Application of Preference Measurement: The Case of Riskalyze

By: Fangfei Dong; David Eil, PhD; Ethan Pew, PhD; and Noah Smith, PhD

May 10, 2015

Executive Summary

- Measuring risk preferences is an important task for financial advisers. In recent years, technologies have been created to measure preferences more accurately and in a way that is more applicable to portfolio selection. The company Riskalyze has created one such technology.

- Basic economic theory says that preferences should be represented by a utility function over levels of wealth. Experimental economics has developed methods for measuring utility functions by asking people to choose between different "lotteries," or gambles.

- The ideal lottery choice task should be simple and specific (quantitative), should have real consequences for the decision maker, should be relevant to the domain of interest, and should have as many questions as is practical. Riskalyze's lottery choice task fulfills these ideals well.

- A common alternative method of measuring preferences, the psychological quiz, is simple and easy to use, and useful for research purposes. However, this method is difficult to apply to portfolio selection, since its quantitative implications are ambiguous.

- To apply preference measurements to portfolio choice, it is necessary to have quantitative measurements of the risk and return of candidate portfolios. Riskalyze's technology estimates risk and expected return using techniques that are standard in the finance industry. Combining preferences with portfolio risk and return characteristics depends on the adviser's available tools and the type of portfolio the adviser is managing for an investor.

- Non-standard theories of risk preferences, such as reference-dependent utility and loss aversion, can also be accommodated by Riskalyze's preference measurement method. The adviser must simply alter the conditions under which she re-administers the lottery choice task. Data from investor preference

measurements, gathered through technologies like Riskalyze's, can potentially enhance the academic field of behavioral finance, by allowing simultaneous observation of preferences and decision-making.

1 Introduction

When they model financial behavior, economists typically assume that investors simply implement their own preferences when making choices. In reality, however, investment decisions are often too complicated for the typical investor to process, and portfolio management is too time-consuming a chore. Thus, investors often hire financial advisers.

Riskalyze, a California-based company, has built new technology to assist financial advisers in two tasks:

1. measuring the risk preferences of investors, and
2. applying these preference measurements to portfolio selection.

Riskalyze has commissioned us to examine their technology and place it in the broad context of academic research on this topic. Our precondition for this work was absolute academic independence and integrity. Riskalyze was contractually barred from instructing us, the authors, to state or include anything in this report that might violate their academic integrity.

The creation of the Riskalyze technology is part of a broader movement in the financial industry toward application of research from the fields of behavioral and experimental finance. The empirical behavioral finance literature, at least as far back as Barber and Odean (2000), has attempted to identify suboptimal choices made by individual investors in the real world. Research in human judgment and decision-making, which goes back even farther, has developed a number of techniques for measuring risk preferences in controlled settings (e.g. Kahneman and Tversky 1979). Riskalyze's technology represents an attempt to use the insights of experimental finance to correct the problems identified by empirical behavioral finance.

In this paper, we consider the validity of preference measurement methods (henceforth PMMs) based on what we call "quasi-hypothetical" lottery choices, and the applications of these methods to financial advising. Taking Riskalyze as an example, we assess the company's PMM in the context of the academic finance and economics literature, and contrast it with a more classic PMM, the psychometric risk tolerance quiz. We also assess Riskalyze's quantitative methodology for measuring the risk and expected returns of portfolios, and discuss how quantitative data should best be combined with preference data.

Finally, we discuss prospects for improving Riskalyze's technology, and the place of behavioral finance research within the financial advising industry.

2 The Financial Adviser's Problem

There are two reasons why a financial adviser needs to measure an investor's risk preferences. These have to do with the two ways that financial advisers add value for investors.

The first way an adviser adds value is to act as the agent for the investor, faithfully carrying out the investor's wishes. Many investors don't have the time or the proper tools to manage a portfolio by themselves, so they delegate the job to an adviser. In order to carry out the investor's wishes, an adviser must have an understanding of the investor's risk preferences.

The second way an adviser adds value is to improve the investor's decisions. Investors may be financially unsophisticated, or subject to certain biases, as documented in Loos et al. (2014). By adding knowledge and subtracting emotion, advisers can improve investors' portfolio choices.

More subtly, an investor may not be able to implement her own risk preferences when choosing a portfolio. An investor who avoids the stock market is avoiding both the risk of the market and the uncertainty of investing in an asset whose risk she does not fully understand. The "money doctor" hypothesis of Gennaioli, Shleifer, and Vishny (2015) states that trusted advisers may act to reassure investors that there is no hidden, unobserved source of risk, allowing investors to take on more risk (and thus earn higher returns) than they would on their own.

The investor's inability to perfectly choose her own portfolio has two very important implications for financial advisers. First, it increases the importance of advisers, since they must find a portfolio that pleases an investor more than the one the investor would choose on his own. Second, it makes the adviser's task more difficult, since the adviser cannot perfectly infer the investor's preferences from his past decisions. The adviser therefore needs tools to better gauge the investor's preferences. Financial advisers can measure investors' preferences informally, through casual conversation, or formally, as with the quiz-type PMMs that have become more common in recent years. It stands to reason that more accurate and directly applicable measurements of financial risk preferences would allow advisers to better serve investors.

3 Riskalyze and its Technology

Riskalyze is a software company. Its customers are financial advisers, who in turn serve investors. Advisers use Riskalyze's technology to measure investors' preferences and to measure the risk and return characteristics of candidate portfolios that they are considering selecting for the investors. Advisers then use the information provided by Riskalyze to inform their judgment in matching investors with portfolios.

To measure investor risk preferences, Riskalyze gives investors a questionnaire. The questionnaire offers investors a series of 50-50 gambles representing portfolio choices, and asks the investor to choose between each gamble and a certainty equivalent. The numbers used in the questionnaire are based on a range defined by three values:

1. the investor's current assets in the portfolio in question,
2. a lower bound specified by the investor as a "disastrous" loss, and
3. an upper bound specified by the investor as an amount he would be "content" with.

As the investor makes choices in the questionnaire, the certainty equivalent of the gambles is adjusted - if the investor accepts more gambles than he rejects, the certainty equivalent is raised in subsequent questions, while if he rejects more gambles than he accepts, the certainty equivalent is lowered in subsequent questions. When an approximate indifference point is found, that combination of gamble and certainty equivalent allows Riskalyze to plot one point on a utility curve. This procedure is then repeated with a different gamble, with different probabilities and gain/loss amounts. The questionnaire continues until three to five utility points have been plotted in different regions of wealth (at least one on each side of the investor's current assets). These points are added to the "disaster" and "comfort" levels, which are assigned utility values of zero and one. Those five to seven points allowing Riskalyze to sketch the investor's utility curve. The utility curve is specified according to a proprietary functional form that has five parameters. The functional form is therefore flexible enough to accommodate both risk-averse and risk-seeking behavior, or an inflection point between the two.

This PMM can be described as a "quasi-hypothetical" lottery choice. The investor knows that his answers will affect the real portfolio choices being made on his behalf, without further consent on his part being required. In addition, the values in the lotteries in the questionnaire correspond approximately to the investor's own potential wealth levels after the portfolio decision is made. So although Riskalyze's lottery questionnaire does not ask the investor to make an actual portfolio decision on the spot, the lottery choices are both more consequential and more realistic than a completely hypothetical exercise would be.

As an alternative, for situations where time is limited, Riskalyze has a short-form questionnaire consisting of three questions. These questions ask investors to specify the amount of portfolio risk (specified as a Value at Risk (5%)) that they would accept in exchange for a given expected return. In this paper, we will consider the long-form questionnaire described in the paragraph above. When we refer to Riskalyze's "questionnaire", we refer to the long-form questionnaire, not to the short-form one.

The output of the questionnaire includes an "individual risk number," which summarizes an investor's mean-variance risk aversion on a 99-point scale, and a "risk fingerprint," which provides more comprehensive information about the investor's preferences. The risk number is a monotonic function of the six-month Value at Risk (5%) of a portfolio optimized with respect to the utility curve sketched by the answers to the questionnaire. In other words, the more risky the optimal portfolio for an investor with a given utility curve, the higher the risk number.

To measure the quantitative risk of a candidate portfolio, Riskalyze uses past data to calculate the portfolio's expected return and variance. Riskalyze then assigns each portfolio a "portfolio risk number", which is a monotonic function of the portfolio's six-month Value at Risk (5%). This is the level of return that the portfolio can be expected to underperform 5 percent of the time. Since Riskalyze assumes normal distributions for all assets, this number is simply calculated as the estimated six-month expected return minus 1.64 times the estimated six-month volatility.

The functions used to obtain risk numbers from portfolio variances are the same for individual risk numbers and portfolio risk numbers.

4 Quasi-Hypothetical Lottery Choices for Measuring Risk Preferences

In this section we discuss the ideal method for measuring risk preferences with lottery choices, reviewing some key concepts about risk preferences from economic theory, as well as evidence from experiments. We then discuss how these preferences might be measured for a real-life investor. Finally, we evaluate Riskalyze's quasi-hypothetical lottery choice method and compare it to the ideal.

4.1 Expected Utility Theory and Risk Aversion

Before discussing how best to measure risk preferences, we must first define what risk preferences are. Saying one person is "more risk averse" than another could have any number of meanings. It could mean that the more risk averse person is willing to pay less money to play a gamble than the less risk-averse person. Or that whenever the more risk averse person chooses a gamble as opposed to the sure thing, the less risk averse person will definitely choose the gamble as well. Or that for the more risk averse person, the pain of dollars lost is high relative to the pleasure of a dollar gained, while for the less risk averse the two are more equal.

An important result from economic theory is that these three types of risk aversion are actually all equivalent. Whoever is more risk averse by one definition is also more risk averse by the other two. The way that economists typically measure risk preferences is by the third concept - the ratio between someone's pain from losing money and their pleasure from gaining it. If someone really hates losing money and doesn't really care that much about gaining more of it, clearly they will be loath to risk any losses.

More formally, for economists, measuring risk preferences entails estimating a "utility of wealth function" that describes how happy an individual is with each level of wealth. This function can then be used to calculate how happy or sad the individual would be about moving from one wealth position to another. How risk averse a person is can be described by how concave the function is - in other words, how much the subjective value of each additional dollar declines as wealth increases. This in turn determines, at any given wealth level, how much more valuable a dollar lost is relative to a dollar gained.

Now we can ask the question of how to estimate this function. It turns out that for technical reasons, you can pick a very low wealth level and assign it a utility value of zero, and pick a very high level and assign it a utility value of one, without changing risk preferences. Risk preferences are captured in how "bowed" the function is between those two points. We discover this by offering the individual a choice between taking some gamble or getting some other amount for sure (called a "certainty equivalent"), and observing whether he chooses the gamble or the sure thing.

For instance, suppose we set a wealth of $1,000 as being worth zero "utils", and set a wealth of $1,000,000 as being worth one "util". Then we ask the individual which of these they prefer:

(A) a 50-50 gamble between having $1,000 or having $1,000,000, or

(B) having $300,000 for sure.

This is called a "lottery choice."

We know the expected utility of the gamble in (A). It is 0.5 (the probability of the \$1,000) times 0 (the utility of that outcome), plus 0.5 (the probability of the \$1,000,000) times 1 (the utility of the \$1,000,000). This is equal to 1/2. So if the individual chooses the gamble, we know that the utility of \$300,000 must be less than 1/2. If he chooses the \$300,000 instead, we know that its utility must be more than 1/2. This simple binary choice gives us either an upper or a lower bound on the utility value of \$300,000. More binary choices allow us to establish narrower bounds. The more such lottery choices we see the individual make, the tighter an estimate we can obtain for his utility function.

4.2 Measuring Risk Preferences With Lottery Choices

The procedure described in the previous section is a basic PMM. It describes one of the most common ways that most economists measure risk preferences in experiments (Holt and Laury 2002). So now we ask, what are the ideal characteristics of the lottery choices an economist would like to offer someone in the real world, in order to gauge his risk preferences?

First, the choices should be consequential, meaning that the subject will actually get the option that he chooses, at least with some probability. Economics holds that true preferences can be known only by revealed preference, i.e. by observing the real choices that people make.

Second, the choices should be simple. If the subject does not understand the gambles offered to her, he stands little chance of making consistent choices that reflect his underlying risk preferences.

Third, the choices should also be specific. If a PMM includes only a generic question like "Do you like to take more risky options or do you usually prefer to have something for certain?", this does not help very much in estimating a utility function. We wouldn't know the utility of the lottery to which the certainty equivalent is being compared, and even if we did, we wouldn't know where on the utility function to put our estimate of the value of the certainty equivalent.

Fourth, the choices should span monetary outcomes that are relevant for the decisions to which the economist would like to then apply the estimated utility function. For example, if the economist were interested in a subject's investment decisions, she would be interested in the subject's risk preferences over large fractions of that investor's wealth. In other words, if an investor has \$100,000 and could conceivably end the investment period with between \$50,000 and \$200,000, then a questionnaire that involved wealth levels of \$100 to \$900 would be inappropriate. If, however, the economist were interested in the subject's purchases of insurance for consumer electronics, she would instead be interested in risk preferences for gambles that involve a few hundreds of dollars, so the \$100-to-\$900 range might be appropriate.

Finally, there should be as many choices as possible. As discussed above, the more choices the subject makes, the more closely the economist can guess the subject's utility function. In addition, having a lot of choices means that if a subject makes a "mistake" and chooses a lottery that he doesn't really want, there will be lots of other "correct" choices to dilute that mistake. This means that the survey will present a more precise and accurate estimate of the subject's true preferences, which in turn means that these estimates will remain more stable over time.

The perfect survey would incorporate all of these considerations. However, it would be very difficult and expensive to administer. And so, not wanting to let the perfect be the enemy of the good, economists generally have to make some compromises when they want to measure someone's risk preferences. For instance, in large surveys where incentivization is logistically impossible, economists must resort to hypothetical choices. While subjects tend to exhibit less risk aversion in hypothetical choices than incentivized choices (Holt and Laury 2002), these choices still correlate well with both choices in incentivized tasks (Dohmen et al. 2011). Hypothetical lottery choices also correlate well with real financial risk taking behavior (Barsky et al. 1997). Psychologists who do research in decision making often do not incentivize choices, but even so, many of the main patterns observed first under hypothetical choice (Kahneman and Tversky, Johnson and Thaler, Lichtenstein and Slovic 1971) have held up when tested by economists using incentivized choices (Grether and Plott, Thaler, Tversky, Kahneman, and Schwartz).

Researchers also generally do not have the funding to offer lotteries involving huge amounts of money, so must settle for offering choices over smaller amounts. Subjects often exhibit levels of risk aversion over these smaller amounts that would imply implausibly large levels of risk aversion over larger amounts (Rabin 2000). Even so, it is still possible to estimate reasonable utility functions for payoffs that span an order of magnitude (Holt and Laury 2002). Also, incentivized choices between small gambles correlate with financial risk taking (Dohmen et al. 2011).

4.3 Riskalyze's Method for Measuring Risk Preferences

The Riskalyze PMM faces many of these same tradeoffs, but is able to mitigate some of the problems faced by researchers.

Riskalyze's PMM is extremely simple and specific. Its choices are binary, using numbers that are familiar to investors. This allows precise estimation of indifference points, and minimizes the risk that an investor will make mistakes while taking the questionnaire.

Riskalyze's questionnaires are, in the strictest sense, hypothetical, since the exact probabilities and payoffs specified in the questionnaire are not the ones that his eventual portfolio will face (in fact, they couldn't be, since there are multiple questions and only one eventual portfolio). But the choices do carry real consequences, since the dollar amounts are relevant to the investor, and it is clear that the choices will be used to make actual investment decisions on the investor's behalf. This is an improvement on the surveys used in academic research, in which choices are completely inconsequential to the survey taker. It allows Riskalyze the benefit of being able to offer choices over the large sums of money that will really be at risk in the investor's portfolio, while still offering the investor incentives to think hard about what her preference really is. Even better, since the financial adviser can ask how much wealth the investor has, the survey can hone in on the areas of the utility function most relevant to the investor's investment decisions.

Ideally, the number of indifference points obtained would be much more than the number of parameters in the utility curve to be estimated; in Riskalyze's case, the number of parameters is five, while the number of indifference points is five to seven. In theory, Riskalyze's technology could provide the repeated sampling that would be needed to reduce noise in the measurement. However, Riskalyze has chosen not to exercise this capability, which given the severely limited time and attention span of investors, may be an unavoidable choice.

5 A Comparison to Psychometric Risk Taking Scales

There are many possible ways to measure risk preferences. While it may seem as though there ought to be a single "best" way of measuring risk, there is no single perfect measure. A recent review paper on elicitation methods (Charness, Gneezy, and Imas 2013) identifies five different types of PMMs, ranging from the simple to the complex, and discussed the advantages and disadvantages of each. One of these PMMs, which they call the "questionnaire" method, is what we refer to in this paper as the "quiz" method. Except for the quiz method, all of the PMMs described in the paper are incentivized - that is, the subjects receive real money according to the choices they make. Three of the incentivized methods they describe measure expected utility using real lottery choices, similar to the ideal method that we described in the previous section.

Additionally, the Society for Judgment and Decision Making has compiled an inventory of Risk Attitude Measures that includes over 40 different individual scales for measuring risk (SJDM.org). So while there is no one perfect measure, selecting a PMM must be based on the question being asked and the capability of respondents to answer meaningfully.

Psychometric scales and risk tolerance quizzes are simple methods of eliciting risk preferences. They have the virtues of being simple, quick, and cheap to administer. Perhaps for this reason, quizzes have long been commonplace in much of the finance industry (see for example Bodie, Kane and Marcus 2013, ch. 6).

For understanding what kinds of people take risks or for assessing how men and women differ in risk preferences, psychometric scales do really well. That is, psychometric scales are able to detect relative differences in risk taking as a function of individual difference or demographic measures. For these types of questions, the risk elicitation method need not be incentivized, and requires no mathematical sophistication. However, this does not address the more specific question of whether a risk tolerance quiz is the right elicitation method for a financial adviser to use with an investor. We address this next.

5.1 Risk Tolerance Quizzes

Consider a standard risk tolerance quiz. It might include questions like "I consider myself highly knowledgeable when it comes to investments." with answer choices ranging from 1="Strongly Disagree" to 7="Strongly Agree". The set of questions are then scored such that lower values indicate less comfort with risk and higher values indicate more comfort with risk. In this system, different score ranges correspond to risk profile categories such as "Highly Conservative", "Conservative", "Moderate", "Aggressive", or "Highly Aggressive". While it's possible to use the scores to compare two different investors to determine person A is less risk tolerant than person B, it is not clear (based on the questionnaire results alone) what portfolios are ideally suited to match the risk preferences for person A and and person B, and certainly there are more than five categories of portfolios. Advisers would generally like to do more than simply to match an investor to one of a handful of generic risk descriptions.

In practice, financial advisers tend not to be interested in comparing the risk tolerance levels for two people on a relative basis. Instead, their objective is to measure risk preferences quantitatively, in order to assess whether or not an individual's portfolio is appropriately designed. While risk tolerance scales are simple and straightforward to administer, psychometric scales cannot be used for estimating a utility function.

In part, the inability to estimate a utility function from a psychometric scale is because most risk tolerance quizzes are assessed without considering an important factor: an investors level of wealth. Since the middle part of the 20th century, economists have recognized that utility should be modeled as reference dependent; however, this element remains conspicuously absent in most risk quizzes.

Consider two investors, one with a \$300,000 portfolio and another with a \$3,000,000 portfolio. These two individuals are likely to differ on many dimensions, including, age, investment goals, and other sources of income. However, these factors are not incorporated into the risk measurement calculation under a traditional quiz-type PMM.

So a 70-year old investor with a stable pension might be risk loving, and define a devastating loss as \$200,000, despite having a \$300,000 portfolio. In contrast, a 28-year old with a \$3,000,000 portfolio might be risk averse, because his investing goal is to generate income while he writes a book. As such, he would define a devastating loss as \$50,000.

While a traditional risk tolerance quiz would identify differences in risk preferences between the two individuals, the result would not provide quantifiable clarity as to what type of portfolio would be most appropriate for each set of circumstances. In practice, traditional quizzes and technologies relying on these inputs are likely to have difficulty handling cases that deviate from the stereotypical norm the classic assumption that younger people have less money and are more aggressive, while older people have more money and are more conservative.

Of course, many financial advisers do use risk tolerance quizzes, which raises the question of how the results inform decisions related to portfolio construction. In practice, quizzes do not elicit a precise measure of risk preferences; rather they provide a means of opening up a conversation about risk. In reviewing the results with an investor, the adviser is able to develop a sense of what an investor perceives as risky. In using the quiz to engage an investor in a systematic conversation, the adviser can translate his or her sense of that conversation into constructing a portfolio that matches the investor's preferences, and perhaps guide the investor to make better financial decisions. However, this practice requires a leap of faith on both sides. The adviser must have faith in his or her ability to adequately read the "tea leaves" revealed in the conversation, and the investor must have faith that the adviser is making choices that reflect his true sensitivity to risk.

In addition, quiz-type PMMs involve a somewhat rigid assumption. They treat risk preferences as a fixed, or at least very slow-changing personality trait. If this is the case, there is little reason to re-administer risk tolerance quizzes in response to things like market events. This makes them somewhat inflexible in terms of real-world use. In contrast, utility-based PMMs like that of Riskalyze are more flexible - because they assume that utility functions are also functions of unobserved parameters, an adviser can re-administer them at any time without raising questions about the method's validity.

In contrast to a coarse measurement that must be interpreted, Riskalyze's PMM can be used to calculate a precise estimate of risk tolerance. By eliciting risk in a way that permits an accurate estimate for utility of wealth, it is possible to directly compare an individual's preferences with the risk profile of a specific portfolio. The Riskalyze methodology also allows the adviser to quantitatively and precisely measure the riskiness of an investor's portfolio. In using the Riskalyze metric to assess both risk tolerance and portfolio riskiness, the comparison between the two becomes meaningful for an investor. This frees the adviser to focus on creating value for investors by helping them to make better decisions, because the investor's preferences are instantly clear.

6 Quantitative Methodology

In order to assist advisers with the task of applying risk preferences to portfolio choice, a technology like Riskalyze's needs to measure the quantitative risk and expected return of candidate portfolios. In this section, we describe and evaluate Riskalyze's quantitative portfolio evaluation methodology. We find that their methods for estimating expected returns and variances of portfolios conform to the current industry standard, and offer advisers flexibility when combining their beliefs with the data. We suggest some ways in which these methods could be augmented with cutting-edge quantitative techniques.

6.1 Overview

In order to evaluate a portfolio using Modern Portfolio Theory, or any other quantitative method, the expected return and volatility of the portfolio should be estimated. The choice of estimation procedure reflects a tradeoff between sophistication on the one hand, and ease of computation on the other. Computationally easy estimation methods also tend to be useful for communication within the financial industry, since limited computational power and mathematical sophistication among the majority of firms means that simple approaches are widespread. The ideal approach will balance these considerations. In addition, the ideal approach should have a way of combining forward-looking information with historical data, including subjective beliefs.

Riskalyze's technology assists advisers with one specific portfolio decision - a six-month portfolio allocation. Therefore, Riskalyze needs to estimate expected return and volatility over a six month period. They do this by estimating monthly expected returns and variances, and then multiplying each estimate by six to yield a six-month value. This assumes normally distributed returns, which is a good approximation for monthly equity returns (Egan 2007). It also ignores error propagation.

6.2 Estimation of Portfolio Expected Returns

To estimate monthly expected returns, Riskalyze uses a Carhart 4-factor model, as described in Carhart (1997). The risk factors in this model are:

1. the return of the S&P 500 (the "market return"),
2. a size factor,
3. a value factor, and
4. a momentum factor.

The estimation of this model is described in the appendix.

Factor models, based on Arbitrage Pricing Theory, are appealing models of risk and return for several reasons. They are parsimonious, easy to estimate, and have a natural and intuitive economic explanation. Unlike equilibrium models such as the Capital Asset Pricing Model, they allow for some securities to have alphas, which allows for the possibility of market outperformance by an active manager. The Carhart 4-factor model, which is based on the 3-factor model of Fama and French (1993), has become the standard model of risk and return throughout much of the finance industry (Hanauer 2014).

6.3 Time Horizon for Expected Return Estimation

The expected returns of the market portfolio and the factor portfolios are estimated using the most recent 37.5 years of monthly data. The factor betas of assets in the candidate portfolio, however, are estimated using monthly data going back only to January 1, 2008.

Riskalyze uses relatively long-term data when calculating the expected returns of factor portfolios, because it believes that these expected returns are relatively stable. It uses short horizons for calculating factor betas because it believes that these betas are not stable over the long term.

6.4 Estimation of Portfolio Variance

To calculate the variance of a candidate portfolio, Riskalyze estimates the portfolio's covariance matrix using the sample covariance matrix. The sample covariance matrix is a standard choice of estimator, for several reasons. It is a natural estimator in the absence of prior information, and it is both consistent and unbiased (Bai and Shi 2011). It is popular and widely used in the finance industry, allowing for ease of communication. Finally, it is straightforward and easy to construct.

6.5 Time Horizon and Missing Data

Sample covariance matrices are calculated using monthly data starting from January 1, 2008. Riskalyze chooses to use a short sample period because they believe that correlations and volatilities are not stable over longer periods of time. They also believe that U.S. financial markets have displayed a range of behavior since 1/2008 that is representative of market behavior in the near future; this includes a large crash in 2008-9, a sideways market in 2011, and a bull market in the years since 2011. Thus, the choice of sample period represents Riskalyze's own prior beliefs about the nature of market volatility and correlation.

When data for an asset in the covariance estimation is not complete over the sample period, Riskalyze proxies for the missing data points, using the index itself as a proxy. This procedure is described in the appendix.

6.6 Alternatives to Sample Covariance Estimation

Riskalyze's current method for estimating portfolio variance, although it conforms to industry norms, has clear drawbacks. The sample covariance can be problematic when the number of instruments is not much less than the number of historical returns. In this case, the covariance matrix can be close to singular (Ledoit and Wolf 2003). It is also unable to account for features like regime switching, factor structure, etc. (Fan et al. 2006, Huang 2006). Here we briefly describe three alternative approaches that we believe could replace sample covariance estimation.

The first alternative is to bootstrap backfilled data. This procedure, which is described in more detail in the appendix, would be more efficient and less biased than the current backfill procedure, although it would increase computational complexity.

The second alternative is to use shrinkage estimation. This procedure, described in Ledoit and Wolf (2003), estimates the covariance as the weighted sum of the sample covariance and a structured covariance matrix F. This procedure would also allow Riskalyze or its clients, financial advisers, to impose priors on the portfolio variance estimates - for example, the prior that the true risk of a large market crash is greater than the last few years might suggest (Wang and Pillai 2011).

The third alternative is to use a factor model itself for estimating the covariance matrix. This approach has become increasingly popular in recent years (Fan et al. 2006). Its main appeal is that it dramatically reduces the number of parameters to be estimated; since Riskalyze uses monthly data, that is an especially attractive feature. Riskalyze already uses a Carhart 4-factor model for estimating expected returns, so it would be natural to use the same model for the variance calculation. We cover the basics of factor-based covariance estimation in the appendix; for a detailed summary, see Fan et al. (2006).

6.7 Interest Rate Sensitivity

Riskalyze gives advisers the option to evaluate any security as a bond. In this "interest rate sensitive" mode, the monthly returns of those assets are regressed on the 10-year Treasury interest rate, yielding an interest rate sensitivity (basically a duration) for the asset. This allows advisers to conduct "stress tests" or sensitivity tests, on the portfolio by inputting scenarios for interest rate changes.

6.8 Incorporating Forward-Looking Information

Estimates of expected returns and variances using sample means and sample covariances are subject to well-known problems. Due to the low information-noise ratio, the sample mean has little out-of-sample predictive power. Other problems with sample covariance estimators have already been discussed.

One basic drawback of these estimators is that they are purely backward-looking. Methods such as the Black-Litterman model have been developed in order to combine forward-looking information with historical information in a formal, mathematically consistent manner. For Riskalyze, this sort of procedure is currently infeasible. However, Riskalyze does allow advisers to replace the historical estimates of expected returns and variances with their own beliefs. One option for an adviser is to examine the historical estimate and to then enter a desired value; this allows advisers to combine their own priors with historical information in an intuitive Bayesian way.

7 Applying Risk Measurements to Portfolio Choice

Once preferences have been assessed and the risk and expected return of candidate portfolios have been estimated, it becomes possible to combine the two - to use preferences to help choose a portfolio with the amount of risk and expected return that suit an investor's needs.

How best to do this? It depends on the nature of the portfolio that the adviser is managing on behalf of the investor. If the investor keeps all of his or her riskless assets in a separate account, for example, Modern Portfolio Theory would dictate that the adviser simply choose the portfolio with the highest expected Sharpe ratio, and let the investor decide how much to allocate to risk-free bonds. If the adviser is managing the investor's complete portfolio, however, theory would instruct the adviser to choose the portfolio that maximizes the investor's utility. There are many other cases, such as the possibility that an investor delegates part of her risky portfolio to the adviser.

Riskalyze's technology assists financial advisers with the problem of matching risk preferences with the quantitative risk of a portfolio. The preference measurement technology itself does not perform the portfolio selection task for advisers; that is, it does not search the universe of investable assets for the optimal portfolio. Riskalyze does provide advisers with an additional tool that does perform this optimization task, but it is optional. If advisers want, they can ignore the optimizer tool, and search for an optimal portfolio themselves. In this case, Riskalyze will still give an estimate of the portfolios expected return and variance.

Riskalyze also does not force advisers to choose a portfolio whose ratio of reward to risk is optimal given the measured risk tolerance of the investor. This is because the portfolio the adviser is managing may or may not be the investor's complete portfolio, and because of various other constraints the adviser many have. Because of this diversity of cases, Riskalyze does not force the adviser into any particular optimization procedure; its numbers are a guideline, not a decision rule. In the next section we discuss possibilities for how the numbers might be used to make decisions.

7.1 Modern Portfolio Theory

A key result in Modern Portfolio Theory is the two-fund theorem. This states that when a risk-free asset is available, an investor should choose a portfolio of assets with the highest available Sharpe ratio, and then allocate his wealth between that optimal risky portfolio and the risk-free asset (see, for example, Bodie, Kane and Marcus 2013).

Riskalyze's technology gives an adviser the ability to do this either automatically (with the optimizer tool) or manually. The estimates of risk and return provided by the quantitative tools can allow an adviser to do a manual search for candidate portfolios with high Sharpe ratios. Riskalyze's PMM then yields an estimate of the degree of risk aversion, allowing the adviser to solve for the optimal allocation between risky and risk-free assets.

In the event that a risk-free asset is not present (i.e. the adviser does not believe that, say, Treasuries are truly risk-free), the portfolio should be chosen to maximize utility over the efficient frontier of risky assets. Riskalyzes quantitative tools and PMM give an adviser the ability to calculate an investors utility from holding a candidate portfolio, and can thus allow utility maximization. If they wish, advisers can use the optional optimizer tool to search for a utility-maximizing portfolio.

In particular, there is the question of whether it is optimal for an adviser to match an investor's portfolio risk number with his individual risk number. This is optimal, from the perspective of Modern Portfolio theory, IF:

1. the adviser has access to the same universe of investable assets that Riskalyze does, and
2. the adviser uses the same estimates of expected return and covariance that Riskalyze uses.

In this case, the adviser can maximize the investor's utility by setting the portfolio risk number equal to the individual risk number, and then choosing the available portfolio with the highest expected return, given the constraint. This is a much simpler procedure than a global search over all possible portfolios.

7.2 Readministering Questionnaires

The optimal frequency and timing of the administration of PMMs depends on several factors.

First of all, it depends on the investor's wealth level. If an investor's utility is defined over wealth levels, then a significant change in his assets will imply a significant change in his risk preferences.

Second, it depends on whether utility is assumed to be reference-dependent, and how the reference point is determined. For example, if the reference point is assumed to always be the investor's current assets under management at the time that the PMM is administered, then a PMM that defines utility over gains and losses may not need to be readministered very frequently. But if the reference point is the investor's expected return, then the PMM may need to be administered frequently, as market conditions - and therefore, expectations - change.

This flexibility allows Riskalyze itself to be agnostic about reference points. Since Riskalyze's PMM can accomodate loss aversion, and since advisers can adjust the timing of the questionnaires based on what kind of reference-dependent utility they believe in, advisers using Riskalyze can measure preferences that, for example, follow the Prospect Theory of Kahneman and Tversky (1979).

8 Conclusion

In this paper, we evaluated Riskalyze's PMM by comparing it to theory and to related empirical research. However, any true evaluation of a PMM must take a stand on the reason for measuring investors' preferences in the first place. In other words, there is no way to get around the question of the financial adviser's problem. Advisers, and PMM providers like Riskalyze, must answer this question for themselves, since no definitive answer exists yet in the finance literature.

If investors always act in their own best interest when making real-world investment decisions, then the best PMM is the one that most accurately predicts what investors would do if they had the time and the tools to manage their own portfolios. If this is the case, then the best way to prove the worth of a PMM is to measure whether it actually predicts the decisions that investors make when they do have the time and the tools.

If, however, investors have biases or fundamental uncertainty that prevents them from being able to maximize their own utility, then this is not the measurement to make. A PMM that perfectly predicted investor decisions would not be a good guide for advisers in this case, since those decisions would be suboptimal.

In this latter case, a better measurement of a PMM's effectiveness may be the success of the investor-adviser relationship itself. If the market for financial advisers is reasonably efficient, then advisers who maximize their investors' utility will tend to be fired less than advisers who take too much risk or who don't get enough average return to justify the risk they take. Thus, the true test of a PMM may simply be the market test - if a technology helps advisers retain their clients, it is a good one.

We believe (without having a thorough knowledge of all the products that exist) that Riskalyze's technology is likely to prove to be a good one. However, there are changes that could improve it.

One way we believe Riskalyze could improve its methodology would be to use data to give advisers suggestions about when to readminister risk tolerance questionnaires to investors. Currently, this is left to the advisers' judgment. However, some economic theories, including Prospect Theory (Kahneman and Tversky 1979), hold that risk tolerance can change when an investor's portfolio crosses a certain "reference point." Behavioral finance researchers do not yet have a good general understanding of reference points. But if data on investors' independent decisions can be correlated with measured preferences, it will allow Riskalyze's system to issue automatic guidance to advisers regarding when to retest their investors' preferences. This would also allow academic finance researchers to improve their understanding of reference-dependent utility.

Riskalyze's quantitative methods, in the meantime, can best be improved by increasing the sophistication of the methods used to estimate covariance matrices (an improvement we suspect would be useful across much of the financial industry).

In general, we believe that the movement toward technologies that take economic theory seriously is a good one. The idea of utility functions as quantifiers of investors' desire to seek return and avoid risk has plenty of support in the experimental economics and finance literatures, but has been slower to gain traction in the world of industry, probably because of the difficulty of measuring utility.

Technologies like Riskalyze's have the potential to make utility measurement easier, more accurate, and more complete, and thus has the potential to improve investors' satisfaction with both their choices and the results of those choices.

References

[1] Jushan Bai and Shuzhong Shi. Estimating high dimensional covariance matrices and its applications. 2011.

[2] Brad M Barber and Terrance Odean. Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The journal of Finance*, 55(2):773–806, 2000.

[3] Robert B Barsky, Miles S Kimball, F Thomas Juster, and Matthew D Shapiro. Preference parameters and behavioral heterogeneity: an experimental approach in the health and retirement survey. Technical report, National Bureau of Economic Research, 1995.

[4] Zvi Bodie et al. *Investments*. Tata McGraw-Hill Education, 2009.

[5] Mark M Carhart. On persistence in mutual fund performance. *The Journal of finance*, 52(1):57–82, 1997.

[6] Gary Charness, Uri Gneezy, and Alex Imas. Experimental methods: Eliciting risk preferences. *Journal of Economic Behavior & Organization*, 87:43–51, 2013.

[7] DMIDI. Risk Attitude Measures. `http://www.sjdm.org/dmidi/Risk_Attitude.html`, 2015. [Online; accessed 02-March-2015].

[8] Thomas Dohmen, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G Wagner. Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550, 2011.

[9] Paul W Drake. *Money doctors, foreign debts, and economic reforms in Latin America from the 1890s to the present*. Number 3. Scholarly Resources Incorporated, 1994.

[10] William J Egan. The distribution of s&p 500 index returns. *Available at SSRN 955639*, 2007.

[11] Jianqing Fan, Yingying Fan, and Jinchi Lv. High dimensional covariance matrix estimation using a factor model. *Journal of Econometrics*, 147(1):186–197, 2008.

[12] David M Grether and Charles R Plott. Economic theory of choice and the preference reversal phenomenon. *The American Economic Review*, pages 623–638, 1979.

[13] Matthias Xaver Hanauer. *Risk factors and capital market anomalies*. PhD thesis, Universität München, 2014.

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

[15] Jianhua Z Huang, Naiping Liu, Mohsen Pourahmadi, and Linxu Liu. Covariance matrix selection and estimation via penalised normal likelihood. *Biometrika*, 93(1):85–98, 2006.

[16] Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, pages 263–291, 1979.

[17] Tze Leung Lai and Haipeng Xing. *Statistical models and methods for financial markets*. Springer, 2008.

[18] Olivier Ledoit and Michael Wolf. Honey, i shrunk the sample covariance matrix. *UPF Economics and Business Working Paper*, (691), 2003.

[19] Sarah Lichtenstein and Paul Slovic. Reversals of preference between bids and choices in gambling decisions. *Journal of experimental psychology*, 89(1):46, 1971.

[20] Matthew Rabin. Risk aversion and expected-utility theory: A calibration theorem. *Econometrica*, 68(5):1281–1292, 2000.

[21] Richard H Thaler and Eric J Johnson. Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management science*, 36(6):643–660, 1990.

[22] Richard H Thaler, Amos Tversky, Daniel Kahneman, and Alan Schwartz. The effect of myopia and loss aversion on risk taking: An experimental test. *The Quarterly Journal of Economics*, pages 647–661, 1997.

[23] Roman Vershynin. How close is the sample covariance matrix to the actual covariance matrix? *Journal of Theoretical Probability*, 25(3):655–686, 2012.

[24] Hao Wang and Natesh S Pillai. On a class of shrinkage priors for covariance matrix estimation. *Journal of Computational and Graphical Statistics*, 22(3):689–707, 2013.

[25] Kent L Womack and Ying Zhang. Understanding risk and return, the capm, and the fama-french three-factor model. *Tuck Case*, (03-111), 2003.

Appendix

A Estimation of a Carhart 4-Factor Model

The Carhart 4-factor model assumes the following functional form for the return of a security i:

$$r_t^{(i)} = \alpha + \beta_1^{(i)} f_t^{(1)} + \ldots + \beta_4^{(i)} f_t^{(4)} + \epsilon_t^{(i)},$$

where

$r_t^{(i)}$ is the monthly return on asset i;

$f_t^{(1)}$ is the excess return of the market (rmrf), measured as the return of the S&P 500 minus the one-month Treasury bill rate;

$f_t^{(2)}$ is the monthly Fama-French Small Minus Big factor (smb), computed as the average return for the 30% of stocks in the S&P 500 with the smallest market capitalization, minus the average return of the 30% with the largest market capitalization in that month. A positive value for smb in a month indicates that small cap stocks out-performed large cap stocks in that month; a negative smb in a given month indicates that large caps outperformed;

$f_t^{(3)}$ is the monthly Fama-French High Minus Low factor (hml), computed as the average return for the 50% of stocks in the S&P 500 with the highest book-to-market ratio minus the average return of the 50% of stocks with the lowest book-to-market ratio in that month. A positive hml in a month indicates that value stocks outperformed growth stocks in that month; negative hml in a given month indicates that growth stocks outperformed;

$f_t^{(4)}$ is a Momentum factor up minus down (umd), which is computed as the average return on the two high prior return portfolios minus the average return on the two low prior return portfolios;

and $\epsilon_t^{(i)}$ are i.i.d. random variables assumed to be normally distributed with pdf $N(0, (\sigma^{(i)})^2)$.

B OLS Estimation

For each asset i, the coefficients of the factor model are estimated by minimizing the sum of squared errors:

$$\hat{\theta} = \arg min \sum_{t \in T^{(i)}} {\epsilon_t^{(i)}}^2 = argmin \sum_{t \in T^{(i)}} (r_t^{(i)} - (\alpha + \beta_1^{(i)} f_t^{(1)} + \ldots + \beta_4^{(i)} f_t^{(4)}))^2,$$

where $\hat{\theta} = \left(\hat{\alpha},\ \hat{\beta}_1^{(i)}, \ldots,\ \hat{\beta}_4^{(i)} \right)$, and $\hat{\epsilon}_t^{(i)}$ are the sample errors.

This optimization yields the estimator:

$$\left(\hat{\alpha},\ \hat{\beta}_1, \ldots,\ \hat{\beta}_4 \right)' = \left(F'F\right)^{-1} F'y,$$

where

$$F = \begin{pmatrix} 1 & f_1^{(1)} & \ldots & f_1^{(4)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & f_T^{(1)} & \ldots & f_T^{(4)} \end{pmatrix}, y = \begin{pmatrix} r_1^{(i)} \\ \vdots \\ r_T^{(i)} \end{pmatrix}$$

The variance-covariance matrix of the sample coefficients is

$$V(b) = \left(\sigma^{(i)}\right)^2 \left(F'F\right)^{-1}.$$

Therefore, the estimator is

$$\hat{V}(b) = (\widehat{\sigma^{(i)}})^2 \left(F'F\right)^{-1}.$$

$100 \times (1 - \alpha)\%$ confidence intervals can also be calculated for θ: $\hat{\theta} \pm t^*_{n-p;1-\frac{\alpha}{2}} \sqrt{\hat{V}(b)}$

C Estimation of Covariance Matrix

Using the four factor model, the theoretical covariance matrix is,

$$\Sigma = \theta' \Sigma_F \theta + D_\epsilon,$$

where Σ_F is the covariance matrix of factors, and D_ϵ is the diagonal covariance matrix of ϵ.

Therefore, the estimator is,

$$\hat{\Sigma} = \hat{\theta}'\hat{\Sigma}_F\hat{\theta} + \hat{D}_\epsilon,$$

where $\hat{\Sigma}_F$ is the sample covariance matrix of the factors, and $\hat{D}_\epsilon$ is the diagonal sample covariance matrix of ϵ.

D Backfilling Data Using a Factor Model

Here we explain how to use a factor model to backfill missing data. We denote the returns of the asset i by $(r_t^{(i)})_{t\in T}$. $T^{(i)}$is the set of times that $r^{(i)}$ is accessible, and $\bar{T}^{(i)}$ is the set of times that is not accessible.

1. $T^{(i)} \bigcap \bar{T}^{(i)} = \varnothing$
2. $T = T^{(i)} \cup \bar{T}^{(i)}$, for all i.

Applying the factor model to the available data, we have

$$r_t^{(i)} = \alpha + \beta_1^{(i)} f_t^{(1)} + \ldots + \beta_4^{(i)} f_t^{(4)} + \epsilon_t^{(i)},$$

where $f_t^{(j)},\ j = 1, \ldots, d$ are the factors, which have no missing data, and $\epsilon_t^{(i)}$ are i.i.d. random variables assumed to follow a normal distribution $N(0, (\sigma^{(i)})^2)$.

For each asset i, fit the factor model by minimizing the squared error as described above, to obtain the estimators, $\hat{\theta} = \left(\ \hat{\alpha},\ \hat{\beta}_1^{(i)},\ \ldots,\ \hat{\beta}_4^{(i)}\right)$, and $\hat{\epsilon}_t$.

There are two possible methods for filling in missing data. These are:

1. Without Bootstrapping:

For $t \in \bar{T}^{(i)}$, let the filled-in data be

$$\hat{r}_t^{(i)} = \hat{\alpha} + \hat{\beta}_1^{(i)} f_t^{(1)} + \ldots + \hat{\beta}_4^{(i)} f_t^{(4)}.$$

Riskalyze currently uses this method to backfill missing data. However, this method overweights the contribution of the backfilled data points. A better method is to use bootstrapping.

2. With Bootstrapping:

For $t \in \bar{T}^{(i)}$, generate a sample error $\hat{\epsilon}_t^{(i)}$ from $N(0, (\hat{\sigma}^{(i)})^2)$, where $(\hat{\sigma}^{(i)})^2$ be the sample variance of the error. Then the filled-in data will be

$$\hat{r}_t^{(i)} = \hat{\alpha} + \hat{\beta}_1 f_t^{(1)} + +\hat{\beta}_d f_t^{(d)} + \hat{\epsilon}_t^{(i)}.$$