq M(1 - \mathbf{z}). \tag{5}$$

Applying Propositions 1 and 2, the update rule by solving (2) with the loss function (3) is reformulated as

$$\begin{aligned}
\min_{r, \mathbf{x}, \mathbf{u}, \mathbf{z}} \ & \tfrac{1}{2}\|r - r_{t-1}\|^2 + \eta_t \|\mathbf{y}_t - \mathbf{x}\|^2 \\
s.t. \ & A\mathbf{x} \geq \mathbf{b}, \\
& \mathbf{u} \leq M\mathbf{z}, \\
& A\mathbf{x} - \mathbf{b} \leq M(1 - \mathbf{z}), \\
& Q_t\mathbf{x} - r\mathbf{c}_t - A^T\mathbf{u} = 0, \\
& \mathbf{x} \in \mathbb{R}^n, \mathbf{u} \in \mathbb{R}_+^m, \mathbf{z} \in \{0, 1\}^m,
\end{aligned} \qquad \text{IPO}$$

where $\mathbf{z}$ is a vector of binary variables used to linearize KKT conditions, and M is an appropriate number used to bound the dual variable $\mathbf{u}$ and $A\mathbf{x} - \mathbf{b}$. Clearly, IPO is a mixed integer second order conic program (MISOCP), and denoted by the *Inverse Problem*.

4. Data Sets

We explain the data sets consisting of seven mutual funds in our paper to help readers better understand the risk-preference learning formulations to be introduced in later sections.

4.1. Observed Portfolio Data Y_t

We collect asset-level portfolio holdings of seven mutual funds (see Table 2) from quarterly reports publicly available on SEC website. Starting from March 2010, we collect portfolio Y_t in each quarter t till March 2020. The total number of observations is $T = 40$. In a mutual fund portfolio, assets excluded or newly included in 10 years duration are all included, and if at a specific time t it does not appear in holding, its weight is 0. As we will test our methods on the seven portfolios across all the experiments, we briefly introduce them in the following:

- VFINX and SWPPX are S&P 500 funds broadly diversified in large-cap market.
- FDCAX's portfolio mainly consists of large-cap U.S. stocks and a small number of non-U.S. stocks.
- VHCAX is an aggressive growth fund invests in companies of various sizes that the managers believe will grow in time but may be volatile in the short-term.
- VSEQX is an actively managed fund holding mid- and small-cap stocks that the advisor believes have above-average growth potential.
- VSTCX is also an actively managed fund that offers broad exposure to domestic small-cap stocks believed to have above-average return potential.
- LMPRX mainly invests in common stocks of companies that the manager believes will have current growth or potential growth exceeding the average growth rate of companies included in S&P 500.

4.2. Asset Price Data P_t

We first query twenty years of historical asset level price data from Jan 1st, 2000 to March 31st, 2020 using AlphaVantage API (AlphaVantage Accessed: 2020). The monthly profit ratio of individual asset $p_{i,t}$ is calculated as a ratio: the numerator is the subtraction of month start opening price from the month end closing price of asset i; the denominator is the month start opening price at time t. Then, we denote $\mathbf{p}_t$ the vector of monthly profit ratio, where each element corresponds to each individual asset at month t. We let $P_t = [\mathbf{p}_1, ..., \mathbf{p}_t]$.

At each month t, we consider the backward history of w months as the observation of market signals. The covariance matrix is calculated as

$$Q_t = \text{cov}([\mathbf{p}_{t-w}, ..., \mathbf{p}_t], [\mathbf{p}_{t-w}, ..., \mathbf{p}_t]), \tag{6}$$

Mutual Fund	Description	No. of Assets
VFINX	Vanguard SP 500 Index Fund	566
SWPPX	Schwab SP 500 Index	564
FDCAX	Fidelity Capital Appreciation Fund	737
VHCAX	Vanguard Capital Opport Fund Admiral	219
VSEQX	Vanguard Strategic Equity Fund	1440
VSTCX	Vanguard Strategic SmallCap Equity Fund	1037
LMPRX	ClearBridge Aggressive Growth Fund	99

Table 2   Seven mutual funds used as illustrating examples

Sector Name	VFINX	SWPPX	FDCAX	VHCAX	VSEQX	VSTCX	LMPRX
Basic Materials	2.0585	2.0696	1.8866	0.00017	2.7600	3.3926	0.5647
Communication Services	10.6418	10.6255	9.6793	3.7649	3.3698	2.1419	25.7736
Consumer Cyclical	9.5455	9.4568	11.0045	8.5593	11.1654	8.2437	0.2692
Consumer Defensive	8.0814	8.0692	4.6294	0.0489	3.2521	3.5085	-
Energy	2.6315	2.6254	1.5087	1.3174	1.7545	1.9985	0.6166
Financial Services	13.8400	13.8892	9.6072	7.2355	13.2327	15.3046	0.9835
Healthcare	15.1290	15.0931	14.9797	33.7680	14.0414	14.0102	32.0470
Industrials	8.0863	8.0675	2.1276	10.4407	14.7876	14.7677	4.6743
Real Estate	2.9809	2.9807	3.1544	0	8.0989	8.0817	-
Technology	21.4863	21.4337	27.1432	27.9279	18.2084	13.3433	29.4258
Utilities	3.5461	3.5375	0.9650	0	4.8055	3.8035	-

Table 3   Sector level holdings (percentage) of Mutual Funds reported by March 2020

and the expected monthly asset-level return $\mathbf{c}_t$ is set to be the medium of profitable monthly returns during the observation period. Formally, the expected portfolio-level return of n assets $\mathbf{c}_t$ is defined as

$$
\begin{aligned}
p_{i,t} &> 0, \forall i \in [n], \\
c_{i,t} &= \text{medium}([p_{i,t-w}, ..., p_{i,t}]), \\
\mathbf{c}_t &= [c_{1,t}, ..., c_{i,t}, ..., c_{n,t}]^T.
\end{aligned}
\tag{7}
$$

REMARK 1. The reason we construct the expected return specifically as (7), instead of simply taking the average of historical profit $[p_{t-w,i}, ..., p_{t,i}]$ is that the simple average may lead to negative expect returns. Since our formulation does not allow shorting, each element of $\mathbf{x}$ is non-negative and the linear term $\mathbf{c}^T\mathbf{x}$ in the objective function of PO can be negative. When this happens, the learned risk tolerance value r is forced to be 0 in order to minimize the second term of objective. In fact, this does not mean that the portfolio has zero risk, since the negative term leads to a wrong interpretation of *risk-tolerance*. To avoid such caveat in our analysis, we restrict the expected asset-level returns to be non-negative, and set them to medium values of non-negative actual monthly returns.

4.3.   Time-varying Observations

As illustrated in Figure 1, we align portfolio Y_t and price P_t on the same timeline. We start our learning from Mar 2015, which corresponds to $t = 183$ for P_t and $t = 20$ for Y_t. At each learning point t, we use the historical price observed in a certain look-back window $p_{t-w}, ..., p_t$, where w is
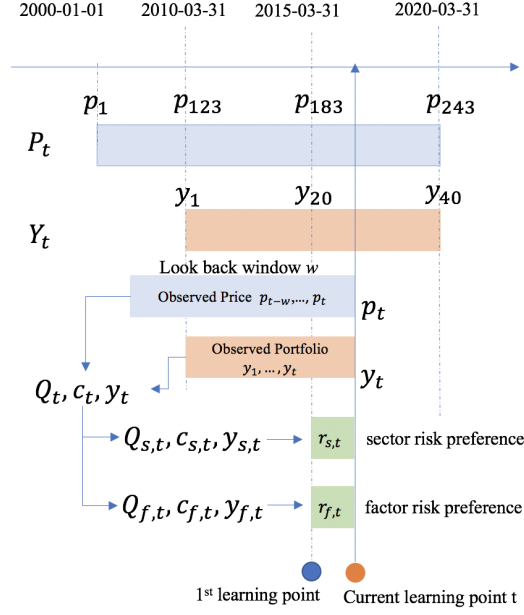
Figure 1   Time-varying observations for risk preference learning.

the window size, and is set to 120 (10 years) by default. At the first learning point, we use price data from 2005-03 to 2015-03 to estimate the covariance matrix Q_t and the expected profit return $\mathbf{c}_t$. Then, we use historical portfolio observations $\mathbf{y}_1, ..., \mathbf{y}_t$ (starting from 2010-03-01) to learn r_t. For the next learning point, we keep the windows size fixed, and shift forward three months to obtain $Q_{t+1}, \mathbf{c}_{t+1}$. The number of portfolio observations increases by t. For example, when $t = 20$, we use $y_1, ..., y_{20}$ to estimate r_t. When $t = 21$, estimation is done by observing $y_1, ..., y_{21}$. So the more recent estimations of r_t have more observations of y_t. The algorithms need an initial guess of r_0, and we use the same initial guess at different t. In other words, the estimation of previous r_t has no impact on next r_{t+1}.

In mutual funds, portfolio $\mathbf{y}$ might have more than 1000 assets, making the problem in IPO very challenging to solve. According to our practical experience, solving a single iteration of IPO involving 1000-dimensional $\mathbf{y}$ can take more than 20 hours. To make the computation feasible, we propose two different approaches to reduce the computational burden. In the first approach, we aggregate the asset-level portfolio and price data to 11 industrial sectors. Table 3 shows examples of the sector-level portfolio weights of 7 mutual funds in March 2020. In the second approach, we project portfolios into a number of factors. We use eigenvalue decomposition on historical price covariance matrix to decompose a number of uncorrelated factors representing market volatility. Figure 2 illustrates the top 5 principal components of monthly profit of all 2236 assets studied in this paper. In this setting, each factor is a linear combination of individual assets, and portfolio holdings assigned on factors correspond to aggregated linear combinations of assets.
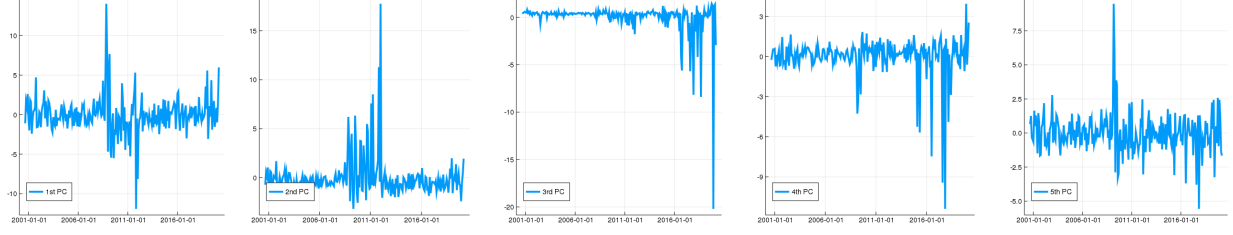
Figure 2 Profitability curve represented by top 5 factors. The factors are learned from profit covariance matrix of 2236 assets in the past 20 years.

5. Portfolio Risk Tolerance Learning

For portfolios composed by more than one thousand assets, we propose in this section reformulations of IPO for sector-based and factor-based risk preference learning to improve the computational efficiency.

5.1. Sector Space

5.1.1. Forward Problem Let $\mathcal{S}$ denote the set of sectors. Similar to PO, we consider the following mean-variance portfolio optimization problem in sector space:

$$\min_{\mathbf{x}} \tfrac{1}{2}\mathbf{x}_s^T Q_s \mathbf{x}_s - r_s \mathbf{c}_s^T \mathbf{x}_s \qquad \text{PO-Sector}$$
$$s.t. \ A_s \mathbf{x}_s \geq \mathbf{b}_s,$$

where $Q_s \in \mathbb{R}^{|\mathcal{S}| \times |\mathcal{S}|} \succeq 0$ is the sector level covariance matrix, $\mathbf{x}_s \in \mathbb{R}^{|\mathcal{S}|}$ is the portfolio allocation aggregated in sector space, $\mathbf{c}_s \in \mathbb{R}^{|\mathcal{S}|}$ is the expected sector level portfolio return, r_s is the risk tolerance factor (in sector space), $A_s \in \mathbb{R}^{m_s \times |\mathcal{S}|}$ is the structured constraint matrix, and $\mathbf{b}_s \in \mathbb{R}^{m_s}$ is the corresponding right-hand side in constraints.

Similar to Section 4.3, time-varying $Q_{s,t}$ and $\mathbf{c}_{s,t}$ are defined as follows:

$$\begin{aligned}
Q_{s,t} &= \mathrm{cov}([\mathbf{p}_{s,t-w}, ..., \mathbf{p}_{s,t}], [\mathbf{p}_{s,t-w}, ..., \mathbf{p}_{s,t}]), \\
\hat{p}_{i,t} &= \sum_{j \in \mathcal{S}_i, p_{j,t} \geq 0} w_{j,t} p_{j,t}, \\
c_{i,t} &= \mathrm{medium}([\hat{p}_{i,t-w}, ..., \hat{p}_{i,t}]), \\
\mathbf{c}_{s,t} &= [c_{1,t}, ..., c_{i,t}, ..., c_{|\mathcal{S}|,t}]^T,
\end{aligned} \qquad (8)$$

where $\mathbf{p}_{s,t}$ is a vector of sector-level returns, $\mathcal{S}_i$ is the set of assets in i-th sector, $w_{j,t}$ is the asset holding weight for $j \in \mathcal{S}_i$, and $\hat{p}_{i,t}$ is the time-varying sector monthly profit calculated by weighted sum of asset level profits in i-th sector.

5.1.2. Inverse Problem Similar to IPO, given the observed portfolio $\mathbf{y}_{s,t}$ in sector level, $Q_{s,t}$, and $\mathbf{c}_{s,t}$, the online update rule of learning sector-level risk tolerance r_s is formulated as

$$\begin{aligned}
\min_{r_s, \mathbf{x}_s} \ & \tfrac{1}{2}\|r_s - r_{s,t-1}\|^2 + \eta_t \|\mathbf{y}_{s,t} - \mathbf{x}_s\|^2 \\
s.t. \ & A_s \mathbf{x}_s \geq \mathbf{b}_s, \\
& \mathbf{u} \leq M\mathbf{z}, \\
& A_s \mathbf{x}_s - \mathbf{b}_s \leq M(1 - \mathbf{z}), \qquad \text{IPO-Sector} \\
& Q_{s,t} \mathbf{x}_s - r_s \mathbf{c}_{s,t} - A_s^T \mathbf{u} = 0, \\
& \mathbf{x}_s \in \mathbb{R}^{|\mathcal{S}|}, \mathbf{u} \in \mathbb{R}_+^{m_s}, \mathbf{z}_s \in \{0, 1\}^{m_s},
\end{aligned}$$

where $\mathbf{u}_s$ is the dual variable, and $\mathbf{z}_s$ is a vector of binary variables used to linearize KKT conditions for PO-Sector.

5.1.3.   Risk Tolerance Learning Algorithm in Sector Space Algorithm 1 illustrates the process of applying online IOP to learn risk preference using sector level data. The number of iteration T varies by observed portfolio.

Algorithm 1 Risk Tolerance Learning in Sector Space

Input: (time-series portfolio and price data) Y_t, P_t

Initialization: $r_{s,0}$ (guess), λ, M (hyper-parameter)

  1: **for** $t = 1$ to T **do**

  2:      receive (Y_t, P_t)

  3:      $(\mathbf{y}_{s,t}, \mathbf{p}_{s,t}) \leftarrow (\mathbf{y}_t, \mathbf{p}_t)$

  4:      $(Q_{s,t}, \mathbf{c}_{s,t}) \leftarrow \mathbf{p}_{s,t}$

  5:      $\eta_t \leftarrow \lambda t^{-1/2}$ $\triangleright$ get updated η_t by decaying factor

  6:      Solve $r_{s,t}$ as IPO-Sector in Section 5.1.2 $\triangleright$ update guess of r_s

  7: **end for**

Output: Estimated r_s

5.1.4.   Hyper-parameter and constraints Solving the inverse problem in IPO-Sector requires hyper-parameters M and η_t. In this paper, we set the constraints $A_s \mathbf{x}_s \geq \mathbf{b}_s$ to be

$$0 \leq x_i \leq 1, \quad i \in \mathcal{S}, \tag{9}$$

$$\sum_{i=1}^{|\mathcal{S}|} x_i = 1, \tag{10}$$

as we do not consider short-selling positions. Specifically, when considering 11 sectors, the structural constraint coefficient matrix A_s is a 22×11 matrix, and $\mathbf{b}_s$ is a 22×1 vector:

$$A_s = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & 1 \\ -1 & & & & \\ & -1 & & & \\ & & \ddots & & \\ & & & -1 & \\ & & & & -1 \\ 1 & 1 & \dots & & 1 \\ -1 & -1 & \dots & & -1 \end{bmatrix}, \quad \mathbf{b}_s = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ -1 \\ -1 \\ \vdots \\ -1 \\ -1 \\ 1 \\ -1 \end{bmatrix}. \tag{11}$$

5.2. Factor Space

5.2.1. Forward Problem Without abuse of notation, we refer to the asset-level covariance matrix as Q, and the asset-level return as $\mathbf{c}$. We perform eigendecomposition of the asset-level covariance matrix Q such that

$$Q \approx F\Sigma F^T, \tag{12}$$

where Σ is a $K \times K$ diagonal matrix whose entries are the K largest eigenvalues of Q, and F is a $n \times K$ matrix whose columns are the eigenvectors correspoinding to the K largest eigenvalues of Q. The relationships between asset-level allocation $\mathbf{x}$ and factor-level allocation $\mathbf{x}_f$ are:

$$F^T\mathbf{x} = \mathbf{x}_f, \tag{13}$$

$$\mathbf{x} = F\mathbf{x}_f, \tag{14}$$

where (13) is projecting asset allocations to factor space, and (14) is reconstruct allocations in sector space from factor space.

PROPOSITION 3. *The forward problem in factor space is*

$$\begin{aligned} \min_{\mathbf{x}_f} \quad & \tfrac{1}{2}\mathbf{x}_f^T\Sigma\mathbf{x}_f - r_f\mathbf{c}_f^T\mathbf{x}_f \\ s.t. \quad & AF\mathbf{x}_f \geq \mathbf{b}, \end{aligned} \qquad \text{PO-Factor}$$

where $\mathbf{c}_f = F^T\mathbf{c} \in \mathbb{R}^K$ can be seen as the factor-level expected profit vector.

Proof. Applying (12) and (14) to the objective function in PO yields

$$\begin{aligned} & \tfrac{1}{2}\mathbf{x}_f^T F^T Q F\mathbf{x}_f - r_f\mathbf{c}^T F\mathbf{x}_f, \\ = & \tfrac{1}{2}\mathbf{x}_f^T F^T F\Sigma F^T F\mathbf{x}_f - r_f(F^T\mathbf{c})^T\mathbf{x}_f, \\ = & \tfrac{1}{2}\mathbf{x}_f^T\Sigma\mathbf{x}_f - r_f(F^T c)^T\mathbf{x}_f, \end{aligned} \tag{15}$$

The second equation follows from the fact that $F^T F = I$. $\square$

REMARK 2. In our experiments, Q is a 2236×2236 covariance matrix of all underlying assets included in seven mutual funds. All mutual fund holdings are represented by a 2236-dimensional vector $\mathbf{y}$, where non-holding assets are set to 0. The number of factors K in all our experiments is set to 5.

5.2.2. Inverse Problem The structure of the inverse optimization problem in factor space is similar to that in asset space and sector space. The main difference occurs on the constraint matrix AF. Since Q changes over time t, the diagonal matrix Σ and F also are variables depending on t.

Given the observed portfolio $\mathbf{y}_t$ in asset level, we obtain the portfolio in factor level by (13): $\mathbf{y}_{f,t} = F_t^T\mathbf{y}_t$. Now, we are ready to state the online update rule of learning factor-level risk tolerance r_f.

THEOREM 1. *The online update rule of learning factor-level risk tolerance r_f can be formulated as*

$$\min_{r_f, \mathbf{x}_f} \frac{1}{2}\|r_f - r_{f,t-1}\|^2 + \eta_t \|\mathbf{y}_{f,t} - \mathbf{x}_f\|^2$$
$$s.t. \quad AF_t\mathbf{x} \geq \mathbf{b}_f,$$
$$\mathbf{u}_f \leq M\mathbf{z}_f, \qquad\qquad\qquad \text{IPO-Factor}$$
$$AF_t\mathbf{x}_f - \mathbf{b}_f \leq M(1-\mathbf{z}),$$
$$\Sigma_t\mathbf{x}_f - r_f\mathbf{c}_{f,t} - (AF_t)^T\mathbf{u}_f = 0,$$
$$\mathbf{x}_f \in \mathbb{R}^K, \mathbf{u}_f \in \mathbb{R}_+^K, \mathbf{z}_f \in \{0,1\}^K.$$

Proof. The theorem is a direct result of applying Propositions 1, 2, and 3. □

REMARK 3. Specifically, the constraint matrix AF_t and the right-hand side $\mathbf{b}_f$ in IPO-Factor are defined as follows:

$$AF_t = \begin{bmatrix} F_t^T \\ -F_t^T \\ \sum_{j=1}^n f_{j,t} \\ -\sum_{j=1}^n f_{j,t} \end{bmatrix}, \quad \mathbf{b}_f = \begin{bmatrix} \varepsilon\mathbf{1} \\ -\mathbf{1} \\ 1 \\ -1 \end{bmatrix}, \tag{16}$$

where F_t is the $N \times K$ eigenvector matrix obtained in (12), and $f_{j,t}$ is its j-th column. Unlike the sparse structural matrix in (11), AF_t is a dense matrix. Moreover, the first module in $\mathbf{b}_f$ is a vector of constant value ε, while it is a vector of all zeros in $\mathbf{b}$. In IPO-Factor, we use $\mathbf{b}_f$ instead of $\mathbf{b}$ because solving IPO-Factor with a dense constraint matrix AF_t is very challenging as the linear systems in the constraints are often infeasible. In practice, we find that relaxing constraints from $0 \leq AF_t\mathbf{x} \leq 1$ to $\varepsilon \leq AF_t\mathbf{x} \leq 1$, where ε is a small positive constant, can make the problem much easier to solve. Under this factor space setting, the allocation weights $\mathbf{x}_f$ on factors can be positive or negative values whose absolute values are larger than 1. When re-projecting them to the original asset space as $F_t\mathbf{x}_f$ in (14), however, all weights are still between ε and 1, and their sum is equal to 1.

5.2.3. Hyper-parameter and constraints Solving IPO-Factor requires three hyper-parameters: M, λ and ε. The number of eigenfactor K is set to 5 in all experiments. For the dense constraint matrix AF_t, in practice, we do not recommend using larger K because it would make IPO-Factor very challenging to compute. If the factor coefficients assigned on asset level are sparse and consequently the constraint matrix AF_t is also sparse, however, the proposed method would still be able to learn from a large number of factors efficiently.

5.2.4. Risk Tolerance Learning Algorithm in Factor Space The learning algorithm in factor space is illustrated in Algorithm 2. We note that, in sector space, each mutual fund has different sector-based holdings and $Q_{s,t}, \mathbf{c}_{s,t}$ are computed separately for each mutual fund. In factor space, however, $Q_t, \mathbf{c}_t, F_t, \Sigma_t$ are the same while $\mathbf{y}_{f,t}$ are different inputs for all mutual funds at each observation point t.

Algorithm 2 Risk Tolerance Learning in Factor Space

Input: (time-series portfolio and price data) Y_t, P_t

Initialization: $r_{f,0}$ (guess), λ, M, ε (hyper-parameter)

 1: **for** $t = 1$ to T **do**

 2:     receive (Y_t, P_t)

 3:     $(Q_t, \mathbf{c}_t) \leftarrow \mathbf{p}_t$

 4:     $(F_t, \Sigma_t) \leftarrow Q_t$ ▷ eigendecomposition

 5:     $(\mathbf{c}_{f,t}) \leftarrow F_t^T \mathbf{c}_t$

 6:     $(\mathbf{y}_{f,t}) \leftarrow F_t^T \mathbf{y}_t$

 7:     $\eta_t \leftarrow \lambda t^{-1/2}$ ▷ get updated η_t by decaying factor

 8:     Solve $r_{f,t}$ as IPO-Factor in Section 5.2.2   ▷ update guess of r_f

 9: **end for**

Output: Estimated r_f

---

6. Validation of Risk Preference Learning
6.1. Validation of Risk Preference Point Estimation

To validate the effectiveness and accuracy of the proposed algorithms, we benchmark hyper-parameters M, λ, ε using grid search and Forward-Inverse validation. For each combination of hyper parameters, we use four known risk tolerance values $r_i^s \in [0.5, 5, 10, 20]$ in the forward problem to generate portfolio data. Then, we start with 10 different initial guess value $r_j^g \in [1, .., 10]$, using a specific group of hyper-parameters to estimate back the known risk tolerance value, denoted by r_{ij}^e, by solving the inverse problem. The total sum of square error between r^s and r^e for portfolio data generated by a specific r^s

$$\frac{1}{40} \sum_{i=1}^{4} \sum_{j=1}^{10} \frac{(r_i^s - r_{ij}^e)^2}{(r_i^s)^2},\tag{17}$$

is considered as average error of risk tolerance point estimation given a specific set of hyper-parameters. The grid search space is defined as follows:

$$M \in [100, 500, 1000, 5000, 10000],$$

$$\lambda \in [100, 500, 1000, 5000, 10000],$$

$$\varepsilon \in [0.005, 0.01, 0.05, 0.1] (\text{only for factor space}).$$

According to evaluation results shown in Figure 3, Forward-Inverse validation of seven mutual fund price data shows similar patterns about the optimal hyper-parameter settings in sector space. We select $M = 500, \lambda = 10000$ as the optimal hyper-parameters because the average squared errors
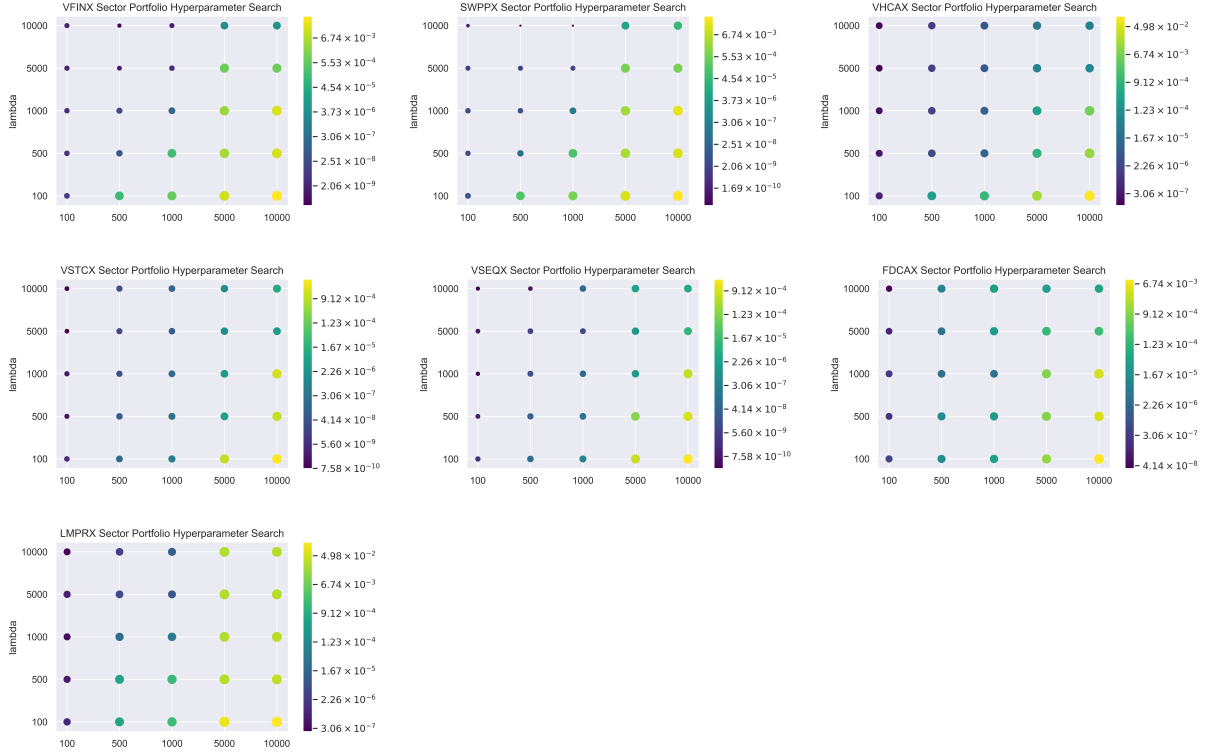
Figure 3   Point Estimation of Risk Tolerance Estimation in Sector Space. The hyper-parameter search is performed on 5 × 5 two dimensional grid. The size and color of dots illustrated in the figures are proportional to the average error defined in (18). In all experiments, the price data P_t is observed at March 31st, 2015 ($t = 183$). For each mutual fund, assets are selected according to their portfolio holdings disclosed on the same quarter. Then, asset-level price data is further aggregated on sector level, and $Q_s, \mathbf{c}_s$ are obtained to investigate error via Forward-Inverse validation.

in (17) are overall the lowest among all mutual funds validations. In factor space, the impact of hyper-parameters is more significant. We select $M = 500, \lambda = 1000, \varepsilon = 0.01$ as the optimal hyper-parameter setting in our factor space learning tasks.

6.2.   Validation of Ordered Risk Preference Learning

We further generate a series of portfolios using a number of sorted risk tolerance values $r_1 \leq, ..., \leq r_m$. Our purpose is to validate whether the learned risk tolerance values can preserve the true order in Forward-Inverse benchmark using real market data.

$$\{\mathbf{y}_1, ..., \mathbf{y}_d\} \leftarrow \text{Forward Problem}(\{r_1, ..., r_d\}),$$

$$\{\hat{r}_1, ..., \hat{r}_d\} \leftarrow \text{Inverse Problem}(\{\mathbf{y}_1, ..., \mathbf{y}_d\}).$$

The validation consists of three steps:

• First, we plot the efficient frontiers (EF) by increasing r in the *Forward Problems* defined in PO-Sector and PO-Factor, respectively. Each data point on the EF curve represents a (risk, profit)
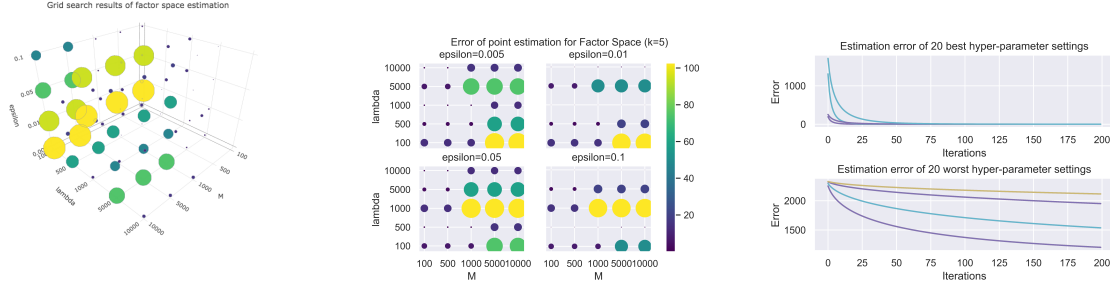
Figure 4 **Point Estimation of Risk Tolerance in Factor Space using similar experimental setup as Figure 3. The hyper-parameter benchmark is performed on the 5 × 5 × 4 three dimensional grid (left figure). In the middle figure, 3D grid is reduced to four 2D grids varied at different ε values. The size and color of dots are proportional to the average error defined in (19), the smaller the better. The right figure shows 200 iterations of 20 best hyper-parameter settings at the top, and 20 worst hyper-parameter settings at the bottom. We see that optimal hyper-parameter settings allow the algorithm to reduce learning errors effectively.**

pair determined by a predefined r. Smaller r are positioned closer to origin points on the EF curve, whereas larger values locate far from the origin on the EF curve.

• Second, we uniformly sample 30 risk tolerance values along the curve, and use the optimal allocation results as inputs to solve the *Inverse problems* defined in IPO-Sector and IPO-Factor, respectively.

• Third, the 30 estimated values are compared with true values. We compare both the orders and the absolute values along diagonal lines. Ideally, perfect estimations should align all points on the diagonal.

According to our experiments illustrated in Figure 5 and Figure 6, in the range of benchmark region, estimations are accurate and estimated values and orders are perfectly aligned with true values.
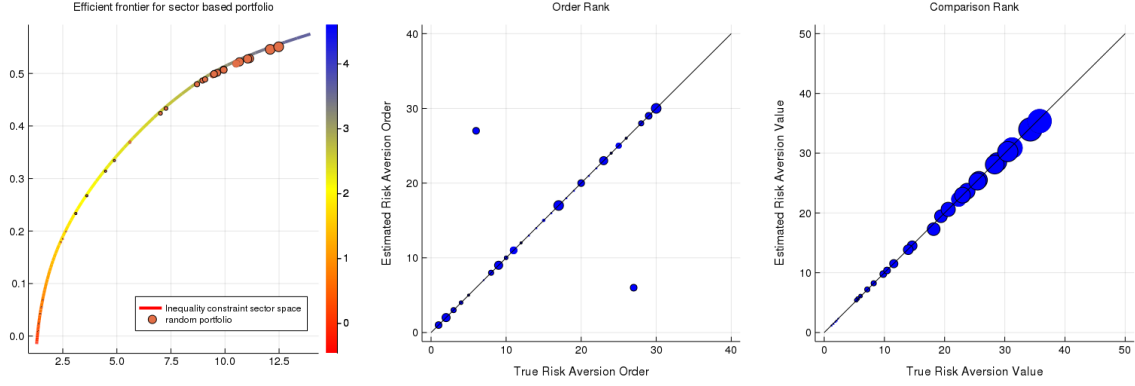
Figure 5 Order Estimation of Risk Tolerance Values in Sector Space. The plot is generated using VFINX asset data observed from Jan 1st 2000 to March 31st 2015 aggregated to sector level. First, the efficient frontier curve in the left figure is generated by solving forward problem using 100 evenly sampled r values on logspace from -0.2 to 2, corresponding to $r = [10^{-0.2}, ..., 10^2]$. The coordinates of curve are (Risk, Profit) values represented by $(\mathbf{x}^T Q \mathbf{x}, \mathbf{c}^T \mathbf{x})$ where $\mathbf{x}$ are obtained by solving Forward problem. The color of curve is mapped to the logarithm of risk tolerance values as indicated in color map. In the same log space, 30 random risk tolerance values, indicated by red circles, are sampled in uniform distributions. The sizes of circles are proportional to their values. Then, portfolios generated by these samples in Forward Problem are used as observations to estimate back r in inverse problem. The middle figure compares the order of true values (x-axis) with the order of estimated values (y-axis). The right figure compares the true values of sampled risk tolerances (x-axis) with the estimated values (y-axis). The sizes of blue circles are proportional to the logarithms of true values.
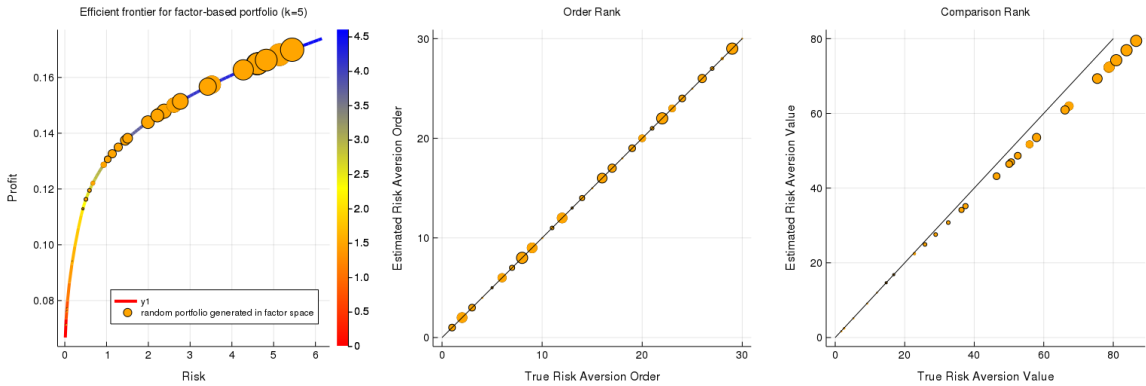

Figure 6 Order Estimation of Risk Tolerance Values in Factor Space. The experimental setup is similar to Sector Space experiments mentioned in Figure 5. The market data contains 2236 asset prices observed from 2000 to 2015, and the number of eigenvector K is set to 5. The efficient frontier curve is generated by solving (PO-Factor) using r values evenly distributed from 1 to 100. To validate the ordered estimation, 30 random risk tolerance values are sampled uniformly between 2 and 99. The middle and right figures are comparisons of true orders/values (x-axis) with estimated orders/values (y-axis), respectively.

7.   Analyzing S&P 500 Portfolios and Discussions

We first apply the IPO formulation to analyze some artificial portfolios composed by subset of stocks in S&P 500. From the holdings of Vanguard S&P 500 Index Fund (VFINX) we select a number of stocks and construct a number of pseudo portfolios. The collected 10-year VFINX holdings have 669 stocks, and we rank all these stocks according to their averagely returns and variances, using the 20-year monthly pricing data discussed in Section 4. Our goal is to validate and interpret the outcomes of proposed risk preference learning algorithm on some *profitable* and *non-profitable* portfolios, and also on *risky* portfolios and *non-risky* ones. Among the 669 S&P 500 stocks, we select the top *profitable* and top *non-profitable* stocks according to the averaged profits. We also rank stocks by their variances, and select the top *risky* stocks (large variances) and *non-risky* ones (small variances). Then, using the true holdings weights of theses stocks in VFINX quarterly portfolios, we re-normalize and create new portfolios, assuming investors only invest on particular selections of S&P 500 stocks, and apply proposed algorithm to learn the risk preference coefficients. The results show some interesting findings:

• Risk preferences coefficients learned on *risky* portfolios (larger variances) are significantly larger than *non-risky* (smaller variances) and randomly selected portfolios, which is consistent with expectations because portfolios showing more volatility involve more risk, so the risk tolerance coefficients should be large.

• Risk preferences coefficients learned on *profitable* portfolios are larger than random and *non-profitable* portfolios.

• After 2017, the learned risk preference scores have a sharp jump on *profitable* and *risky* portfolios.

• More importantly, the scale of learned risk preference coefficient is dependent on the dimension of portfolio, which means we cannot directly compare risk preferences of portfolios containing different number of stocks in asset space. According to our experiments, larger portfolios tend to have smaller risk preference scores, which shows an interesting reconfirmation of the value of diversification in risk control. It also suggests that to compare portfolios containing different number assets, we need to aggregate or project them to sector space or factor space as proposed in Section 5.
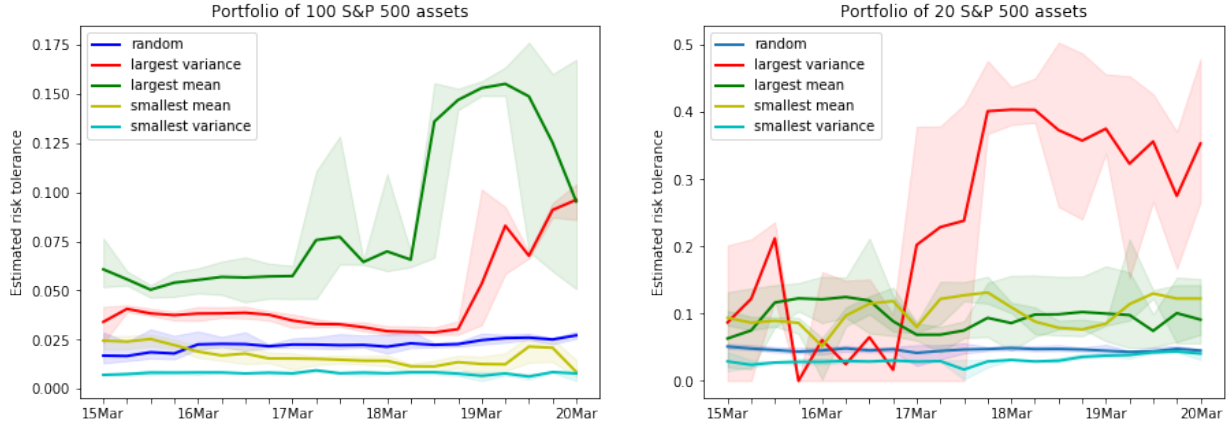
Figure 7   Comparison of Risk Tolerance Values on portfolios constructed by subset of S&P 500 stocks. A total number of 669 stocks included in VFINX holdings from 2010 till 2020 are ranked, based on 20 years of market price data (Jan 2000 to March 2020), according to their averaged profitability and variance. We select the top 20/100 most profitable stocks and the top 20/100 non-profitable stocks (100 stocks shown in the left figure, 20 stocks shown in the right figure), and create portfolios denoted as largest mean and smallest mean. We also select the to 20/100 most volatile stocks, denoted as largest variance, and less volatile stocks called smallest variance to compose sample portfolios. In addition, we include a portfolio consists of randomly select 20/100 stocks called random. We assume the holding weights of these stocks are proportional to their original allocation weights in VFINX, and we re-normalize their weights so their sum is 1. For each pseudo portfolio, 21 time-varying risk tolerance values are estimated, corresponding to 21 quarterly observed portfolio holdings (renormalizing the allocation weights of selected stocks) from March 2015 to March 2020. At each quarterly observation point t at x-axis, a sequence of portfolio holdings from March 2010 to t is considered as $\mathbf{y}_t$, and a sequence of monthly historical price data from $t - w$ to t is considered as an observation of $\mathbf{p}_t$. To create the confidence intervals, we change the length of look back history w from 100 months to 140 months by 2 months, and create 21 observations of $\mathbf{p}_t$. We then construct Q_t and $\mathbf{c}_t$, and combine them with $\mathbf{y}_t$ to estimate risk tolerance r. The mean values and deviations of 21 estimated r, per each observation point t, are illustrated as confidence intervals. Figure at the left side shows results estimated from portfolios composed by 100 stocks, and figure on the right shows results estimated on portfolios composed by 20 stocks. Notice the scales of estimated values (y-axis) are different between left and right because its dependency on the dimensionality of portfolios.
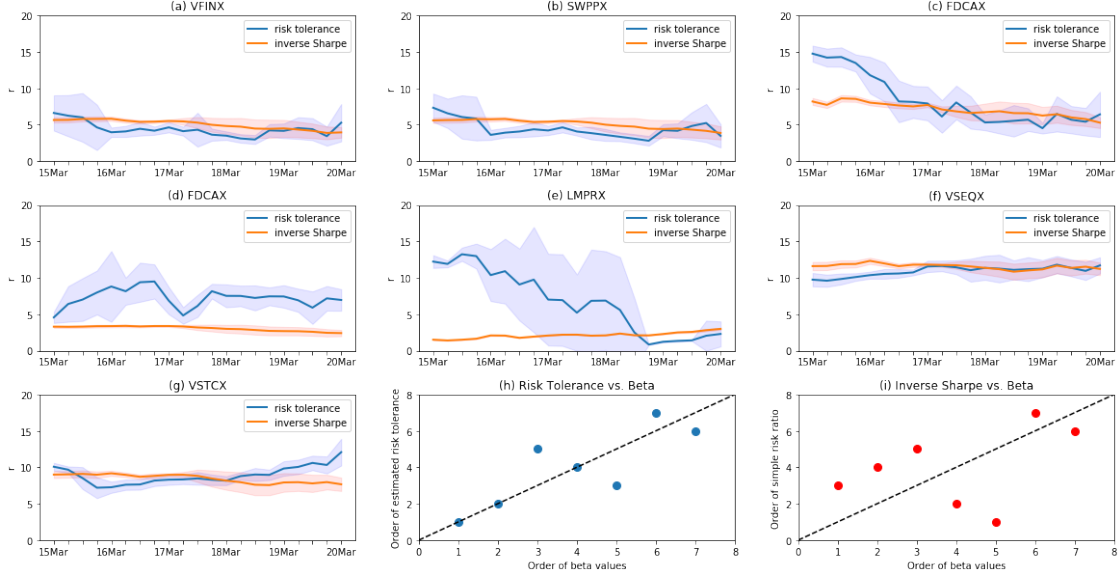
Figure 8   **Comparison of Mutual Fund Risk Tolerance Values in Sector Space. Estimated risk tolerance parameters (blue curves) of seven mutual funds are illustrated from (a) to (g). For each fund, 21 time-varying risk tolerance values are estimated, corresponding to 21 quarterly observed portfolio holdings from March 2015 to March 2020. At each quarterly observation point t at x-axis, a sequence of portfolio holdings from March 2010 to t is considered as $\mathbf{y}_t$, and a sequence of monthly historical price data from $t - w$ to t is considered as an observation of $\mathbf{p}_{s,t}$. To create the confidence intervals, we change the length of look back history w from 100 months to 140 months by 2 months, and create 21 observations of $\mathbf{p}_{s,t}$. We then construct $Q_{s,t}$ and $\mathbf{c}_{s,t}$, and combine them with $\mathbf{y}_{s,t}$ to estimate risk tolerance r_s using Algorithm 1. The mean values and deviations of 21 estimated r_s, per each observation point t, are illustrated as confidence intervals. Furthermore, we calculate the inverse Sharpe ratios according to 19 for comparison. Since the inverse Sharpe ratio is not estimated via online learning, we can calculate a ratio value for each observed historical quarterly holding $\mathbf{y}_t$. For 21 observations of $\mathbf{p}_t$, we calculate corresponding inverse Sharpe ratio for all settings and plot the means and deviations as confidence intervals. In figure (h), we rank the seven mutual funds by the order of average risk tolerance values estimated over all 21 quarters, from small to large, and compare them with the order of beta values shown in Table 4. In figure (g), we compare the order ranked by inverse Sharpe ratios with order ranked by beta values.**

8.   Analyzing Real Multiple Mutual Fund Portfolios and Discussions

We apply the proposed algorithm on seven time-series mutual fund portfolio holdings. By aggregating high dimensional portfolios into sector space and factor space, we are able to efficiently estimate their time-varying risk tolerance values over 21 quarters. To further validate interpretations of estimated risk tolerance values, we compare them with two standard metrics evaluating investment risk used in financial domain: **Inverse of Sharpe Ratio** and **Mutual Fund beta values** .
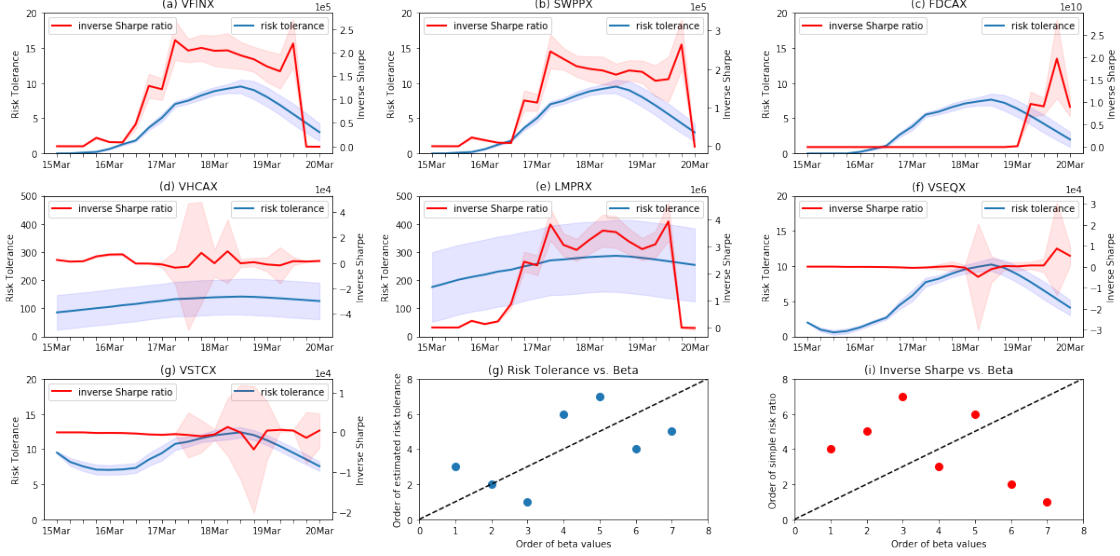
Figure 9   Comparison of Mutual Fund Risk Tolerance Values in Factor Space. The experimental settings of obtaining $Q_t, \mathbf{c}_t, \mathbf{y}_t$ are almost identical to what mentioned in Figure 8. The only difference is that we perform eigenvalue decomposition on Q_t and select 5 eigenvectors corresponding to 5 largest eigenvalues as factors. Since the inverse Sharpe ratio values can be large positive and negative numbers, we use a secondary vertical axis at the right of subplots (a) to (g) to represent their values. As all funds are analyzed in the same covariance factor space, the overall trend of risk tolerance values over the past 5 years are similar. For mutual funds VFINX, SWPPX, VSTCX, the risk tolerance values estimated in Factor space and Sector space are very close (between 5 and 10). For VHCAX and LMPRX, factor space analysis leads to very large risk tolerance values. When comparing the average of the estimations with the standard beta value, the results are still reasonable. VFINX and SWPPX have the lowest risk tolerance comparing with the other five aggressively managed mutual funds.

8.1.   Comparisons with Two Investment Risk Evaluation Metrics

8.1.1.   Comparison with Inverse of Sharpe Ratio In finance, the Sharpe ratio (Sharpe 1965) measures the performance of a portfolio compared to a risk-free asset, after adjusting the risk it bears. It is defined as the difference between the expected return of the asset return $E(R)$ and the risk-free return $E(R_f)$, divided by the standard deviation of the investment:

$$S = \frac{E(R) - E(R_f)}{\sigma}. \tag{18}$$

Note that the Sharpe ratio is proportional to *risk-aversion*. In our paper, we use a modified *Inverse Sharpe ratio* to compare with learned *risk tolerance* values:

$$IS_t = \frac{\mathbf{x}_t^T Q_t \mathbf{x}_t}{\mathbf{c}_t^T \mathbf{x}_t}. \tag{19}$$

In (19), $\mathbf{x}_t, Q_t, \mathbf{c}_t$ are variables defined over asset-level, and change over time. To avoid redundant discussions, we derive similar Inverse Sharpe ratios in sector space and factor space following those in Section 5.

Mutual Fund	β (by 2019-12)	β (by 2020-04)
VFINX	1.0	1.0
SWPPX	1.0	1.0
FDCAX	1.03	1.01
VHCAX	1.14	1.13
VSEQX	1.17	1.30
VSTCX	1.22	1.38
LMPRX	1.17	1.07

Table 4   Beta values of mutual funds used in our approach.

According to Figure 8, estimated risk-tolerance values and inverse Sharpe ratios show similar trends and ranges on five out of seven mutual funds. Our estimations are more volatile over time, and the confidence intervals are generally wider than inverse Sharpe ratios. On two mutual funds (FDCAX and LMPRX), risk tolerance estimations are much larger than the inverse Sharpe ratio. However, in factor space, the Inverse Sharpe ratio becomes very unreliable. One of the main reasons is that factor-based allocation $\mathbf{x}_f$ and expected return $\mathbf{c}_f$ are unbounded (in asset space $\mathbf{x}, \mathbf{c}$ are non-negative, but after multiplying eigenvector F they are unbounded real values). Estimated risk tolerance values, in contrast, show stable trends. Among seven mutual funds, five of them have time-varying risk tolerance values between 0 and 10. Two exceptions are LMPRX and VHCAX, where estimated risk tolerance jump above 100 and have large uncertainties.

8.1.2.   Comparison with Mutual Fund beta values In Capital Asset Pricing Model (CAPM) (Treynor 1961, Sharpe 1964, Lintner 1969, Mossin 1966), the *beta* (beta coefficient) of a portfolio is a measure of the risk arising from exposure to general market movements as opposed to idiosyncratic factors.

$$\beta = \frac{E(R) - E(R_f)}{E(R_m) - E(R_f)}, \tag{20}$$

where $E(R_m)$ is the expected return of market (e.g., benchmark portfolio) and others are the same as defined in (18). The *beta* measures how much risk the investment is compared to the market. If a stock is riskier than the market, it will have a *beta* greater than one. If a stock has a *beta* of less than one, the formula assumes it will reduce the risk of a portfolio.

The *beta* values of mutual funds are public data and change over time. We are unable to find complete historical *beta* values for given mutual funds. Thus, we only use the recent two beta values in Table 4.

We sort the averaged risk-tolerance values of 7 mutual funds over time from small to large, and compare their orders with the order of beta values averaged from Table 4. In Figure 8 (h) and Figure 9 (h), most of the scatter plots are close to diagonal, indicating that the order of estimated risk tolerance is consistent with the order of beta values. Therefore, risk indicators estimated by

proposed methods are intuitively rational because they are consistent with common knowledge in the financial domain. In addition, our proposed method has some significant advantages:

Readiness for Automated and Personalized Portfolio Construction Our risk tolerance parameters can be directly used as inputs in the mean-variance framework to drive automated portfolio advice, known as Robo-advising. Most existing Robo-advising systems construct risk parameters according to *one-time interaction* with the client (Capponi et al. 2019), typically profiled based on client's financial objectives, investment horizons, and demographic characteristics. Such an approach has two main limitations: First, risk tolerance measured by interaction-based survey is typically categorized as Aggressive/Moderate/Conservative through regression or classification analysis. An advisor then translates the client's feedback into a numerical score, herein referred to as the client's risk preference parameter. However, what exact value should r be assigned for a conservative investor? If $r = 5$, does it mean aggressive or conservative? Such questions are practically tough to handle with an exact quantification and separation of systematic component vs. idiosyncratic component or investment risk. The second limitation is that treating risk-tolerance as a constant value over a long investment time horizon is considered insufficient because the subjective risk tolerance should change over time, and be influenced by significant lifetime events and consumer decisions.

Our proposed method shares similar visions with some approaches that adapt mean-variance optimization criterion and assume risk tolerance as a stochastic parameter that needs to be updated at all times (Capponi et al. 2019). Those approaches assume the client knows her risk tolerance parameters at all times, while only communicates them to the Robo-advisor at specific updating times(Capponi et al. 2019). In contrast, we do not assume the investor explicitly knows her risk parameter. Instead, we assume a successful investor is rational in maintaining an appropriate level of risk compared to some benchmark portfolios. Our experiments show that the time-varying risk tolerances of two S&P 500 mutual funds, which are considered a market benchmark portfolio, range from 5 to 10. We also show that some actively managed funds have larger risk tolerance values. For example, if the client is aware that S&P 500 fund has risk tolerance $r = 5$ at current market period and an active fund has risk tolerance $r = 10$, she may prefer to have a personalized risk tolerance that is between the S&P 500 fund and the active fund. The Robo-advisor can directly suggest a personalized risk tolerance score $r = 7$ and start with a personalized portfolio construction process. Later, when the market has changed, risk tolerance values of S&P 500 fund and VSTCX may also change accordingly. The Robo-advisor should update the process with a new parameter to reflect the time-varying risk preference.

Readiness to Extend Black-Litterman Model Our proposed method can be considered as an extension to Black-Litterman model (Black and Litterman 1992), which overcomes the problem

of reliably estimating expected returns in Modern Portfolio theory. Black-Litterman model assumes the initially expected returns can be learned from the market benchmark portfolio. An investor is only required to blend her personal assumptions about expected returns, and Black–Litterman method computes the desired (mean-variance efficient) asset allocation using the new personalized expected returns. Our approach, analogously, can be used to construct a process that does not require an investor's inputs about her risk preference. We can provide several time-varying risk preference values estimated from popular mutual funds. These allow an investor to set her personal risk tolerance higher or lower than reference risk values, and construct a personalized portfolio. As pointed out in (Bertsimas et al. 2012a), Black-Litterman actually can be interpreted as an inverse optimization process. In fact, our proposed inverse optimization framework can not only learn r, but also be customized to learn $\mathbf{c}$ or r and $\mathbf{c}$ together in PO-Sector. Based on these insights, one of the major focuses of our future work is to extend Black-Litterman model to incorporate personalized risk tolerance as well as expected returns.

Generalization of Factor-based Formulation The proposed factor-based formulations has shown to effectively reduce the computational burden of IPO on large dimensional portfolios. More importantly, our formulations can be generalized to incorporate customized factor settings. In our paper, we obtain factor coefficient F from eigendecompoistion. In fact, F can be generalized to different forms, e.g., a discrete clustering membership form as follows:

$$F_{ij} = \begin{cases} \frac{1}{\sqrt{n_j}}, & \text{if } x_i \in C_j, \\ 0, \end{cases} \tag{21}$$

where F is a $n \times K$ matrix and n is total number of assets, K is number of clusters, n_j is number of assets in j-th cluster C_j. The i-th row and j-th column element of F is $\frac{1}{\sqrt{n_j}}$ if asset i belongs to the j-th cluster and 0 otherwise. The construction of F_{ij} ensures $F^T F = I$, and equations in (15) are still valid (covariance matrix is no longer diagonal) because F is a linear transformation. Therefore, Algorithm 2 can also be applied to analyze risk tolerance properties on clustered portfolios.

Computational Efficiency We compare the computational efficiency of two different formulations, IPO-Sector and IPO-Factor, in estimating risk preferences from real high-dimensional market price and portfolio data. As illustrated in Table 5, the proposed algorithms are capable to efficiently estimate risk preference from large scale problems containing portfolios up to 500 assets/sectors. Notice the values shown here only represent solving a single iteration of inverse problem, while in online learning setting the total required time scales up with the number of observations in the *(price, portfolio)* sequence. So in average estimating risk preference from 20 portfolios containing 500 assets usually take half an hour, and only take about 3 minutes if the data is projected to 5 factors.

Formulation	Problem Size	CPU-time(seconds)
IPO-Sector	100 sectors	0.466
	200 sectors	5.908
	300 sectors	32.396
	400 sectors	46.380
	500 sectors	79.485
IPO-Factor	400 assets 5 factors	3.524
	400 assets 6 factors	23.429
	400 assets 7 factors	166.250
	400 assets 8 factors	283.050
	400 assets 9 factors	293.150
	400 assets 10 factors	297.826
	500 assets 5 factors	7.365
	500 assets 6 factors	40.524
	500 assets 7 factors	221.100
	500 assets 8 factors	228.550
	500 assets 9 factors	300.350
	500 assets 10 factors	300.672

Table 5   **Comparison of computational time solving a single iteration of IPO-Sector or IPO-Factor. Each CPU time value shown in the table is the average number of 10 iterations consist of randomly sampled matrices. The performance is evaluated on a workstation equipped with AMD Ryzen Threadripper 1950X 16-Core and 64G memory. The modeling scripts are written using Julia 0.6, using Gurobi v8.1 linux version as the solver, and fully parallelized on Linux Mint 19. Notice that in reality we don't have 500 sectors, but the complexity of solving a 500 sector problem is the same as solving a formulation of 500 asset problem, so we mainly compare asset/sector based formulation vs. factor based formulation.**

9. Conclusions and Future Work

In this paper, we present a novel approach to learn risk preference parameters from portfolio allocations. We formulate our learning task in a generalized inverse optimization framework (Dong et al. 2018). We consider portfolio allocations as outcomes of an optimal decision-making process governed by mean-variance portfolio theory, thus risk preference parameter is a learnable decision variable through IOP. The proposed approach has been validated and exhibited in real problems that consists of a 20-year history of market data and 10-year history of portfolio holdings collected in seven mutual funds. We also demonstrate the efficiency and effectiveness of the proposed approach on forward and inverse asset allocation problems containing more than 41 decision periods, 2200 assets, and complex constraints. The proposed risk tolerance values estimated from mutual funds have a high ordinal correlation with well-known industry metrics and could be deployed as real-time reference indices in Robo-advising systems that facilitate automation and personalization.

Our future efforts will focus on the following directions:

• The proposed method can be applied to individual investors' portfolios to understand their risk profiles in personalized investment advice. A potential challenge is that personal investors, unlike fund managers, usually make non-optimal or sub-optimal decisions. Also, personal investments are

usually not as diversified as mutual funds, and speculations blend their expectations of returns. Our inverse optimization paradigm can be applied to extend Black-Litterman model to investigate expected return and risk tolerance from reference portfolios (mutual funds). Then, in conjunction with some personalization mechanisms, updating personal investors' expectations and risk profiles in securities chosen by specific factors is a practical and fruitful area for future research.

- The proposed method provides an effective solution in Robo-advising where risk-return trade-off needs to be dynamically updated based on the risk profile communicated by the client. Risk preference values learned from our approach can be used directly as input parameters, or as references of market risk preference, in Robo-advising portfolio construction.

- Our inverse optimization approach is able to handle learning task that consists of hundreds of assets in portfolio. For portfolios composed by more than thousands of assets, we propose Sector-based and Factor-base projection to improve the computational efficiency. In particular, to the best of our knowledge, it is the first time that factor analysis is introduced in inverse optimization approach on portfolio allocation. We demonstrate the proposed algorithms on a large scale time-series portfolio and price data that contains 2236 stock daily price in 20 years, and 7 mutual fund portfolio holdings in 10 years. We aim to compare the proposed methodology on different categories of financial instruments, and to setup a public repository of relevant data sets and indices.

Acknowledgements

Ahuja, Ravindra K, James B Orlin. 2001. Inverse optimization. *Operations Research* **49**(5) 771–783.

Aït-Sahalia, Yacine, Andrew W Lo. 2000. Nonparametric risk management and implied risk aversion. *Journal of Econometrics* **94**(1-2) 9–51.

AlphaVantage. Accessed: 2020. AlphaVantage API Documentation. `https://www.alphavantage.co/documentation/`.

Alsabah, Humoud, Agostino Capponi, Octavio Ruiz Lacedelli, Matt Stern. 2020. Robo-Advising: Learning Investors' Risk Preferences via Portfolio Choices*. *Journal of Financial Econometrics* doi:10.1093/jjfinec/nbz040. URL `https://doi.org/10.1093/jjfinec/nbz040`. Nbz040.

Ang, Andrew. 2014. *Asset Management: A Systematic Approach to Factor Investing*. Oxford University Press.

Arrow, Kenneth J. 1971. The theory of risk aversion. *Essays in the theory of risk-bearing* 90–120.

Aswani, Anil, Zuo-Jun Shen, Auyon Siddiq. 2018. Inverse optimization with noisy data. *Operations Research* .

Bärmann, Andreas, Sebastian Pokutta, Oskar Schneider. 2017. Emulating the expert: Inverse optimization through online learning. *ICML*.

Barsky, Robert B, F Thomas Juster, Miles S Kimball, Matthew D Shapiro. 1997. Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *The Quarterly Journal of Economics* **112**(2) 537–579.

Bertsimas, Dimitris, Vishal Gupta, Ioannis Ch. Paschalidis. 2012a. Inverse optimization: A new perspective on the black-litterman model. *Operations research* **60 6** 1389–1403.

Bertsimas, Dimitris, Vishal Gupta, Ioannis Ch Paschalidis. 2012b. Inverse optimization: A new perspective on the Black-Litterman model. *Operations Research* **60**(6) 1389–1403.

Bertsimas, Dimitris, Vishal Gupta, Ioannis Ch Paschalidis. 2015. Data-driven estimation in equilibrium using inverse optimization. *Mathematical Programming* **153**(2) 595–633.

Birge, John R, Ali Hortaçsu, J Michael Pavlin. 2017. Inverse optimization for the recovery of market structure from market outcomes: An application to the miso electricity market. *Operations Research* **65**(4) 837–855.

Black, Fischer, Robert Litterman. 1992. Global portfolio optimization. *Financial Analysts Journal* **48**(5) 28–43. doi:10.2469/faj.v48.n5.28. URL `https://doi.org/10.2469/faj.v48.n5.28`.

Brennan, Thomas J, Andrew W Lo. 2011. The origin of behavior. *The Quarterly Journal of Finance* **1**(01) 55–108.

Capponi, Agostino, Sveinn Ólafsson, Thaleia Zariphopoulou. 2019. Personalized robo-advising: Enhancing investment through client interactions. *Wealth Management eJournal* .

Chan, Timothy CY, Tim Craig, Taewoo Lee, Michael B Sharpe. 2014. Generalized inverse multiobjective optimization with application to cancer therapy. *Operations Research* **62**(3) 680–695.

Chan, Timothy CY, Neal Kaw. 2020. Inverse optimization for the recovery of constraint parameters. *European Journal of Operational Research* **282**(2) 415–427.

Chan, Timothy CY, Taewoo Lee. 2018. Trade-off preservation in inverse multi-objective convex optimization. *European Journal of Operational Research* **270**(1) 25–39.

Chan, Timothy CY, Taewoo Lee, Daria Terekhov. 2019. Inverse optimization: Closed-form solutions, geometry, and goodness of fit. *Management Science* .

Chetty, Raj. 2006. A new method of estimating risk aversion. *American Economic Review* **96**(5) 1821–1834.

Cohen, Alma, Liran Einav. 2007. Estimating risk preferences from deductible choice. *American economic review* **97**(3) 745–788.

Cohn, et al, Richard A. 1975. Individual Investor Risk Aversion and Investment Portfolio Composition. *Journal of Finance* **30**(2) 605–620. URL `https://ideas.repec.org/a/bla/jfinan/v30y1975i2p605-20.html`.

Deisenroth, Marc, Carl E Rasmussen. 2011. Pilco: A model-based and data-efficient approach to policy search. *ICML.*

Dong, Chaosheng, Yiran Chen, Bo Zeng. 2018. Generalized inverse optimization through online learning. *NeurIPS*.

Dong, Chaosheng, Bo Zeng. 2018. Inferring parameters through inverse multiobjective optimization. *arXiv preprint arXiv:1808.00935* .

Dong, Chaosheng, Bo Zeng. 2020a. Expert learning through generalized inverse multiobjective optimization: Models, insights, and algorithms. *ICML*.

Dong, Chaosheng, Bo Zeng. 2020b. Wasserstein distributionally robust inverse multiobjective optimization. *arXiv preprint arXiv:2009.14552v1* .

Esfahani, Peyman Mohajerin, Soroosh Shafieezadeh-Abadeh, Grani A Hanasusanto, Daniel Kuhn. 2018. Data-driven inverse optimization with imperfect information. *Mathematical Programming* **167**(1) 191–234.

Fabozzi, Frank J., Petter N. Kolm, Dessislava A. Pachamanova, Sergio Focardi. 2007. *Robust Portfolio Optimization and Management*. John Wiley & Sons.

Guiso, Luigi, Monica Paiella. 2008. Risk aversion, wealth, and background risk. *Journal of the European Economic association* **6**(6) 1109–1150.

Hanna, Sherman, Michael Gutter, Jessie Fan. 1998. A theory based measure of risk tolerance. *Proceedings of the Academy of Financial Services* **15** 10–21.

He, Guangliang, Robert B. Litterman. 2002. The intuition behind black-litterman model portfolios.

Henderson, Peter, Riashat Islam, Philip Bachman, Joelle Pineau, Doina Precup, David Meger. 2018. Deep reinforcement learning that matters. *AAAI*.

Hey, John D. 1999. Estimating (risk) preference functionals using experimental methods. *Uncertain Decisions*. Springer, 109–128.

Holt, Charles A, Susan K Laury. 2002. Risk aversion and incentive effects. *American economic review* **92**(5) 1644–1655.

Iyengar, Garud, Wanmo Kang. 2005. Inverse conic programming with applications. *Operations Research Letters* **33**(3) 319–330.

Keshavarz, Arezou, Yang Wang, Stephen Boyd. 2011. Imputing a convex objective function. *Intelligent Control (ISIC), 2011 IEEE International Symposium on*. IEEE, 613–619.

Linnér, Richard Karlsson et al. 2019. Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences1. *bioRxiv* doi: 10.1101/261081. URL `https://www.biorxiv.org/content/early/2019/01/08/261081`.

Lintner, John. 1969. The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets: A Reply. *The Review of Economics and Statistics* **51**(2) 222–224.

Lobo, Miguel Sousa, Maryam Fazel, Stephen Boyd. 2007. Portfolio optimization with linear and fixed transaction costs. *Annals of Operations Research* **152**(1) 341–365.

Lu, Jingfeng, Isabelle Perrigne. 2008. Estimating risk aversion from ascending and sealed-bid auctions: The case of timber auction data. *Journal of Applied Econometrics* **23**(7) 871–896.

Markowitz, Harry. 1952. Portfolio selection. *The Journal of Finance* **7**(1) 77–91.

Mcgraw, A. Peter, Jeff Larsen, David Schkade. 2010. Comparing gains and losses. *Psychological science* **21** 1438–45. doi:10.1177/0956797610381504.

Mossin, Jan. 1966. Equilibrium in a capital asset market. *Econometrica* **34** 768.

O'Donoghue, Ted, Jason Somerville. 2018. Modeling risk aversion in economics. *Journal of Economic Perspectives* **32** 91–114. doi:10.1257/jep.32.2.91.

Payne, Brian, Jazmin Brown-Iannuzzi, Jason Hannay. 2017. Economic inequality increases risk taking. *Proceedings of the National Academy of Sciences* **114** 201616453. doi:10.1073/pnas.1616453114.

Post, Thierry, Martijn J Van den Assem, Guido Baltussen, Richard H Thaler. 2008. Deal or no deal? decision making under risk in a large-payoff game show. *American Economic Review* **98**(1) 38–71.

Pratt, John W. 1964. Risk aversion in the small and in the large. *Econometrica* **32**(1/2) 122–136.

Rabin, Matthew, Richard H Thaler. 2001. Anomalies: risk aversion. *Journal of Economic perspectives* **15**(1) 219–232.

Roszkowski, Michael, G.E. Snelbecker, S.R. Leimberg. 1993. Risk tolerance and risk aversion. *The Tools and Techniques of Financial Planning* 213–225.

Schaefer, Andrew J. 2009. Inverse integer programming. *Optimization Letters* **3**(4) 483–489.

Schooley, Diane K., Debra Drecnik Worden. 1996. Risk aversion measures: Comparing attitudes and asset allocation. *Financial Services Review* **5** 87–99.

Sharpe, William. 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* **19**(3) 425–442. URL `https://EconPapers.repec.org/RePEc:bla:jfinan:v:19:y:1964:i:3:p:425-442`.

Sharpe, William. 1965. Mutual fund performance. *The Journal of Business* **39**. URL `https://EconPapers.repec.org/RePEc:ucp:jnlbus:v:39:y:1965:p:119`.

Sokol-Hessner, Peter, Ming Hsu, Nina G Curley, Mauricio R. Delgado, Colin F. Camerer, Elizabeth A. Phelps. 2009. Thinking like a trader selectively reduces individuals' loss aversion. *Proceedings of the National Academy of Sciences* **106** 5035 – 5040.

Starmer, Chris. 2000. Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature* **38**(2) 332–382. doi:10.1257/jel.38.2.332. URL `https://www.aeaweb.org/articles?id=10.1257/jel.38.2.332`.

Szpiro, George G. 1986. Relative risk aversion around the world. *Economics Letters* **20**(1) 19–21.

Thomas, P.J. 2016. Measuring risk-aversion: The challenge. *Measurement* **79** 285 – 301. doi:https://doi.org/10.1016/j.measurement.2015.07.056. URL `http://www.sciencedirect.com/science/article/pii/S0263224115004091`.

Treynor, Jack L. 1961. Market value, time, and risk.

Utz, Sebastian, Maximilian Wimmer, Markus Hirschberger, Ralph E Steuer. 2014. Tri-criterion inverse portfolio optimization with application to socially responsible mutual funds. *European Journal of Operational Research* **234**(2).

von Neumann, J., O. Morgenstern. 1947. *Theory of games and economic behavior*. Princeton University Press.

Wang, H., T. Zariphopoulou, X. Zhou. 2020. Reinforcement learning in continuous time and space: A stochastic control approach.

Wang, Lizhi. 2009. Cutting plane algorithms for the inverse mixed integer linear programming problem. *Operations Research Letters* **37**(2) 114–116.

Warren, Geoffrey. 2018. Choosing and using utility functions in forming portfolios. *SSRN Electronic Journal* doi:10.2139/ssrn.3207865.