Preferences

This version: 10-08-2012

Abstract

Investors generally dislike risk. Optimal portfolio choice takes into account how investors are averse to different types of risk. While mean-variance utility treats gains and losses symmetrically, other models of preferences allow investors to seek safety first, to weight the pain of losses more heavily than the benefit of gains, and to have their utility depend on their past consumption (habit) and the returns or actions of other investors.

1. Picking Up Nickels and Dimes In Front of a Steamroller

Short volatility strategies are like the sly fox in the tale of the little gingerbread man. Just as the fox slickly persuaded the tasty gingerbread man to ride on his back to cross a river, the steady and high profits generated by short volatility strategies during stable times often lure investors into complacency. Many investors pile into short volatility strategies after they have seen stable profits for many years. Then, when volatility spikes during market crashes, short volatility strategies suddenly turn around and give a nasty bite to investors. Most of the profits earned during normal times are given back at a quick snap.

Figure 1 graphs the cumulated wealth of $1 invested in a short volatility strategy at the end of March 1989 to December 2011. Monthly returns on the volatility strategy are given by the MLHFEV1 index produced by Merrill Lynch. Volatility strategies are implemented in derivative markets using options, variance swaps, and even derivatives on derivatives like options on swaps. (Merrill Lynch's index is produced by trading variance swaps.) Volatility traders also buy and sell securities with embedded options such as convertible notes and bonds.

[Figure 1 here]

The returns graphed in Figure 1 are for an investor who has sold volatility – this investor collects a premium during normal times for selling volatility protection. On the other side, the people paying the premiums have purchased protection against increases in volatility. When volatility jumps, the investor with the short volatility position experiences a loss, and these losses represent gains to those who have bought volatility protection. The seller's position is similar to that of an insurance company selling fire insurance: the insurer collects premiums when there is no fire. When there is a fire, the insurance company experiences a loss, which is the payout to the policyholder ignoring expenses and costs. (In this chapter I will always refer to volatility strategy profits from the point of view of the investor selling volatility insurance.)

The losses in short volatility strategies are devastating. In Figure 1, there are some blips in the cumulated wealth in the volatility strategy like the 1998 emerging markets crisis, the recession and turbulence following the 9/11 attacks of 2001, and even a little wobble during the beginning of the subprime crisis in 2007, but overall returns marched steadily upwards – until late 2008, when the global financial crisis caused volatility to skyrocket and stocks around the world crashed. From September to December of 2008, the volatility strategy lost more than 70%. Then in 2009 when policymakers stabilized the financial markets, the volatility strategy's profitability recovered. For comparison, Figure 1 also overlays the cumulated wealth of $1 invested in the S&P 500. Stock returns have exhibited more swings than the volatility strategy and from 2000 to 2010, stock returns ended the first decade of the 21st century flat (the *Lost Decade*). They rode a rollercoaster getting there though, dipping down through the early 2000s recession, climbing upwards during the mid-2000s and then also screamed downwards, but not to the same extent as the volatility strategy, during the financial crisis.

How much should we allocate to volatility strategies? Given that there are pronounced (left-hand) tail risks, what sort of investor should pick up the small nickels and dimes in front a roaring steamroller?

2. Choices

Economists want agents (people, firms, or institutions) to have as many choices as possible, and to guide agents to make optimal choices.[1] These choices reflect asset owners' preferences. Preferences are ultimately a collection of trade-offs: how do I balance a set of risks against a set of rewards? Trade-offs apply at a point in time, as in the choice between holding boring T-bills or picking up nickels in front of the oncoming steamroller in the volatility strategy. Trade-offs also occur across time, as in deciding how much to save today for retirement in the future. Preferences are unique to each asset owner, and ethics and morals (and sometimes lack thereof) are important factors. Each asset owner also has different goals, plans, and lifestyle choices. Psychological and family or institutional dynamics also play a role, and choices are influenced as well by the decisions of his peers and other social factors.

We represent preferences by *utility*. Utility is a number – an index numerically describing preferences in the sense that decisions that are made by ranking or maximizing utilities fully coincide with the asset owner's underlying preferences. (Economists call utility values "*utils*".) Utility is about choice, and we build a *utility function* to measure how satisfied an agent is by making choices.

Good choices in the context of optimal asset management policy start with the Ancient Greek maxim "know thyself." Specifically, we build advice premised on "how do you feel during bad

[1] My colleague Sheena Iyengar (2010) at Columbia Business School shows that in some contexts, too much choice ends up being detrimental to consumers in making choices. I do not go down this route. Framing and other psychological contexts matter in adhering to optimal portfolio choice strategies as I discuss in Section 2.6.

times." To capture this notion quantitatively, we define how the asset owner perceives risk and how the asset owner responds to it.

2.1 Risk

In his book, *Against the Gods: The Remarkable Story of Risk*, Peter Bernstein (1998) chronicles our astonishing progress in measuring risk through the development of statistics and probability starting with the invention of Arabic numerals to today's (cloud) computers.[2] We model all risk, and rewards, through these statistical methods. Our notion of risk today is extremely broad – it is not just the likelihood of a set of returns. We think of risk today in terms of probability functions encompassing many different kinds of events (or even parallel universes, if you are a theoretical physicist).[3] We even think of risk as the probability of sets of probabilities changing over time, as I explain below.

Risk in financial economics is inherently a subjective concept. We do not have controlled laboratory conditions in finance and must estimate the distribution of asset returns. The procedure used to estimate probabilities, and the model that is behind the estimation, imparts subjectivity. Nevertheless, we refer to direct estimates of the behavior of returns from data as being "*real world measures*" or "*objective measures.*" In contrast, "*subjective measures*" are probabilities that an asset owner believes in, and that differ from real-world probabilities. A coin flip, for example, has a 50 percent objective probability of landing heads, but an investor might believe the coin is biased and assign a subjective probability of 60 percent to the coin coming up heads. Probabilities of returns are harder to estimate than the probabilities of coin flips because there are an infinite number of return outcomes ranging from losing all your money (a return of -

[2] The 2008-9 financial crisis still shows that we are far from mastering systemic risk.
[3] This is measure theory. Even the foundation of Brownian motion, which we use in option pricing and dynamic portfolio choice (see Chapter XX), rests on probabilities specified over function spaces (called "functionals").

100 percent) to becoming a bazillionaire (a return of nearly infinity), and everything in between. (If you invest with leverage, your return may even be lower than -100 percent.)

In Panel A of Figure 2, I estimate objective probabilities for the volatility strategy at the monthly frequency from April 1989 to December 2011. The top half presents a histogram of returns and the bottom half graphs the strategy's *probability distribution function* (pdf). In both graphs, the *x*-axis lists the returns (what can happen), and the *y*-axis lists the probabilities of the return outcomes (how often they happen). The huge occasional loss that occurs in volatility investing, which happened during 2008-9, imparts a very long left-hand tail to its probability distribution. The very long left-hand tail is the statistical representation of picking up the nickels and dimes right before the steamroller crushes you.

[Figure 2 here]

The estimated pdf of the volatility strategy has a much thinner body and longer tail than a fitted normal distribution. We call these distributions *leptokurtic*, which is Greek for "thin body," since they have much more slender distributions at the center than a normal distribution. When your distribution has a more slender body, it also has longer tails than a normal distribution. Figure 2, Panel B repeats the same exercise for the S&P 500 for comparison. Equity returns over this time period are also leptokurtic, but much less so than the volatility strategy. While there is a noticeable left-hand tail for equities, the normal distribution is not a bad approximation. (While it is a much closer approximation for equity returns, an enormous literature formally rejects that equity returns are normally distributed.)

We summarize the pdf by *moments*. Moments measure the shape of the pdf, like where it is centered (*mean*), how disperse the returns can be (*variance* or *standard deviation*, which is the square root of the variance), how far the left-hand or right-hand tails extend (*skewness*), and how fat are the tails are (*kurtosis*). The first four moments for the volatility strategy and S&P 500 are:

	Volatility Strategy	Equities
Mean	9.9%	9.7%
Standard Deviation	15.2%	15.1%
Skewness	-8.3	-0.6
Kurtosis	104.4	4.0

The volatility strategy and equities have approximately the same mean, around 10%. We see this in Figure 1 where the cumulated wealth of both strategies results in approximately the same ending value. While the standard deviation (or volatility – and I will use these two terms interchangeably) of the short volatility strategy and equities are similar at 15%, the long-left tail of the volatility strategy gives it very large negative skewness, of -8.3, and a humungous kurtosis of 104. The S&P 500 has small skewness of -0.6 and relatively small kurtosis of 4.0. In comparison, the normal distribution has a skewness of zero and kurtosis of 3.0. Hence, in Figure 2, Panel B, the normal distribution is a fairly close match for the distribution of equity returns.

2.2 Risk Aversion

While the probability distribution of wealth shows us what returns are possible, and how often, investors do not directly use these returns in making decisions (unless they are risk neutral). The utility function defines how a return outcome is felt as a bad time or good time by the investor. Simply put, utilities convey the notion of "how you feel." The utility function defines bad times for the investor.

Let us define the utility as a function of wealth, W, so that bad times are times when we are poor. (Technically we write $U(W)$, where utility, U, is a function of wealth, W.) The investor's utility is not final wealth itself – the utility transforms wealth to an investor's subjective value of wealth. Figure 3 plots three commonly used utility functions as a function of wealth: exponential, logarithmic, and constant relative risk aversion (CRRA), which is also called *power utility*. I have purposely left the units off the *x*- and *y*-axes. These utility functions are *ordinal*, so that the actual numerical value of utility has no meaning. The numbers merely denote ranks: we can replace the utility function U by $a + bU$ for any positive constants a and b without changing the preference ordering. (That is, utility orderings are unchanged by *affine transformations*.)

[Figure 3 here]

All of the utility functions in Figure 3 increase with wealth, which realistically reflects the fact that asset owners are generally greedy. Furthermore, utility functions are all *concave* over wealth. Concavity is a measure of how much investors value an extra $1 of wealth. When investors are poor, we sit close to the *y*-axis in Figure 3, and the slopes of all the curves are very steep. When you only have $1, you really value going from $1 to $2. When wealth is high, the utility curves flatten out: when you already have $10 million, there is little to be gained from going from $10,000,000 to $10,000,001 in wealth. The utility curves exhibit *diminishing marginal utility* in wealth. This is an appealing property of concave utility functions.

The slope of the utility function (the changing marginal utility as wealth increases) is how we measure just how bad, or good, the investor feels. Utilities which decrease very fast as wealth decreases correspond to investors for whom losses really hurt. Asset owners who have utilities that are very flat as wealth increases don't care much for increasing their wealth when wealth is already very high. Times of high marginal utility are bad times for the asset owner. These are times where the utility curve is very steep and the asset owner really, really wants another $1

("Brother can you spare a dime?"). Good times are when marginal utility is low, which in Figure 3 correspond to when the asset owner is wealthy and the utility function is very flat.

The investor's *degree of risk aversion* governs how painful bad times are for the asset owner. Technically, risk aversion controls how fast the slope of the utility function increases as wealth approaches zero, and how slowly the utility function flattens out as wealth approaches infinity. Risk aversion controls the degree of concavity in the asset owner's utility function.

Everyone wants the bird in hand. Risk aversion measures just how much the investor prefers the sure thing. Consider the utility function in Figure 4. There are two outcomes: *X* (low) and *Y* (high). Suppose each of them occurs with equal probability. The two vertical lines drawn at *X* and *Y* correspond to the asset owner's utility at these low and high outcomes of wealth, *U(X)* and *U(Y)*, respectively. Consider what happens at the wealth outcome $\frac{1}{2}X+\frac{1}{2}Y$, which is represented by the center vertical line. If wealth is equal to $\frac{1}{2}X+\frac{1}{2}Y$ for certain, then the asset owner's utility is given by the point on her utility curve denoted by $U(\frac{1}{2}X+\frac{1}{2}Y)$. This point is marked with the star.

[Figure 4 here]

Now consider the straight, diagonal line connecting *U(X)* and *U(Y)*. If future wealth is *X* with probability $\frac{1}{2}$ and *Y* with probability $\frac{1}{2}$, then the expected utility $\frac{1}{2}U(X)+\frac{1}{2}U(Y)$ lies on the diagonal line connecting *U(X)* and *U(Y)* and is shown by the triangle. Utility is higher for the certain outcome than for the random amount:[4]

$$U(\tfrac{1}{2}X+\tfrac{1}{2}Y)\geq \tfrac{1}{2}U(X)+\tfrac{1}{2}U(Y).$$

[4] Or $U(E[X])\geq E[U(X)]$ which is a consequence of Jensen's inequality for the concave function.

The greater the difference between the star and the circle in Figure 4, the more risk averse the investor. Put another way, the more risk averse the asset owner becomes, the more she wants the sure thing. The more concave the utility function, the more risk averse the investor.

In the special case that the investor is *risk seeking*, the utility function becomes convex rather than concave. If an investor is *risk neutral* the utility function is linear. These cases are rare; most investors are risk averse.

2.3 How Risk Averse Are You?

Risk aversion can change as wealth changes, and for most individuals risk aversion decreases as wealth increases – individuals generally take on more risk as they become more financially secure. For this reason, I will talk about (relative) risk aversion, where risk aversion is measured relative to wealth and for short hand I will drop the word "relative" from now on.[5]

Constant relative risk aversion (CRRA) utility takes the following form:

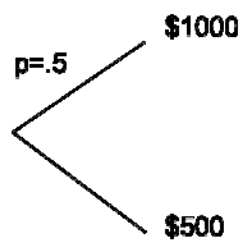
$$U(W) = \frac{W^{1-\gamma}}{1-\gamma}, \tag{0.1}$$

where γ is the asset owner's risk aversion coefficient. CRRA utility is widely used in portfolio choice theory. A very attractive property of CRRA utility is that it leads to portfolio weights that do not depend on wealth: it doesn't matter whether one is managing \$10 million or \$100 million – the technology of asset management is exactly the same – and this is the beauty of the scalability of the wealth management industry, at least for the fund manager. (This property is

[5] Formally, there is the absolute risk aversion coefficient introduced by Arrow (1971) and Pratt (1964) defined as $-U''(W)/U'(W)$ and the relative risk aversion coefficient defined as $-WU''(W)/U'(W)$. Note these definitions of risk aversions are invariant to affine transformations just as utility ordering is itself invariant to affine transformations. Levy's (1994) subjects exhibited decreasing absolute risk aversion as they become wealthier and approximately constant relative risk aversion.

called *wealth homogeneity*). Under CRRA utility, risk aversion is the same for all levels of wealth – so the advantage for this utility function is also an inherent disadvantage, because investors are generally more risk averse over losses than over gains (see Section 4). But it is convenient that we can summarize all behavior with just one risk aversion parameter. With this caveat in mind, what risk aversion levels do investors have?

This question is pretty meaningless for most people. I don't describe myself using CRRA utility in equation (1) and rattle off a risk aversion parameter (but yes, some of my academic colleagues can). We can, however, estimate risk aversion levels indirectly. Suppose I gave you the following lottery:



where you can win $1,000 with 50% probability or win $500 with 50% probability. That is, you will get $500 for sure, but you have the possibility of winning $1,000. How much would you pay for this opportunity? The real-life counterparts to this lottery are financial advisors' questionnaires which try to infer clients' risk aversion.[6]

We can map what an investor would pay against the investor's degree of risk aversion, γ. (We compute this through a *certainty equivalent*, and we detail this concept further when we discuss mean-variance utility below.) The following table allows us to read off risk aversion levels as a function of the amount you would pay to enter the gamble:

[6] Some of these are covered, along with other techniques to elicit clients' risk aversions, in "Stay the Course," Columbia CaseWorks, ID#110309, 2011. See also Grable and Lytton (1999).

Risk Aversion γ	Amount You Would Pay
0	750
0.5	729
1	707
2	667
3	632
4	606
5	586
10	540
15	525
20	519
50	507

An individual not willing to risk anything – you get $500 for sure – is infinitely risk averse. An individual willing to pay the fair value of the gamble, which is $750, is risk neutral and has a risk aversion of $\gamma = 0$. If an investor is willing to risk more than the fair value then she is risk seeking.

Most individuals have risk aversions between one and 10; it's very rare to have a risk aversion greater than 10. That is, most people would be willing to pay between $540 and $710 to enter this lottery. These estimates come from a very large body of experimental and survey evidence. Some of the papers in the literature are quite creative. Andrew Metrick (1995) studied participants on the game show Jeopardy. The stakes on this show are high, and there is tremendous pressure because only the winning player gets to return. (Jeopardy participants are also really smart, so you can't say they don't know what they're doing!) Metrick finds that risk aversion levels are low and near risk-neutrality. More recent estimates are provided by Aarbu and Schoryen (2009), who surveyed Norwegians and found risk aversion estimates around four. Kimball, Sahm and Shapiro (2008) also perform surveys and their estimates are higher, at around eight. Note that these are surveys, however, rather than real financial choices. My colleagues at

Columbia Business School, Paravisini, Rappoport and Ravini (2010), examined actual financial decisions made by investors in an online person-to-person lending platform deciding who gets funding, and how much. They estimated investors have risk aversions around three.

2.4 Expected Utility

Expected utility combines probabilities of outcomes (how often do these events occur?) with how investors feel about these outcomes (are they a bad time for me?):

$$U = E[U(W)] = \sum_s p_s U(W_s). \tag{0.2}$$

In equation (2), the subscript *s* denotes outcome *s* happening, so expected utility multiplies the probability of it happening, p_s, with the utility of the event in that state, $U(W_s)$. Technically, we say that expected utility involves *separable* transformations on probabilities and outcomes. The probabilities measure how often bad times occur, and the transformations of outcomes capture how an investor feels about these bad times.

Expected utility traces back to the 1700s, to the mathematicians Gabriel Cramer and Daniel Bernouilli, but its use as a decision making tool exploded after John van Neumann and Oskar Morgenstern formalized expected utility in *The Theory of Games and Economic Behavior* in 1944. Von Neumann was one of the greatest mathematicians in modern times, and he also made seminal contributions in physics and computer science. He and the economist Morgenstern created a new field of economics – game theory – which studied how agents interact strategically with each other, as in the famous Prisoner's Dilemma game. Von Neumann and Morgenstern initially developed expected utility to apply on games of chance and expected utility theory was just an accompaniment to the overall theory of games. The stock market is a game of chance –

but unlike poker, stock market probabilities are ultimately unknown. In an important extension, Savage (1954) showed that expected utility still applied when the probabilities were subjective and derived from agents' actions, rather than being taken as given. Thus, the probabilities in equation (2) can be objective or subjective.[7]

To make choices, the asset owner *maximizes expected utility*. The problem is formally stated as:

$$\max_{\theta} E[U(W)] \tag{0.3}$$

where θ are choice (or control) variables. Some examples of choice variables include asset allocation decisions (portfolio weights), spending/savings plans, and production plans for a firm, etc. The maximization problem in equation (3) is often solved subject to constraints. For an asset owner, these constraints often involve constraints on the *investment universe*, *position constraints* like *leverage constraints*, no *short-sale constraints*, or the inability to hold certain sectors or asset classes, and when the asset allocation problem is being done over time there may be constraints that depend on an investor's past asset positions (like *turnover constraints*).

A very special class of the expected utility framework is the set of *rational* expected utility models. Traditionally, rational expected utility models were derived to satisfy certain axioms (like the independence axiom, which states roughly that if I prefer *x* to *y*, then I also prefer the possibility of getting *x* to the possibility of getting *y* when *x* and *y* are mixed with the same lottery; violation of this maxim led to loss aversion utility we describe in Section 4). Not surprisingly many of these axioms are violated in reality.[8] Rejecting the whole class of expected utility as a decision-making tool solely on this restrictive set of rational expected utility models is

[7] See Schoemaker (1982) for a summary on the expected utility model and its variants.
[8] An excellent book that offers insight into investors' behavioral tendencies is Kahneman's (2011) *Thinking Fast and Slow*.

misguided; expected utility remains a useful tool to guide asset owners' decisions – even when these investors have behavioral tendencies, as most investors tend to do in the real world. In fact, the broad class of expected utility nests behavioral models, as I describe below.

2.5 What Choice Theory is Not About

How agents make decisions, and giving them advice to make good decisions, is not about

1. Wealth

   Of course, greater wealth allows asset owners to have more choices. But wealth per se is not the objective. Figure 1 illustrates that how we get to a given level of wealth – through slow and steady increases or by extreme ups and downs – can matter just as much as the end level of wealth. The entire concept of a utility function captures the notion that investors do not value each dollar of wealth one-for-one.

2. Happiness

   Happiness is an important emotion and maximizing happiness could be a criterion for making a decision. Happiness is also correlated with wealth in empirical studies.[9] One popular interpretation of an asset owner maximizing utility by taking a series of actions is that she achieves bliss (or as close to it as possible) by choosing a particular portfolio. This is not strictly correct. The asset owner balances risk and return and dislikes certain risks, such as the risk of large and ruinous losses, more than others, such as the risk of small fluctuations. The preferences of the asset owner capture both risk and return, which pure happiness-seeking does not.

3. Rationality vs. Behavioral Approaches

[9] Although wealthier people are happier than poorer people within countries, Easterlin (1974) found rich countries are not happier, on the whole, than poorer countries. There was no change in happiness over time, even though there have been tremendous increases in GDP per capita. Recently Stevenson and Wolfers (2008) have disputed some of these findings.

> A healthy debate rages between finance academics over whether high returns for a particular style of investing (value stocks, for example, have higher returns on average than growth stocks, see Chapter XX) are due to rational or behavioral stories. This is less relevant for the asset owner, who should treat the rational and behavioral frameworks as ways to make optimal decisions. Most people do not have rational expected utility. But that does not mean that certain rational expected utility models are not useful in making decisions.

2.6 The Normative vs. Positive Debate

An important distinction is between normative economics (what is the best portfolio to hold?) versus positive economics (what do people actually hold?). The normative vs. positive debate is ultimately philosophical, but the asset owner needs to spend some time thinking about this issue.

The normative approach starts with characterizing the asset owner. This approach assumes you "know thyself" and can describe the set of bad times and how you feel about those bad times. Combined with estimating the probability of bad times, expected utility provides a way for us to set an optimal allocation policy. This advice is what you should do. This is normative economics.

But what if the investor does not follow this "optimal" advice? Then, we could conclude that our theory is not describing what the investor is actually doing. If we want a better description of what how the investor is acting, perhaps we should go back and re-visit the utility function, or go back and re-estimate estimate the probabilities perceived by the investor. We can now find the actual utility function the investor has, rather than the utility function we originally assumed. Since our investor failed to act in the way we prescribed, let's better characterize how the investor perceives risk. This is positive economics.

I concentrate on normative asset management. That is, I give prescriptive advice for what asset owners should do. It turns out that most people and institutions do not, unfortunately, invest the way optimal theory tells them. For example, the theory that we develop advocates:

- Diversifying widely

  But many individuals hold concentrated positions in their employers' stocks. Many investors also fail to invest overseas and suffer from home bias. (We take up these issues in the next chapter.)

- Rebalancing

  But many asset owners fail to rebalance, as it involves buying assets that have lost value. Why, they wonder, should I buy a loser? (See Chapter XX.)

- Dissaving after retirement

  Many individuals actually continue to save, rather than dissave, when they enter retirement years as predicted by lifecycle models (see Chapter XX). Annuities are also predicted to be very good investments for retirees, but sadly few retirees hold annuities.

- Using factors, not asset class labels, in investing

  Putting a label on a type of financial intermediary, like putting "private" in front of "equity" does not make "private equity" an asset class. Calling a collection of funds "absolute return strategies" does not make them so. Few investors delve into the underlying factor drivers of their portfolio returns (see Chapter XX).

- Recognizing that your asset manager is not your friend

  Agency issues permeate the asset management industry. People often do not give asset managers the proper incentives or monitor them correctly (see Chapter XX).

Many investors fail to follow normative advice because of behavioral tendencies, poor governance, or the inability to be time consistent. Couching normative recommendations in ways that mitigate these effects is a newly developing science that takes into account the way people respond to framing and incentives.[10] We will touch on some of these issues in the chapters to come. Asset owners are advised that the best normative frameworks take into account the behavioral or institutional settings that hinder asset owners from following advice, especially advice given in traditional ways. Good normative structures also take into account the possibility that the asset owner will fail, and give advice specifically for these contingencies.

2.7 Non-Monetary Considerations

Choices reflect our underlying personalities – who we are, our ethics, the culture of our institutions, and our core beliefs. These considerations should enter our utility framework as well. In my exposition, utility functions represent how investors quantitatively measure bad times. Violations of core principles are some of the worst times – moral and social indignation can be translated to extremely high utility values, too.

For certain institutions, stakeholders expect asset management to be conducted a certain way to support the institution. The Norwegian sovereign wealth fund, CalPERS, CalSTRS, and other prominent large funds refuse to hold tobacco stocks. Norway goes further and automatically excludes all companies making cluster munitions and nuclear weapons from its portfolio. It has divested from companies for human rights violations (Wal-Mart), destroying the environment

[10] See Thaler and Sunstein (2009), among others.

(Rio Tinto and Freeport McMoRan), and other violations of ethical norms (Potash Corporation of Saskatchewan). Many of these decisions have generated substantial attention in the press.[11]

For Norway's sovereign wealth fund, ethical investing reflects the preferences of the asset owner – the Norwegian people. Doing this confers legitimacy on the funds in the eyes of the stakeholders. In many other funds, the asset management style choice rightly reflects the stakeholder preferences of the funds. In the extreme case, risking the existence of an institution by failing to manage assets in a certain way is actually bringing on the worst bad time possible – annihilation.

Utility is a representation of how asset owners make choices. We use utility to (normatively) derive asset management policies. Thus, if an asset owner needs to make choices that reflect non-monetary considerations, asset management should reflect those choices, subject to the considerations on whether our optimal policies reflect normative or positive considerations. The most important thing is to take a stand on who we are.

2.8 Summary

Expected utility is a tool for asset owners to quantitatively assess how they feel about bad times. Risk is captured by the (potentially subjective) probabilities of bad times occurring, and utility functions allow an investor to perceive how they feel about outcomes during these bad times. The more risk averse the asset owner, the more these bad times hurt and the more the asset owner prefers sure things.

While the simplest expected utility models capture bad times by the probabilities and marginal utilities of low wealth outcomes, utilities are a shorthand to represent the underlying preferences

[11] See "The Norwegian Government Pension Fund:The Divestiture of Wal-Mart Stores Inc.," Columbia CaseWorks, ID#080301, 2010.

of the asset owner. There are other definitions of bad times than just low wealth. Preferences are also affected by ethics, psychological tendencies, peer decisions, and other social factors. We return to some of these utility functions in Section 4.

3. Mean-Variance Utility

With mean-variance utility, asset owners care only about means (which they like), and variances (which they dislike). Mean-variance utility defines bad times as low means and high variances.

Mean-variance utility is given by:

$$U = E(r_p) - \frac{\gamma}{2}\mathrm{var}(r_p), \qquad (0.4)$$

where r_p is the return of the investor's portfolio and γ is her coefficient of risk aversion.

Mean-variance utility is closely related to CRRA utility (see equation (1)). In fact, we can consider expected utility using CRRA utility and mean-variance to be approximately the same and we will do so for many purposes in this book.[12] Thus, the γ in equation (4) has the same interpretation as the risk aversion in Section 2.3, and we expect most investors' γs to lie between 1 and 10. The parameter of $\frac{1}{2}$ in equation (4) in simply a scaling parameter; it is of no consequence and can be ignored.

Mean-variance utility is the workhorse utility of the investment industry. It was introduced by Harry Markowitz in his pathbreaking work in 1952, but Markowitz did not fully rationalize using mean-variance utility in an expected utility framework until 1979 in a paper co-authored with

[12] The portfolio choices made with CRRA utility and mean-variance utility are very close and will converge to each other under certain conditions, like employing (log) normal distributions and sampling intervals that tend to zero.

Haim Levy. Levy and Markowitz approximate any expected utility function by a mean-variance utility function:

$$E[U(1+r_p)] \approx U(1+E(r_p)) + \tfrac{1}{2}U''(1+E(r_p))\,\mathrm{var}(r_p), \tag{0.5}$$

where $U''(\bullet)$ denotes the second derivative of the utility function. Utility functions are concave, so their second derivatives are negative and the second term in equation (5) is negative. Hence, equation (5) shows that an investor maximizing expected utility is approximately the same as an investor maximizing mean for a given level of variance, and thus equation (5) takes the same form as mean-variance utility in equation (4).

The intuition behind the Levy-Markowitz approximation in equation (5) of mean-variance utility to any expected utility function is that the two most important effects are where the returns are centered (the mean) and how disperse they are (the variance). There is a trade-off between these two effects, which is captured by the investor's risk aversion. This is precisely the mean-variance utility setup of equation (4). Mean-variance utility is ubiquitous in the investment industry (much less so, thankfully, in academia). It has been relentlessly criticized, quite rightly, but mean-variance techniques bring us a great deal of intuition that extends to more complex situations. And the two most important things are indeed means and variances. Often other things also matter – and unfortunately mean-variance utility does not allow us to capture anything other than the mean and variance.

Mean-variance utility does not assume returns are normal. Often people confuse using mean-variance utility with assuming normally distributed returns. This error comes about because with normal distributions, there are only two parameters – the mean and the variance and these completely describe the normal distribution (they are called *sufficient statistics*). With normal distributions only means and variances matter, so mean-variance utility is perfect. Levy and Markowitz showed that using mean-variance utility is often a good approximation with non-normal returns. But sometimes the approximation can be lousy. Let's go back to Figure 1 showing the cumulated wealth of the volatility strategy and the S&P 500. Both end up with approximately the same return, and they have approximately the same mean (around 10%) and the same standard deviation (around 15%). But investing in the volatility strategy is like picking up the nickels and dimes before an oncoming steamroller – the skewness is -8 and kurtosis is 104. The S&P 500 is much closer to a normal distribution. Using only the mean and variance would imply that both the volatility strategy and the S&P 500 would be approximately the same. But clearly the volatility strategy and the S&P 500 are not the same. Would you want to use the mean and variance exclusively?

Having said that, we'll stick with using means and variances for now and return to more realistic utility functions in Section 4.

3.1 Indifference Curves

We represent mean-variance utility pictorially by *indifference curves*. One particular indifference curve represents one particular level of utility. Figure 5 plots three different indifference curve for different utility levels for an asset owner with a risk aversion of $\gamma = 3$. We plot these in mean-standard deviation space on the *y*-axis and *x*-axis, respectively. Along one particular indifference curve, an investor is indifferent to all the mean-volatility combinations. For example, on the indifference curve with the triangles in Figure 5, which corresponds to a utility of $U = 0.15$, the asset owner does not care which mean-volatility combination she picks; all of them result in the same number of utils. Investors seek the highest possible utility. Higher indifference curves lie to the left. Higher indifference curves have higher means and lower volatilities, which correspond to higher utility from equation (7). While Figure 5 plots just three indifference curves, there is in fact an infinite family of indifference curves for every investor. These indifference curves lie parallel to each other.[13]

[Figure 5 here]

In Section 2.3, we interpreted the shape of a utility function as being summarized by the risk aversion coefficient. The same intuition applies to indifference curves. The more risk averse an investor, the steeper the slope of his indifference curves. In Figure 6, I plot indifference curves for three levels of risk aversion. For the risk aversion of $\gamma = 3,$ the indifference curve has the steepest slope. If we increase risk (standard deviation), this investor needs to be compensated the most so the indifference curve shoots sharply upwards as risk increases. In contrast, the investor with $\gamma = 0.5$ is relatively risk tolerant: this investor does not need to be compensated as much with high expected returns when risk is increased. Thus, her indifference curves are relatively flat. It is important to note that an investor has one risk aversion parameter and family of indifference curves (as drawn in Figure 5). Figure 6 compares indifference curves of different investors.

[Figure 6 here]

3.2 Certainty Equivalent

The utility certainty equivalent is the sure amount of wealth, or the risk-free return, that makes the investor feel the same as holding a risky asset position.[14] The certainty equivalent is the compensation the investor requires that would make him relinquish a risky asset.

[13] They are parabolas because equation (6) is quadratic in σ_p .

For mean-variance utility, the indifference curve provides a convenient way to compute the certainty equivalent. An investor has the same utility along all combinations of risk and return on an indifference curve, by definition. The intersection at the *y*-axis of an indifference curve represents an asset with a positive return and no volatility – a risk-free asset. Thus, we can trace the indifference curve going through a risky asset position to the *y*-axis intercept to find the certainty equivalent.

Figure 7 plots the mean and standard deviation of the volatility strategy, which is marked with the square, and the indifference curve that goes through the volatility strategy for a $\gamma = 3$ investor. The certainty equivalent of the volatility strategy is 6.45%. This means that if the investor could not invest $1 in the volatility strategy, the investor would need to be compensated 6.45c on that $1; the volatility strategy is equivalent to a risk-free return of 6.45%. The certainty equivalent is also called *willingness-to-pay*. An asset owner would be willing to pay 6.45c per dollar of wealth (risk-free) to be able to have the same utility as investing in the volatility strategy (which is risky).

Mean-variance utility ignores the large negative skew of the volatility strategy in computing this certainty equivalent. The mean and standard deviation of the volatility strategy are almost the same as the stock market, so this would be approximately the same certainty equivalent of the S&P 500 as well.

[Figure 7 here]

When we compute the certainty equivalent with mean-variance utility, we make use of the fact that the mean-variance utility function in equation (7) has economic meaning stated in terms of

[14] Formally the certainty equivalent is the amount *C* satisfying $U(C) = E[U(X)]$ for the risky lottery *X*.

its level. In equation (7), the utility level itself is the certainty equivalent value. That is, mean-variance utility exhibits *cardinality*, as opposed to the formulation of expected utility in Section 2 which was only ordinal.

Certainty equivalent values are extremely useful in portfolio choice. We will use them, for example, to gauge the cost of not diversifying (see Chapter XX), to compute the cost of holding illiquid asset positions and to use certainty equivalents to estimate illiquidity premiums (see Chapter XX), and to estimate how much compensation an investor would require in exchange for foregoing the right to withdraw capital from a hedge fund during a lock-up period (see Chapter XX).

3.3 The Risk Aversion of a Typical Pension Fund

A typical pension fund holds 40% in fixed income and 60% in more risky assets, a category that includes equities, property, and alternative assets such as hedge funds and private equity. What risk aversion level does this 40% fixed income and 60% (more) risky asset position correspond to?

Figure 8, Panel A graphs all possible combinations of risk (standard deviation) and return that can be achieved by holding equities and bonds. I use U.S. data from January 1926 to December 2011 from Ibbotson as my proxies for fixed income and riskier assets. The square and circle plot the volatility and mean of stocks and bonds, respectively. The curve linking the two represents all portfolio positions holding both stocks and bonds between 0% and 100%. It is called the *mean-variance frontier*. (The next chapter will delve deeper into the economics of mean-variance frontiers for multiple assets.) In this example, the mean-variance frontier represents all the possible risk-return combinations that we can obtain by holding stock and bond positions (without employing leverage).[15]

[Figure 8 here]

Bonds lie on an *inefficient* part of the mean-variance frontier. By holding some equities, we move in a clockwise direction and strictly increase the portfolio's expected return and lower the portfolio's volatility. Thus, we would never hold a portfolio of 100% bonds. Equity diversifies a bond portfolio because equities have a low correlation with bonds; the correlation of equities with bonds in this sample is just 11%. Thus, the investor optimally will hold some equities because equities have a chance of paying off when bonds do poorly. But how much equity should the asset owner hold?

The optimal portfolio holdings depend on the asset owner's degree of risk aversion. Panel B of Figure 8 overlays indifference curves for an investor with risk aversion $\gamma = 2$. She wishes to maximize her utility:

$$\max_{w} E(r_p) - \frac{\gamma}{2}\text{var}(r_p) \tag{0.6}$$

where we choose the optimal portfolio combination *w*. For this problem, specifying a weight in one asset (equities) automatically determines the weight in the other asset (bonds) as the weights

[15] For two assets, the mean-variance frontier is traced by taking all combinations of the two assets A and B. Denoting the mean and volatility of asset A by $E(r_A)$ and σ_A, respectively, the mean and volatility of asset B by $E(r_B)$ and σ_B, respectively, and the correlation between the returns of asset A and asset B by ρ, the portfolio mean and volatility for the mean-variance frontier for two assets are given by

$$E(r_p) = wE(r_A) + (1-w)E(r_B)$$

$$\sigma_p = \sqrt{w^2\sigma_A^2 + (1-w^2\sigma_B^2) + 2w(1-w)\rho\sigma_A\sigma_B},$$

where we vary the portfolio weight *w* held in asset A. The charts in Figure 8 restrict *w* to lie between 0 and 1.

sum to one. Panel B of Figure 8 visually depicts the optimization problem in equation (6). Maximizing utility is equivalent to finding the highest possible indifference curve.

The highest indifference curve is tangent to the mean-variance frontier in Panel B at the triangle. At this point, the asset owner holds 78% in equities. There are higher indifference curves graphed, but these indifference curves do not intersect the mean-variance frontier – these indifference curves are *infeasible*. We have to lie on the mean-variance frontier as these represent the only set of portfolios that can be obtained by holding positions in equities and bonds. Thus, the triangle is the highest utility that can be obtained.

In Panel C, I show indifference curves for a more risk-averse investor with a risk aversion level of $\gamma = 7$. (Note that these indifference curves are steeper than the ones in Panel B.) The tangency point is marked in the triangle and is closer to the 100% bonds point on the mean-variance frontier. For the $\gamma = 7$ asset owner, a 32% equities position is optimal.

So what risk aversion level corresponds to a 60% equities position? Panel D plots the equities weight as a function of risk aversion. A 60%-40% equities-bond position corresponds to a risk aversion level of $\gamma = 2.8$. Thus, typical pension funds – like most individual investors – are moderately risk averse. This method of backing out risk aversion from portfolios that are observed is called *revealed preference*.

3.4 The Market's Risk Aversion

The *Capital Allocation Line* (CAL) describes portfolios that can be chosen when there is one risky asset, say equities, and a risk-free asset. For now, interpret the latter as U.S. T-bills, even though these have a tiny bit of (sovereign) risk (see Chapter XX), which we will ignore for the time being. We assume that a period is known – say a quarter – so that we know what the risk-free rate is. We denote the risk-free rate by r_f and the expected return and standard deviation of the risky asset as $E(r)$ and σ, respectively.

The CAL is described in portfolio mean-standard deviation space ($E(r_p), \sigma_p$) by the following line:

$$E(r_p) = r_f + \frac{E(r) - r_f}{\sigma} \sigma_p. \tag{0.7}$$

The CAL traces out all the possible portfolios of a risk-free asset and a risky asset. Figure 9 plots the CAL for a risk-free rate of $r_f = 1\%$ and U.S. equities using Ibbotson data over January 1926 to December 2011. U.S. equities are marked in the square.

[Figure 9 here]

The slope of the CAL is the Sharpe ratio, named after William Sharpe, one of the founders of the CAPM and winner of the 1990 Nobel Prize:

$$\text{Sharpe Ratio} = \frac{E(r) - r_f}{\sigma}. \tag{0.8}$$

In Figure 9, the Sharpe ratio for U.S. equities was 0.53, using the average equity return over the sample and a 1% risk-free rate.[16] The Sharpe ratio is actually a zero-cost trading strategy: in the context of Figure 9 it is long equities and short T-bills, divided by the risk of equities. It is the reward (mean excess return) per unit of risk (standard deviation). I will also use the term "raw Sharpe ratio" which refers to

[16] T-bill rates do move around in data, so the empirical Sharpe ratio of excess returns (stock returns in excess of T-bills) is $0.0766 / 0.1918 = 0.40$.

$$\text{Raw Sharpe Ratio} = \frac{E(r)}{\sigma}, \tag{0.9}$$

which does not subtract the risk-free rate. In Figure 9, the raw Sharpe ratio for equities was 0.58.

The problem of optimal asset allocation over one risky asset and one risk-free asset is solved by finding the tangency point of the highest indifference curve with the CAL. Figure 9 shows two tangency points for risk aversion levels of $\gamma = 7$ (Panel A) and $\gamma = 2$ (Panel B). In the $\gamma = 7$ case, the investor holds 40% equity and 60% T-bills. The tangency point on the CAL is to the left of the equity position marked with the square. In the $\gamma = 2$ case, the asset owner holds a levered portfolio. This more risk tolerant investor holds a much more aggressive equity position. He shorts the risk-free asset, or borrows money. He holds -39% in risk-free assets and a 139% position in equities. The tangency point of the indifference curve is to the right-hand side of the equities position marked with the square.

The analytical solution in this simple setting of one risky asset and one risk-free asset for the optimal weight in the risky asset, w^*, is

$$w^* = \frac{1}{\gamma}\frac{E(r) - r_f}{\sigma^2}, \tag{0.10}$$

and $1 - w^*$ held in the risk-free asset. In equation (10), note that as the investor becomes more risk averse ($\gamma \to \infty$), the risky asset weight goes to zero and the investor holds entirely T-bills. As the risky asset becomes more attractive ($E(r) - r_f$ increases or σ decreases) the optimal weight in the risky asset also increases.

What is the market's risk aversion coefficient? At the end of December 2011, the combined stock market capitalizations on the NYSE and NASDAQ were \$11.8 and \$3.8 trillion, respectively – a total of \$15.6 trillion in equities. At the same date there were \$1.5 trillion in T-bills issued by the U.S. Treasury. Taking only T-bills as the risk-free asset, we have a weight of

$$w^* = \frac{15.6}{15.6+1.5} = 91.2\%$$

invested in equities by a "representative" investor who is assumed to represent the U.S. equity and T-bill markets. (I ignore all the non-U.S. risky markets and the risky fixed income markets for the purposes of this question.) Rearranging the optimal weight in equation (10) at a risk-free rate of 1% and using the equity premium and equity volatility over January 1926 to December 2011, we have

$$\gamma = \frac{1}{w^*}\frac{E(r)-r_f}{\sigma^2} = \frac{1}{0.9123}\frac{0.1019}{(0.1915)^2} = 3.0,$$

and thus the typical participant in the equity market has a modest degree of risk aversion.

The mean-variance solution in equation (10) turns out to be the same as CRRA utility (see equation (1)) if returns are log-normally distributed. This is one sense that mean-variance and CRRA are the same.

4. Realistic Utility Functions

There are several shortcomings of mean-variance utility:

1. The variance treats the upside and downside the same (relative to the mean)

But asset owners generally exhibit asymmetric risk aversion, feeling losses more powerfully than any equivalent gain.

2. Only the first two moments matter

   People prefer positive skewness, like the chance of a big lottery payoff, and dislike negative skewness, like the barreling steamroller crushing you as you pick up nickels and dimes.

3. Subjective probabilities matter

   The way people perceive probabilities is different to the actual distribution of returns. In particular, people tend to overestimate the probability of disasters.

4. Bad times other than low means and high variances matter

   Bad times aren't just times of low wealth. Mean-variance utility rather simplistically portrays bad times as consisting only of low means and high variances. But being rich or poor in absolute terms might not matter so much as whether you are rich or poor relative to your neighbor, or whether you are rich or poor now relative to whether your were rich or poor in the past. That is, your utility could be relative.

There are many richer models of utility that incorporate all of these considerations. Some of these models fit into the expected utility framework of the previous sections, so all of the previous intuition and economic machinery apply. In economics, there are many utility functions that realistically describe how people behave. In the asset management industry, unfortunately, only one utility model dominates – by a long shot, it's the restrictive mean-variance utility model. It would be nice if we had a commercial optimizer where one could toggle between various utility functions, especially those incorporating downside risk aversion. It would be even better if an application could map a series of bad times, and how the risk of these bad times is perceived by an investor, to different classes of utility functions. Sadly, there are no such asset allocation applications that I know of as the time of writing, in 2012, that can do this. And yet all the economic theory and optimization techniques are already published.

What follows are some examples of asset allocation in the context of the volatility strategy graphed in Figure 1, where the risk in the left-hand tail is especially pronounced and a risk-free asset pays 1%.

4.1 Safety First

In the safety first utility framework, investors do exactly what the name suggests – they seek safety first. Roy's (1952) utility is very simple: it is just zero or one depending on whether a portfolio return is greater or less than a pre-determined level. If it is less than this level, a disaster results. If the return is greater than this level, the disaster is averted. Safety first minimizes the chance of a disaster; it is ideal for agents for whom meeting a liability is crucial. Not surprisingly, since safety first takes care of the downside first, asset allocation in the safety-first approach is very straightforward: you conservatively hold safe assets up until the pre-determined level is satisfied, and then you take on as much risk as you can after the safety level is met.

A related approach, formulated by Manski (1988) and Rostek (2010), is quantile utility maximization. In this approach, agents care about the worst outcomes that can happen with a given probability. Probability cutoffs of a pdf are called quantiles. Quantiles are intuitive measures of investor pessimism. An asset owner might look at the worst outcome that can occur in 90% of all situations, for example. In this case, the relevant outcome is what happens at the 0.1 quantile (which is the first decile).[17] Another asset owner might look at the worst outcome

[17] Value-at-Risk that is so commonly used in risk management is a quantile measure. People usually pick the 0.01 or 0.05 quantiles.

that might happen 50% of the time, that is what happens at the 0.5 quantile (which is the median). In Manski-Rostek's utility, the quantile is a parameter choice, like risk aversion, and it is a measure of the investor's attitude to downside risk. Because the quantile parameter choice is an alternative measure of investors' downside risk aversion, the asset allocation from this framework is similar to loss aversion, which we now discuss.

4.2 Loss Aversion or Prospect Theory

Developed in a landmark paper by Kahneman and Tversky in 1979, loss aversion was originally created as an alternative to pure, rational expected utility. It is, however, a variant of expected utility, broadly defined, because it involves measuring risk by separable functions of (subjective) probabilities and utility functions of returns. Daniel Kahneman won the 2002 Nobel Prize for his work combining psychology and economics, especially for his work on prospect theory. His long-time collaborator, Amos Tversky, unfortunately did not live long enough to be awarded the same honor. Kahneman and Tversky formulated prospect theory based on how people actually make decisions (it was a positive theory).[18]

There are two parts to prospect theory:

1. Loss Aversion Utility

   Investors find the pain of losses to be greater than the joy from gains. Loss aversion utility allows the investor to have different slopes (marginal utilities) for gains and losses.

2. Probability Transformations

   Kahneman and Tversky move beyond probabilities – even subjective probabilities. They transform probabilities to *decision weights*, which do not necessarily obey probability

[18] For a summary of applications of prospect theory, and other behavioral theories, to finance, see Barberis and Thaler (2003).

laws. They do not, for example, have to sum to one like probabilities. Decision weights allow investors to potentially severely overweight low probability events – including both disasters and winning the lottery.

We can write prospect theory as a form of expected utility similar to equation (2):

$$U = \sum_s w(p_s) U(W_s), \tag{0.11}$$

for a probability weight function $w(\bullet)$ and a loss aversion function $U(\bullet)$, which operate over different states *s* which occur with probability p_s.

Let us focus on loss aversion utility. Figure 10 graphs a typical loss aversion utility function. Utility is defined relative to a reference point, which is the origin on the graph. The reference point can be zero (absolute return), or it could be a risk-free rate (sure return), or a risky asset return (benchmark return). Gains are defined as positive values on the *x*-axis relative to the origin and losses are defined as negative values. The utility function has a kink at the origin, so there is asymmetry in how investors treat gains and losses. The utility function over gains is concave, so investors are risk averse over gains. The shape of the utility function for losses is *convex*, which captures the fact that people are *risk-seeking over losses* – that is, they are willing to accept some risk to avert a certain loss. The utility function for losses is also steeper than the utility function over gains. Thus, investors are more sensitive to losses than to gains.

[Figure 10 here]

The fact that investors are risk-seeking over losses is borne out in many experiments that involve choosing between bets like:

A: Losing \$1000 with probability of 0.5

B: Losing \$500 for sure

People overwhelmingly choose A over B. They are willing to take on some risk to avoid a sure loss.

How investors respond to losses relative to gains is governed by the coefficient of *loss aversion*, which we denote by the parameter λ. Experimental evidence points to estimates of λ around two, so losses are penalized about as twice as much as what gains add in investors' utilities. Tversky and Kahneman's (1992) estimate of λ is 2.25.

In Figure 11, I plot portfolio weights for optimal investment in the volatility strategy. The graph is produced starting for a regular CRRA investor holding 100% in the volatility strategy. The CRRA investor is a special case of the loss averse investor with no risk asymmetry and $\lambda = 1$. This investor has a risk aversion of $\gamma = 1.68$.[19] Now we start overweighting losses relative to gains, and we use a reference point of the risk-free rate of $r_f = 1\%$. As the asset owner places greater weight on downside outcomes, he lowers the weight in the volatility strategy. At $\lambda = 2$, he holds approximately only 20% in the volatility strategy and then when he weights downside outcomes more than 2.2 times the weight given gains, he foregoes the volatility strategy entirely and chooses a portfolio consisting wholly of risk-free assets.

[Figure 11 here]

[19] For comparison, a mean-variance investor with a risk aversion of $\gamma = 3.86$ would hold 100% in the volatility strategy. In this case, mean-variance is not such a good approximation to CRRA – as expected because of the very long left-hand tail of the volatility strategy. Note that that the CRRA investor needs to be more risk tolerant (lower γ) because CRRA utility takes into account the higher moments (see equation (1)), whereas mean-variance utility does not.

4.3 Disappointment Aversion

Disappointment aversion is the rational cousin to loss aversion. Like loss aversion, the asset owner cares more about downside vs. upside outcomes. It is "rational" because it is axiomatically derived (originally by Gul in 1991) and thus it is appealing for those who like the rigor of formal decision theory. This does bring benefits – the disappointment utility function is mathematically well defined and so always admits a solution (while Kahneman-Tversky's specification sometimes sends optimal portfolio weights to unbounded positions), it can be extended to dynamic contexts in a consistent fashion (see Chapter XX), and it also gives economic meaning to concepts like reference points that are quite arbitrary in classical loss aversion theory.[20]

The upside outcomes in disappointment aversion are called "elating" and the downside outcomes are called "disappointing." This is just relabeling; investors overweight losses relative to gains, just like in loss aversion utility. What is different to loss aversion is that the reference point is *endogenous* and is equal to the certainty equivalent (see Section 3.2).

The disappointment aversion parameter, A, is the equivalent of the loss aversion parameter, λ, in disappointment aversion utility. In Figure 12, I plot the optimal weight in the volatility strategy as a function of A. Disappointment aversion nests CRRA as a special case, so I start with a 100% investment in the volatility strategy, which corresponds to a risk aversion of $\gamma = 1.68$ (similar to the loss aversion plot in Figure 11). The disappointment utility investor weights disappointing outcomes more than elating outcomes by $1/A$, so values of A smaller than 1 indicate the investor is averse to downside risk. Figure 12 shows that as we decrease A (or we become more

[20] See Ang, Bekaert and Liu (2005) for an asset allocation application of disappointment aversion utility. Disappointment aversion is generalized by Routledge and Zin (2010) so that the reference point is not restricted to be the certainty equivalent.

downside risk averse), she lowers the weight of her portfolio allocated to the volatility strategy. At a disappointment aversion level of $A = 0.45$, she holds entirely risk-free assets. This corresponds to her weighting disappointing outcomes approximately twice as much as elating outcomes. This is quite similar to the loss aversion weight of λ being approximately two where the loss aversion investor has also completely divested from the volatility strategy.

[Figure 12 here]

4.4 Habit Utility

Habit utility falls into a class of utility functions where bad times are defined not just by wealth outcomes, but also by an investor's environment. Specifically, it is not wealth (or more correctly how wealth is perceived by the investor) that matters, it is wealth relative to a reference point that is important. With habit utility, bad times are defined by an investor's wealth coming close to habit.[21]

Agents quickly get used to a particular level of consumption. When you're stuck at the back of the plane, you're just glad that you're travelling somewhere exotic; the hard seat, having your knees crammed into the tray table in front of you, and battling your neighbor over the armrest aren't that bothersome (in most cases). Then you get an upgrade to business class. Whoa, is it hard to go back! And then you get the lucky bonanza of a first-class upgrade on one of those spiffy international airlines whose first-class patrons enjoy showers and seats that lie flat, plus unlimited champagne. Stepping back down to cattle class really hurts, even though by any objective standard you are better off – and certainly no worse off – than had you never escaped from coach.

---

[21] Some important papers in this large literature include Sundaresan (1989), Constantinides (1990), and Campbell and Cochrane (1999).

The “habit” in habit utility can be interpreted as a “subsistence” level of required wealth – perhaps literally subsistence consumption, but more generally as the level required to live a certain lifestyle. As wealth comes close to habit, the investor acts in a much more risk-averse manner. Thus, with habit utility, *risk aversion is endogenous*. There is an overall curvature of an agent’s utility function, but local curvature is always changing.

Habit, too, evolves over time. Your income could skyrocket so you really don’t have to go back to economy class. Or your investment bank employer could blow up and then you are stuck at the back of the plane. Habit can be external – driven by factors outside an investor’s control, such as macro factors – or internal, where it depends on the past consumption and wealth of an individual.

Asset allocation with habit utility requires you to state your portfolio’s returns in relation to your habit. In the case where you are close to habit, the asset owner’s implied risk aversion will be very high – this is a bad time for her and she wants safe assets. But in good times when her wealth is far above habit, she will be very risk tolerant and hold large amounts of equities. Habit investors are not good volatility strategy investors – when volatility spikes, prices tend to crash, and these are times when their wealth creeps closer to their habits. The volatility strategy loss comes at a particularly bad time.

4.5 Catching Up with the Joneses

Catching up with the Joneses utility defines bad times relative to other investors. It is not your performance in absolute terms that matters; what matters is your performance relative to your peers. We say that utilities exhibit *externalities* because they depend on both your own returns as well as those of your peers.[22] You want to "keep up with the Joneses."

Catching up with the Joneses is a realistic utility function for individuals who don't want to feel left out when their golf club buddies are jumping on the hottest investment tips. The same individuals feel schadenfreude their neighbors stumble. It is especially relevant for endowment managers, mutual fund managers, and others who are explicitly benchmarked against their peers. If the standard expected utility framework in Section 2 involving wealth can be described in terms of greed, catching up with the Jones is really about envy.

One of the consequences of catching up with the Joneses is that investors herd. Demarzo, Kaniel and Kremer (2005) show that when one set of investors (the Joneses) in a local community take a position in certain stocks, for rational reasons like hedging or even for irrational reasons, then other investors hold the same stocks to keep up with the Joneses. If portfolio managers are benchmarked to each other, they will want to hold the same stocks. Herding arises endogenously, even if there is no explicitly set benchmark, in catching up with the Joneses, because there is large risk for investors in deviating from the pack. The positive vs. normative implications really bite with this utility formulation. Take the board of trustees of an endowment. Do you explain (ex post) the fact that you are going with the crowd because your endowment manager has used a catching-up-with-the-Joneses utility function (positive), or are you adopting a catching-up-the-

[22] These are also called interdependent preferences and consumption externalities in utilities, which were first introduced by Dusenberry (1952). Catching up with the Joneses utility was first formulated by Abel (1990) and Gali (1994). Note whether you "catch up" or "keep up" with the Joneses is just a matter of timing and irrelevant for the purposes of this discussion.

Joneses utility function because you really are benchmarked against your other endowment peers, in which case the optimal asset allocation policy is to go with the pack (normative)?[23]

4.6 Uncertainty Aversion

We measure how often bad times occur with a pdf (see Section 2.1). But what if agents do not have a single probability distribution in mind, but have beliefs about a whole family of possible probability distributions? For example, an emerging country has newly liberalized. There's no return history. If the country is stable, returns could be drawn from a "good" distribution. But if there are multiple coups, or if investors' capital is confiscated, returns would be drawn from a "bad" distribution. There could be all possible distributions in between. There is a distinction between a single probability distribution – the traditional measure of risk – and a range of possible distributions. We call the latter *Knightian uncertainty*, for Frank Knight (1921), after terminology introduced by Gilboa and Schmeidler (1989).

Under uncertainty aversion, investors' utility depends on both risk and uncertainty. The more precise information is available about the set of possible distributions that can occur, the higher the agent's utility. Agents have an *uncertainty aversion* parameter, just as they have a risk aversion parameter. Uncertainty aversion is also called *ambiguity aversion*, as the existence of multiple distributions is also known as ambiguity. Researchers also talk about *robust utility*, which is closely related and also deals with multiple sets of probability distributions.[24] With robust utility, agents exhibit a preference ("concern") for how robust their decisions are under different probability distributions. (My MBA students might call this a glorified version of

[23] Goetzmann and Oster (2012) argue that universities vigorously compete with each other resulting in herding in asset allocation policies.
[24] The reference tome on robustness is Hansen and Sargent (2007).

sensitivity analysis.) It is also referred to as *maxmin utility* because uncertainty averse agents take the worst possible probability distribution of risk (min) and then maximize utility (max).

A comprehensive review of uncertainty aversion and portfolio choice was done by Guidolin and Rinaldi (2010). I have the following brief remarks:

1. Uncertainty aversion causes agents to effectively behave in a much more risk-averse fashion than just using plain old risk aversion. You will hold a much smaller weight in equities when you are uncertainty averse.
2. Many of the tools in expected utility can be used – and more important, the same intuition as in this chapter applies in many cases. There is an explicit connection even to mean-variance utility.[25] (Clearly in a very simple case, expected utility is uncertainty aversion with only one probability distribution.)
3. There are data on agents' different probability beliefs – surveys! Some of the more recent papers in the literature nicely exploit these to show how the set of beliefs evolve over time and how they can be used to infer future asset prices.[26]

5. Picking Up Nickels and Dimes In Front of a Steamroller Redux

Utility functions are a way to measure bad times. Times of low wealth are bad times. Bad times can also be defined relative to an absolute level of safety, a reference point defining gains and losses – where investors can exhibit different attitudes to risk across gains and losses, past wealth (habit) of an investor, or the relative performance of other investors. During bad times, marginal utility is very high and the investor considers every spare dollar especially precious.

[25] In Maenhout (2004), you can re-interpret the classical CRRA risk aversion coefficient as the sum of the CRRA risk aversion coefficient and the uncertainty risk aversion coefficient. See also Trojani and Vanini (2004).
[26] See Ulrich (2011).

Mean-variance utility is very restrictive in that bad times are represented only by low portfolio returns and very disperse portfolio returns (low means, high variances). The workhorse utility model of the financial industry is mean-variance. The volatility strategy has a mean and variance very close to the S&P 500, but it has much more pronounced crash risk and fatter tails than the stock market. The volatility strategy's probability distribution function exhibits a very skewed left-hand tail (skewness of -8) and enormous kurtosis (over 100). Mean-variance would not be an appropriate framework to evaluate this strategy, except for those investors who only have preferences over means and variances. Investors exhibiting downside risk aversion would hold much less of this volatility strategy relative to a mean-variance (or approximately CRRA) investor. If deciding between the volatility strategy and a risk-free asset, downside risk averse investors who weight losses about twice as much as gains would optimally choose to forego the volatility strategy and invest only in T-bills instead.

Figure 1

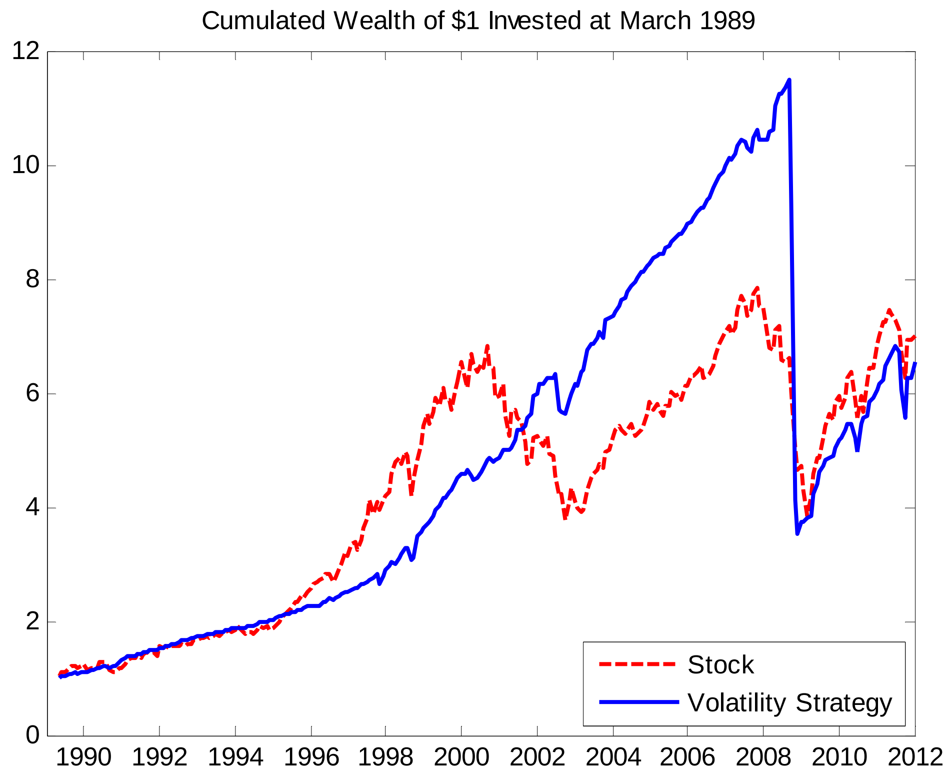

Cumulated Wealth of $1 Invested at March 1989
12
10
8
6
4
2
0
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
Stock
Volatility Strategy

Figure 2 Panel A

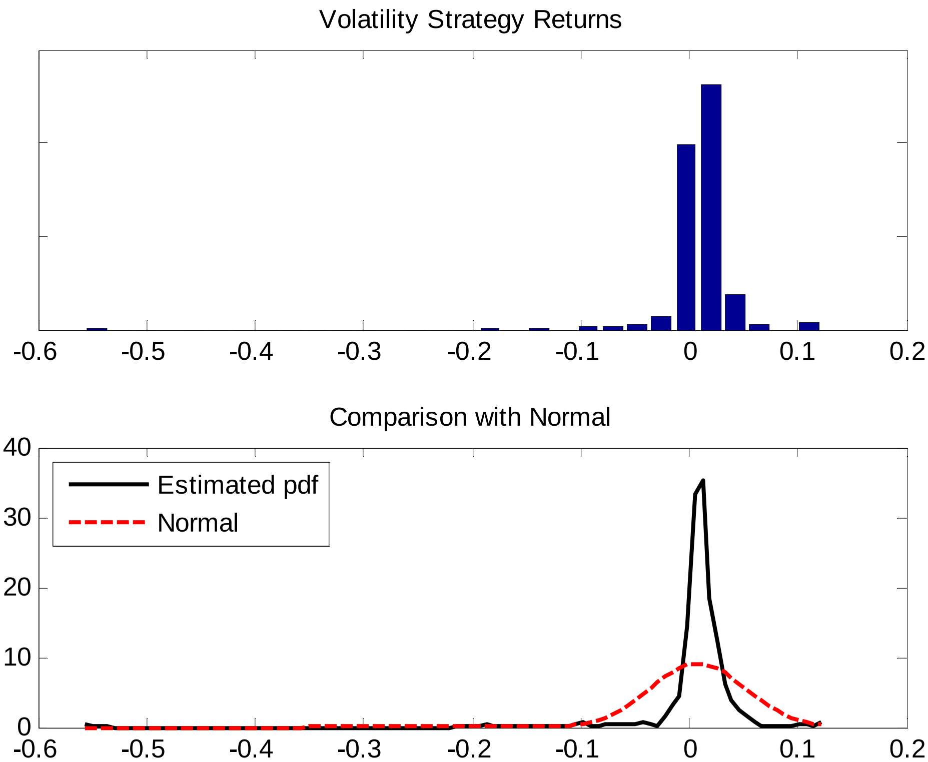


Figure 2 Panel B

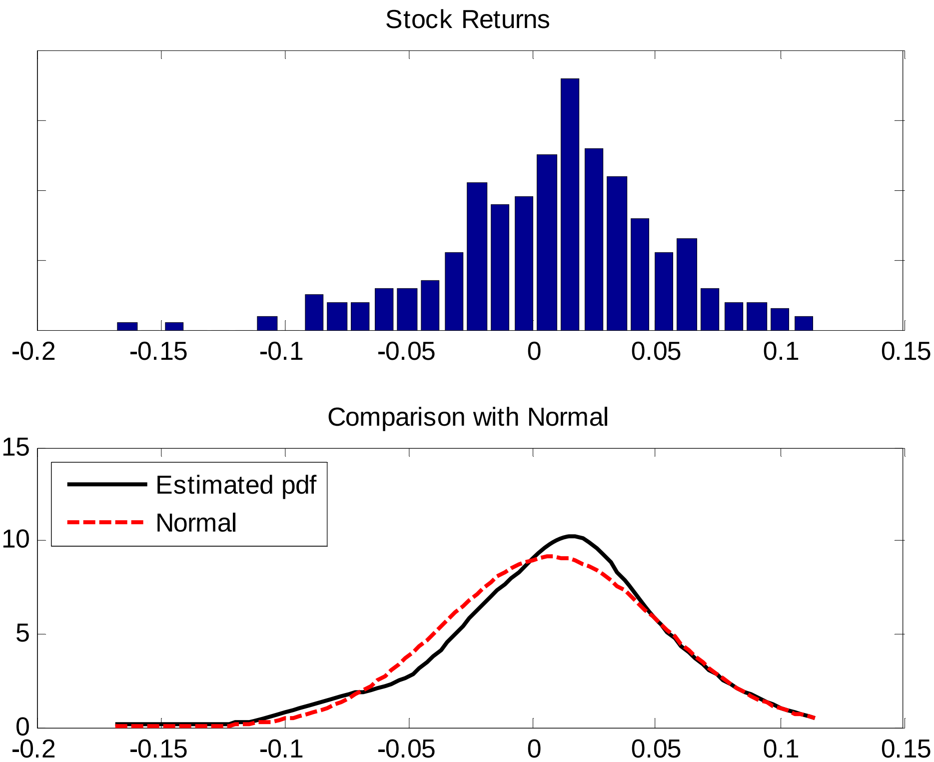


Figure 3

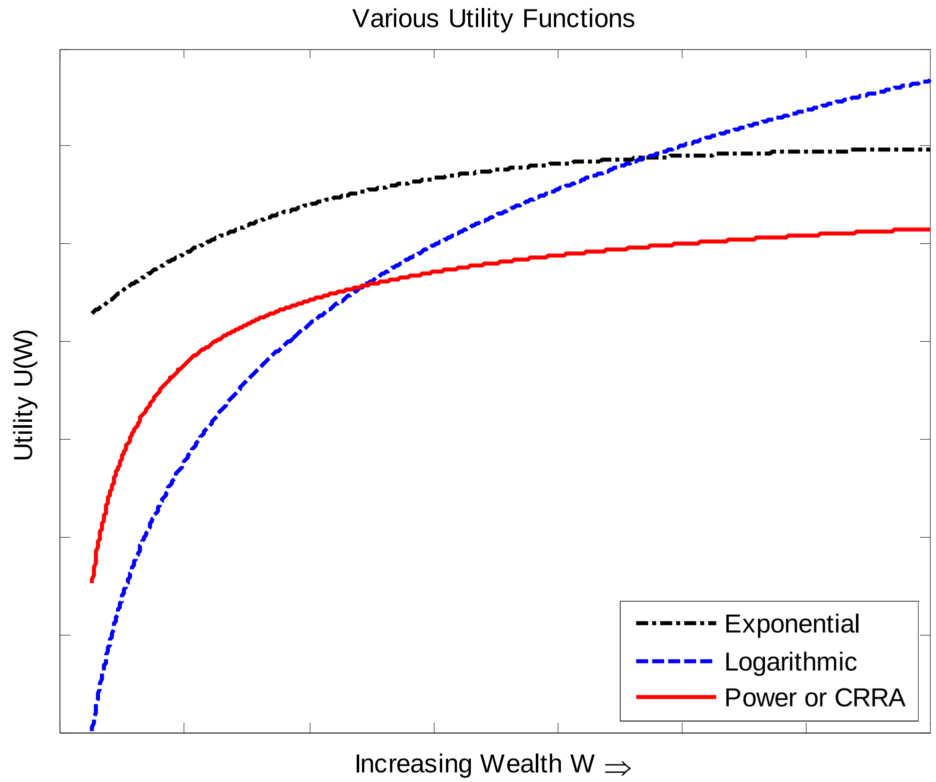

Various Utility Functions
Utility U(W)
Increasing Wealth W ⇒
Exponential
Logarithmic
Power or CRRA

Figure 4

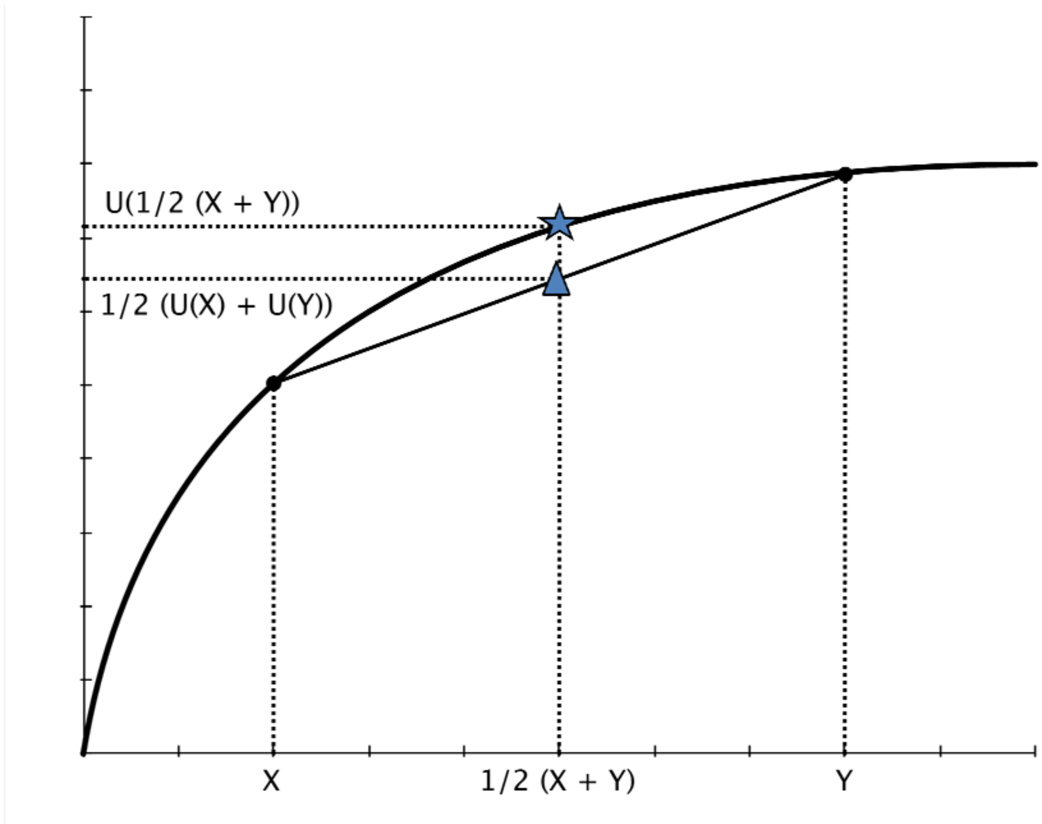


Figure 5

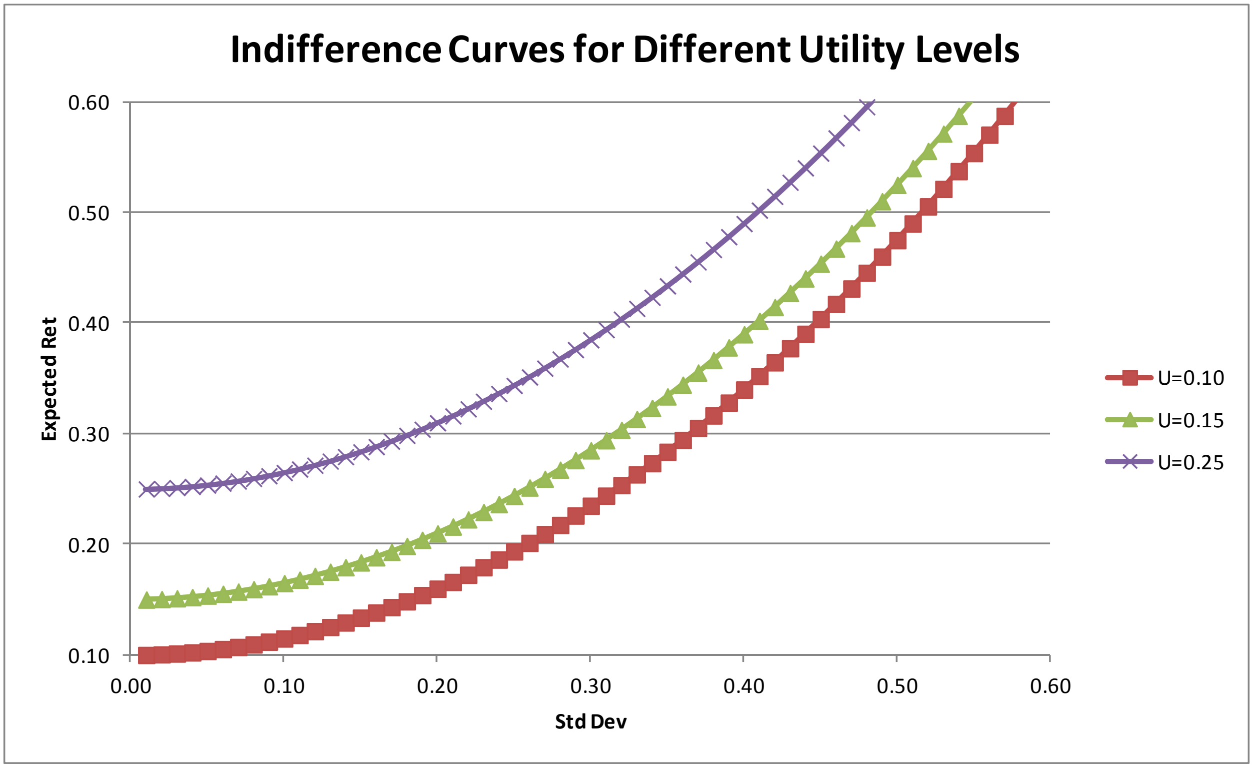

Indifference Curves for Different Utility Levels
Expected Ret
Std Dev
0.10
0.20
0.30
0.40
0.50
0.60
0.00
0.10
0.20
0.30
0.40
0.50
0.60
U=0.10
U=0.15
U=0.25

Figure 6

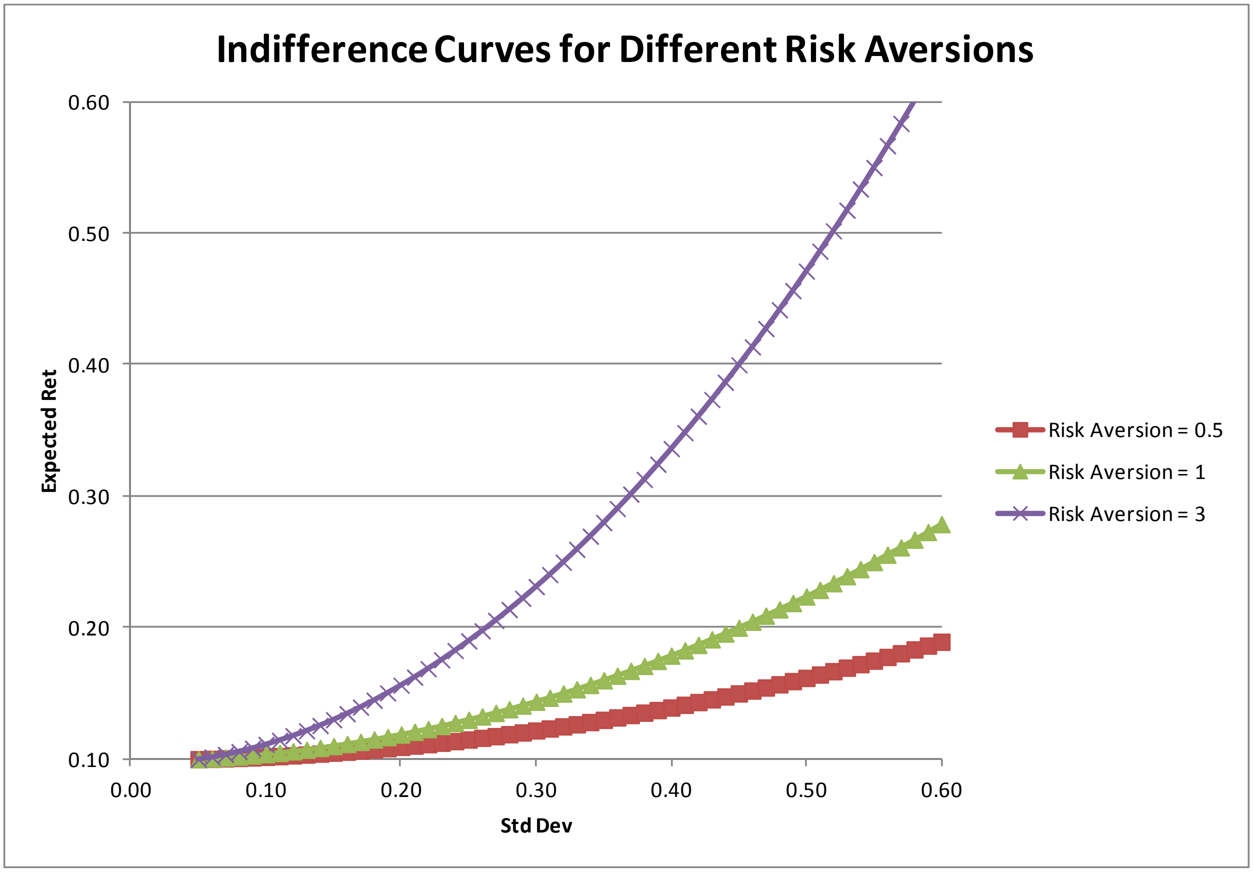

Indifference Curves for Different Risk Aversions
Expected Ret
Std Dev
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0.10
0.20
0.30
0.40
0.50
0.60
Risk Aversion = 0.5
Risk Aversion = 1
Risk Aversion = 3

Figure 7

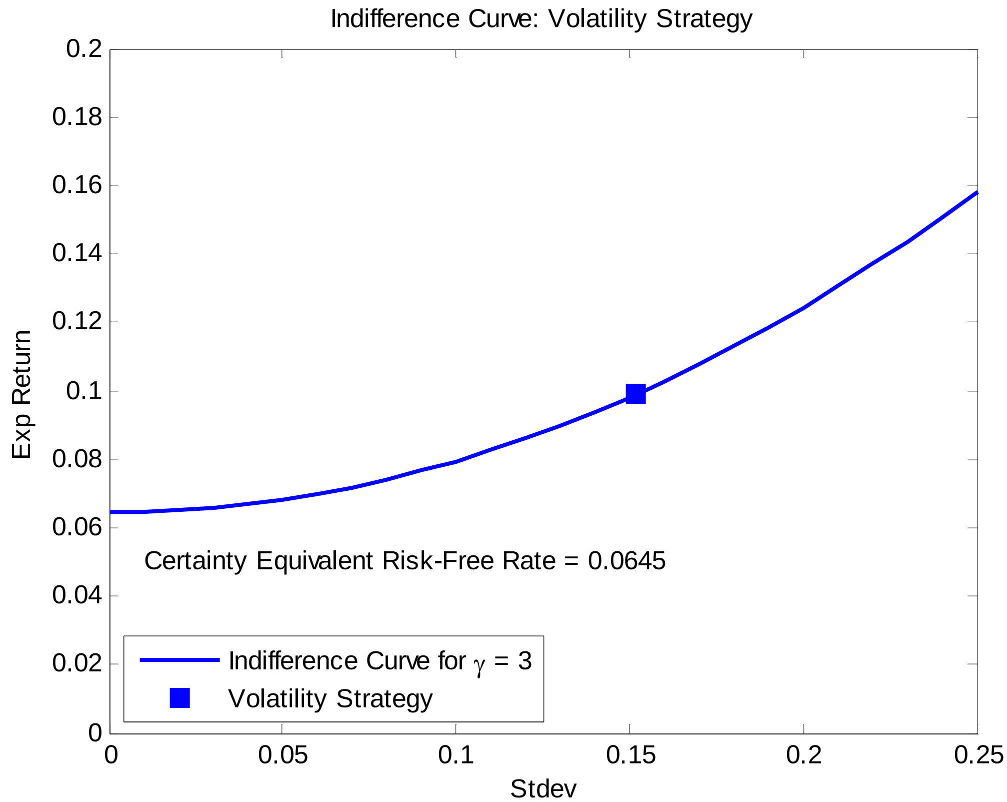

Indifference Curve: Volatility Strategy
Exp Return
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
Certainty Equivalent Risk-Free Rate = 0.0645
Indifference Curve for γ = 3
Volatility Strategy
0
0.05
0.1
0.15
0.2
0.25
Stdev

Figure 8

Panel A

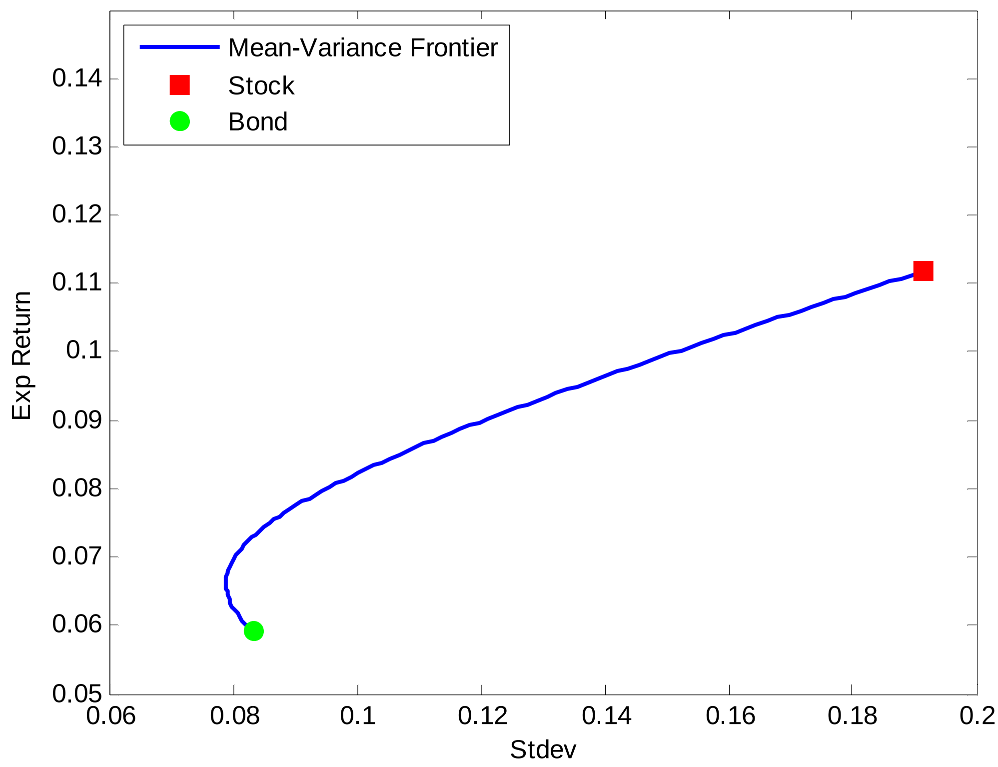


Figure 8

Panel B

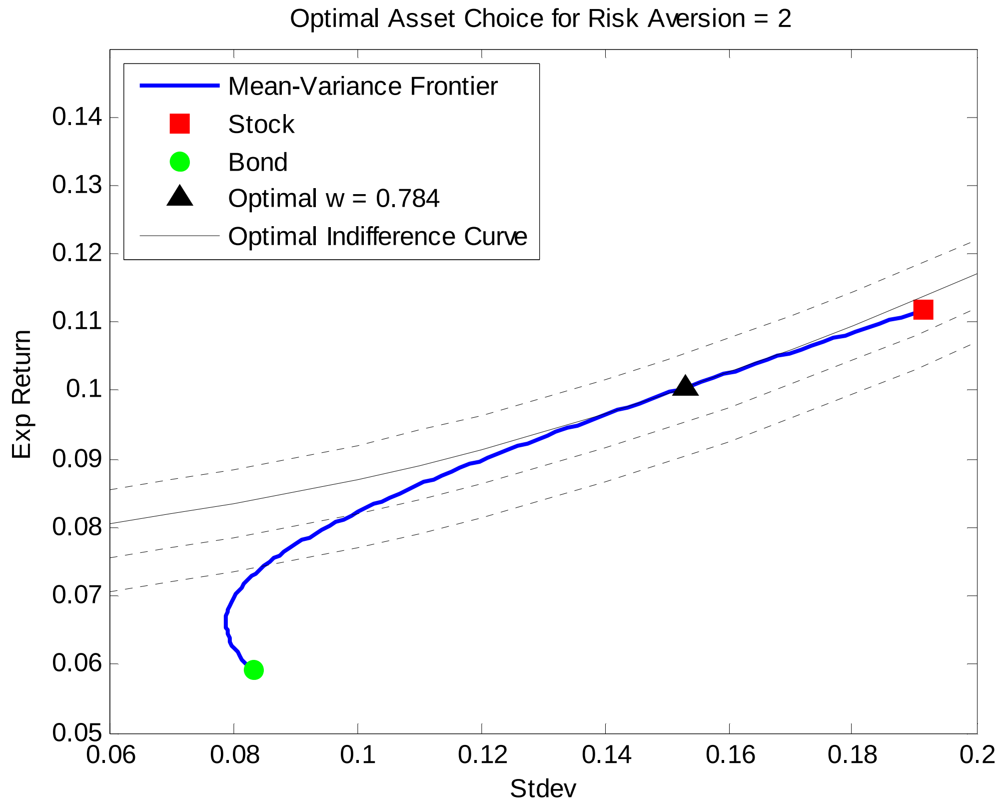


Figure 8

Panel C

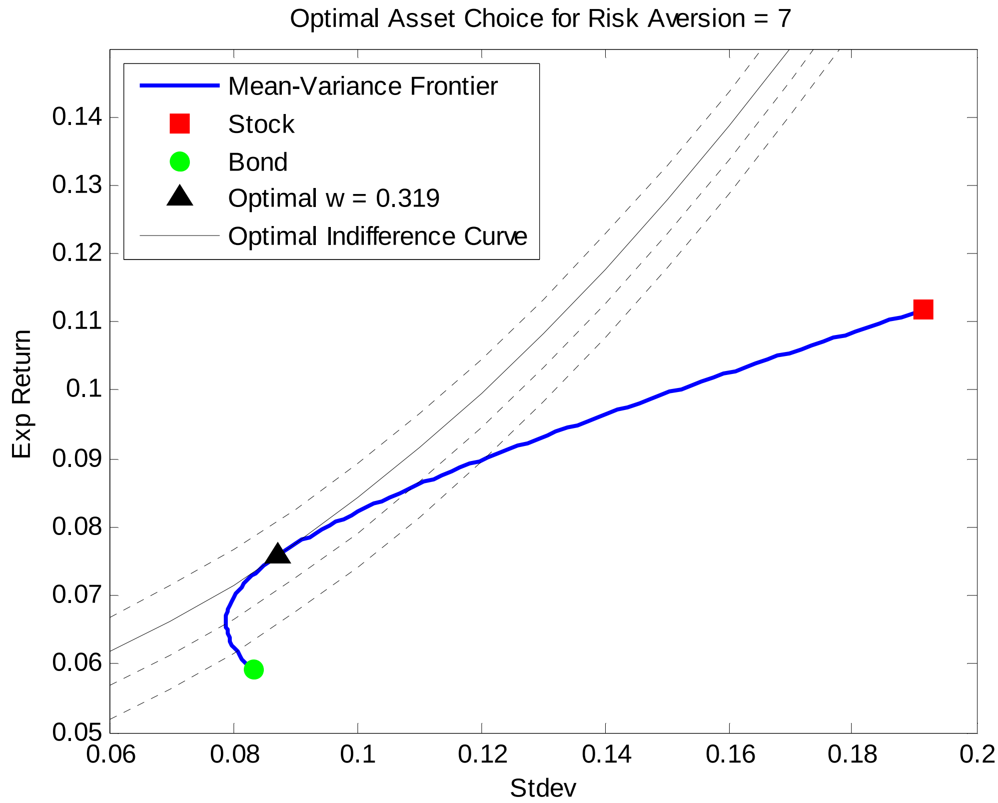


Figure 8

Panel D

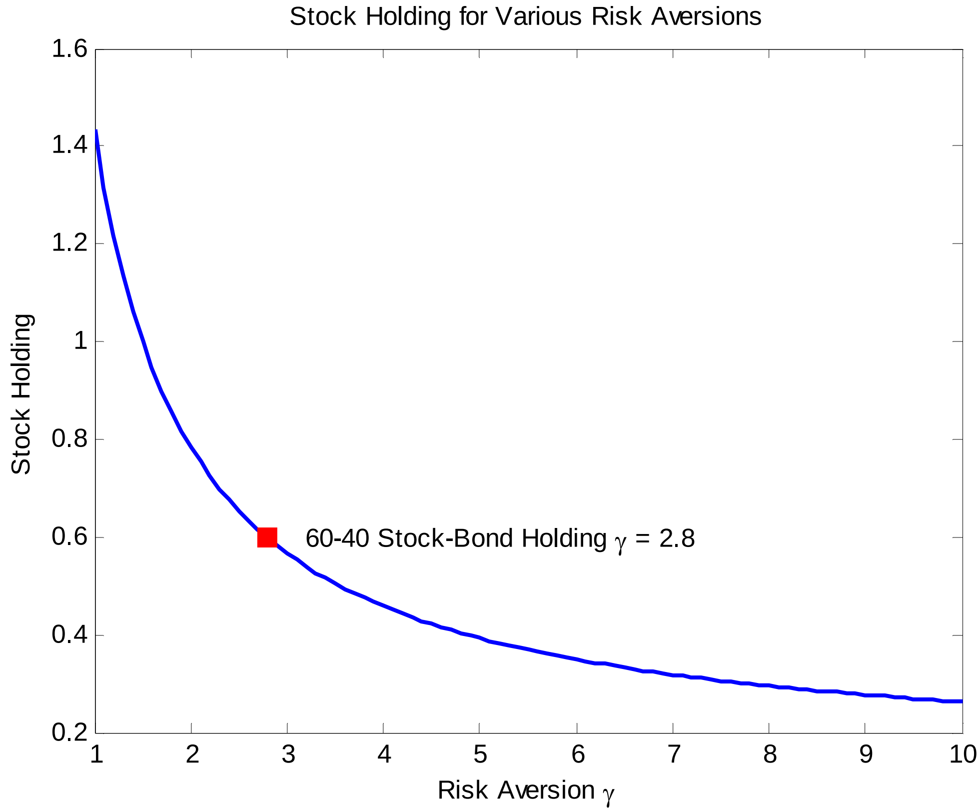


Figure 9

Panel A

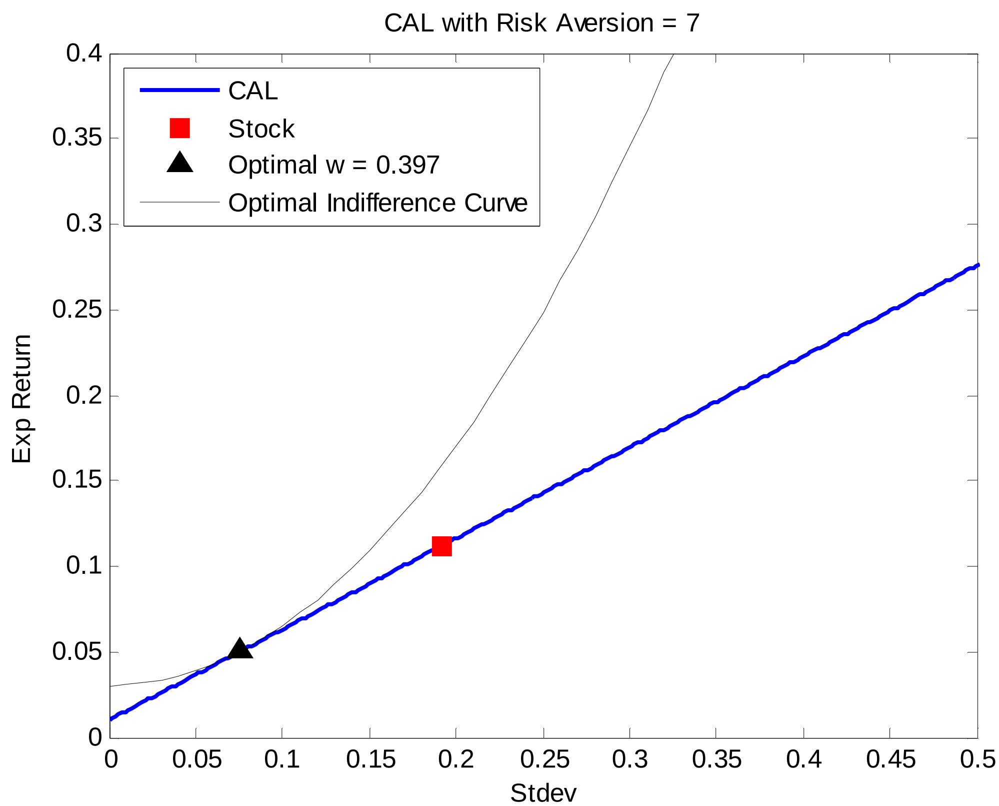


Figure 9

Panel B

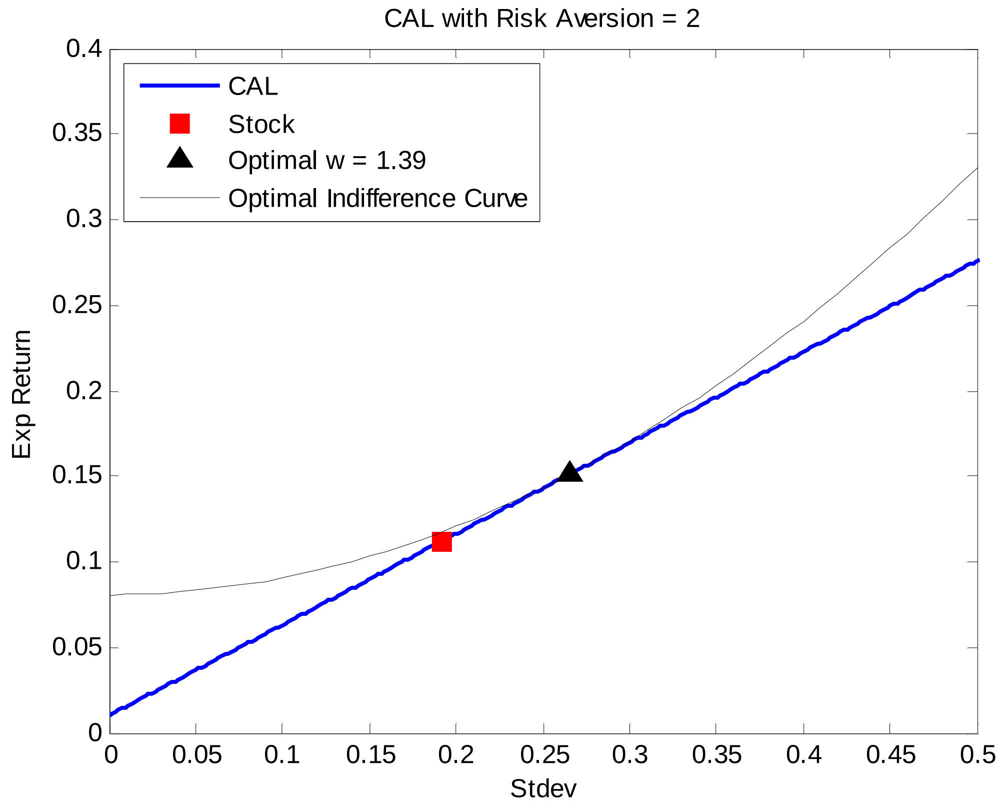


Figure 10

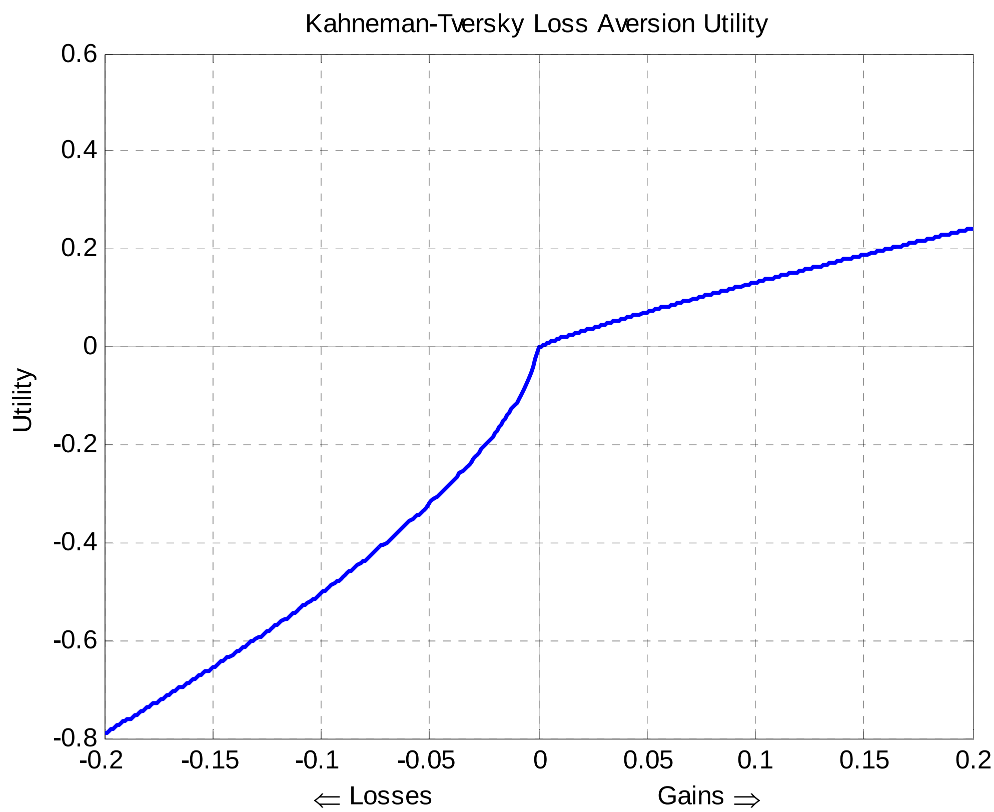

Kahneman-Tversky Loss Aversion Utility
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
Utility
-0.2
-0.15
-0.1
-0.05
0
0.05
0.1
0.15
0.2
⇐ Losses
Gains ⇒

Figure 11

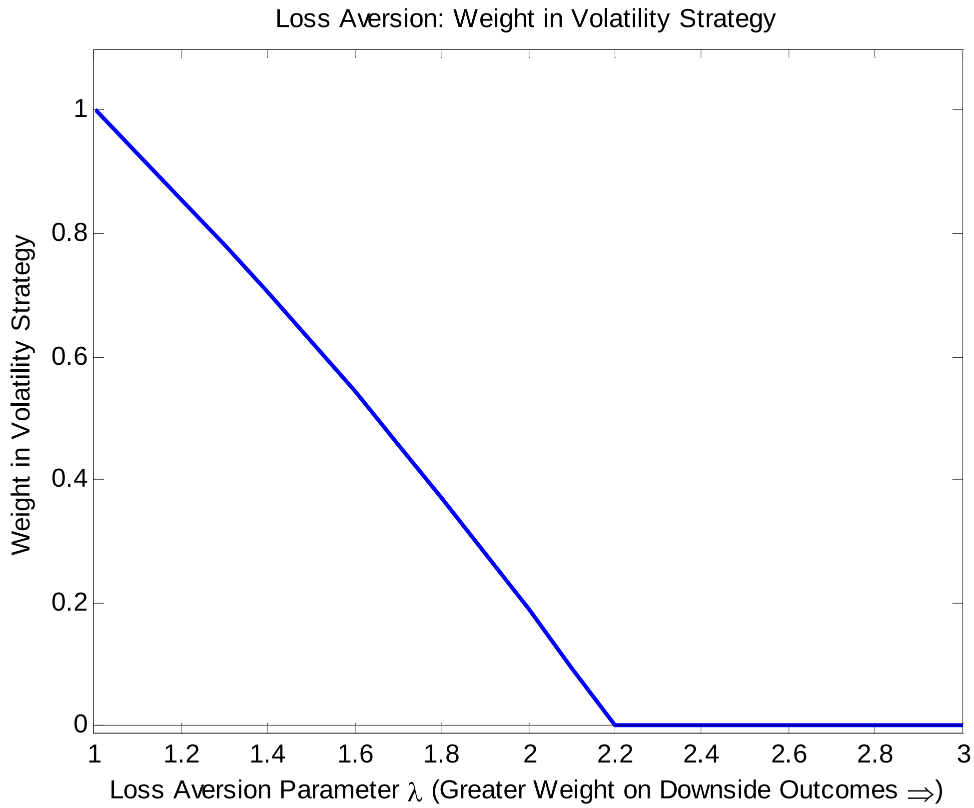
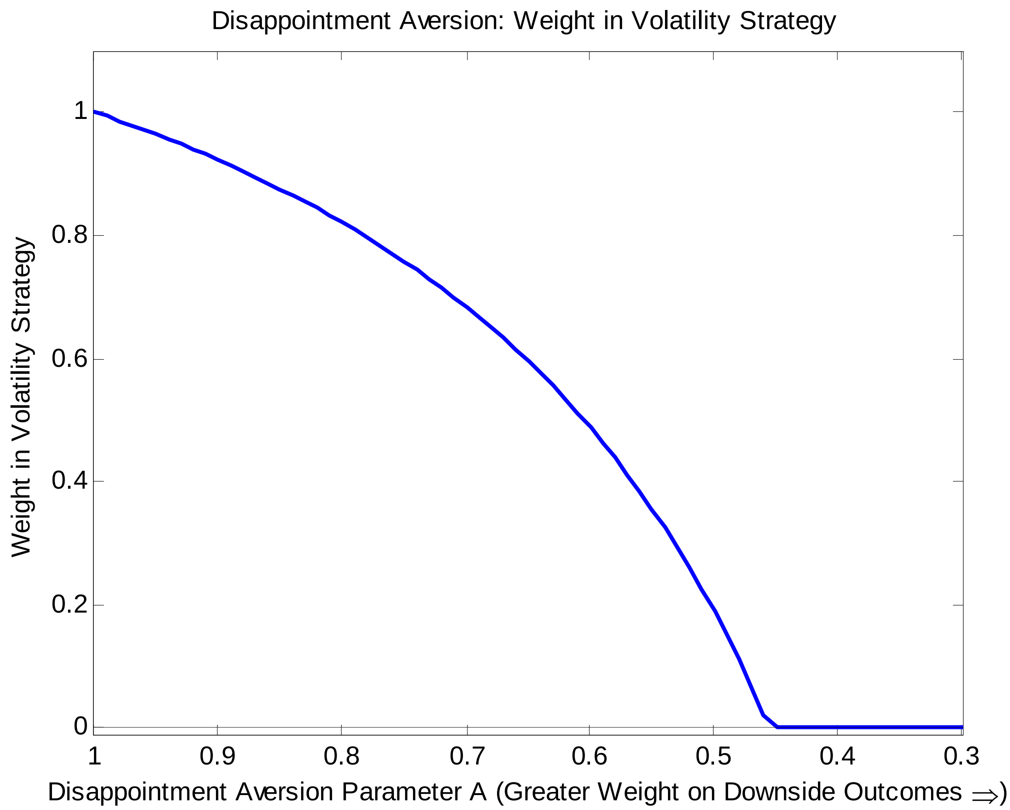

Loss Aversion: Weight in Volatility Strategy
Weight in Volatility Strategy
1
0.8
0.6
0.4
0.2
0
1
1.2
1.4
1.6
1.8
2
2.2
2.4
2.6
2.8
3
Loss Aversion Parameter λ (Greater Weight on Downside Outcomes ⇒)

Figure 12


Disappointment Aversion: Weight in Volatility Strategy
Weight in Volatility Strategy
1
0.8
0.6
0.4
0.2
0
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
Disappointment Aversion Parameter A (Greater Weight on Downside Outcomes ⇒)