A Decision-Making Support Tool for Generating Profitable, Practical, Capital Non-Intensive Investment Strategies

A thesis presented
by
Adam Arthurs

To
Applied Mathematics
in partial fulfillment of the honors requirement
for the degree of
Bachelor of Arts
Harvard College
Cambridge, Massachusetts

March 30, 2012

Abstract

We present a realistic capital non-intensive investment strategy that was found to consistently outperform the S&P 500 by up to 38% in average annual returns. This strategy is created by combining pre-Hilbert space theory, a well-established financial result, real-world investment restrictions, and quantitative decision-making methods of our own creation. Specifically, the strategy is created in four steps. First, using the formal mathematical framework of the appropriate pre-Hilbert space we prove in a way, that to the best of our knowledge is original, the well-known financial result that given a fixed set of assets, known expectations, variances, and pairwise covariances of the returns thereof, and a fixed overall return, there always exists a unique portfolio of these assets that is expected to achieve this overall return with minimum variance. Second, we create an investment decision-making support tool by combining the ability to create such fixed-return minimum-variance portfolios with a general investment procedure and an accurate method of calculating real-world returns that accounts for all fees and investment regulations. This tool, upon specifying values for a set of parameters, produces a comprehensively defined investment strategy. Subsequently, third, we devise several different parameter-choosing strategies, all of which both learn from the past to draw inferences about the future and learn from the past in order to determine exactly how much of the past to learn from. We combine the decision-making support tool with these parameter-choosing strategies to create several different investment strategies that all dynamically learn from historical data. Finally, we backtest these investment strategies to see how well they would have performed in past years using only data that would have been available at the time and then inspect the results to determine which of the investment strategies is best.

Acknowledgments

My inspiration for proving the fixed-return minimum-variance result using the pre-Hilbert space method presented here was a homework problem for Math 116 written by Paul Bamberg, the course's instructor. Math 116 covers the dual approximation theorem and Paul had the original idea of applying it to construct minimum-variance portfolios given fixed portfolio expectation and known individual asset expectations, variances, and covariances. Paul, thanks for the idea. I've done my best to turn it into something lucrative.

If it weren't for my adviser, Mauricio Santillana, I wouldn't have written a thesis at all. I submitted a significantly less well-developed and less thorough version of this paper as a final project for his class, Applied Math 120, and Mauricio convinced me to turn it into a thesis, offering to advise the project himself. Since then, Mauricio has been immensely helpful and generous with his time. He has suggested countless improvements, which I have done my best to implement, to both this paper and the investment decision-making support tool itself. Mauricio, thanks for convincing me that this would be worth pursuing. It certainly has been.

Contents

0 Introduction

Consider someone who has very little financial knowledge and somewhere between \$2000 and \$5000 that he is seeking to invest. Perhaps he is in his 20s and has a real job and thus disposable income for the first time in his life; perhaps he is in his 70s with plenty of cash saved away but has never invested before. He may seek modest returns that are just a bit higher than what he could earn by putting his money in the bank, aggressive returns that would substantially increase his disposable cash, or anything in between. Regardless, our potential investor is not content with doing nothing with his cash as inflation erodes its value.

The fact that the investor has very little financial knowledge classifies him as an "unsophisticated" investor. Specifically, he has little, if any, understanding of financial markets, is familiar with stocks and perhaps bonds but no other assets, and wouldn't know how to begin making an informed, strategic investment decision. Furthermore, the fact that he only has a few thousand dollars of initial capital to invest precludes him from hiring an investment manager or the like which generally require larger initial investments.

Help has arrived: in this paper we construct from scratch an investment decision-making support tool and use it to devise a profitable and capital non-intensive investment strategy that is practical for an unsophisticated investor to implement. This strategy would have returned, on average and including all fees and regulatory considerations, 64% in 2009, 26% in 2010, and 30% in 2011, beating the S&P 500,[1] a standard benchmark, by 38%, 11%, and 28%, respectively.

[1] The S&P 500 is an index that tracks the combined performance of 500 large-capitalization U.S. stocks. The capitalization of a stock is its share price multiplied by the number of shares outstanding and is one measure of a company's value.

0.1 Investment Decision-Making Tool, Parameter-Choosing Strategies, and Investment Strategies

The investment decision-making support tool consists of a general investment method that is only uniquely specified with the specification of a number of parameters. So, only upon combining the decision-making tool with a parameter-choosing strategy is an investment strategy defined. The utility of the distinction is that, whereas there is one and only one decision-making tool, various different parameter-choosing strategies can be devised. Combining the decision-making tool with different parameter-choosing strategies results in different investment strategies. In this way, the decision-making tool can be used to generate different investment strategies.

We only formally present a set of parameter-choosing strategies – the ones we devised that result in the best-performing investment strategies. We then narrow this optimal set down to a single best parameter-choosing strategy by selecting the one that results in the single best-performing investment strategy. This best-performing investment strategy is the main deliverable of this paper. We also briefly discuss a number of inferior parameter-choosing strategies that we devised en route to the above-mentioned optimal set of parameter-choosing strategies.

0.2 Financial Understanding

Given our lack of expertise in financial markets, we neither claim nor attempt to understand financial markets well enough to predict anything having to do with future asset prices. In fact, it can be argued that, because financial markets are systems driven by human behavior, it is impossible to truly and fundamentally understand them. Specifically, the price of an asset does not necessarily accurately reflect its value; rather, any asset price is driven by the possibly-irrational decisions of all the market actors that trade the asset. The asset price is simply the price at which market actors agree to buy and sell the asset. If more people buy the asset than sell it then the price goes up; if more people sell the asset than buy it than the price goes down. So, a true fundamental understanding of financial markets would require an accurate understanding of how people make investment decisions. Whether this is impossible or merely difficult, it is certainly well-beyond our expertise.

Playing to our strengths, we will devise and present an investment decision-making support tool and resultant investment strategy that don't rely on any financial or economic arguments. Instead, We use only mathematical and logical reasoning.

0.3 Backtesting

Regardless of what sort of reasoning leads to an investment strategy and regardless of how sound this reasoning is, the strategy still needs to be tested in order to determine whether or not it's successful and consequently whether or not the underlying reasoning is justified. We will test the investment strategy that we devise by observing how well it would have performed in past years, a technique known as backtesting.

It is important to note that backtesting does not cheat by effectively looking into the future and using now-known information about past years' asset returns to retroactively change the investment strategy and improve how well it would have performed. For example, backtesting results for 2009 are computed as follows. Even though we now know exactly what asset returns were in 2009, we hide this data from the investment strategy: we construct the portfolio that the investment strategy would have suggested using only information that would have been available to the investor on the first day of 2009, which is all data from 2008 or earlier. Only after this 2009-blind construction do we use 2009 data to compute how well this portfolio would have performed in 2009.

As described above, only upon confirming that our strategy would have performed well during the past several years are the decisions that went into constructing it justified. This is more than "guess and check" because we use math and logic to make very specific, educated decisions, but the "check" is still paramount: we can only say that the various elements of the strategy work because we can test the strategy and observe that it would have been sufficiently profitable in each of the past several years.

0.4 Learning from History

The fundamental assumption behind both the decision-making tool and the resultant investment strategy is that the past is indicative of the future and thus past data can be used to make predictions about the future. This assumption is logical in the sense that historical data is all we have to learn from so we might as well attempt to learn from it, but, as described in the previous section's discussion of backtesting, this logic alone is not enough to make the assumption necessarily true. Rather, we simply logically speculate that the past may be indicative of the future and then test this hypothesis by backtesting the investment strategy. Because it turns out that the investment strategy would have performed consistently well the past several years, the ways in which the strategy uses past data to draw inferences about the future are valid and we can conclude that, at least when done in specific ways, the past can be used to predict the future.

So, the investment strategy is dynamic in that it learns from the past to choose investments for the future. It suggests a different portfolio this week than it will next week because more information is available next week. Furthermore, not only does the investment strategy learn from the past, it also learns from the past in order to determine how much of the past to learn from. In summary, it is a dynamic tool that analyzes historical data to decide for the investor exactly how much he should invest in what.

0.5 Roadmap

In section 1, General Mathematical Theory, we present the formal mathematics that will be necessary for creating the investment decision-making support tool. A close reading of this entire section, though instructive, is not essential to understanding the rest of the paper; the reader can simply take the final result of the section, the dual approximation theorem (section 1.6), as proved and move on.

In section 2, Mathematical Theory of Minimum-Variance Portfolios, we define the appropriate pre-Hilbert space and apply the dual approximation theorem to prove that, given a fixed set of assets and known expectations, variances, and pairwise covariances of the returns thereof, for any fixed return there is a unique portfolio of these assets that is expected to achieve said return with minimum variance. We will use this result in section 4 as the basis of the investment decision-making support tool. It is important to note that this result is not new: it is a well-known concept of modern portfolio theory. Our proof of the result with the dual approximation theorem, on the other hand, is, to the best of our knowledge, novel.

In section 3, Building Towards a Practical Application, we work through various technical aspects that must be addressed before we can weaponize the minimum-variance result proved in section 2 into the decision-making support tool and subsequently a profitable, practical, and capital non-intensive investment strategy. As a result of some of these technical aspects being unrelated, parts of section 3 have no logical flow; the reader is warned that the section as a whole reads like a list.

In section 4, Investment Strategies, we devise and present the decision-making support tool using the fixed-return minimum-variance result from section 2 and various technical aspects from section 3. We also devise and present the optimal set of parameter-choosing strategies. The combinations of the decision-making support tool with these optimal parameter-choosing strategies result in an optimal set of investment strategies.

In section 5, Results, we present the results obtained by backtesting each investment strategy in the optimal set of investment strategies to see how well they would have performed in each of the last several years. We then inspect these results to select a single optimal investment strategy and we confirm that this is the profitable, practical, and capital non-intensive strategy for an unsophisticated investor we seek by verifying that it outperforms various benchmarks.

In section 6, Conclusions, we summarize the accomplishments of this project and discuss further applications of the investment decision-making support tool.

1 General Mathematical Theory

This section presents the formal mathematics that will be necessary for constructing the investment decision-making support tool. A close reading of this entire section, though instructive, is not essential to understanding the rest of the paper; the reader can simply take the final result of the section, the dual approximation theorem (section 1.6), as proved and move on.

David G. Luenberger's *Optimization by Vector Space Methods*[1] was used as a guide but all the proofs here are written in original language and some are of our own creation though are likely not new.

1.1 Norms

In a normed vector space V, the norm of any vector $x \in V$, denoted $||x||$, is a generalization of distance or length. For example, $\mathbb{R}^3$ has the well-known Euclidean norm $||(x_1, x_2, x_3)|| = ({x_1}^2 + {x_2}^2 + {x_3}^2)^{1/2}$. The following properties are necessary and sufficient for defining a norm:

1. $||x|| \geq 0 \ \forall x \in V$ and $||x|| = 0$ only if x is the zero vector,
2. $||\alpha x|| = |\alpha| \cdot ||x|| \ \ \forall \alpha \in \mathbb{R}, x \in V$,
3. $||x + y|| \leq ||x|| + ||y|| \ \ \forall x, y \in V$.

Note that the specific norm used in any given space is a property of the space itself. For example, $\mathbb{R}^3$ is the space of all 3-component vectors of real numbers with the Euclidean norm. One could also define another space that is the space of all 3-component vectors of real numbers with the norm $||(x_1, x_2, x_3)|| = ({x_1}^5 + {x_2}^5 + {x_3}^5)^{1/5}$, which is, indeed, a valid norm according to the properties above.

[1] [1] in the sources section.

1.2 Inner Products

Along with a norm, a vector space V may have an inner product defined on it. As with the norm, the specific inner product is an intrinsic property of the space itself. If there is an inner product defined, it takes any two vectors in V and returns a scalar: any inner product is a function from $V \times V$ to $\mathbb{R}$. The inner product of x and y is denoted $(x \,|\, y)$. The following properties are necessary and sufficient for defining an inner product on any vector space for which the underlying field is the reals[2]:

1. $(x \,|\, y) = (y \,|\, x) \; \forall x, y \in V$,
2. $(x + y \,|\, z) = (x \,|\, z) + (y \,|\, z) \; \forall x, y, z \in V$,
3. $(\alpha x \,|\, y) = \alpha (x \,|\, y) \; \forall \alpha \in \mathbb{R},\; x, y \in V$,
4. $(x \,|\, x) \geq 0 \; \forall x \in V$ and $(x \,|\, x) = 0$ only if x is the zero vector.

1.3 Pre-Hilbert Space

Theorem: Cauchy-Schwarz Inequality. For any vector space V with an inner product, $\forall x, y \in V$, it holds that

$$|(x \,|\, y)|^2 \leq (x \,|\, x) \cdot (y \,|y)\,. \tag{1.1}$$

Proof. If y is the zero vector, the inequality holds trivially as the equality $0 = 0$. Otherwise, note that the properties of an inner product imply that $0 \leq (x - \alpha y \,|\, x - \alpha y) = (x \,|\, x) - 2\alpha (x \,|\, y) + \alpha^2 (y \,|\, y) \;\; \forall \alpha \in \mathbb{R},\; x, y \in V$. Substituting $\alpha = (x \,|\, y)/(y \,|\, y)$ into the previous expression gives $0 \leq (x \,|\, x) - (x \,|\, y)^2/(y \,|\, y)$ which can be rearranged into (1.1), the desired result.

Theorem. For any valid inner product, the following function is a valid norm:

$$||x|| = \sqrt{(x \,|\, x)}\,. \tag{1.2}$$

Proof. The numbers below correspond to the norm properties from section 1.1.

1. Follows trivially from inner product property 4.
2. Follows from inner product property 3: $||\alpha x|| = \sqrt{(\alpha x \,|\, \alpha x)} = \sqrt{\alpha^2 (x \,|\, x)} = |\alpha| \cdot ||x||$.
3. Follows from (1.1) and inner product properties 1 and 2: $||x + y||^2 = (x + y \,|\, x + y) = (x \,|\, x) + 2(x \,|\, y) + (y \,|\, y) \leq ||x||^2 + 2||x||\,||y|| + ||y||^2 = (||x|| + ||y||)^2 \;\; \Rightarrow \;\; ||x + y|| \leq ||x|| + ||y||$.

[2] If the underlying field is the complex numbers as opposed to the reals, the first property changes to $(x \,|\, y) = \overline{(y \,|\, x)}$ and the other properties are unchanged.

By definition, a vector space is a pre-Hilbert[3] space if and only if it is normed, has an inner product, and the inner product invokes the norm as per (1.2).

1.4 Finite-Dimensional Pre-Hilbert Space Projection Theorem

Theorem. $\forall$ finite-dimensional real pre-Hilbert space H, subspace $M \subset H$, and vector $x \in H$, it holds that $\exists m_0 \in M$ such that $||x-m_0|| \leq ||x-m|| \; \forall m \in M$ and this m_0 is unique. In other words, in finite-dimensional pre-Hilbert space, for any subspace and vector x, there is a unique vector in the subspace that is closest to x.

Proof. $\forall y, z \in H$, if $(y \,|\, z) = 0$ then

$$||y+z||^2 = (y+z \,|\, y+z) = ||y||^2 + 2(y \,|\, z) + ||z||^2 = ||y||^2 + ||z||^2 \, . \tag{1.3}$$

Now, using the notation presented in the theorem statement above, suppose that $\exists m_0 \in M$ such that $(x - m_0 \,|\, m) \; \forall m \in M$. Then, $\forall m \in M$ it holds that

$$||x-m||^2 = ||x - m_0 + m_0 - m||^2 = ||x-m_0||^2 + ||m_0 - m||^2 \tag{1.4}$$

where the second equality is from (1.3) which can be invoked because $(m_0 - m) \in M$ as a result of M being a subspace and so the fact that $(x - m_0 \,|\, m) \; \forall m \in M$ implies that $(x - m_0 \,|\, m_0 - m) = 0$. Noting that $||m_0 - m|| > 0$ if $m \neq m_0$ and $||m_0 - m|| = 0$ if $m = m_0$, it holds that $||x-m|| > ||x-m_0|| \; \forall m \in M$ for which $m \neq m_0$, and thus m_0 is the unique vector in M that is closest to x. So, if $\exists m_0 \in M$ such that $(x - m_0)$ is orthogonal to every vector in M then this m_0 is the unique vector in M that is closest to x. Thus, if we have a fail-safe method for constructing such a m_0 then the finite-dimensional pre-Hilbert space projection theorem is proved.

H is finite-dimensional and thus any subspace $M \subset H$ must also be finite dimensional so M must have a finite number n of basis vectors $\{y_1, ..., y_n\}$. By definition of a basis, every vector in M can be expressed as a linear combination of these basis vectors, and thus $(x - m_0)$ is orthogonal to every vector in M if and only if it is orthogonal to each of $\{y_1, ..., y_n\}$. So, $\forall i \in \{1, ..., n\}$,

$$(x - m_0 \,|\, y_i) = 0 \quad \Rightarrow \quad (m_0 \,|\, y_i) = (x \,|\, y_i) \, . \tag{1.5}$$

[3] A Hilbert space has the additional requirement of being complete (every Cauchy sequence converges to an element of the space). As it turns out, the space S that we will be working with is not only pre-Hilbert but also Hilbert. However, we avoid proving that it is Hilbert because we can get all the necessary results with fewer proofs using only pre-Hilbert space.

But $m_0 \in M$ so it, too, can necessarily be expressed as a linear combination of the basis vectors: $\exists\{\alpha_1, ..., \alpha_n\} \subset \mathbb{R}$ such that

$$m_0 = \sum_{j=1}^{n} \alpha_j \, y_j \, . \tag{1.6}$$

Substituting (1.6) into the left side of (1.5) for any fixed $i \in \{1, ..., n\}$ gives

$$\sum_{j=1}^{n} \alpha_j (y_j \,|\, y_i) = (x \,|\, y_i) \, . \tag{1.7}$$

The set of equations (1.7) for each $i \in \{1, ..., n\}$ can be written as

$$\underbrace{\begin{bmatrix} (y_1 \,|\, y_1) & (y_1 \,|\, y_2) & \dots & (y_1 \,|\, y_n) \\ (y_1 \,|\, y_2) & (y_2 \,|\, y_2) & \dots & (y_2 \,|\, y_n) \\ \vdots & \vdots & & \vdots \\ (y_1 \,|\, y_n) & (y_2 \,|\, y_n) & \dots & (y_n \,|\, y_n) \end{bmatrix}}_{G} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} (x \,|\, y_1) \\ (x \,|\, y_2) \\ \vdots \\ (x \,|\, y_n) \end{bmatrix} \tag{1.8}$$

where the matrix on the left is known as the Gram matrix and is denoted G. For any basis $\{y_1, ..., y_n\}$ for M, it is possible to perform the Gram-Schmidt procedure on these basis vectors in order to create an orthonormal basis $\{e_1, ..., e_n\}$ for M such that, $\forall i, j \in \{1, ..., n\}$,

$$(e_i \,|\, e_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \, . \tag{1.9}$$

Substituting $\{e_1, ..., e_n\}$ into G instead of $\{y_1, ..., y_n\}$ gives the identity matrix I which is invertible. Because it is always possible to use Gram-Schmidt to choose an orthonormal basis for M, it is true that G is always invertible.

Because G is invertible then there is necessarily a unique set of $\{\alpha_1, ..., \alpha_n\}$ that solves (1.8), and via (1.6) we have a fail-safe method of constructing an $m_0 \in M$ such that $(x - m_0)$ is orthogonal to every vector in M. Thus, this m_0 is the unique vector in M that is closest to x and the finite-dimensional pre-Hilbert space projection theorem is proved.

1.5 Modified Finite-Dimensional Pre-Hilbert Space Projection Theorem

Theorem. $\forall$ finite-dimensional real pre-Hilbert space H, subspace $M \subset H$, and vector $x \in H$, construct the linear variety $N \subset H$ that is M translated by x:

$$N = M + \{x\} = \left\{ m + x \;\middle|\; m \in M \right\} . \tag{1.10}$$

It holds that there is a unique vector in N of minimum norm and that this vector is orthogonal to every vector in M.

Proof. According to the above-proved regular version of the finite-dimensional pre-Hilbert space projection theorem, there is a unique vector in M that is closest to $(-x)$: $\exists$ a unique $m_0 \in M$ such that, $\forall m \in M$,

$$||m_0 + x|| \leq ||m + x|| \,. \tag{1.11}$$

Define $n_0 = (m_0 + x) \in N$; then, $\forall n \in N$, $\exists m \in M$ such that

$$||m_0 + x|| = ||n_0|| \leq ||n|| = ||m + x|| \tag{1.12}$$

where the first equality is by definition of n_0, the inequality is via (1.11), and the last equality and the existence of m are via (1.10). So, via (1.12) and m_0 being unique, it holds that n_0 is the unique vector in N for which $||n_0|| \leq ||n||\ \forall n \in N$. Furthermore, recall that in the proof of the regular version of the theorem it was proved that $(x - m_0)$ is orthogonal to every vector in M. Here, we're using $(-x)$ instead of x so it holds that $(-x - m_0)$ is orthogonal to every vector in M. Switching the sign preserves orthogonality, so $n_0 = (x + m_0)$, the unique vector in N of minimum norm, is orthogonal to every vector in M and the modified version of the theorem is proved.

1.6 Dual Approximation Theorem

Theorem. Given a real pre-Hilbert space H of finite dimension, a set of k linearly independent vectors $\{y_1, ..., y_k\} \subset H$, and a set of k real numbers $\{\alpha_1, ..., \alpha_k\} \subset \mathbb{R}$, form the set K of vectors $x \in H$ that satisfy the k constraints $(y_1 \,|\, x) = \alpha_1, ..., (y_k \,|\, x) = \alpha_k$:

$$K = \left\{ x \in H \;\middle|\; (y_i \,|\, x) = \alpha_i \ \forall i \in \{1, ..., k\} \right\} . \tag{1.13}$$

It holds that

$$\arg\min_{x \in K} ||x|| = \sum_{i=1}^{k} \beta_i \, y_i \tag{1.14}$$

where

$$\begin{bmatrix} (y_1 \,|\, y_1) & (y_1 \,|\, y_2) & \dots & (y_1 \,|\, y_k) \\ (y_1 \,|\, y_2) & (y_2 \,|\, y_2) & \dots & (y_2 \,|\, y_k) \\ \vdots & \vdots & & \vdots \\ (y_1 \,|\, y_k) & (y_2 \,|\, y_k) & \dots & (y_k \,|\, y_k) \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix} . \tag{1.15}$$

In other words, the vector of minimum norm subject to a set of inner product constraints must be a linear combination of the constraint vectors. Furthermore, this solution must exist and is unique: the matrix on the left side of

(1.15) is necessarily invertible.

Proof. Let H be n-dimensional and let M be the k-dimensional subspace that has $\{y_1, ..., y_k\}$ as a basis. The orthogonal subspace to M, denoted $M^\perp$, is the $(n-k)$-dimensional subspace of all vectors in H that are orthogonal to M:

$$\begin{aligned} M^\perp &= \left\{ p \in H \;\middle|\; (m \,|\, p) = 0 \;\forall m \in M \right\} \\ &= \left\{ p \in H \;\middle|\; (y_i \,|\, p) = 0 \;\forall i \in \{1, ..., k\} \right\} . \end{aligned} \tag{1.16}$$

Comparing (1.13) to the right side of (1.16), it's clear that if $\alpha_1 = \alpha_2 = ... = \alpha_k = 0$ then $K = M^\perp$. With arbitrary $\{\alpha_1, ..., \alpha_k\}$, it holds that K is a translation of the subspace $M^\perp$: $\exists v \in H$ such that $K = M^\perp + \{v\} = \left\{ p + v \;\middle|\; p \in M^\perp \right\}$. As it turns out, any $v \in K$ does the job. This can be proved by showing that, $\forall$ fixed $v \in K$, it holds that $K \subseteq (M^\perp + \{v\})$ and $(M^\perp + \{v\}) \subseteq K$ and thus it must be the case that $K = (M^\perp + \{v\})$.

$\forall v, x \in K$, define $p = x - v$ such that $x = p + v$. Then, $\forall i \in \{1, ..., k\}$, it holds that $(y_i \,|\, p) = (y_i \,|\, x - v) = (y_i \,|\, x) - (y_i \,|\, v) = \alpha_i - \alpha_i = 0$ where the second to last equality is via (1.13), and thus $p \in M^\perp$ via (1.16). So, $x = p + v$ where $p \in M^\perp$ which means that $x \in (M^\perp + \{v\})$. Hence, $K \subseteq (M^\perp + \{v\})$.

$\forall v \in K$, $x \in (M^\perp + \{v\})$, it holds by definition of $(M^\perp + \{v\})$ that $\exists p \in M^\perp$ such that $x = p + v$. Then, $\forall i \in \{1, ..., k\}$, it holds that $(y_i \,|\, x) = (y_i \,|\, p + v) = (y_i \,|\, p) + (y_i \,|\, v) = 0 + \alpha_i = \alpha_i$ where the second to last equality is via (1.16) and (1.13), and thus $x \in K$ via (1.13). Hence, $(M^\perp + \{v\}) \subseteq K$.

So, K is the linear variety $K = M^\perp + \{v\}$ where v is any vector in K and M is the subspace that has $\{y_1, ..., y_k\}$ as a basis. We can now apply the modified finite-dimensional pre-Hilbert space projection theorem to K: there exists a unique vector in K of minimum norm and this vector is orthogonal to every vector in $M^\perp$. By definition, the space of vectors orthogonal to $M^\perp$ is M. Thus, the unique vector in K of minimum norm, call it x_0, must be in M and so must be a linear combination of $\{y_1, ..., y_k\}$, which proves (1.14). Furthermore, $x_0 \in K$ so x_0 must satisfy the constraints in (1.13). Substituting (1.14) into these constraints gives, $\forall i \in \{1, ..., k\}$,

$$(y_i \,|\, x) = \left(y_i \;\middle|\; \sum_{j=1}^{k} \beta_j \, y_j \right) = \sum_{j=1}^{k} (y_i \,|\, y_j) \beta_j = \alpha_i \tag{1.17}$$

which is equivalent to (1.15). The dual approximation theorem is now proved.

1.7 Geometric Intuition of the Dual Approximation Theorem

Using the notation established in the proof of the dual approximation theorem, let $H = \mathbb{R}^2$ where the inner product is the well-known dot product such that orthogonal vectors are at right angles to each other. As per the proof, for any fixed $v \in K$ it holds that $K = M^\perp + \{v\}$. To establish geometric intuition, choose $v \in M \cap K$ such that $M^\perp$, $K = M^\perp + \{v\}$, and v are as shown in Figure 1.1. In this example, M is one-dimensional and thus M must be the subspace that has $\{v\}$ as a basis. Accordingly, there can only be one constraint that defines K and it must be of the form $(v \mid x) = \alpha$. We seek the x that satisfies this constraint (i.e. is in K) of minimum norm. From the diagram, it's clear that the vector in K that is closest to the origin is v. So, in this example, $x_0 = v$ is, indeed, a linear combination of the constraint vectors $\{v\}$.

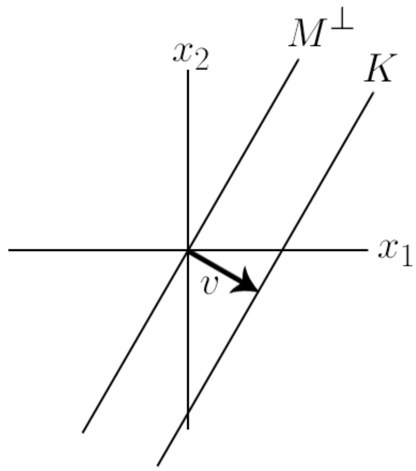


Figure 1.1: *Illustration of the dual approximation theorem with one constraint embedded in two dimensions.*

This geometric intuition generalizes to arbitrary dimensions. $K = M^\perp + \{v\}$ where we choose, without loss of generality, $v \in M \cap K$. M is defined as the subspace that has the constraint vectors as a basis and thus v must be a linear combination of the constraint vectors. Furthermore, it's clear geometrically that, always, $x_0 = v$, and thus the vector in K of minimum norm is a linear combination of the constraint vectors.

2 Mathematical Theory of Minimum-Variance Portfolios

This section works through the details of how the appropriate pre-Hilbert space and the dual approximation theorem can be used to show that, given fixed $r \in \mathbb{R}$, a fixed set of assets, and known expectations, variances, and pairwise covariances of the returns thereof, there always exists a unique portfolio of these assets that is expected to return r with minimum variance. This result is important to us because we will ultimately use it as the basis for the investment decision-making support tool.

It must be noted that the existence and uniqueness of the minimum-variance isn't original: it is known as the efficient portfolio frontier[1] in modern portfolio theory. What is original, to the best of our knowledge, is our proof of the result using the dual approximation theorem.

2.1 Portfolio Space

Suppose it is 12:01am on the first business day of time period 1 and there is a set of N assets that the investor has somehow chosen to invest in. The investor will enter into a specific portfolio of these assets when the markets open at 9:30am, hold this portfolio all time period, and then cash out of the portfolio when markets close on the last business day of the time period. The investor seeks the portfolio that is expected to achieve a return of r over this time period with minimum variance.

Let X_y^i be the random variable that is the rate of return of the i-th asset

[1] See "An Analytic Derivation of the Efficient Portfolio Frontier" by Robert Merton, [2] in the sources section.

during time period m:

$$X_m^i = \frac{\text{closing price of asset } i \text{ on the last business day of time period } m}{\text{opening price of asset } i \text{ on the first business day of time period } m}\,. \tag{2.1}$$

Given the date it holds that X_m^i is known $\forall m \in \{0, -1, -2, ...\}$ and unknown $\forall m \in \{1, 2, 3...\}$. However, assume that $\mathrm{E}(X_1^i)$ and $\mathrm{cov}(X_1^i, X_1^j)$ are known $\forall i, j \in \{1, ...n\}$. As will be discussed in section 3.2, this is a reasonable assumption because we can compute these values from X_m^i $\forall m \in \{0, -1, -2, ...\}$ under the premise that the past may be indicative of the future.

Represent the investor's portfolio as the N-dimensional vector

$$x = [x_1\ x_2\ \ldots\ x_N]^T \tag{2.2}$$

where x_i is the fraction of the total dollar amount invested that is invested in the i-th asset.[2] Accordingly, it must hold that

$$\sum_{i=1}^{N} x_i = 1\,. \tag{2.3}$$

The random variable Y_m that is the rate of return of the entire portfolio during time period m can be written

$$Y_m = \sum_{i=1}^{N} x_i\, X_m^i\,. \tag{2.4}$$

We are concerned with Y_1, the return over the upcoming time period. Let S be the space of all N-dimensional portfolio column vectors such that $x \in S$. Let C be the N by N covariance matrix of asset returns: the element in the i-th row and j-th column of C is

$$C_{i,j} = \mathrm{cov}(X_1^i, X_1^j)\,. \tag{2.5}$$

Via (2.4) and the formula for the variance of a sum of random variables:

$$\mathrm{var}(Y_1) = \mathrm{var}\left(\sum_{i=1}^{N} x_i\, X_1^i\right) = \sum_{i=1}^{N}\sum_{j=1}^{N} x_i\, x_j\, \mathrm{cov}(X_1^i, X_1^j) = x^T C x\,. \tag{2.6}$$

Theorem. The following function is a valid inner product over S:

$$(y\,|\,z) = y^T C z \quad \forall y, z \in S\,. \tag{2.7}$$

[2]Note that each x_i can be negative (as well as positive, of course) because the investor can short any asset. Shorting an asset is the exact opposite of buying it. If you buy an asset: if its price increases by \$1 then you make \$1 and if its price decreases by \$1 then you lose \$1. If you short an asset: if its price increases by \$1 then you lose \$1 and if its price decreases by \$1 then you make \$1. The mechanics of shorting an asset are as follows: the investor borrows the asset from a lender, immediately sells it to the market, and then buys it back from the market at a later time and returns in to the lender.

Proof.

- $\text{cov}(X_1^i, X_1^j) = \text{cov}(X_1^j, X_1^i)$ and thus, via (2.5), C is symmetric, so $[(z\,|\,y)]^T = (z^T Cy)^T = y^T Cz = (y\,|\,z)$. Any 1 by 1 matrix (a scalar) is symmetric so $[(z\,|\,y)]^T = (y\,|\,z)$ and thus $(y\,|\,z) = (z\,|\,y)\ \forall y, z \in S$.
- $(x + y\,|\,z) = (x + y)^T Cz = (x^T + y^T)Cz = x^T Cz + y^T Cz = (x\,|\,z) + (y\,|\,z)\ \ \forall x, y, z \in S$.
- $(\alpha x\,|\,y) = (\alpha x)^T Cy = \alpha x^T Cy = \alpha(x\,|\,y)\ \ \forall \alpha \in \mathbb{R},\ x, y \in S$.
- $(x\,|\,x) = x^T Cx = \text{var}(Y_1)$ via (2.6) and variances are nonnegative so $(x\,|\,x) \geq 0\ \forall x \in S$. Furthermore, $\text{var}(Y_1) = 0$ if and only if Y_1 is a constant. It is assumed that there is no nontrivial linear combination of the $\{X_1^i\}$ that is a constant[3] and thus $\text{var}(Y_1) = 0$ if and only if $x_1 = x_2 = ... = x_n = 0$, which is equivalent to x being the zero vector. Thus, $(x\,|\,x) = 0$ if and only if x is the zero vector.

Thus S is a pre-Hilbert space with the norm

$$||x|| = \sqrt{(x\,|\,x)} = \sqrt{x^T Cx} = \sqrt{\text{var}(Y_1)}\,. \tag{2.8}$$

2.2 Constraints

There are two constraints on x. The first is (2.3). The second is that the investor wants a return of exactly r on average:

$$\text{E}(Y_1) = \text{E}\left(\sum_{i=1}^{N} x_i\, X_1^i\right) = \sum_{i=1}^{N} \text{E}(X_1^i)\, x_i = r \tag{2.9}$$

where the first equality is via (2.4) and the second equality is via the linearity of expectation. In order to facilitate an imminent application of the dual approximation theorem, these constraints are rewritten as inner product constraints in pre-Hilbert space S.

For (2.3), we seek $u \in S$ such that

$$(u\,|\,x) = u^T Cx = \sum_{i=1}^{N} x_i \tag{2.10}$$

at which point the constraint could be written as $(u\,|\,x) = 1$. Examining the second equality in (2.10), it's clear that we require $u^T C = [1\ 1\ \ \ldots\ \ 1]$ and thus

[3] If none of the N assets is a financial derivative of one of the others then this "asset independence" is a very safe assumption: there are, in general, no perfect linear relationships between returns of distinct assets so it is impossible to linearly combine distinct assets to create a risk-free return. Furthermore, no asset is risk-free (none of the X_1^i is a constant) so a single-asset portfolio can never be risk-free.

$(u^T C)^T = C^T u = Cu = [1\ 1\ \ldots\ 1]^T$ which gives $u = C^{-1}[1\ 1\ \ldots\ 1]^T$. So, the first constraint can be written in inner product form as

$$(u \,|\, x) = \alpha_1\,,\ \ u = C^{-1}[1\ 1\ \ldots\ 1]^T\,,\ \ \alpha_1 = 1\,. \tag{2.11}$$

For (2.9), we seek $v \in S$ such that

$$(v \,|\, x) = v^T C x = \sum_{i=1}^{N} \mathrm{E}(X_1^i)\, x_i \tag{2.12}$$

at which point the constraint could be written as $(v \,|\, x) = r$. Following the same steps as above, we require $v^T C = [E(X_1^1)\ E(X_1^2)\ \ldots\ E(X_1^N)]$ and thus $(v^T C)^T = C^T v = Cv = [E(X_1^1)\ E(X_1^2)\ \ldots\ E(X_1^N)]^T$ which gives $v = C^{-1}[E(X_1^1)\ E(X_1^2)\ \ldots\ E(X_1^N)]^T$. So, the second constraint can be written in inner product form as

$$(v \,|\, x) = \alpha_2\,,\ \ v = C^{-1}[E(X_1^1)\ E(X_1^2)\ \ldots\ E(X_1^N)]^T\,,\ \ \alpha_2 = r\,. \tag{2.13}$$

Because S is pre-Hilbert there does not exist a nonzero vector $x \in S$ such that $Cx = 0$ (else $(x \,|\, x) = x^T C x = 0$ with $x \neq 0$) and thus C^{-1} necessarily exists.

Note that it's technically possible for u and v to be multiples of each other, which would make the constraints (2.11) and (2.13) either redundant or contradictory. However, this only occurs if $[1\ 1\ \ldots\ 1]$ and $[E(X_1^1)\ E(X_1^2)\ \ldots\ E(X_1^N)]$ are multiples of each other which, because different assets have different average returns, is extremely unlikely.

2.3 Minimum-Variance Portfolio

We seek the portfolio $x \in S$ that minimizes $\mathrm{var}(Y_1) = ||x||^2$ while satisfying the constraints (2.11) and (2.13). The square root function is monotonically increasing so minimizing $||x||$ is equivalent. Thus, the investor seeks the $x \in S$ of minimum norm subject to two inner product constraints.

This is a straightforward application of the dual approximation theorem. Let y be the minimum-variance r-returning portfolio that we seek. Via (1.14), a simple rearrangement of (1.15), and the definition of the portfolio space inner product (2.7), it holds that

$$y = \beta_1 u + \beta_2 v\,,\ \ \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} u^T C u & u^T C v \\ u^T C v & v^T C v \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ r \end{bmatrix} \tag{2.14}$$

where C is defined by (2.5) and constructed from the known expectations and covariances of the N assets, u and v are defined in (2.11) and (2.13) respectively, and r is the average rate of return we demand. Then, via (2.8), the minimum variance is

$$\min \mathrm{var}(Y_1) = ||y||^2 = y^T C y\,. \tag{2.15}$$

2.4 Lagrange Multipliers

This solution could have been determined solely with Lagrange multipliers without any mention of norms, inner products, pre-Hilbert space, etc. Accordingly, the pre-Hilbert space method is necessary not because it provides a method for finding the solution if one exists, but rather because it proves that one always exists. Lagrange multipliers are not up to this task because they provide a necessary but not sufficient condition for an extremum.

3 Building Towards a Practical Application

In this section we work through various technical topics that must be addressed before we can weaponize the minimum-variance result proved in section 2 into a decision-making support tool and ultimately a profitable, practical, and capital non-intensive investment strategy. As a result of some of these technical topics being unrelated, parts of this section has no logical flow; rather, it reads like a list.

3.1 Variance as Risk

Given a set of n assets, known values of $E(X_1^i)$ and $\text{cov}(X_1^i, X_1^j)\ \forall i, j \in \{1, ..., N\}$, and a requested return r, we now know how to construct the unique portfolio of these N assets that is expected to return r during time period 1 with minimum variance. But what is variance? $\text{var}(Y_1)$ measures how far the random variable Y_1 is likely to be from the mean of Y_1, which is r. Accordingly, variance is a good measure of the financial risk in the time period 1 return: a low variance indicates that Y_1 is likely to be near its mean, which we know to be r, whereas a higher variance indicates more uncertainty in that Y_1 is more likely to be far from its mean. In the extreme, a variance of zero means no risk: we know ahead of time that the return during time period 1 is going to be r. So, from now on, the words "variance" and "risk" will be used interchangeably.

3.2 Calculating Expectations and Covariances From Historical Data

It would be a mistake to assume that past asset returns are necessarily indicative of future returns; for two reasons it is nonetheless reasonable to use historical data to calculate each $\mathrm{E}(X_1^i)$ and $\mathrm{cov}(X_1^i, X_1^j)$. First, in the absence of a strong understanding of financial markets, historical data is all we have to go by. Second, we can simply speculate that the past may be indicative of the future and then test this hypothesis by checking how well the resultant investment strategies perform (see section 5.1).

Specifically, we compute asset expectations and covariances[1] using the previous k time periods of data:

$$\mathrm{E}\left(X_1^i\right) = \mathrm{mean}\left(\left\{X_{-k+1}^i, ..., X_0^i\right\}\right) , \quad (3.1)$$

$$\mathrm{cov}\left(X_1^i, X_1^j\right) = \mathrm{cov}\left(\left\{X_{-k+1}^i, ..., X_0^i\right\}, \left\{X_{-k+1}^j, ..., X_0^j\right\}\right) . \quad (3.2)$$

Note that (3.1) and (3.2) constitute one way in which any resultant investment strategy learns from the past to draw inferences about the future. Specifically, we use past data to compute metrics (expectations and covariances) that give us insight into both the future values of the $\{X_1^i\}$ and the pairwise relationships thereof.

3.3 Investment Restrictions

We assume that our investor has roughly \$5000 that he is able to invest,[2] access only to an online stock brokerage,[3] and has only very limited financial knowledge.

Online brokerages have transaction costs of roughly \$10 per trade for most assets,[4] meaning that entering into a portfolio of n assets costs \$10n and cashing out of that portfolio costs another \$10n, resulting in total transaction costs of \$20n. For our tastes, transaction costs should be no more than a percent or two of the initial cash invested per year: we don't want transaction costs to detract any more than 2% per year of any return. Noting that $20 \cdot 5/5000 = 2\%$, we see that one way to keep transaction costs less than or equal to 2%/year is to set $n = 5$ and trade only once per year: the investor sinks \$50 in transaction costs to enter into a portfolio of $n = 5$ assets, holds this portfolio for exactly one year without doing any other trades, and then sinks another \$50 in transaction costs to cash out of the portfolio. We fix $n = 5$ and annual trading for the rest of the paper. So, the investor's true return is the nominal return he earns on his assets during the year minus transaction fees of 2%.

Of course, $n = 5$ with annual trading is not the only strategy that results in transaction costs less than or equal to 2%/year. The investor could let

[1] And therefore also variances: $\mathrm{var}(X_1^i) = \mathrm{cov}(X_1^i, X_1^i)$.

[2] The strategies presented here will work given as little as \$2000 initial capital, but, as will be described, lower initial capital implies lower returns due to fixed transaction costs.

[3] E-Trade (www.etrade.com), for example.

[4] See sources [3], [4], and [5].

$n = 4$ and still trade only once per year resulting in lower transaction costs, or he could let $n = 2$ and trade twice per year, etc. Furthermore, the restriction that transaction costs cannot exceed 2%/year is not well-justified; as aforementioned, it's simply personal preference. Accordingly, the restriction of the portfolio to $n = 5$ assets traded annually is rather arbitrary and we have no reason to believe that it is optimal. However, as will be shown in the results section, this restriction, along with other elements of the investment decision-making support tool and the appropriate parameter-choosing strategy, gives profitable results: it is a configuration that works and that is justification enough.

Another result of the $5000 investment limit is that fixed income assets are off limits: almost all fixed income assets (bonds & CDs) that are available to trade through online brokerages have a $5000 investment minimum[5] and we need to invest in more than one asset ($n = 5$). Furthermore, because the investor has very limited financial knowledge, we will assume that the only assets he understands, or at least feels comfortable investing in, are stocks and bonds. Because bonds are off limits we are left with only stocks. We proceed given the restrictions that the investor must invest in $n = 5$ stocks, hold this portfolio for one year, and then cash out.

3.4 Time Scale

As just described, the investor enters into a portfolio, holds it for a year, and then cashes out. Accordingly, we will express the return of any portfolio as percentage per year: it doesn't make sense to use a shorter time scale (i.e. per month) because the investor cannot invest for less than a year and it doesn't make sense to use a longer time scale because the investor doesn't have to invest for any longer than a year. Furthermore this facilitates easy comparison between the strategies presented here and various benchmarks because returns are almost always expressed yearly.

However, we define the length of one time period to be one month: X_m^i is the rate of return of the i-th asset during month m. We do so for two reasons.

- As will be described in the subsequent sections, returns of portfolios are path dependent. Suppose, as we will be doing here, that you enter into a portfolio today, hold it for a year, and then cash out. Path dependence means that you cannot determine the return of the portfolio simply from asset prices today and asset prices in one year. The path the asset prices follow over the next year matter. To be completely accurate we would need daily asset prices, but, as will become evident in the subsequent sections, this would be too computationally intensive. Using monthly asset prices is a good balance between accuracy and computational efficiency.

[5]See sources [3], [4], and [5].

- Recall from section 3.2 that we calculate each $\mathrm{E}(X_1^i)$ and $\mathrm{cov}(X_1^i, X_1^j)$ using the last k time periods of historical data. If we set each time period to be one year then we would be constrained to using the last integer number of years of historical data. However, what if it is optimal to use some non-integer number of years of historical data? To account for this, we require finer than yearly resolution. Again, monthly resolution is a good compromise between computational feasibility and accuracy: we will calculate each $\mathrm{E}(X_1^i)$ and $\mathrm{cov}(X_1^i, X_1^j)$, the expectations and pairwise covariances of asset returns over the next month, using the last k months of data as per (3.1) and (3.2).

So, the investment decision-making support tool that we present in section 4 will produce investment strategies that involve calculating the minimum-variance portfolio for the next month (month 1), immediately entering into this portfolio, and then holding it for an entire year (months 1 through 12). Given that we restrict trading to once per year, this is justifiable: we have no knowledge of the future so the minimum-variance portfolio for the next month must be identical to the minimum-variance portfolio for the next year.

For the purpose of constructing the minimum-variance portfolio and actually entering into it, it doesn't matter what year it is; we use the last k months of historical data to construct the minimum-variance r-returning portfolio for the upcoming month (month 1) and then hold this portfolio for 12 months independent of what year it happens to be. However, in the upcoming sections we compare how different investment strategies would have performed in different calendar years so we need notation to distinguish between years. Accordingly, let the year be t: we suppose that it is currently 12:01am on the first business day of month 1 of year t. For our purposes, t always denotes the calendar year such that the first business day of month 1, which is when we enter into a portfolio, is always the first business day of January. However, the strategies presented here could be used to invest from, say, April 22nd 2012 to April 21st 2013 simply by redefining a year to start on April 22nd. So, we let t denote the calendar year without loss of generality.

3.5 Long and Short Positions

The following function is well-defined:

$$y_t(G, r, k) = [x_1 \ x_2 \ \dots \ x_N]^T \tag{3.3}$$

which returns the unique minimum-variance portfolio (as per (2.14) along with (2.5), (2.11), and (2.13)) that:

- uses the set of N stocks G,
- is for the upcoming month 1 of year t,

- is expected to return r during the time period it is intended for, which is the upcoming month 1 of year t,
- uses only the last k months of historical data to calculate the $\mathrm{E}(X_1^i)$ and $\mathrm{cov}(X_1^i, X_1^j)$ (as per (3.1) and (3.2)).

For example, $y_{t=2011}(G, r, k = 3)$ returns the minimum-variance r-returning portfolio for January 2011 using only data from October, November, and December of 2010 and using the N stocks in set G.

Given a portfolio $y = y_t(G, r, k)$, define P to be the number of components of y that are positive; formally:

$$P = \left|\left\{i \in \{1, ..., N\} \,\middle|\, x_i > 0\right\}\right| \tag{3.4}$$

where the absolute value of a set is the number of elements in it. Note that $(N - P)$ is the number of non-positive components of y. Define α_i to be the index of the i-th positive component of y and β_i to be the index of the i-th non-positive component of y; formally:

$$\left|\left\{j \in \{1, ..., \alpha_i\} \,\middle|\, x_j > 0\right\}\right| = i \;\; \forall i \in \{1, ..., P\}\,, \tag{3.5}$$

$$\left|\left\{j \in \{1, ..., \beta_i\} \,\middle|\, x_j \leq 0\right\}\right| = i \;\; \forall i \in \{1, ..., N - P\}\,. \tag{3.6}$$

So, $\{x_{\alpha_1}, ..., x_{\alpha_P}\}$ are the dollar proportional of the investor's long positions in assets $\alpha_1, ..., \alpha_p$ and $\{x_{\beta_1}, ..., x_{\beta_{N-P}}\}$ are the dollar proportional of the investor's short positions in assets $\beta_1, ..., \beta_{N-P}$. Define ℓ_m to be the total value at the end of month m of all assets bought long and s_m to be the total value at the end of month m of all assets sold short, both per overall dollar invested. The end of month 0 is the beginning of month 1 so ℓ_0 (or s_0) is, simply, the sum of the investor's initial long (or short) positions:

$$\ell_0 = \sum_{i=1}^{P} x_{\alpha_i}\,, \tag{3.7}$$

$$s_0 = \sum_{i=1}^{N-P} x_{\beta_i}\,. \tag{3.8}$$

Note that (2.3) along with (3.7) and (3.8) imply that

$$\ell_0 = s_0 + 1\,. \tag{3.9}$$

The value at the end of month m is equal to the value at the end of month $(m - 1)$ plus the change that occurred during month m, so, $\forall m \in \{1, ..., 12\}$:

$$\ell_m = \ell_{m-1} + \sum_{i=1}^{P} x_{\alpha_i} X_m^{\alpha_i} = \ell_0 + \sum_{n=1}^{m}\sum_{i=1}^{P} x_{\alpha_i} X_n^{\alpha_i}\,, \tag{3.10}$$

$$s_m = s_{m-1} + \sum_{i=1}^{N-P} x_{\beta_i} X_m^{\beta_i} = s_0 + \sum_{n=1}^{m}\sum_{i=1}^{N-P} x_{\beta_i} X_n^{\beta_i} \tag{3.11}$$

where ℓ_0 and s_0 are defined by (3.7) and (3.8), respectively.

3.6 Cash, Debt, and Equity

Let c_m be the non-negative amount of usable cash the investor has in his account at the end of month m and let d_m be the non-negative amount of debt to the online brokerage that the investor has accumulated by the end of month m. Define the investor's equity at the end of month m as

$$q_m = \ell_m + c_m - d_m \tag{3.12}$$

which is the dollar value per overall dollar invested of the assets in the portfolio that the investor actually owns. The intuition is as follows: if the investor has no cash, a long position of ℓ_m, but he funded this long position by borrowing from the brokerage such that he now owes d_m, then he only actually owns $(\ell_m - d_m)$ of the long position. In other words, if he were to pay back his loans at the end of month m then he would have exactly q_m of value remaining. Define the investor's percentage equity at the end of month m as

$$p_m = \frac{q_m}{\ell_m + c_m} = \frac{\ell_m + c_m - d_m}{\ell_m + c_m} \tag{3.13}$$

which can be interpreted as the percentage of his entire portfolio that the investor actually owns. It will become clear in the next section why we need to define these quantities. Rearranging (3.13) gives

$$d_m = (1 - p_m)(\ell_m + c_m)\,. \tag{3.14}$$

Note that both (3.12) and (3.13) seem to ignore the investor's short position because s_m appears in neither. This is not the case: as will be described in section 3.7, the "short covering requirement" requires the investor to immediately realize any gains or losses of the short position in cash, and thus the value of the short position is encapsulated in c_m and d_m.

3.7 Regulations and Requirements

Whereas in section 3.3 we discussed soft, somewhat arbitrarily imposed restrictions that result from our investor's particular situation ($5000 to invest, access only to an online brokerage, very limited financial knowledge), in this section we discuss regulations and requirements that apply to all investors who choose to use an online brokerage. Some of these regulations and requirements are laws imposed by the Financial Industry Regulatory Authority (FINRA); others are rules that all online brokerages impose on their customers; all are firm and carry consequences if broken.[6]

[6]We gleamed the information in this section from countless phone calls with E-Trade customer support representatives as well as sources [6] and [7].

3.7.1 Margin Requirements for Borrowing

The investor can borrow cash from the brokerage either just to have on hand or for the purpose of buying stocks subject to the following restrictions. In order to take out a new loan at the end of month m, thus increasing d_{m-1} by $(d_m - d_{m-1})$ to d_m, it must be the case that p_m would be greater than or equal to 50% after taking out the new loan. So, via (3.14) we're allowed to increase d_{m-1} to $d_m > d_{m-1}$ by taking out a new loan at the end of month m if and only if

$$d_m \leq 0.5(\ell_m + c_m)\,. \tag{3.15}$$

This is the 50% "initial margin requirement" for borrowing. Furthermore, at all times (not just in order to borrow more cash), it must hold that p_m is greater than or equal to 30%. So, via (3.14), we require

$$d_m \leq 0.7(\ell_m + c_m) \tag{3.16}$$

$\forall m \in \{1, ..., 12\}$ whether or not a new loan is being taken out at any particular m. This is the 30% "maintenance margin requirement" for borrowing.

The cash is borrowed at a predetermined compounding frequency and rate of interest. Here, we will use E-Trade's lending terms: daily compounding at an annual rate of 8.44%. Given that there are approximately 21 business days per month resulting in approximately 252 business days per 12 months, we simulate daily compounding of the monthly debt by, at the end of each month m after all other transactions, multiplying the debt d_m by

$$\left(1 + \frac{0.0844}{252}\right)^{21}\,. \tag{3.17}$$

3.7.2 Short Covering Requirement

The short covering requirement necessitates that at any time the investor needs to have enough cash in his so-called margin account to buy back (a.k.a. close out; cover) his entire short position. So, when the investor shorts s_0 of stock yielding him s_0 in cash, all of that cash is automatically moved to his margin account such that the margin account contains exactly the amount of cash necessary to buy back his short (which for the moment would cost him s_0). He cannot use any portion of this s_0 to fund his long positions or to do anything else; all cash in the margin account is untouchable, distinct from the investor's c_m (at the end of month m) of usable cash, and so does not contribute to the investor's equity q_m.

So, in general, as time progresses and the value of the shorted assets changes, $\forall m \in \{1, ..., 12\}$ the investor must have s_m of cash (the amount necessary to cover the short at the end of month m) in his margin account at the end of month m. If the value of the shorted assets decreases then some portion of the margin account cash is freed up and transferred to the investor's usable cash c_m. For example, if $s_m < s_{m-1}$ then, at the end of month m, $(s_{m-1} - s_m)$ of the s_{m-1} of cash in the margin account becomes usable and is transferred to the investor's free cash so $c_m = c_{m-1} + s_{m-1} - s_m$, and the other s_m of cash in the margin account has to stay in the margin account in order to cover the short position which is currently worth s_m.

On the other hand, if the value of the shorted assets increases then the investor must somehow input cash into the margin account in order to cover his short position. For example, if $s_m > s_{m-1}$ then, at the end of month m, the investor needs to somehow add $(s_m - s_{m-1})$ of cash into the margin account which would combine with the s_{m-1} of unusable cash already there to give a total of s_m of (unusable) margin account cash, exactly the amount necessary to cover the short position. If $c_{m-1} \geq s_m - s_{m-1}$ then it's easy: the investor simply moves $s_m - s_{m-1}$ of his free cash to his margin account, resulting in the short position being covered and $c_m = c_{m-1} - (s_m - s_{m-1})$. However, if the investor does not have enough free cash to move into his margin account to cover the short position then he must borrow from the brokerage.

Accordingly, before he is allowed to take a short position in the first place, the investor must agree to let his brokerage automatically lend him cash at a predetermined interest rate if he ever needs additional cash to cover a short position. As per the above, this happens if $s_m > s_{m-1}$ and $c_{m-1} < s_m - s_{m-1}$. In this case the brokerage would automatically lend the investor $s_m - s_{m-1} - c_{m-1}$ directly into his margin account at the end of month m, so $d_m = d_{m-1} + s_m - s_{m-1} - c_{m-1}$, and the investor would eventually have to pay back this amount plus interest. The investor is allowed to borrow as much as he needs to cover his short position for as long as he needs, provided that the initial and maintenance margin requirements for borrowing, as described in section 3.7.1, are met.

In effect, this short covering requirement requires the investor to immediately realize any gains or losses from the short position such that, at all times, the net value of the short position is zero. In other words, when it's time to cash out of the portfolio at the end of month 12, the short covering requirement necessitates that the investor have s_{12} in cash just sitting in his margin account, exactly the amount necessary to close out the short position. So, gains and losses of the short position aren't realized at the end of month 12 when the investor cashes out (as they are for the long position); rather, they are realized monthly as c_m and d_m fluctuate as the short position goes in the investor's favor and he gets margin account cash freed up, increasing c_m, or the short position goes against the investor and he has to borrow to meet the short covering requirement, increasing d_m. Accordingly, the net value of the short position at any given time is $(c_m - d_m)$.

3.7.3 Margin Requirements for Shorting

In addition to this short covering requirement that mandates the investor always have enough cash in his margin account (that can't be used for anything else) to buy back his short position, the brokerage also imposes initial and maintenance margin requirements for shorting assets. As with the margin requirements for borrowing presented in section 3.7.1, the initial requirement is 50% and the maintenance one is 30%. However, these requirements have slightly different meanings in the context of shorting as opposed to borrowing.

The initial margin requirement for short selling is the mandate that, in order to initially short the s_0 dollars of assets, the investor needs to have $0.5s_0$ of equity in his account (separate from the s_0 of cash he must have in his margin account immediately after entering into the short). So, in order to enter into the short position in the first place, we require

$$0.5s_0 \leq q_0 = \ell_0 + c_0 - d_0 \,. \tag{3.18}$$

It will be shown in section 3.8 that (3.18) is always necessarily satisfied and so we don't have to worry about the initial margin requirement for shorting.

The maintenance margin requirement for short selling is that, at all times after a short position has been entered into, the investor needs to have 30% of the value of the shorted assets of equity in his account (separate from the 100% of the value of the shorted assets in cash that must be in his margin account). So, $\forall m \in \{1, ..., 12\}$, we require

$$0.3s_m \leq q_m = \ell_m + c_m - d_m \,. \tag{3.19}$$

Note that, whereas the short covering requirement can only be satisfied by cash in the margin account that cannot be used for anything else, these additional margin requirements can be satisfied by all of the investor's equity including his long position. So, the combination of the short covering requirement and the maintenance margin requirement for shorting is the mandate that, at any time, were the investor to liquidate his long position and pay back any debt he owes to the brokerage, he would have at least 130% of the cash necessary to close out his short position. He needs at least 100% of this at least 130% in cash in his margin account and the rest is the cash left over after liquidating the long position and paying back all debt.

If the maintenance margin requirement for shorting is ever not satisfied then the brokerage automatically lends the investor the amount of cash necessary to satisfy the requirement, quite similarly to what happens if the short covering requirement is ever not satisfied as described in section 3.7.2. The only difference is as follows: in the short covering case the additional funds are loaned from the brokerage directly into the investor's margin account (where he is required to have them), whereas in this maintenance margin requirement for shorting case the additional funds are lent to the investor as usable cash. Explicitly, if (3.19) is not satisfied at the end of month m then the brokerage would automatically lend $(0.3s_m - q_m)$ to the investor, instantaneously adding this amount to both c_m and d_m such that it now holds that $0.3s_m = q_m$. Of course, this lending is subject to the margin requirements for borrowing explained in 3.7.1.

3.7.4 Margin Call

What happens if either the investor needs to borrow cash (to satisfy the short covering requirement and/or the maintenance margin requirement for shorting) but the necessary loan would violate (3.15) or (3.16)? Or what if (3.16) is suddenly violated without a new loan being taken out? In either case, an event termed a “margin call” occurs. In a margin call, the brokerage will notify the investor that he needs to immediately deposit enough cash into his account such that (3.15) and (3.16) are once again both satisfied. The investor cannot borrow the additional cash because either (3.15) was violated, which by definition means that the investor cannot borrow any more, or (3.16), which is more lenient than (3.15) was violated which would imply that (3.15) is also violated. So, the investor needs outside funds. If the investor is able to procure enough outside cash and deposit it then all is well. If he is not able to then the brokerage can, without further notice, liquidate the investor’s entire portfolio and pay itself back any debt the investor owed.

Accordingly, if a margin call ever occurs, we assume that the entire portfolio is liquidated, all debt that is owed is paid back to the brokerage, and the investor doesn’t trade for the remainder of the 12 months. We do this because we don’t want to require the investor to have any extra cash on hand outside his account: when we say the investor needs \$5000 of initial cash, we mean that he only needs \$5000 of initial cash, margin call or no margin call.

3.8 Cash and Debt Management Strategy

On the first day of month 1, the investor buys long a proportion ℓ_0 dollars of assets and sells short a proportion s_0. The short covering requirement as detailed in section 3.7.2 mandates that the investor cannot use any part of the s_0 cash proceeds from the short to buy assets. However, this does not necessarily mean that the investor needs a proportion ℓ_0 of cash to start with. Let b be the so-called borrowing parameter such that, on the first day of month 1, the investor borrows $\ell_0 b$ of cash which he combines with $\ell_0(1-b)$ of his own cash in order to enter into the ℓ_0 long position. This means that the investor immediately accumulates $\ell_0 b$ of debt: $d_0 = \ell_0 b$. We set $c_0 = 0$ because having extra cash initially would be inefficient returns-wise (see (3.21) in the next section), so the initial margin requirement for borrowing (3.15) mandates that

$$\ell_0 b \leq 0.5\ell_0 \;\; \Rightarrow \;\; b \leq 0.5\,. \tag{3.20}$$

The investor cannot borrow a negative amount of money so $0 \leq b \leq 0.5$. We will seek to optimize over b in upcoming sections.

We are now in position to show that the initial margin requirement for shorting (3.18) is necessarily satisfied. Substituting c_0 and $d_0 = \ell_0 b$ into (3.18) shows that the requirement is equivalent to $0.5s_0 \leq (1-b)\ell_0$. Because $b \leq 0.5$ as described above and $s_0 \leq \ell_0$ as per (3.9), this equivalent condition is necessarily satisfied and thus we don't have to worry about the initial margin requirement for shorting.

We also specify what the investor will do with any extra cash: if $c_m > 0$ after any required cash has been moved to the margin account (to cover the short) and if the shorting maintenance margin requirement is satisfied then the investor pays back as much debt as possible. Note that paying back debt will never violate the borrowing maintenance margin requirement if it wasn't already violated because if d_m and c_m are decreased by the same amount then the left side of (3.16) decreases by more than the right side decreases.

3.9 Calculating Returns

As previously discussed, the investor somehow chooses r, k, and a group G of $N = 5$ stocks (how he chooses these parameters is the subject of the next several sections), calculates the minimum-variance portfolio $y_t(G, r, k)$, buys into this portfolio on the first business day of year t, and cashes out of it on the last business day of year t. If we observe the returns of the investor's 5 stocks over this year, we should be able to calculate the investor's overall return. Specifically, given known values of $X^i_m \; \forall m \in \{1, ..., 12\}$, $i \in \{1, ..., 5\}$, we seek to calculate the investor's annual return subject to all the regulations and requirements presented in section 3.7. The investor's true annual return using borrowing parameter b is

$$\begin{aligned} a_b &\equiv \frac{\text{(output cash on last day of month 12)} - \text{(input cash on first day of month 1)}}{\text{(input cash on the first day of month 1)}} \\ &= \frac{(\ell_{12} + c_{12} - d_{12}) - (1-b)\ell_0}{(1-b)\ell_0} - 0.02 \end{aligned} \tag{3.21}$$

where the first equality is by definition of return and the second is via the facts that the investor initially inputs $(1-b)\ell_0$ of his own cash (as per section 3.8), that at the end of month 12 the investor liquidates his long position and pays back all debt, and that transaction costs necessarily detract 2% from the return as discussed in section 3.3.

Note that a_b as defined in (3.21) is qualitatively different from the return Y_m defined in (2.4) because (2.4) does not take into account regulations and transaction fees. Thus, even if the portfolio does perform exactly according to expectation, it doesn't actually return r per month. Still, it's justifiable to use the minimum-variance portfolio that is expected to "return" r because, as will be shown later, this strategy works.

Note that the smaller the $\{\ell_m\}$ are and the larger the $\{s_m\}$ are the worse it is for the investor. Accordingly, to conservatively account for the fact that we only use monthly resolution but a_b is technically path-dependent on daily asset prices, we do the following. Define, $\forall m \in \{1, ..., 12\}$,

$$\bar{\ell}_m = \min(\ell_{m-1}, \ell_m)\,, \tag{3.22}$$

$$\bar{s}_m = \max(s_{m-1}, s_m)\,, \tag{3.23}$$

and assume that:

- an instant after markets open on the first business day of month m, the value of the long position instantaneously changes from ℓ_{m-1} to $\bar{\ell}_m$ and the value of the short position instantaneously changes from s_{m-1} to $\bar{s}_m$;
- the long position is worth $\bar{\ell}_m$ and the short position is worth $\bar{s}_m$ for the rest of month m;
- an instant before the markets close on the last business day of month m, the value of the long position changes from $\bar{\ell}_m$ to ℓ_m and the value of the short position changes from $\bar{s}_m$ to s_m.

In other words, we know that the value of the investor's long and short positions go from ℓ_{m-1} and s_{m-1}, respectively, at the beginning of month m to ℓ_m and s_m, respectively, at the end of month m, and to account for the fact that we don't know what happens in between we assume that the worst endpoint for each position holds all month.

Note that, though this is very conservative, it is not a strict lower bound. It is possible, though unlikely, that the value of the long position could, an instant after the markets open on the first business day of month m, drop down from ℓ_{m-1} to a value lower than $\bar{\ell}_m$, remain there all month, and then, an instant before the markets close on the last business day of month m, shoot up to ℓ_m.

An algorithm for calculating a_b for any portfolio y given X_m^i $\forall m \in \{1, ..., 12\}$, $i \in \{1, ..., 5\}$ is as follows, where ℓ_m, s_m, $\bar{\ell}_m$, and $\bar{s}_m$ are defined for any y as per (3.10), (3.11), (3.22), and (3.23).

1: $c_0 \leftarrow 0$
2: $d_0 \leftarrow \ell_0 b$
3: **for** $m = 1 \rightarrow 12$ **do**
4: $\quad c_m \leftarrow c_{m-1} - (\bar{s}_m - \bar{s}_{m-1})$
5: $\quad$ **if** $c_m > 0$ **then**
6: $\qquad$ **if** $d_m > 0.7(\bar{\ell}_m + c_m)$ **then**
7: $\qquad\quad$ margin call: **return** $\dfrac{\bar{\ell}_m + c_m - d_m - (1-b)\ell_0}{(1-b)\ell_0} - 0.02$
8: $\qquad$ **end if**
9: $\qquad$ **if** $c_m > d_m$ **then**
10: $\qquad\quad c_m \leftarrow c_m - d_m$
11: $\qquad\quad d_m \leftarrow 0$
12: $\qquad$ **else** $c_m \leq d_m$
13: $\qquad\quad d_m \leftarrow d_m - c_m$
14: $\qquad\quad c_m \leftarrow 0$
15: $\qquad$ **end if**
16: $\quad$ **else** $c_m \leq 0$
17: $\qquad d_m \leftarrow d_m - c_m$
18: $\qquad c_m \leftarrow 0$
19: $\qquad$ **if** $d_m > 0.5\bar{\ell}_m$ **then**
20: $\qquad\quad$ margin call: **return** $\dfrac{\bar{\ell}_m + c_m - d_m - (1-b)\ell_0}{(1-b)\ell_0} - 0.02$
21: $\qquad$ **end if**
22: $\quad$ **end if**
23: $\quad$ **if** $0.3\bar{s}_m > \bar{\ell}_m + c_m - d_m$ **then**
24: $\qquad d_m \leftarrow d_m + 0.3\bar{s}_m - (\bar{\ell}_m + c_m - d_m)$
25: $\qquad c_m \leftarrow c_m + 0.3\bar{s}_m - (\bar{\ell}_m + c_m - d_m)$
26: $\qquad$ **if** $d_m > 0.5(\bar{\ell}_m + c_m)$ **then**
27: $\qquad\quad$ margin call: **return** $\dfrac{\bar{\ell}_m + c_m - d_m - (1-b)\ell_0}{(1-b)\ell_0} - 0.02$
28: $\qquad$ **end if**
29: $\quad$ **end if**
30: $\quad d_m \leftarrow (1 + 0.0844/252)^{21} d_m$
31: **end for**
32: $c_{12} \leftarrow c_{12} + \bar{s}_{12} - s_{12}$
33: **return** $\dfrac{\ell_{12} + c_{12} - d_{12} - (1-b)\ell_0}{(1-b)\ell_0} - 0.02$

Algorithm 3.1: *Algorithm for calculating a_b for any portfolio y.*

In words the algorithm is as follows, where the numbers are the lines of the algorithm.

- 1-2: Set the initial cash and initial debt as per section 3.8.
- 3: Loop through the twelve months of the year and do the following for each year.

- 4: Transfer between margin account and cash to exactly satisfy the short covering requirement as described in section 3.7.2.
- 5: If this results in a positive cash balance then no borrowing is required to cover the short.
- 6-7: If the maintenance margin requirement for borrowing (3.16) isn't satisfied then liquidate the portfolio and terminate the algorithm, resulting in a return analogous to (3.21). However, to calculate the return we use $\bar{\ell}_m$ instead of ℓ_m because the investor liquidates his position as soon as the margin call occurs, whether or not that's at the end of the month.
- 9-11: As per section 3.8, if there is enough cash to payoff all the debt then do so.
- 12-14: As per section 3.8, if there is only enough cash to payoff some of the debt then payoff as much as possible.
- 16-18: If transferring between margin account and cash to exactly satisfy the short covering requirement results in a negative cash balance, then borrow exactly the difference such that the short covering requirement is satisfied.
- 19-20: If this act of borrowing violates the initial margin requirement for borrowing (3.15) then liquidate the portfolio and terminate the algorithm, resulting in a return identical to the one described in the above description of lines 6-7.
- 23-25: If the shorting maintenance requirement (3.19) is violated then borrow exactly the amount of cash necessary to satisfy it.
- 26-27: If this act of borrowing violates the initial margin requirement for borrowing (3.15) then liquidate the portfolio and terminate the algorithm, resulting in a return identical to the one described in the above description of lines 6-7.
- 30: Increase debt to include a month of daily compounded interest at 8.44% per year as per (3.17).
- 32: Update the amount of cash in hand at the end of the year to account for the facts that $\bar{s}_{12} \geq s_{12}$ as per (3.23) and that the investor liquidates his position at the very end of month 12.
- 33: Liquidate the portfolio and collect the return (3.21) where we use ℓ_{12} not $\bar{\ell}_{12}$ because the investor liquidates his position at the very end of month 12.

So, the following function is well-defined:

$$a_b(y_t(G, r, k)) \tag{3.24}$$

which first calculates $y_t(G, r, k)$ as described in section 3.5 using only data from the k months prior to January of year t and then, using algorithm 3.1 and X_m^i $\forall m \in \{1, ..., 12\}$, $i \in \{1, ..., 5\}$ for year t, calculates what the portfolio $y_t(G, r, k)$ actually would have returned in year t, with all fees and regulations taken into account, had the investor borrowed a percentage b of the cash necessary to enter the portfolio. In other words, given r and k, $a_b(y_t(G, r, k))$ constructs the minimum-variance portfolio of the stocks in G that the investor would have constructed on the first business day of January of year t with only data that was been available to him at that time, and then, given borrowing parameter b, calculates how this portfolio would have actually performed in terms of return, with all regulations and associated costs considered, over the 12 months of year t, which are the 12 months the investor would have held the portfolio for. So, $a_b(y_t(G, r, k))$ is a true, all-things-considered, real-world, unbiased measure of how well the portfolio $y_t(G, r, k)$ would have actually performed during the year it was constructed to be a minimum-variance portfolio for. Note that $a_b(y_t(G, r, k))$ is only defined if it is January 1st of year $(t + 1)$ or later. Accordingly, at the time of writing of this paper, $a_b(y_t(G, r, k))$ is only defined $\forall t \in \{2011, 2010, ...\}$.

4 Investment Strategies

In this section, with $y_t(G, r, k)$ and $a_b(y_t(G, r, k))$ in hand, we devise and present the decision-making support tool using the fixed-return minimum-variance result from section 2 and various technical aspects from section 3. We also present the optimal set of parameter-choosing strategies along with the progression of parameter-choosing strategies that led to this optimal set. The combinations of the decision-making support tool with these optimal parameter-choosing strategies result in an optimal set of investment strategies.

4.1 Decision-Making Support Tool

Suppose that it is the first business day of year t. Given an exogenously chosen group of 5 stocks G and exogenously chosen values for b, r, and k, the minimum-variance portfolio for the upcoming year $y_t(G, r, k)$ is uniquely specified and the investor can simply enter into this portfolio with borrowing parameter b, hold the portfolio all year, and then cash out on the last business day of year t. Furthermore, with a_b, as defined in section 3.9, in hand, the investor can calculate the true, all-things-considered return of any minimum-variance portfolio for any year before t.

We define the decision-making support tool as consisting of the following three elements:

1. $y_t(G, r, k)$: the ability to construct the minimum-variance portfolio of the stocks in G that is expected to return r during month 1 of year t according to historical data from the prior k months;

2. $a_b(y_t(G, r, k))$: the ability to calculate the all-things-considered return of any minimum-variance portfolio assuming that a fraction b of the initial cash required to enter into the portfolio is borrowed from the broker;

3. the general strategy of, given G, b, r, and k, entering into $y_t(G, r, k)$ on the first business day of year t with borrowing parameter b, holding this portfolio all year, and then cashing out on the last business day of year t.

With this investment decision-making tool in hand, all the investor has to do is choose G, b, r, and k and he has a comprehensive and uniquely specified investment strategy. But how should he choose these parameters? We need to devise a parameter-choosing strategy. Then, the combination of the decision-making tool with this parameter-choosing strategy is a comprehensive and uniquely specified investment strategy.

Note that the investment decision-making tool is kept distinct from the parameter-choosing strategy because, with the decision-making tool in hand, we can devise several different parameter-choosing strategies, resulting in several different investment strategies. In fact, this is what we will do in the following section: devise a set of parameter-choosing strategies and hence a set of investment strategies.

While it is the case that, given a set of parameters, only elements 1 and 3, as numbered above, of the decision-making support tool are required to invest, element 2 is useful in that it will help us choose the parameters in the first place. That is, element 2 is an important element of the parameter-choosing strategies presented in the next section.

4.2 Optimal Set of Parameter-Choosing Strategies

Several different parameter-choosing strategies were devised and tested. The ones presented in this section are optimal in that they resulted in the best-performing investment strategies. Section 4.3 describes the progression of parameter-choosing strategies that yielded these optimal ones.

In these parameter-choosing strategies, we define two additional parameters in order to choose the parameters G, r, and k: the initial long position cutoff L and the return cutoff A, both of which will be defined below. Note that b was not mentioned: each of these parameter-choosing strategies selects values for G, r, and k but not b and not L or A either; b, L, and A must be exogenously chosen. Accordingly, we have a different parameter-choosing strategy for each different exogenously chosen triple of (b, L, A) values. This is why we have been referring to these optimal parameter-choosing strategies collectively as the optimal set of parameter-choosing strategies; they are fundamentally identical, only differing in their (b, L, A) triples. So, we have one parameter-choosing strategy defined by ($b = 0$, $L = 2$, $A = 0.3$), another defined by ($b = 0.25$, $L = 3$, $A = 0.6$), etc. They are all optimal in that, essentially regardless of the (b, L, A) triple, each one, when combined with the decision-making support tool, results in an investment strategy that is superior to any of the investment strategies created from any other parameter-choosing strategies.

The parameter-choosing strategy for any (b, L, A) triple is as follows. Denote by $\mathcal{C}_n$ the n-combination function such that $\mathcal{C}_n(Z)$ is the set of all unique subsets of Z that have n elements:

$$\mathcal{C}_n(Z) = \left\{ B \subset Z \;\middle|\; |B| = n \right\} \quad \forall n \leq |Z| \,. \tag{4.1}$$

For example, $\mathcal{C}_2(\{1,2,3\}) = \{\{1,2\},\{1,3\},\{2,3\}\}$.

Exogenously choose a master set $\mathcal{G}$ of more than 5 stocks. Define discrete grids[1] in r and k:

$$\mathcal{R} = \{0.005, 0.01, 0.015, ..., 0.095\} \;\; \text{(in units of monthly return)}\,, \tag{4.2}$$

$$\mathcal{K} = \{5, 10, 15, ..., 95\} \;\; \text{(in units of months of historical data)}\,. \tag{4.3}$$

Then, for each group of 5 stocks that can be created from $\mathcal{G}$, that is $\forall G \in \mathcal{C}_5(\mathcal{G})$, define the G-specific (r, g) pair that gave the highest return last year with borrowing parameter b:

$$(r_G, k_G) = \arg \max_{(r,k) \in \mathcal{R} \times \mathcal{K}} a_b \left(y_{t-1} \left(G, r, k \right) \right) \,. \tag{4.4}$$

If

$$a_b \left(y_{t-1} \left(G, r_G, k_G \right) \right) \geq A \tag{4.5}$$

and $y_t \left(G, r_G, k_G \right)$ is such that

$$\ell_0 \leq L \tag{4.6}$$

(where ℓ_0 is defined[2] by (3.7)) then "accept" G. In words, if last year's G-specific highest return is at least A and if this year's portfolio with last year's G-specific optimal (r, g) pair has an initial long position worth no more than L, then "accept" G. Otherwise, if either (4.5) or (4.6) is violated then "reject" G. Denote by $C_t(b, L, A)$ the set of all year-t accepted groups using the triple of parameters (b, L, A):

$$C_t(b, L, A) = \Big\{ G \in \mathcal{C}_5(\mathcal{G}) \;\Big|\; a_b \left(y_{t-1} \left(G, r_G, k_G \right) \right) \geq A \,, \quad y_t \left(G, r_G, k_G \right) \text{ has } \ell_0 \leq L \Big\} \,. \tag{4.7}$$

Choose G by randomly selecting an element of $C_t(b, L, A)$.

[1]See section 4.2.4 for a discussion of these discrete grids. Also, in the more-easily-interpretable units of annual return, we have $12\mathcal{R} = \{0.06, 0.12, 0.18, ..., 1.14\}$.

[2]Recall from section 3.5 that $\ell_0 \geq 1$ is the per-dollar initial value of all the assets the investor buys long.

4.2.1 Optimal Set of Investment Strategies

Now, if we can exogenously choose $\mathcal{G}$, for any (b, L, A) triple we have an investment strategy that is the combination of the decision-making support tool with the (b, L, A) parameter-choosing strategy described above. This investment strategy is to, recalling that the parameter-choosing strategy chooses G and defines r_G and k_G, randomly enter into $y_t(G, r_G, k_G)$ on the first business day of year t with borrowing parameter b, hold this portfolio all year, and cash out on the last business day of year t. The set of all such investment strategies for all possible[3] (b, L, A) triples is the optimal set of investment strategies.

In the upcoming results and conclusions sections, we will determine the best (b, L, A) triple, which is equivalent to determining the single best parameter-choosing strategy out of the optimal set of parameter-choosing strategies and hence the single best investment strategy out of the above-defined optimal set of investment strategies. We will do so by testing the different investment strategies that result from different (b, L, A) triples and choose the best-performing one.

We will not outline a method for choosing $\mathcal{G}$; it is the only parameter that truly needs to be chosen exogenously. However, the $\mathcal{G}$ used in this paper, as defined in the results section, worked well and thus the interested investor is encouraged to use this same $\mathcal{G}$, though any sufficiently large $\mathcal{G}$ should do the job.

4.2.2 Learning from History

Note that the method by which this parameter-choosing strategy selects G and r is one way in which the resultant comprehensive investment strategy learns from the past to make predictions about the future. Specifically, the strategy uses past data to predict which G and r will be optimal in the future. The same is done with k and thus not only does the strategy learn from the past to make predictions about the future, it also learns from the past in order to decide how much of the past to learn from.

4.2.3 Return and Initial Long Position Cutoffs

The reasoning behind limiting ℓ_0 to L is to prevent from needing too much cash up front. Even if we max out how much the investor initially borrows ($b = 0.5$), he still needs $0.5\ell_0$ of his own cash to enter the portfolio. So, if ℓ_0 is unbounded which it otherwise would be because the minimum-variance portfolio could have any ℓ_0, then the amount of cash the investor needs is unbounded. This, of course, can be detrimental to returns (see (3.21)). That

[3] As will be described in 5.5, it is necessary to limit (b, L, A) triples to those on a discrete, finite grid. So, we don't really consider all possible (b, L, A) triples; only certain ones.

being said, there's no reason to believe a certain L value should be better than any other. We will test several different L values, including $L = \infty$ which is equivalent to having no restriction at all.

The reasoning behind accepting a set G of 5 stocks only if its best return last year was at least A is that last year's best return isn't desirable if, for example, it's -90%. So, $\forall G \in \mathcal{C}_5(\mathcal{G})$ we choose the r and k that gave the best return last year, but we only accept this best return from last year as "good enough" if it is at least some cutoff A. As is the case with L, there's no reason to believe that a certain A value is optimal and so we will test several different values, including $A = -\infty$ which is equivalent to having no restriction at all.

4.2.4 Optimization over a Discrete Grid

In defining r_G and k_G, we lack a closed-form expression for $a_b(y_t(G, r, k))$ that we can analytically maximize, either by differentiation with respect to r and k or by other methods, to find a true (r, k) maximum. So, we must rely on numerical methods: we must optimize over a discrete grid as done in (4.4) as opposed to a continuous region in (r, k)-space, hence the necessity of defining $\mathcal{R}$ and $\mathcal{K}$. Note that we cannot be sure that we've chosen the best grids for $\mathcal{R}$ and $\mathcal{K}$. We could have made the resolution of $\mathcal{K}$ finer to the point of including every positive integer in a certain range, we could have made the resolution of $\mathcal{R}$ infinitely finer, and we could have made the range of each wider. The specific bounds and resolutions in (4.2) and (4.3) were chosen as a balance between computational feasibility, resolution, and range.

4.3 Evolution of Parameter-Choosing Strategies

As previously mentioned, the optimal parameter-choosing strategies presented in the previous section were neither the only nor the first ones developed. This section discusses the progression of parameter-choosing strategies that resulted in this optimal set.

All of these parameter-choosing strategies involve exogenously choosing the master group of stocks; we never devised a method to choose $\mathcal{G}$.

The first type of parameter-choosing strategy that was devised is to set $b = 0$, exogenously choose an (r, k) pair, and randomly choose $G \in \mathcal{G}$. So, unlike the optimal parameter-choosing strategies from the previous section, these strategies don't set G-specific r and k values. This results in a set of parameter-choosing strategies and thus a set of investment strategies with each defined by a (r, k) pair. Not only did these investment strategies perform poorly overall, it was also the case that there was no optimal (r, k) pair: no one strategy clearly outperformed the others.

The next step was to let b vary as well and keep everything else the same, resulting in a set of parameter-choosing strategies and thus a set of investment strategies with each defined by a (b, r, k) triple. The results were similarly disappointing: the investment strategies performed poorly and there was no clearly optimal (b, r, k) triple.

The next step was to introduce the initial long position cutoff L exactly as defined in the optimal parameter-choosing strategies from the previous section, the return cutoff A, almost exactly as defined in said optimal parameter-choosing strategies, and G-specific (b, r, k) triples analogous to the G-specific (r, k) pairs defined in said optimal parameter-choosing strategies. Specifically, the G-specific triples were defined as, analogous to (4.4),

$$(b_G, r_G, k_G) = \arg \max_{(b,r,k) \in \mathcal{B} \times \mathcal{R} \times \mathcal{K}} a_b(y_{t-1}(G, r, k)) \tag{4.8}$$

for some grid $\mathcal{B}$ in b-space, and the return-cutoff was, analogous to (4.5), to accept a $G \in \mathcal{G}$ if and only if

$$a_{b_G}(y_{t-1}(G, r_G, k_G)) \geq A\,. \tag{4.9}$$

Accordingly, L and A were the only exogenously chosen parameters, resulting in a set of parameter-choosing strategies and thus a set of investment strategies with each defined by a (L, A) pair. Note that $L = \infty$, equivalent to having no initial long position cutoff, and $A = -\infty$, equivalent to having no return cutoff, were values included in the (L, A) grid. These investment strategies performed significantly better than those that came before them, but were still not quite as historically profitable as desired.

Finally, we arrived at the optimal set of parameter-choosing strategies and thus the optimal set of investment strategies, each defined by a (b, L, A) triple, by making b exogenously chosen as opposed to G-specific by changing (4.8) to (4.4) and (4.9) to (4.5). These investment strategies performed well enough that we settled with them and didn't devise any further parameter-choosing strategies. However, an even better set of parameter-choosing strategies could exist; we call the set here "optimal" because it was the optimal set that we tested, not because it is necessarily globally optimal.

5 Results

In this section we present the results obtained by backtesting each strategy in the optimal set of investment strategies to see how well they would have performed in each of the last several years. We then inspect these results to select a single optimal investment strategy and we confirm that this is the profitable, practical, and capital non-intensive strategy for an unsophisticated investor we seek by verifying that it outperforms various benchmarks.

5.1 Backtesting

Recall that it is the first business day of year t. We want to test how well some investment strategy performs but we don't want to wait a full year to see what return the strategy achieves during year t. In other words, we don't want to wait a full year until $a_b(y_t(G, r, k))$ is defined. However, $a_b(y_{t-1}(G, r, k))$, $a_b(y_{t-2}(G, r, k))$, $a_b(y_{t-3}(G, r, k))$, etc. are currently defined so we can test the strategy by observing exactly how much it would have returned in each of the past several years. This method is known as backtesting and is what we will employ in the upcoming results section to gauge the effectiveness of the investment strategy. Specifically, for various different triples of (b, L, A) values, we will observe the distribution of possible returns for each of 2008, 2009, 2010, and 2011 by calculating $a_b(y_t(G, r, k))$ $\forall t \in \{2008, 2009, 2010, 2011\}$ and $\forall G \in C_t(b, L, A)$.

More explicitly, in order to check how well the strategy would have performed in year $t \in \{2008, 2009, 2010, 2011\}$ for any triple of (b, L, A) values, we construct the distribution of possible year-t returns:

$$D_t(b, L, A) = \left\{ a_b\left(y_t\left(G, r_G, k_G\right)\right) \;\middle|\; G \in C_t(b, L, A) \right\} \tag{5.1}$$

where, recall, $C_t(b, L, A)$ as defined by (4.7) is the set of all year-t accepted groups of stocks using (b, L, A). $D_t(b, L, A)$ is simply a set of values; we call it a distribution to emphasize the fact that the strategy randomly selects a $G \in C_t(b, L, A)$ to trade with and thus the year-t return of the strategy is a random variable on a uniform distribution with discrete sample space $D_t(b, L, A)$. Accordingly, in the upcoming results section, for each (b, L, A) triple and each $t \in \{2008, 2009, 2010, 2011\}$, we present a number of metrics that summarize the distribution $D_t(b, L, A)$.

Recall from section 3.9 that each $a_b(y_t(G, r, k))$ value is a true, unbiased, all-things-considered measure of historical performance. Even though we now know exactly what asset returns were in, for example, 2009, the function $y_{2009}(G, r, k)$ constructs the minimum-variance portfolio with only information that would have been available to the investor on the first business day of 2009, and only after then does a_b apply the knowledge of what returns actually were in 2009 to see how $y_{2009}(G, r, k)$ would have actually performed in 2009. Furthermore, as also discussed in 3.9, $a_b(y_t(G, r, k))$ takes into account all transaction costs as well as all fees associated with any rules and regulations that apply over the course of holding the portfolio $y_t(G, r, k)$ during year t. So, all in all, we can be completely confident in using $a_b(y_t(G, r, k))$ values as the basis of our results.

5.2 Not Using 2008 as a Guide

The financial crisis of 2008 was a particularly bad and truly anomalous[1] time to be long the stock market,[2] and, as per (2.3), any investment strategy generated by the decision-making support tool is necessarily long the stock market. A result of this, as it turns out, is that 2008 is a poor guide for choosing parameters for 2009: using (4.4) to set the 2009 values for r and k for each stock group $G \in \mathcal{C}_5(\mathcal{G})$ as the best 2008 values of r and k for that G results in a terrible 2009 return distribution $D_{2009}(b, L, A)$ for all (b, L, A) triples that were tested. Accordingly, we adjust the strategy such that, for 2009, it chooses the best r and k from 2007 as opposed to 2008. This manifests in the following adjustment to (4.4) that only applies when $t = 2009$:

$$\begin{aligned}(r_G, k_G) &= \arg \max_{(r,k)\in\mathcal{R}\times\mathcal{K}} a_b\left(y_{t-2}\left(G, r, k\right)\right) \\ &= \arg \max_{(r,k)\in\mathcal{R}\times\mathcal{K}} a_b\left(y_{2007}\left(G, r, k\right)\right) \, . \end{aligned} \tag{5.2}$$

[1] The S&P 500 has yielded a mean annual return of 8.47% with standard deviation 15.69% over the last 62 years. In 2008 the index returned -37%, the lowest return in the last 62 years and nearly three standard deviations away from the mean. Accordingly, it is safe to label 2008 as an anomalous year for the stock market. On the other hand, the index's annual returns during 2009, 2010, and 2011 are, in order, 26.46%, 15.06%, and 2.05%, which are, respectively, roughly one, one half, and one half standard deviations from the mean. So, it is safe to label 2009, 2010, and 2011 as more typical years for the stock market. See source [8].

[2] Among other assets. See source [9].

In the spirit of accurately regarding 2008 as an anomaly that does not help us choose parameters for 2009, we also adjust the return cutoff to skip 2008 such that we only use a group of stocks G in 2009 if the optimal return for that group in 2007, as opposed to 2008, was at least A. This manifests in the following adjustment to (4.5) that only applies when $t = 2009$:

$$a_b\left(y_{t-2}\left(G, r_G, k_G\right)\right) = a_b\left(y_{2007}\left(G, r_G, k_G\right)\right) \geq A\,. \tag{5.3}$$

Because it's after 2009 we can observe various D_{2009} distributions both with and without the adjustments and conclude that the 2009 return distributions are better with them. However, if, for example, it's the first business day of year t and we observed year $(t-1)$ to be an anomaly similar to 2008, we cannot be sure that various D_t distributions will ultimately be more favorable if we adjust the strategy to "skip" year $(t-1)$ as we "skipped" 2008.

Regardless, we make it the official policy of the optimal set of parameter-choosing strategies to do so: as an addendum to section 4.2, in the optimal set of parameter-choosing strategies make adjustments analogous to (5.2) and (5.3) in order to not use any anomalous year $(t-1)$, like 2008, as a guide for choosing r and k and enforcing the return cutoff in year t. The justification for this is that the main underlying premise of the decision-making support tool, the optimal set of parameter-choosing strategies, and thus the optimal set of investment strategies is to use historical data to make educated guesses about the future. Albeit with unit sample size $(t-1 = 2008)$, we observe that, in the past, any anomalous year does not serve as an accurate guide for choosing (r, k) and enforcing the return cutoff in the following year. Thus, from this historical observation, we make the educated guess that in the future, too, any anomalous year will not serve as an accurate guide for choosing (r, k) and enforcing the return cutoff in the following year.

Of course, an exact definition of "anomalous year" remains to be specified. As per footnote 1, perhaps a metric based on how many standard deviations away from the mean the annual S&P 500 return was would work well.

5.3 Distribution Metrics

The metrics we will use in the upcoming sections to summarize any return distribution $D_t(b, L, A)$ are as follows.

We can express the number of returns in $D_t(b, L, A)$, which is equivalent to the number of elements of $\mathcal{C}_5(\mathcal{G})$ that were deemed acceptable groups of stocks, both in absolute terms, as

$$|D_t(b, L, A)|\ , \tag{5.4}$$

and as a fraction of the maximum possible number of acceptable groups of stocks, as

$$|D_t(b, L, A)|_{\%} = \frac{|D_t(b, L, A)|}{|\mathcal{C}_5(\mathcal{G})|} = \frac{|D_t(b, L, A)|}{\binom{|\mathcal{G}|}{5}} \,. \tag{5.5}$$

The fraction of the returns in $D_t(b, L, A)$ that are profitable (don't lose money) is

$$|D_t(b, L, A) \geq 0|_{\%} = \frac{\left|\left\{a \in D_t(b, L, A) \;\middle|\; a \geq 0\right\}\right|}{|D_t(b, L, A)|} \,. \tag{5.6}$$

Denote the mean and standard deviation of $D_t(b, L, A)$ by, respectively,

$$\mu_t(b, L, A) = \text{mean}\left(D_t(b, L, A)\right) \,, \tag{5.7}$$

$$\sigma_t(b, L, A) = \sqrt{\text{var}\left(D_t(b, L, A)\right)} \,. \tag{5.8}$$

The $P\%$ percentile of $D_t(b, L, A)$, denoted by $Q_t^{P\%}(b, L, A)$, is the return in $D_t(b, L, A)$ such that as close as possible to exactly $P\%$ of all returns in $D_t(b, L, A)$ are less than or equal to it; formally $Q_t^{P\%}(b, L, A)$ is the element of $D_t(b, L, A)$ that best approximates

$$\frac{\left|\left\{a \in D_t(b, L, A) \;\middle|\; a \leq Q_t^{P\%}(b, L, A)\right\}\right|}{|D_t(b, L, A)|} \simeq P\% \,. \tag{5.9}$$

$Q^{75\%}$, $Q^{50\%}$, and $Q^{25\%}$ are known, respectively, as the upper quartile, median, and lower quartile. Furthermore, note that

$$Q_t^{100\%}(b, L, A) = \max\left(D_t(b, L, A)\right) \,, \tag{5.10}$$

$$Q_t^{0\%}(b, L, A) = \min\left(D_t(b, L, A)\right) \,. \tag{5.11}$$

$D_t(b, L, A)$ can be summarized by the set of values $Q_t^{P\%}(b, L, A)$ $\forall P \in \{100, 75, 50, 25, 0\}$.

5.4 Benchmarks

In order to accurately evaluate the results we obtain from backtesting and thus label the strategy as "successful" or "unsuccessful", we need benchmark returns for comparison. For example, a yearly return of 50% seems impressive but it's not if the stock market as a whole averaged the same 50% during that year. This section discusses how to create benchmark values that will be used for comparison in upcoming sections.

As described in section 5.1, given a triple (b, L, A) we have a distribution of possible returns $D_t(b, L, A)$ for each year t. Accordingly, we compare this distribution to a benchmark distribution of returns which is created as follows. For each $G \in \mathcal{C}_5(\mathcal{G})$, randomly create a portfolio by randomly and independently choosing x_1, x_2, x_3, and x_4 from a uniform distribution on[3] $(-0.8, 1.2)$ and setting $x_5 = 1 - x_1 - x_2 - x_3 - x_4$ such that (2.3) is satisfied. For each G and for all[4] t, denote this random portfolio by $\widetilde{y}_t(G)$. For each year $t \in \{2008, 2009, 2010, 2011\}$, calculate what each of these portfolios would have returned with borrowing parameter[5] $b = 0$, thus generating a distribution of random returns for each year. Formally, define the year-t distribution of random returns as

$$E_t = \left\{ a_{b=0}\left(\widetilde{y}_t(G)\right) \;\middle|\; G \in \mathcal{C}_5(\mathcal{G}) \right\} . \tag{5.12}$$

Given that we will want to compare the standard deviation of $D_t(b, L, A)$ to a benchmark standard deviation, for each (b, L, A) triple and for each t we need to create a benchmark distribution that has the same number of returns[6] as $D_t(b, L, A)$. So, for each (b, L, A) triple and for each year t, randomly select $|D_t(b, L, A)|$ elements of E_t to put into the benchmark distribution for that (b, L, A) and that t. Formally, call this benchmark distribution $\widetilde{D}_t(b, L, A)$ such that

$$\widetilde{D}_t(b, L, A) = \text{a random element of } \mathcal{C}_{|D_t(b,L,A)|}\left(E_t\right) . \tag{5.13}$$

Define the "benchmark mean" and "benchmark standard deviation" as, respectively,

$$\widetilde{\mu}_t(b, L, A) = \text{mean}\left(\widetilde{D}_t(b, L, A)\right) , \tag{5.14}$$

$$\widetilde{\sigma}_t(b, L, A) = \sqrt{\text{var}\left(\widetilde{D}_t(b, L, A)\right)} . \tag{5.15}$$

Note that the only exogenous choice that needs to be made to execute the investment strategy is choosing the stocks that compose $\mathcal{G}$. Accordingly, to construct the benchmark distributions, we use the same $\mathcal{G}$ but then randomize everything else, paralleling the fact that the investment strategy chooses everything else. So, comparing $D_t(b, L, A)$ to $\widetilde{D}_t(b, L, A)$ is a true test of the investment strategy – it allows us to determine if the decisions the strategy

[3] The uniform distribution is centered around 0.2 such that, on average, each portfolio has uniform proportions: $x_1 = x_2 = x_3 = x_4 = x_5 = 0.2$. The distribution has lower bound -0.8 and upper bound 1.2 such that the portfolio weights vary up to a unit proportion away from the mean of 0.2.

[4] For a fixed G, we use the same random portfolio for each year.

[5] We choose $b = 0$ to construct the distribution of random returns because setting $b \neq 0$ is one of the tricks of the parameter-choosing strategies and thus the investment strategies used here; an unsophisticated investor certainly wouldn't think to set $b \neq 0$ on his own.

[6] This is because if we take n independent samples of a random variable, the variance of the set of samples increases with n. For example, if $n = 1$ then the variance is necessarily zero, whereas if $n > 1$ (and the sample space is non-finite) then the variance is nonzero with probability one.

makes are actually beneficial or if, for the sake of returns, said decisions would have been better off randomized.

Additionally, we include the yearly returns of the S&P 500, a standard risky[7] benchmark return, and 1-year CDs,[8] a standard risk-free benchmark return. There are no distributions of returns for the S&P 500 and 1-year CDs: for each $t \in \{2008, ..., 2011\}$ we know[9] the single value that is the return the S&P 500 achieved during year t and the single value that is the return a 1-year CD provided during year t. To emphasize that these are fixed rates of return as opposed to means of return distributions, denote them by r:

$$r_{\text{S\&P},\,t} = \text{ S\&P 500 annual return during year } t\,, \tag{5.16}$$

$$r_{\text{CD},\,t} = \text{ 1-year CD annual interest rate during year } t\,. \tag{5.17}$$

So, for each $t \in \{2008, ..., 2011\}$ we now have three benchmark returns: $\widetilde{\mu}_t$, $r_{\text{S\&P},\,t}$, and $r_{\text{CD},\,t}$. Note that, because we have known returns for the S&P 500 and 1-year CDs as opposed to distributions, there are no benchmark standard deviations associated with these S&P 500 and 1-year CD benchmark returns. So, whereas $r_{\text{S\&P}}$ and r_{CD} are analogous to $\widetilde{\mu}$, there is neither a S&P 500 nor a 1-year CD analogy to $\widetilde{\sigma}$.

5.5 (b, L, A)-Space

For any pair (b, L), define the maximum A value (in steps[10] of 0.1) for which there is at least one accepted group of stocks for each year:

$$A_{\max}(b, L) = \max\left(\left\{A \in \{-\infty, 0, 0.1, 0.2, 0.3, ...\} \;\middle|\; |D_t(b, L, A)| \geq 1 \;\forall t \in \{2008, ..., 2011\}\right\}\right). \tag{5.18}$$

Note that $A_{\max}(b, L)$ is always finite because $|\mathcal{C}_5(\mathcal{G})| < \infty$ and thus $\exists A < \infty$ such that $a_b(y_{t-1}(G, r_G, k_G) < A \;\forall G \in \mathcal{C}_5(\mathcal{G})$.

[7]It is a risky benchmark because the underlying stocks are risky: the rate of return of any stock over an upcoming time period is unknown, in contrast to a CD (see footnote 8).

[8]Certificate of Deposit; the simplest way to put money in a bank and earn interest. You deposit money in the bank at a fixed, predetermined rate of interest with a certain maturity and only upon maturity can you withdraw your money (plus interest). The interest rate is risk-free in the sense that it is fixed and known: unless the bank defaults, you know exactly how much interest you will be able to withdraw along with your deposit upon maturity. Given that the investment strategy presented here involves holding a portfolio for one year, 1-year CDs are the proper-maturity risk-free instruments to use for comparison.

[9]Historical S&P 500 returns are easily obtainable for free online (see source [8]). We were unable to find a free, online, reliable source for historical 1-year CD rates; instead, we used the Harvard network to access source [10]: Global Financial Data.

[10]The A values tested were integer multiples of 0.1 (as well as $A = -\infty$), so $A_{\max}$ isn't the true maximum value that resulted in accepted stock groups each year but rather the greatest non-negative multiple of 0.1 that resulted in accepted stock groups each year. This is identical to the issue of grid resolution discussed in section 4.2.4.

As it turns out, for each (b, L) pair that was tested (see (5.19), below), greater A values result in strictly higher returns for all years and all of the return-based distribution metrics from section 5.3. Thus, in the results below, for each (b, L) pair we present only results for $A = A_{\max}(b, L)$. The (b, L) pairs tested are the elements of the grid[11]

$$\mathcal{B} \times \mathcal{L} = \{0, 0.1, 0.25, 0.38, 0.5\} \times \{2, 3, \infty\} \tag{5.19}$$

where $\mathcal{B}$ is b-space and $\mathcal{L}$ is L-space such that the (b, L, A) triples for which results are presented are elements of the grid

$$\mathcal{B} \times \mathcal{L} \times \{A_{\max}(b, L)\} . \tag{5.20}$$

5.6 Cumulative Returns

In the results that follow, for each $(b, L, A) \in \mathcal{B} \times \mathcal{L} \times \{A_{\max}(b, L)\}$, we compare the year-by-year cumulative returns resultant from reinvesting in the strategy's mean portfolio each year to the year-by-year cumulative returns resultant from reinvesting in each of the three benchmarks each year (where we use the mean $\widetilde{\mu}_t$ of the benchmark random distribution). Explicitly, for any (b, L, A) triple, the investment strategy's cumulative 2008 through year-t mean return and cumulative 2009 through year-t return are denoted by c and defined as

$$c_{2008,\,t}(\mu) = \left[\prod_{n=2008}^{t} (1 + \mu_n(b, L, A)) \right] - 1 \,, \tag{5.21}$$

$$c_{2009,\,t}(\mu) = \left[\prod_{n=2009}^{t} (1 + \mu_n(b, L, A)) \right] - 1 \,. \tag{5.22}$$

In these equations, replace $\mu_n(b, L, A)$ by:

- $\widetilde{\mu}_n(b, L, A)$ to obtain $c_{2008,\,t}(\widetilde{\mu})$ and $c_{2009,\,t}(\widetilde{\mu})$, the cumulative 2008 and 2009 through year-t returns of the mean of the benchmark random distribution,
- $r_{\text{S\&P},\,n}$ to obtain $c_{2008,\,t}(r_{\text{S\&P}})$ and $c_{2009,\,t}(r_{\text{S\&P}})$, the cumulative 2008 and 2009 through year-t returns of the S&P 500,
- $r_{\text{CD},\,n}$ to obtain $c_{2008,\,t}(r_{\text{CD}})$ and $c_{2009,\,t}(r_{\text{CD}})$, the cumulative 2008 and 2009 through year-t returns of a 1-year CD.

In words, $c_{2008,\,t}(\mu)$ is the total return that results from earning an annual return of μ_{2008} for one year, then reinvesting and earning an annual return of μ_{2009} for one year, etc., and finally reinvesting and earning an annual return of μ_t for one year.

[11] As with all the discrete grids so far, there is no reason to believe that this is the best one to use. It is simply a compromise between computational feasibility, resolution, and range (except for b for which the range is bounded as per (3.20) and thus the b compromise is only between computational feasibility and resolution).

5.6.1 Annualized Cumulative Returns

For each $(b, L, A) \in \mathcal{B}\times\mathcal{L}\times\{A_{\max}(b, L)\}$, in addition to summarizing $D_t(b, L, A)$ $\forall t \in \{2008, ..., 2011\}$, we also summarize the distribution of annualized cumulative returns resultant from reinvesting each year 2008 through 2011 and the distribution of annualized cumulative returns resultant from reinvesting each year 2009 through 2011. Explicitly, for any (b, L, A) triple, define

$$\mu_{2008-2011}(b, L, A) = \left[\prod_{t=2008}^{2011} (1 + \mu_t(b, L, A))\right]^{1/4} - 1\,, \tag{5.23}$$

$$\mu_{2009-2011}(b, L, A) = \left[\prod_{t=2009}^{2011} (1 + \mu_t(b, L, A))\right]^{1/3} - 1\,, \tag{5.24}$$

and analogously for $\widetilde{\mu}_{2008-2011}(b, L, A)$, $\widetilde{\mu}_{2009-2011}(b, L, A)$, $r_{\text{S\&P},\,2008-2011}$, $r_{\text{S\&P},\,2009-2011}$, $r_{\text{CD},\,2008-2011}$, and $r_{\text{CD},\,2009-2011}$. In words, an annualized cumulative return is the annual return that, if compounded annually, is equivalent to the cumulative return. For example, $\mu_{2008-2011}$ is the annual return that, if compounded annually, is equivalent to earning an annual return of μ_{2008} for one year, then reinvesting and earning an annual return of μ_{2009} for one year, then reinvesting and earning an annual return of μ_{2010} for one year, and finally reinvesting and earning an annual return of μ_{2011} for one year.

We can also define annualized cumulative percentile returns: the annualized cumulative returns resultant from reinvesting each year 2008 through 2011 and 2009 through 2011, respectively, in the yearly portfolios of the same percentile. Explicitly, for any (b, L, A) triple, define

$$Q^{P\%}_{2008-2011}(b, L, A) = \left[\prod_{t=2008}^{2011} \big(1 + Q^{P\%}_t(b, L, A)\big)\right]^{1/4} - 1\,, \tag{5.25}$$

$$Q^{P\%}_{2009-2011}(b, L, A) = \left[\prod_{t=2009}^{2011} \big(1 + Q^{P\%}_t(b, L, A)\big)\right]^{1/3} - 1\,. \tag{5.26}$$

5.6.2 Distributions of Annualized Cumulative Returns

A more accurate method of calculating the 2008-2011 and 2009-2011 annualized cumulative means and percentiles would be as follows. For each (b, L, A) triple, create the actual distribution of annualized cumulative 2008-2011 returns

$$\begin{aligned} D_{2008-2011}(b, L, A) = \Big\{ a_{2008}\, a_{2009}\, a_{2010}\, a_{2011} \;\Big|\; &\forall\, a_{2008} \in D_{2008}(b, L, A), \\ &a_{2009} \in D_{2009}(b, L, A),\ a_{2010} \in D_{2010}(b, L, A), \\ &a_{2011} \in D_{2011}(b, L, A) \Big\} \end{aligned} \tag{5.27}$$

and the actual distribution of annualized cumulative 2009-2011 returns

$$D_{2009-2011}(b,L,A) = \Big\{ a_{2009}\, a_{2010}\, a_{2011} \;\Big|\; \forall\, a_{2009} \in D_{2009}(b,L,A), \\ a_{2010} \in D_{2010}(b,L,A),\; a_{2011} \in D_{2011}(b,L,A) \Big\}. \tag{5.28}$$

Then, every metric that is calculated for $D_t(b,L,A)\ \forall t \in \{2008,...,2011\}$ could also be calculated for $D_{2008-2011}(b,L,A)$ and $D_{2009-2011}(b,L,A)$.

However, this method cannot be used for reasons of computational feasibility. For example, as displayed in the results that follow, for $b = 0$, $L = 2$, and $A = A_{\max}(b = 0, L = 2) = 0.5$, we get $|D_{2008}| = 2709$, $|D_{2009}| = 448$, $|D_{2010}| = 2390$, and $|D_{2011}| = 109$ which would give $|D_{2008-2011}| = 2709 \times 448 \times 2390 \times 109 = 3 \times 10^{11}$, making $D_{2008-2011}$ prohibitively large computation-wise. Accordingly, we take (5.23), (5.24), (5.25), and (5.26) as approximations to the respective true values that would be obtained from actually constructing $D_{2008-2011}(b,L,A)$ and $D_{2009-2011}(b,L,A)$, and we don't calculate annualized cumulative values for all the metrics that require an actual distribution: $|D|$, $|D|_{\%}$, $|D \geq 0|_{\%}$, σ, and $\widetilde{\sigma}$.

5.6.3 2008 Returns

We include annualized cumulative 2009-2011 returns along with annualized cumulative 2008-2011 returns because 2008 was an unusually bad year for being long the stock market[12] and, as per (2.3), every investment strategy presented here is necessarily long the stock market. Such is evident in the results presented in the next section: for all (b,L,A) triples that were tested, both $\mu_{2008}(b,L,A)$ and $\widetilde{\mu}_{2008}(b,L,A)$ are severely negative (in the range of -15% to -55%). Accordingly, we present the annualized cumulative 2009-2011 data to show how the investment strategies perform in the long-term over a more typical[13] sequence of years.

That is certainly not to say that the 2008 results should be ignored. Stock market crashes do occur (obviously) so it is instructive to see exactly how badly the strategy performs during one. In fact, given that the market crash in 2008 was part of what is considered the worst financial crisis since the Great Depression,[14] the 2008 results are particularly instructive in that they show how our investment strategies perform in a truly-worst-case scenario.

[12] See footnote 1 and footnote 2.
[13] See footnote 1.
[14] See footnote 2.

5.7 Results

A master group of 21 stocks, all of which are components of the S&P 500,[15] was used to generate the results in this section: $\mathcal{G}$ = {Alcoa (AA), Apple (AAPL), American Express (AXP), Boeing (BA), Bank of America (BAC), Caterpillar (CAT), Cisco (CSCO), Chevron (CVX), DuPont (DD), Disney (DIS), General Electric (GE), Home Depot (HD), Hewlett-Packard (HPQ), International Business Machines (IBM), Johnson & Johnson (JNJ), JPMorgan Chase (JPM), Travelers (TRV), United Technologies (UTX), Verizon (VZ), Wal-Mart (WMT), and Exxon Mobil (XOM)}.

As discussed in section 4.2.1, we will not outline a strategic method for choosing $\mathcal{G}$. Accordingly, we encourage an investor who has no predispositions to certain stocks to use the $\mathcal{G}$ that is defined above because, as will be demonstrated in the following sections, it results in historically profitable portfolios.

For each $(b, L, A) \in \mathcal{B} \times \mathcal{L} \times \{A_{\max}(b, L)\}$, we present:

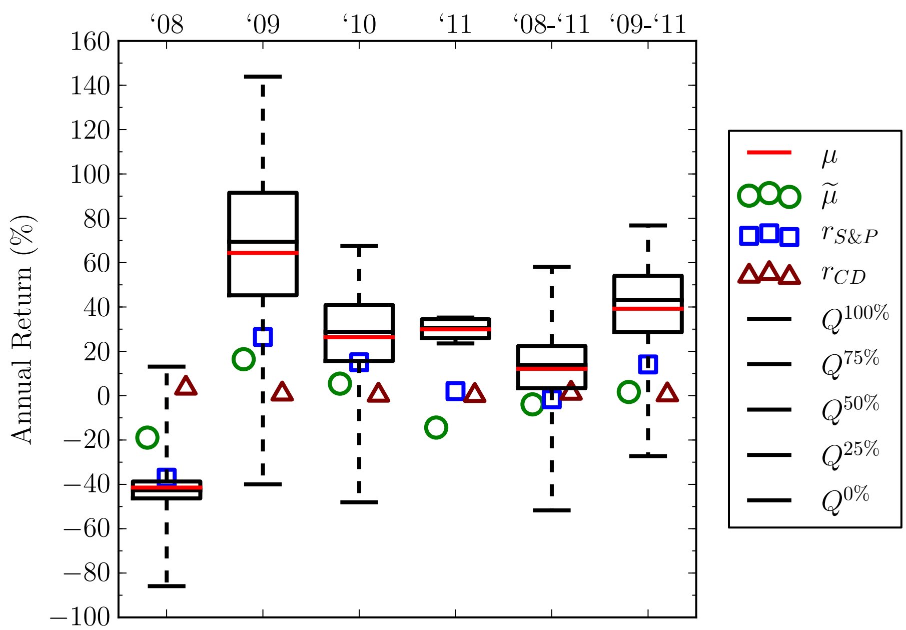
- a table summarizing the annual and annualized cumulative return distributions $D_t(b, L, A)$, $t \in \{2008, ..., 2011, 2008\text{-}2011, 2009\text{-}2011\}$ via the distribution metrics from section 5.3 as well as the benchmark metrics from section 5.4;
- a box and whisker plot presenting the exact same information as the table above with columns for each $t \in \{2008, ..., 2011, 2008\text{-}2011, 2009\text{-}2011\}$ and, in each column,
  - the red line is $\mu_t(b, L, A)$,
  - the green circle is $\widetilde{\mu}_t(b, L, A)$,
  - the blue square is $r_{\text{S\&P},\,t}$,
  - the maroon triangle is $r_{\text{CD},\,t}$,
  - the top black line is $Q_t^{100\%}(b, L, A)$,
  - the top of the black box is $Q_t^{75\%}(b, L, A)$,
  - the middle black line (in the black box) is $Q_t^{50\%}(b, L, A)$,
  - the bottom of the black box is $Q_t^{25\%}(b, L, A)$,
  - the bottom black line is $Q_t^{0\%}(b, L, A)$;
- two plots showing year-by-year cumulative returns starting, respectively, in 2008 (labeled "2008-2011") and 2009 (labeled "2009-2011"), where t is on the horizontal axis and
  - the red curves are $c_{2008,\,t}(\mu)$ and $c_{2009,\,t}(\mu)$,

[15] This fact further justifies using the S&P 500 as a benchmark.

- the green curves are $c_{2008,\,t}(\widetilde{\mu})$ and $c_{2009,\,t}(\widetilde{\mu})$,
- the blue curves are $c_{2008,\,t}(r_{\mathrm{S\&P}})$ and $c_{2009,\,t}(r_{\mathrm{S\&P}})$,
- the maroon curves are $c_{2008,\,t}(r_{\mathrm{CD}})$ and $c_{2009,\,t}(r_{\mathrm{CD}})$.

The tables and plots are as follows. We first present the table and plots for the (b, L, A) triple that resulted in the best-performing investment strategy which was ($b = 0.38$, $L = 2$, $A = 0.9$) as further discussed in section 5.9; we then present tables and plots for the rest of the triples.

	Annual				Annualized Cumulative	
	2008	2009	2010	2011	2008-2011	2009-2011
$\|D\|$	2184	291	1185	4	-	-
$\|D\|_{\%}$	10.73%	1.43%	5.82%	0.02%	-	-
$\|D \geq 0\|_{\%}$	0.14%	92.1%	89.7%	100.0%	-	-
μ	-41.46%	64.37%	26.36%	29.92%	12.11%	39.22%
$\widetilde{\mu}$	-18.92%	16.44%	5.43%	-14.38%	-3.92%	1.68%
$r_{\text{S\&P}}$	-37.0%	26.46%	15.06%	2.05%	-1.65%	14.09%
r_{CD}	3.48%	0.81%	0.3%	0.15%	1.18%	0.42%
σ	9.07%	38.5%	21.4%	4.94%	-	-
$\widetilde{\sigma}$	19.29%	30.33%	15.24%	13.63%	-	-
$Q^{100\%}$	13.11%	143.88%	67.49%	35.23%	58.1%	76.77%
$Q^{75\%}$	-38.71%	92.11%	40.88%	35.23%	22.38%	54.1%
$Q^{50\%}$	-42.69%	69.65%	28.79%	34.21%	13.82%	43.07%
$Q^{25\%}$	-46.34%	45.27%	15.64%	26.69%	3.37%	28.62%
$Q^{0\%}$	-85.91%	-40.0%	-48.08%	23.56%	-51.75%	-27.26%



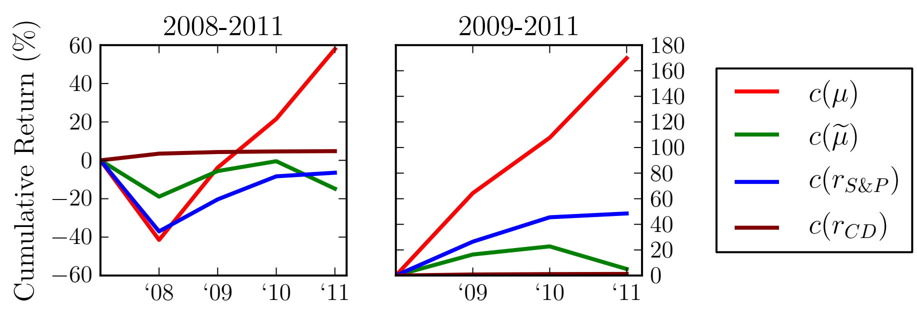
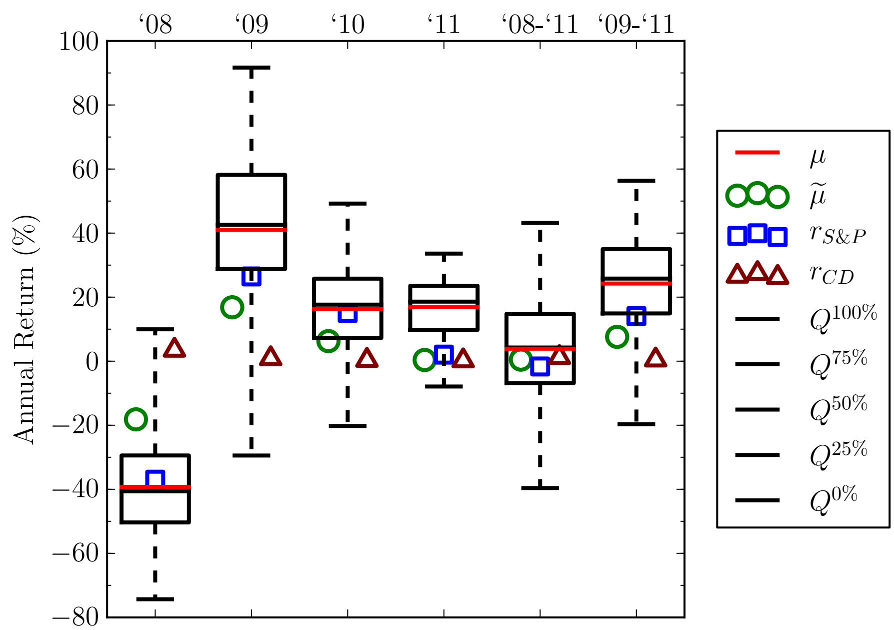


Figure 5.1: (best triple) $b = 0.38$, $L = 2$, $A = A_{\max}(b, L) = 0.9$.
The table summarizes the strategy's return distributions and compares to benchmarks. The middle plot displays the same summaries and benchmark comparisons in box and whisker form. The bottom two plots show cumulative returns, starting in 2008 and starting in 2009, for the strategy and the benchmarks.

	Annual				Annualized Cumulative	
	2008	2009	2010	2011	2008-2011	2009-2011
$\lvert D\rvert$	2709	448	2390	109	-	-
$\lvert D\rvert_{\%}$	13.31%	2.2%	11.75%	0.54%	-	-
$\lvert D \geq 0\rvert_{\%}$	0.41%	92.63%	88.16%	95.41%	-	-
μ	-39.25%	41.04%	16.31%	16.91%	3.9%	24.24%
$\widetilde{\mu}$	-18.16%	16.82%	6.26%	0.37%	0.49%	7.61%
$r_{\text{S\&P}}$	-37.0%	26.46%	15.06%	2.05%	-1.65%	14.09%
r_{CD}	3.48%	0.81%	0.3%	0.15%	1.18%	0.42%
σ	14.74%	23.66%	13.0%	9.34%	-	-
$\widetilde{\sigma}$	19.29%	30.36%	15.34%	13.89%	-	-
$Q^{100\%}$	9.96%	91.67%	49.21%	33.6%	43.17%	56.33%
$Q^{75\%}$	-29.42%	58.33%	25.79%	23.63%	14.79%	35.0%
$Q^{50\%}$	-40.61%	42.64%	17.66%	18.72%	4.27%	25.79%
$Q^{25\%}$	-50.35%	28.86%	7.25%	9.78%	-6.84%	14.91%
$Q^{0\%}$	-74.33%	-29.46%	-20.24%	-7.91%	-39.61%	-19.68%



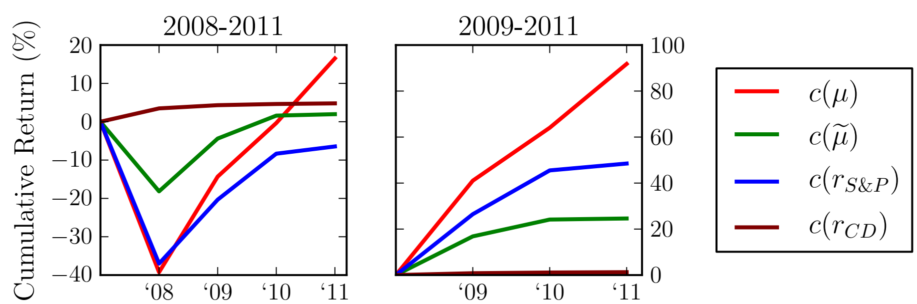


Figure 5.2: $b = 0$, $L = 2$, $A = A_{\max}(b, L) = 0.5$.
The table summarizes the strategy's return distributions and compares to benchmarks. The middle plot displays the same summaries and benchmark comparisons in box and whisker form. The bottom two plots show cumulative returns, starting in 2008 and starting in 2009, for the strategy and the benchmarks.

	Annual				Annualized Cumulative	
	2008	2009	2010	2011	2008-2011	2009-2011
$\|D\|$	3536	669	1463	4	-	-
$\|D\|_{\%}$	17.38%	3.29%	7.19%	0.02%	-	-
$\|D \geq 0\|_{\%}$	6.67%	85.2%	88.17%	100.0%	-	-
μ	-33.18%	31.9%	17.33%	19.03%	5.33%	22.58%
$\widetilde{\mu}$	-18.92%	17.41%	6.39%	-3.94%	-0.68%	6.26%
$r_{\mathrm{S\&P}}$	-37.0%	26.46%	15.06%	2.05%	-1.65%	14.09%
r_{CD}	3.48%	0.81%	0.3%	0.15%	1.18%	0.42%
σ	19.56%	29.56%	13.66%	5.63%	-	-
$\widetilde{\sigma}$	19.39%	30.48%	15.38%	15.56%	-	-
$Q^{100\%}$	24.78%	88.96%	45.66%	26.85%	44.47%	51.71%
$Q^{75\%}$	-19.63%	51.65%	27.17%	26.85%	18.41%	34.74%
$Q^{50\%}$	-35.91%	38.05%	19.07%	19.3%	5.87%	25.15%
$Q^{25\%}$	-48.44%	21.07%	8.05%	19.03%	-5.35%	15.89%
$Q^{0\%}$	-75.61%	-54.1%	-20.24%	10.92%	-43.9%	-25.95%

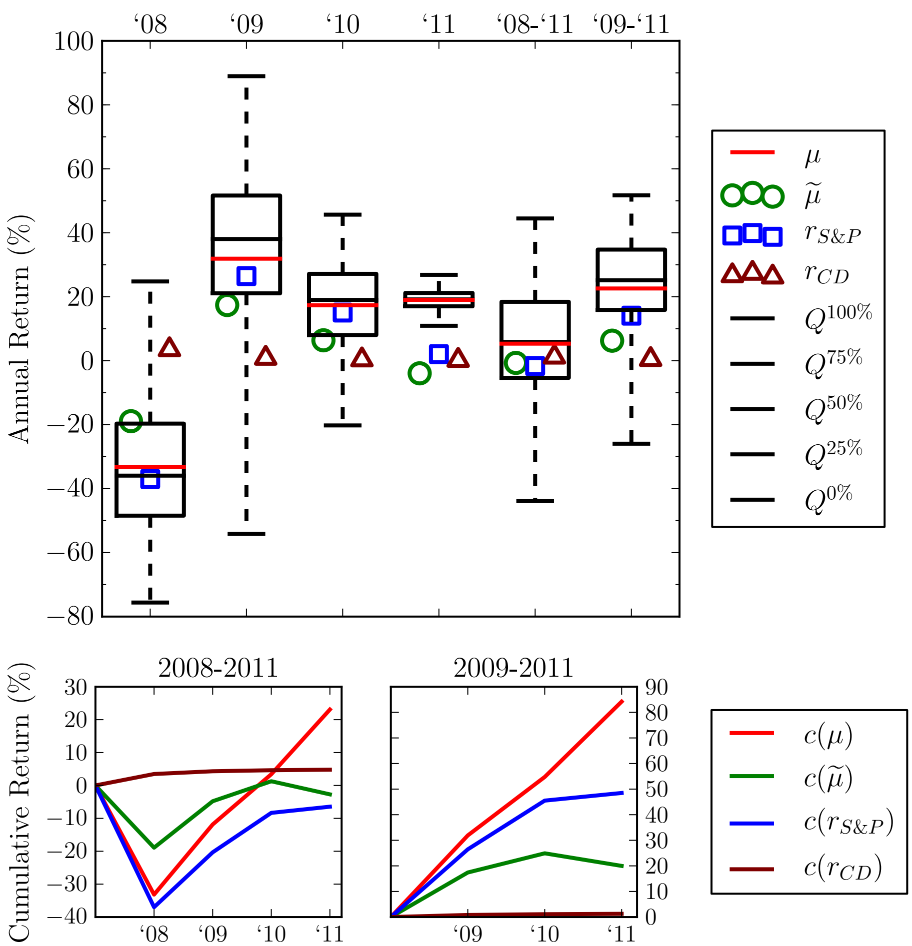


Figure 5.3: $b = 0,\ L = 3,\ A = A_{\max}(b, L) = 0.6$.
The table summarizes the strategy's return distributions and compares to benchmarks. The middle plot displays the same summaries and benchmark comparisons in box and whisker form. The bottom two plots show cumulative returns, starting in 2008 and starting in 2009, for the strategy and the benchmarks.

	Annual				Annualized Cumulative	
	2008	2009	2010	2011	2008-2011	2009-2011
$\lvert D \rvert$	4076	4076	1519	20	-	-
$\lvert D \rvert_{\%}$	20.03%	20.03%	7.46%	0.1%	-	-
$\lvert D \geq 0 \rvert_{\%}$	10.33%	64.67%	88.02%	90.0%	-	-
μ	-30.83%	9.66%	17.33%	18.88%	1.42%	15.22%
$\widetilde{\mu}$	-18.38%	17.58%	6.21%	-4.96%	-0.79%	5.88%
$r_{\text{S\&P}}$	-37.0%	26.46%	15.06%	2.05%	-1.65%	14.09%
r_{CD}	3.48%	0.81%	0.3%	0.15%	1.18%	0.42%
σ	21.2%	32.24%	13.64%	9.98%	-	-
$\widetilde{\sigma}$	19.44%	30.47%	15.2%	15.98%	-	-
$Q^{100\%}$	39.52%	88.96%	46.99%	31.92%	50.37%	54.17%
$Q^{75\%}$	-16.09%	34.96%	27.16%	25.39%	15.92%	29.1%
$Q^{50\%}$	-33.72%	16.59%	19.13%	22.28%	3.0%	19.31%
$Q^{25\%}$	-47.42%	-23.15%	7.94%	15.94%	-15.68%	-1.31%
$Q^{0\%}$	-75.61%	-58.15%	-20.24%	-3.2%	-47.02%	-31.38%

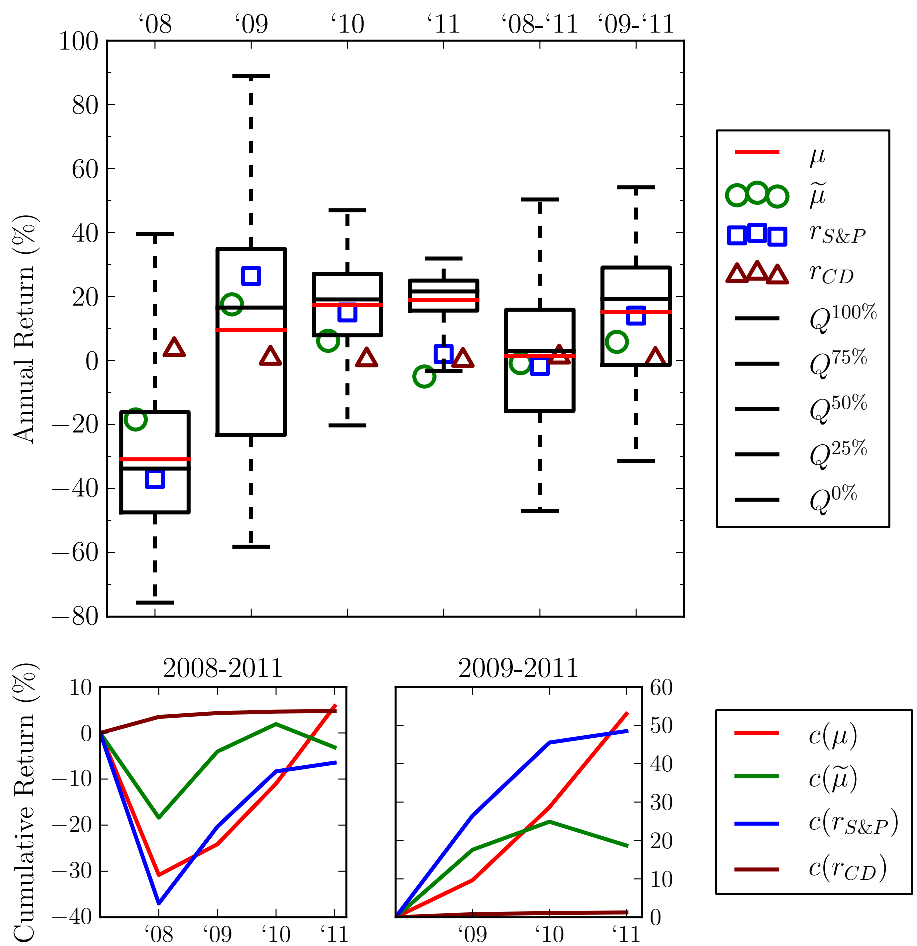


Figure 5.4: $b = 0,\ L = \infty,\ A = A_{\max}(b, L) = 0.6$.

The table summarizes the strategy‘s return distributions and compares to benchmarks. The middle plot displays the same summaries and benchmark comparisons in box and whisker form. The bottom two plots show cumulative returns, starting in 2008 and starting in 2009, for the strategy and the benchmarks.

	Annual				Annualized Cumulative	
	2008	2009	2010	2011	2008-2011	2009-2011
$\lvert D\rvert$	2426	356	1798	33	-	-
$\lvert D\rvert_{\%}$	11.92%	1.75%	8.84%	0.16%	-	-
$\lvert D \geq 0\rvert_{\%}$	0.41%	92.13%	87.1%	100.0%	-	-
μ	-44.6%	46.07%	17.46%	21.69%	3.71%	27.81%
$\widetilde{\mu}$	-18.52%	14.47%	6.15%	-0.87%	-0.47%	6.4%
$r_{\mathrm{S\&P}}$	-37.0%	26.46%	15.06%	2.05%	-1.65%	14.09%
r_{CD}	3.48%	0.81%	0.3%	0.15%	1.18%	0.42%
σ	16.28%	26.48%	14.75%	7.25%	-	-
$\widetilde{\sigma}$	19.7%	29.4%	15.58%	16.99%	-	-
$Q^{100\%}$	10.77%	101.24%	49.34%	37.1%	46.16%	60.31%
$Q^{75\%}$	-33.62%	65.18%	28.17%	26.28%	15.3%	38.61%
$Q^{50\%}$	-46.09%	48.69%	18.91%	22.52%	3.88%	29.27%
$Q^{25\%}$	-56.77%	32.17%	7.14%	18.33%	-7.77%	18.77%
$Q^{0\%}$	-83.21%	-28.89%	-23.24%	2.11%	-44.69%	-17.71%

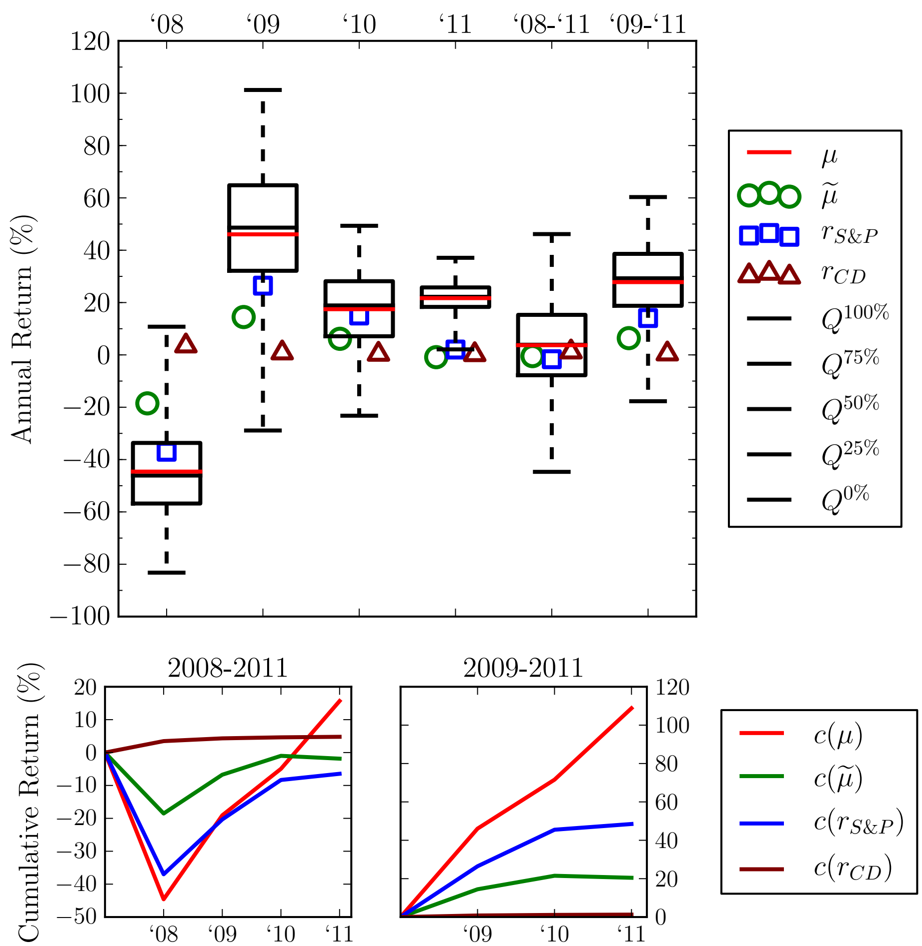


Figure 5.5: $b = 0.1$, $L = 2$, $A = A_{\max}(b, L) = 0.6$.
The table summarizes the strategy‘s return distributions and compares to benchmarks. The middle plot displays the same summaries and benchmark comparisons in box and whisker form. The bottom two plots show cumulative returns, starting in 2008 and starting in 2009, for the strategy and the benchmarks.

	Annual				Annualized Cumulative	
	2008	2009	2010	2011	2008-2011	2009-2011
$\lvert D\rvert$	3906	811	2000	57	-	-
$\lvert D\rvert_{\%}$	19.2%	3.99%	9.83%	0.28%	-	-
$\lvert D \geq 0\rvert_{\%}$	6.37%	84.83%	86.9%	98.25%	-	-
μ	-37.06%	34.61%	17.95%	21.71%	5.02%	24.56%
$\widetilde{\mu}$	-18.56%	16.57%	5.42%	-0.42%	-0.08%	6.96%
$r_{\text{S\&P}}$	-37.0%	26.46%	15.06%	2.05%	-1.65%	14.09%
r_{CD}	3.48%	0.81%	0.3%	0.15%	1.18%	0.42%
σ	21.76%	31.52%	15.12%	8.43%	-	-
$\widetilde{\sigma}$	19.4%	30.14%	15.56%	16.07%	-	-
$Q^{100\%}$	27.16%	98.31%	50.17%	39.88%	51.71%	60.9%
$Q^{75\%}$	-21.52%	54.93%	29.07%	26.55%	18.68%	36.23%
$Q^{50\%}$	-39.79%	40.85%	19.5%	23.54%	5.7%	27.51%
$Q^{25\%}$	-53.94%	22.07%	7.18%	18.08%	-8.17%	15.6%
$Q^{0\%}$	-84.63%	-50.77%	-23.24%	-4.88%	-51.51%	-28.9%

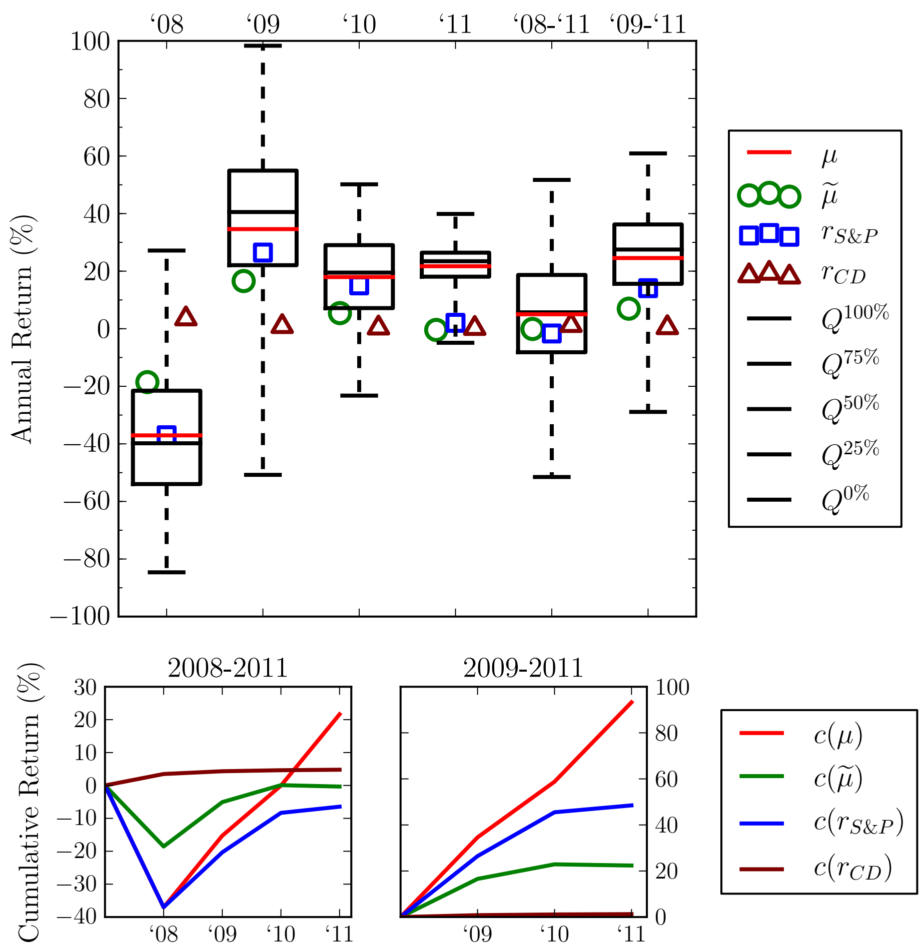


Figure 5.6: $b = 0.1,\ L = 3,\ A = A_{\max}(b, L) = 0.6.$

The table summarizes the strategy‘s return distributions and compares to benchmarks. The middle plot displays the same summaries and benchmark comparisons in box and whisker form. The bottom two plots show cumulative returns, starting in 2008 and starting in 2009, for the strategy and the benchmarks.

	Annual				Annualized Cumulative	
	2008	2009	2010	2011	2008-2011	2009-2011
$\lvert D\rvert$	3774	3774	1130	2	-	-
$\lvert D\rvert_{\%}$	18.55%	18.55%	5.55%	0.01%	-	-
$\lvert D \geq 0\rvert_{\%}$	9.96%	64.52%	87.96%	50.0%	-	-
μ	-35.31%	11.31%	19.52%	10.23%	-1.31%	13.61%
$\widetilde{\mu}$	-18.39%	16.03%	5.6%	8.9%	2.15%	10.09%
$r_{\text{S\&P}}$	-37.0%	26.46%	15.06%	2.05%	-1.65%	14.09%
r_{CD}	3.48%	0.81%	0.3%	0.15%	1.18%	0.42%
σ	23.56%	34.8%	14.96%	12.48%	-	-
$\widetilde{\sigma}$	19.17%	30.6%	15.38%	10.82%	-	-
$Q^{100\%}$	42.48%	98.31%	48.77%	22.71%	50.7%	53.55%
$Q^{75\%}$	-19.19%	38.61%	30.35%	22.71%	15.65%	30.38%
$Q^{50\%}$	-39.19%	18.64%	21.56%	22.71%	1.85%	20.96%
$Q^{25\%}$	-53.56%	-24.97%	9.99%	22.71%	-21.79%	-6.95%
$Q^{0\%}$	-84.63%	-55.97%	-23.24%	-2.26%	-52.53%	-30.87%

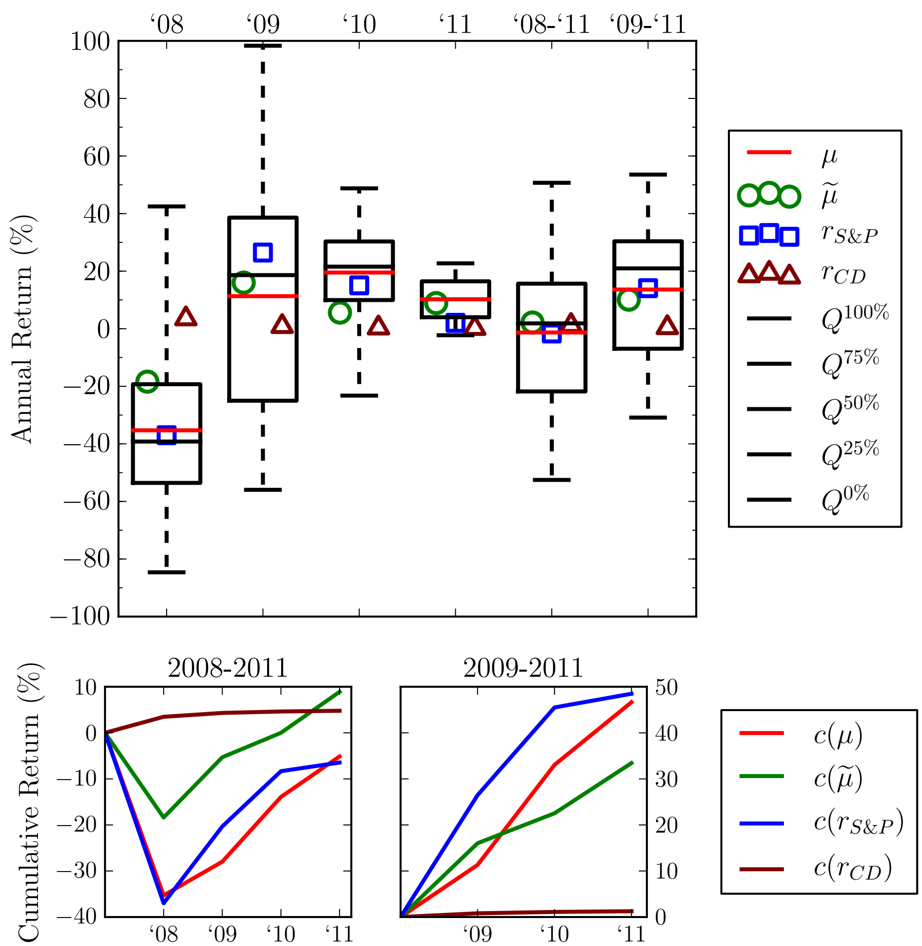
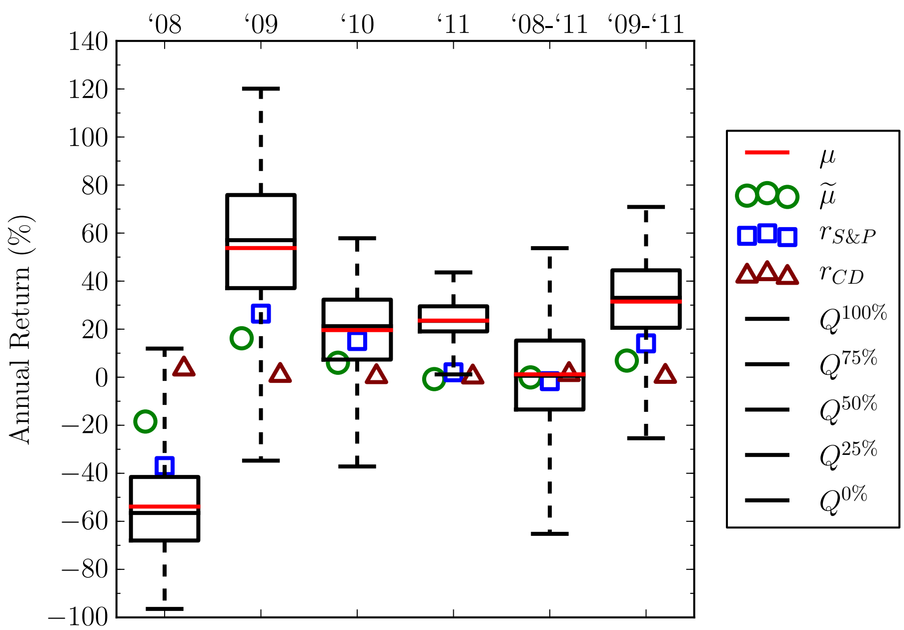


Figure 5.7: $b = 0.1$, $L = \infty$, $A = A_{\max}(b, L) = 0.7$.
The table summarizes the strategy‘s return distributions and compares to benchmarks. The middle plot displays the same summaries and benchmark comparisons in box and whisker form. The bottom two plots show cumulative returns, starting in 2008 and starting in 2009, for the strategy and the benchmarks.

	Annual				Annualized Cumulative	
	2008	2009	2010	2011	2008-2011	2009-2011
$\|D\|$	2456	362	1837	40	-	-
$\|D\|_{\%}$	12.07%	1.78%	9.03%	0.2%	-	-
$\|D \geq 0\|_{\%}$	0.33%	91.71%	86.23%	100.0%	-	-
μ	-53.87%	53.77%	19.61%	23.56%	1.19%	31.47%
$\widetilde{\mu}$	-18.45%	16.21%	6.03%	-0.79%	-0.08%	6.92%
$r_{\text{S\&P}}$	-37.0%	26.46%	15.06%	2.05%	-1.65%	14.09%
r_{CD}	3.48%	0.81%	0.3%	0.15%	1.18%	0.42%
σ	18.73%	31.61%	17.56%	9.22%	-	-
$\widetilde{\sigma}$	19.32%	28.85%	15.5%	13.02%	-	-
$Q^{100\%}$	11.91%	120.12%	57.84%	43.62%	53.73%	70.88%
$Q^{75\%}$	-41.53%	76.36%	32.3%	29.66%	15.25%	44.5%
$Q^{50\%}$	-56.5%	57.06%	21.27%	23.65%	0.6%	33.04%
$Q^{25\%}$	-67.98%	37.14%	7.35%	19.15%	-13.44%	20.58%
$Q^{0\%}$	-96.47%	-34.74%	-37.16%	1.17%	-65.22%	-25.42%



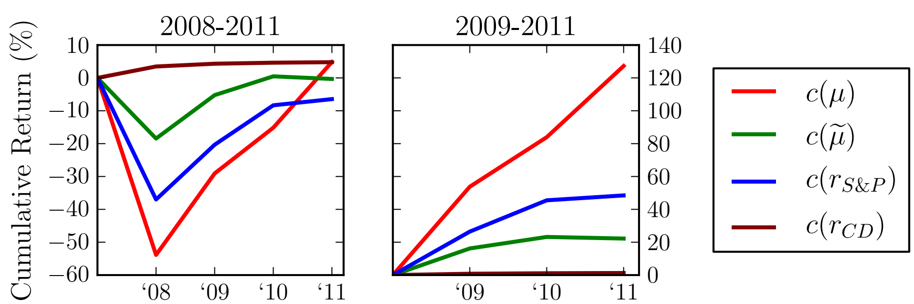


Figure 5.8: $b = 0.25$, $L = 2$, $A = A_{\max}(b, L) = 0.7$.
The table summarizes the strategy‘s return distributions and compares to benchmarks. The middle plot displays the same summaries and benchmark comparisons in box and whisker form. The bottom two plots show cumulative returns, starting in 2008 and starting in 2009, for the strategy and the benchmarks.

	Annual				Annualized Cumulative	
	2008	2009	2010	2011	2008-2011	2009-2011
$\lvert D\rvert$	3937	827	2051	73	-	-
$\lvert D\rvert_{\%}$	19.35%	4.06%	10.08%	0.36%	-	-
$\lvert D \geq 0\rvert_{\%}$	5.69%	85.01%	85.81%	95.89%	-	-
μ	-44.59%	40.73%	20.27%	23.61%	3.77%	27.9%
$\widetilde{\mu}$	-18.93%	15.49%	5.8%	-4.41%	-1.35%	5.31%
$r_{\text{S\&P}}$	-37.0%	26.46%	15.06%	2.05%	-1.65%	14.09%
r_{CD}	3.48%	0.81%	0.3%	0.15%	1.18%	0.42%
σ	24.83%	36.93%	18.05%	12.41%	-	-
$\widetilde{\sigma}$	18.99%	29.22%	15.18%	15.3%	-	-
$Q^{100\%}$	31.44%	116.61%	58.84%	47.02%	60.58%	71.66%
$Q^{75\%}$	-26.93%	64.3%	33.53%	32.04%	20.6%	42.54%
$Q^{50\%}$	-49.14%	47.86%	22.04%	26.92%	3.84%	31.74%
$Q^{25\%}$	-64.37%	25.16%	7.44%	17.44%	-13.4%	16.44%
$Q^{0\%}$	-95.35%	-60.89%	-37.16%	-21.77%	-69.25%	-42.29%

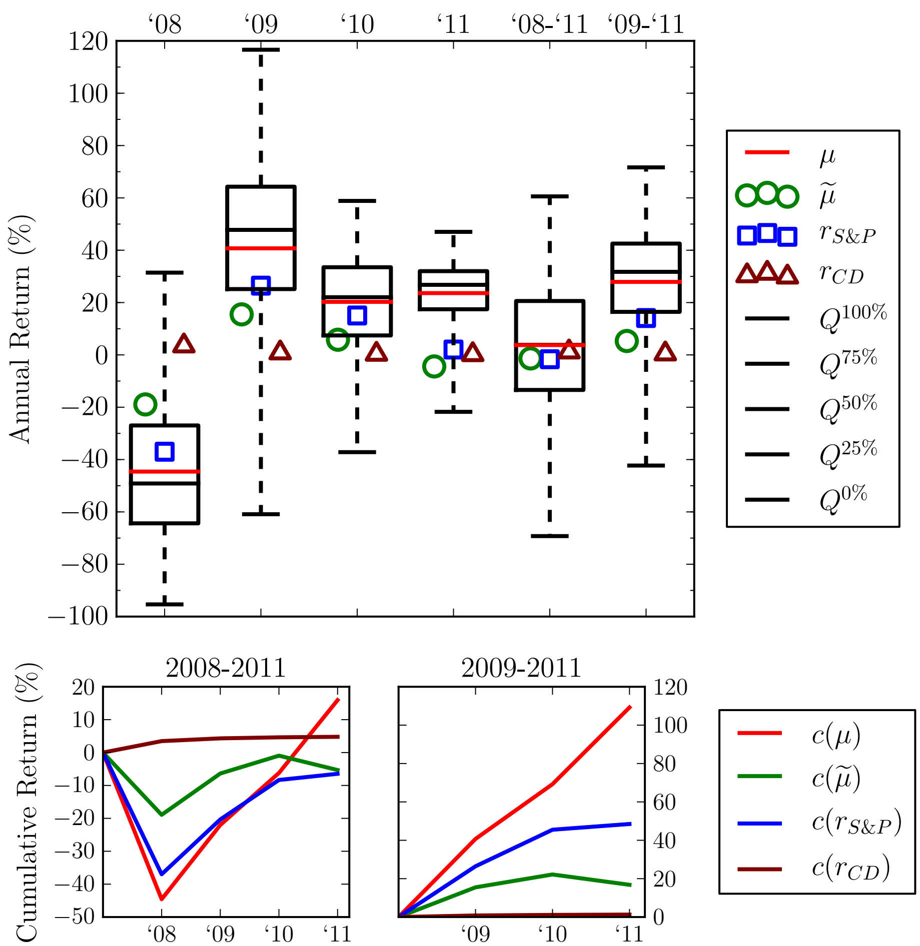


Figure 5.9: $b = 0.25$, $L = 3$, $A = A_{\max}(b, L) = 0.7$.
The table summarizes the strategy‘s return distributions and compares to benchmarks. The middle plot displays the same summaries and benchmark comparisons in box and whisker form. The bottom two plots show cumulative returns, starting in 2008 and starting in 2009, for the strategy and the benchmarks.

	Annual				Annualized Cumulative	
	2008	2009	2010	2011	2008-2011	2009-2011
$\|D\|$	3928	3928	1347	13	-	-
$\|D\|_{\%}$	19.3%	19.3%	6.62%	0.06%	-	-
$\|D \geq 0\|_{\%}$	9.24%	62.53%	86.04%	92.31%	-	-
μ	-42.23%	12.41%	21.14%	27.98%	0.17%	20.34%
$\widetilde{\mu}$	-18.46%	16.73%	5.85%	-0.23%	0.13%	7.22%
$r_{\text{S\&P}}$	-37.0%	26.46%	15.06%	2.05%	-1.65%	14.09%
r_{CD}	3.48%	0.81%	0.3%	0.15%	1.18%	0.42%
σ	27.3%	40.85%	18.24%	10.89%	-	-
$\widetilde{\sigma}$	19.36%	30.38%	15.57%	14.26%	-	-
$Q^{100\%}$	51.27%	116.61%	60.49%	40.83%	64.97%	69.8%
$Q^{75\%}$	-23.58%	44.79%	34.22%	35.98%	19.06%	38.03%
$Q^{50\%}$	-47.9%	20.59%	23.53%	31.59%	0.14%	24.51%
$Q^{25\%}$	-63.97%	-28.48%	8.76%	25.98%	-22.93%	-0.69%
$Q^{0\%}$	-94.75%	-67.12%	-37.16%	-4.07%	-68.06%	-41.69%

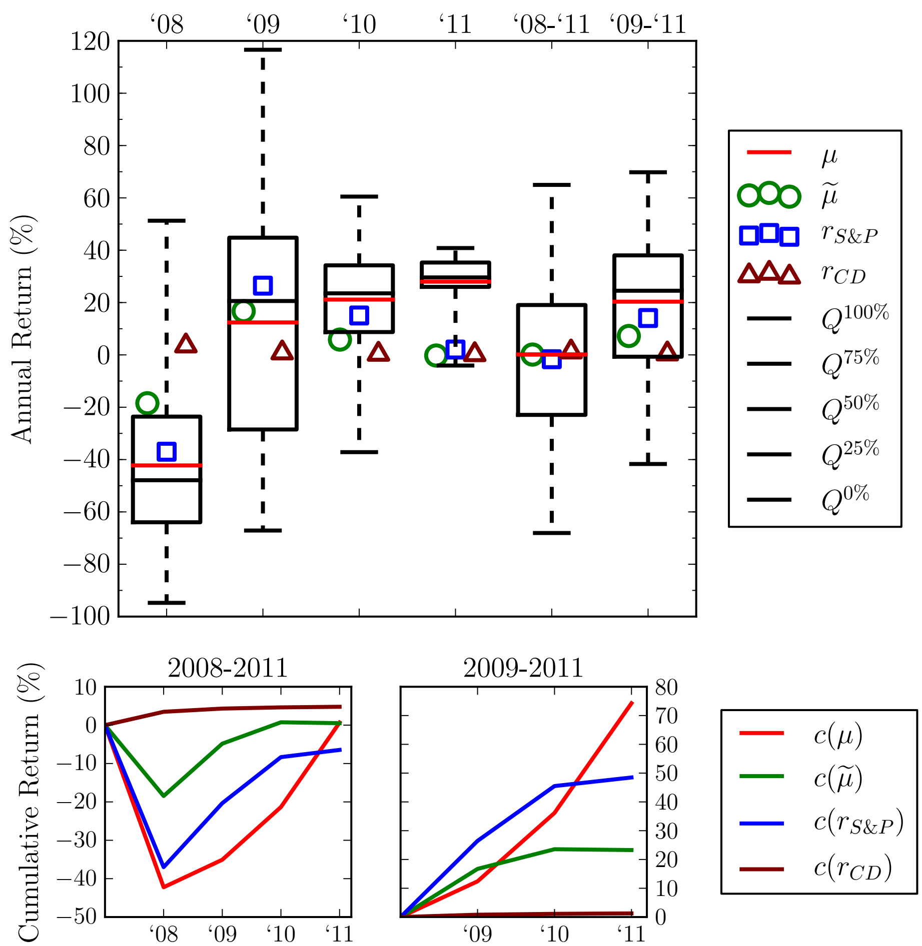


Figure 5.10: $b = 0.25$, $L = \infty$, $A = A_{\max}(b, L) = 0.8$.
The table summarizes the strategy's return distributions and compares to benchmarks. The middle plot displays the same summaries and benchmark comparisons in box and whisker form. The bottom two plots show cumulative returns, starting in 2008 and starting in 2009, for the strategy and the benchmarks.

	Annual				Annualized Cumulative	
	2008	2009	2010	2011	2008-2011	2009-2011
$\|D\|$	3655	714	1347	15	-	-
$\|D\|_{\%}$	17.96%	3.51%	6.62%	0.07%	-	-
$\|D \geq 0\|_{\%}$	4.27%	82.91%	87.9%	100.0%	-	-
μ	-40.04%	47.06%	26.44%	29.5%	9.62%	34.04%
$\widetilde{\mu}$	-18.51%	14.18%	5.67%	-7.57%	-2.37%	3.7%
$r_{\text{S\&P}}$	-37.0%	26.46%	15.06%	2.05%	-1.65%	14.09%
r_{CD}	3.48%	0.81%	0.3%	0.15%	1.18%	0.42%
σ	14.61%	45.98%	22.16%	9.48%	-	-
$\widetilde{\sigma}$	19.14%	29.47%	14.91%	15.36%	-	-
$Q^{100\%}$	36.7%	139.63%	69.58%	44.07%	68.2%	80.23%
$Q^{75\%}$	-39.71%	79.26%	41.9%	35.58%	20.07%	51.07%
$Q^{50\%}$	-43.46%	56.76%	29.6%	29.44%	10.14%	37.56%
$Q^{25\%}$	-47.54%	25.83%	14.5%	26.28%	-1.22%	21.98%
$Q^{0\%}$	-82.54%	-73.69%	-48.08%	9.57%	-59.8%	-46.91%

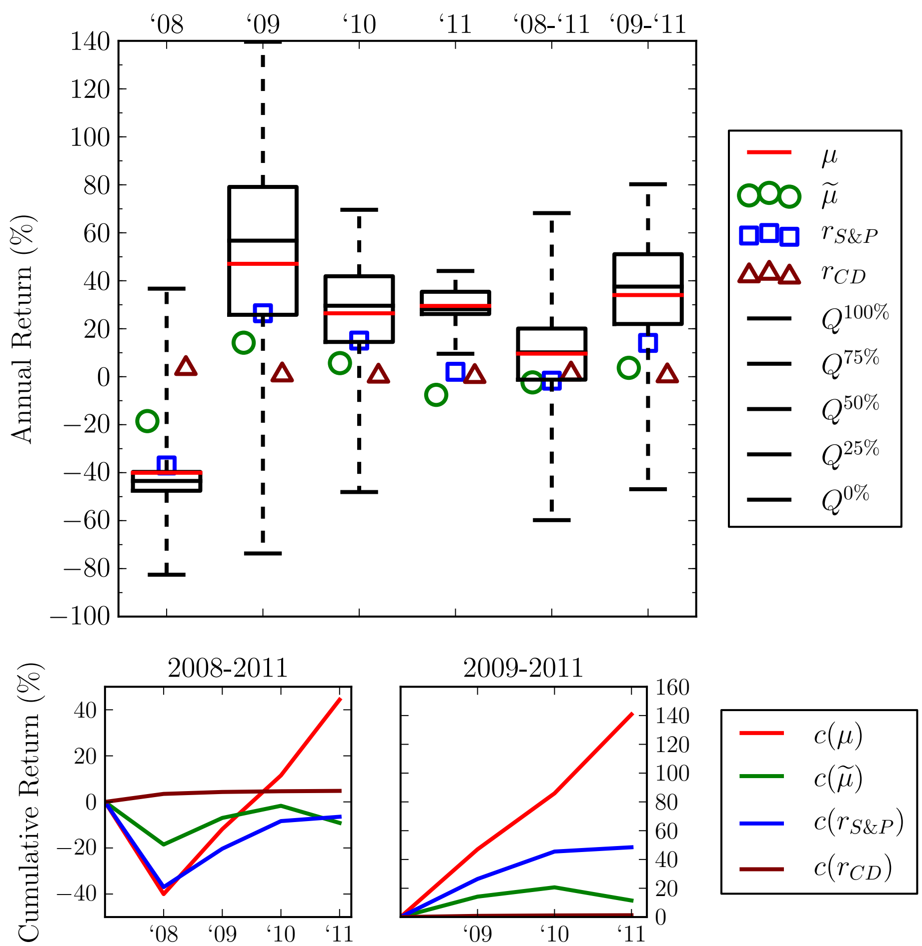


Figure 5.11: $b = 0.38,\ L = 3,\ A = A_{\max}(b, L) = 0.9$.
The table summarizes the strategy‘s return distributions and compares to benchmarks. The middle plot displays the same summaries and benchmark comparisons in box and whisker form. The bottom two plots show cumulative returns, starting in 2008 and starting in 2009, for the strategy and the benchmarks.

	Annual				Annualized Cumulative	
	2008	2009	2010	2011	2008-2011	2009-2011
$\lvert D\rvert$	4204	4204	1413	29	-	-
$\lvert D\rvert_{\%}$	20.66%	20.66%	6.94%	0.14%	-	-
$\lvert D \geq 0\rvert_{\%}$	6.35%	58.33%	86.48%	93.1%	-	-
μ	-38.47%	13.2%	25.95%	25.64%	2.46%	21.45%
$\widetilde{\mu}$	-18.41%	15.79%	6.48%	-0.6%	-0.0%	7.02%
$r_{\text{S\&P}}$	-37.0%	26.46%	15.06%	2.05%	-1.65%	14.09%
r_{CD}	3.48%	0.81%	0.3%	0.15%	1.18%	0.42%
σ	17.49%	47.29%	22.28%	19.16%	-	-
$\widetilde{\sigma}$	19.05%	29.76%	15.4%	11.88%	-	-
$Q^{100\%}$	59.81%	139.63%	71.75%	44.32%	75.53%	81.1%
$Q^{75\%}$	-38.52%	51.6%	41.59%	41.18%	16.28%	43.8%
$Q^{50\%}$	-43.31%	13.46%	29.47%	33.13%	2.25%	24.46%
$Q^{25\%}$	-47.35%	-26.15%	13.58%	20.79%	-14.54%	0.44%
$Q^{0\%}$	-82.54%	-81.22%	-48.08%	-42.03%	-68.48%	-61.62%

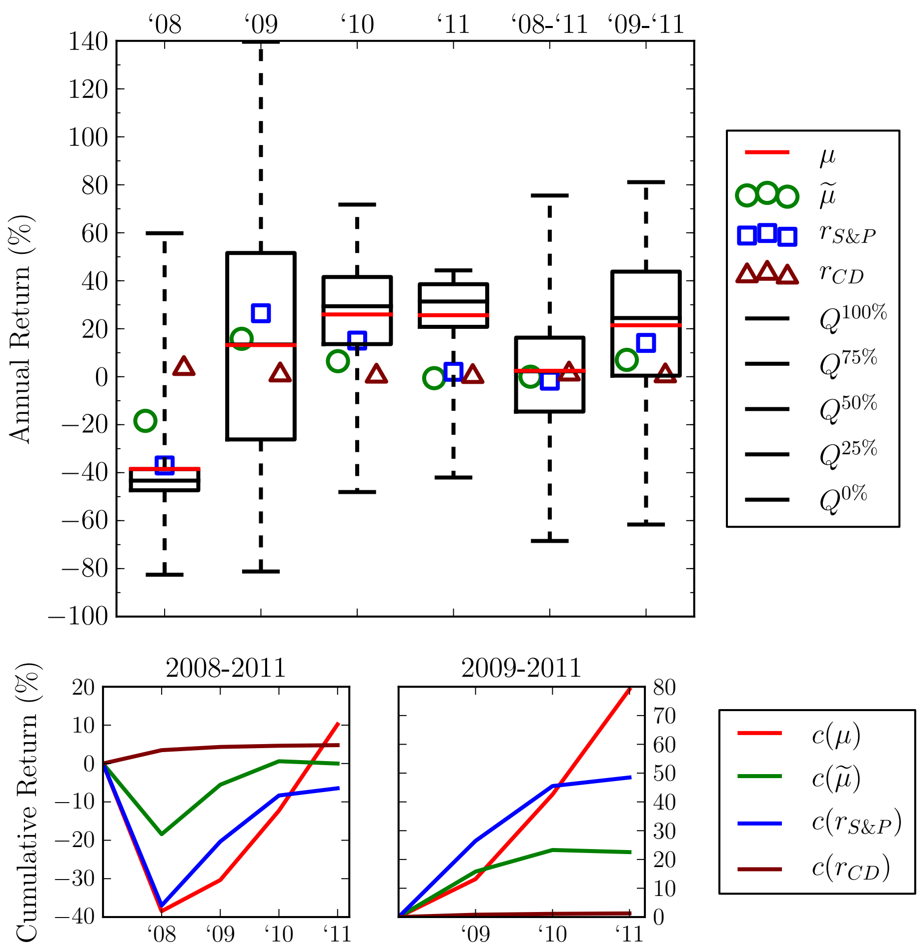


Figure 5.12: $b = 0.38$, $L = \infty$, $A = A_{\max}(b, L) = 0.9$.
The table summarizes the strategy's return distributions and compares to benchmarks. The middle plot displays the same summaries and benchmark comparisons in box and whisker form. The bottom two plots show cumulative returns, starting in 2008 and starting in 2009, for the strategy and the benchmarks.

	Annual				Annualized Cumulative	
	2008	2009	2010	2011	2008-2011	2009-2011
$\|D\|$	1534	494	1082	32	-	-
$\|D\|_{\%}$	7.54%	2.43%	5.32%	0.16%	-	-
$\|D \geq 0\|_{\%}$	0.0%	13.56%	0.55%	87.5%	-	-
μ	-39.92%	-2.22%	-16.95%	4.27%	-15.54%	-5.39%
$\widetilde{\mu}$	-18.53%	17.79%	5.51%	-3.39%	-0.55%	6.29%
$r_{\text{S\&P}}$	-37.0%	26.46%	15.06%	2.05%	-1.65%	14.09%
r_{CD}	3.48%	0.81%	0.3%	0.15%	1.18%	0.42%
σ	10.4%	34.71%	4.78%	3.45%	-	-
$\widetilde{\sigma}$	19.5%	30.07%	14.99%	17.19%	-	-
$Q^{100\%}$	-8.83%	172.97%	49.5%	15.3%	43.92%	67.57%
$Q^{75\%}$	-32.88%	-4.71%	-15.31%	6.58%	-12.86%	-4.93%
$Q^{50\%}$	-40.32%	-10.6%	-17.64%	4.43%	-17.7%	-8.39%
$Q^{25\%}$	-47.88%	-16.24%	-19.28%	2.63%	-22.47%	-11.49%
$Q^{0\%}$	-67.0%	-38.12%	-28.11%	-1.25%	-38.29%	-23.98%

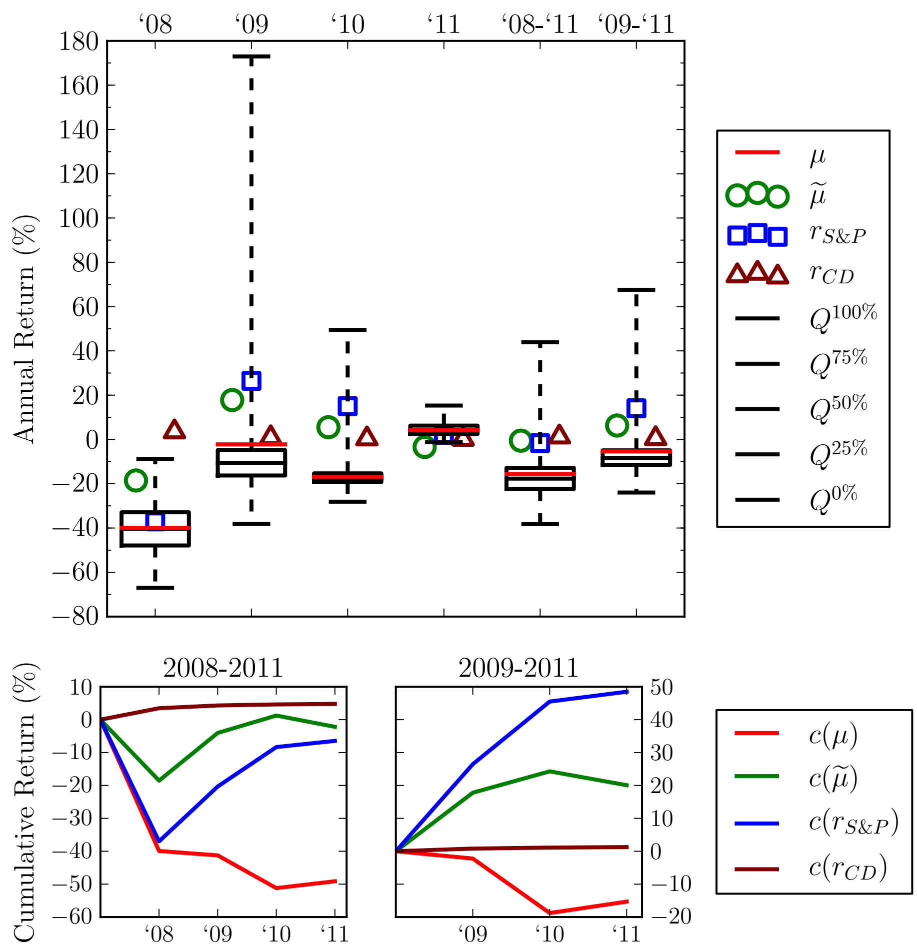


Figure 5.13: $b = 0.5$, $L = 2$, $A = A_{\max}(b, L) = 0.7$.
The table summarizes the strategy's return distributions and compares to benchmarks. The middle plot displays the same summaries and benchmark comparisons in box and whisker form. The bottom two plots show cumulative returns, starting in 2008 and starting in 2009, for the strategy and the benchmarks.

	Annual				Annualized Cumulative	
	2008	2009	2010	2011	2008-2011	2009-2011
$\|D\|$	2261	906	1334	60	-	-
$\|D\|_{\%}$	11.11%	4.45%	6.56%	0.29%	-	-
$\|D \geq 0\|_{\%}$	0.0%	17.11%	0.45%	63.33%	-	-
μ	-43.28%	-0.53%	-17.73%	2.05%	-17.04%	-5.83%
$\widetilde{\mu}$	-18.37%	15.69%	5.78%	-1.67%	-0.45%	6.37%
$r_{\text{S\&P}}$	-37.0%	26.46%	15.06%	2.05%	-1.65%	14.09%
r_{CD}	3.48%	0.81%	0.3%	0.15%	1.18%	0.42%
σ	10.45%	34.61%	4.59%	4.09%	-	-
$\widetilde{\sigma}$	19.42%	30.16%	15.61%	16.74%	-	-
$Q^{100\%}$	-9.74%	171.19%	49.5%	15.3%	43.32%	67.2%
$Q^{75\%}$	-36.9%	-3.73%	-16.13%	4.93%	-14.51%	-5.39%
$Q^{50\%}$	-44.08%	-9.93%	-18.42%	1.81%	-19.59%	-9.22%
$Q^{25\%}$	-51.4%	-15.72%	-19.54%	-1.17%	-24.46%	-12.49%
$Q^{0\%}$	-67.98%	-67.06%	-30.29%	-4.78%	-48.56%	-39.76%

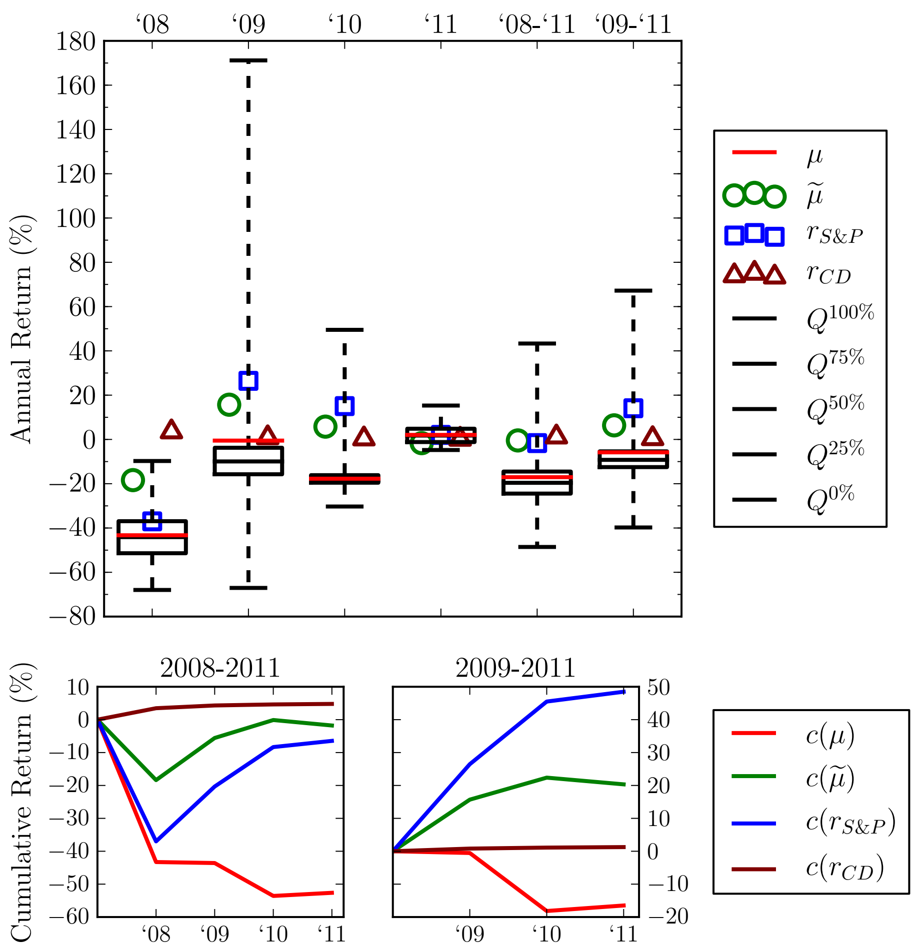


Figure 5.14: $b = 0.5$, $L = 3$, $A = A_{\max}(b, L) = 0.7$.
The table summarizes the strategy's return distributions and compares to benchmarks. The middle plot displays the same summaries and benchmark comparisons in box and whisker form. The bottom two plots show cumulative returns, starting in 2008 and starting in 2009, for the strategy and the benchmarks.

	Annual				Annualized Cumulative	
	2008	2009	2010	2011	2008-2011	2009-2011
$\lvert D\rvert$	2415	2415	1355	5	-	-
$\lvert D\rvert_{\%}$	11.87%	11.87%	6.66%	0.02%	-	-
$\lvert D \geq 0\rvert_{\%}$	0.0%	12.92%	0.3%	20.0%	-	-
μ	-44.15%	-4.88%	-18.05%	-1.53%	-19.08%	-8.44%
$\widetilde{\mu}$	-19.18%	16.42%	5.93%	-3.26%	-0.91%	6.06%
$r_{\text{S\&P}}$	-37.0%	26.46%	15.06%	2.05%	-1.65%	14.09%
r_{CD}	3.48%	0.81%	0.3%	0.15%	1.18%	0.42%
σ	10.43%	27.76%	4.47%	1.35%	-	-
$\widetilde{\sigma}$	19.05%	30.87%	15.53%	9.87%	-	-
$Q^{100\%}$	-6.53%	171.19%	49.5%	0.95%	39.85%	59.95%
$Q^{75\%}$	-37.38%	-5.12%	-16.21%	0.95%	-16.27%	-7.75%
$Q^{50\%}$	-44.39%	-10.26%	-18.49%	-1.25%	-20.55%	-10.52%
$Q^{25\%}$	-51.98%	-16.42%	-19.58%	-2.47%	-25.1%	-13.14%
$Q^{0\%}$	-70.17%	-80.6%	-30.76%	-2.85%	-55.58%	-49.28%

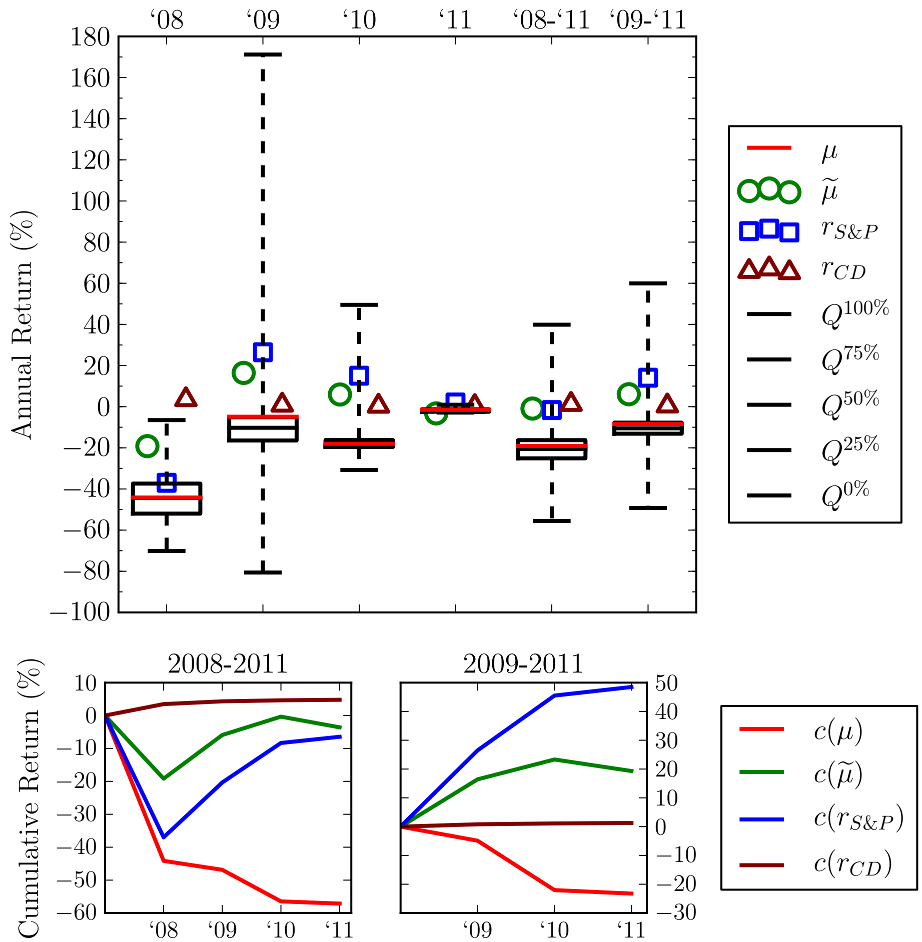


Figure 5.15: $b = 0.5,\ L = \infty,\ A = A_{\max}(b, L) = 0.8.$

The table summarizes the strategy‘s return distributions and compares to benchmarks. The middle plot displays the same summaries and benchmark comparisons in box and whisker form. The bottom two plots show cumulative returns, starting in 2008 and starting in 2009, for the strategy and the benchmarks.

5.8 Results Summary

The following table, as a summary of the results presented in the previous section, displays, for each $(b, L, A) \in \mathcal{B} \times \mathcal{L} \times \{A_{\max}(b, L)\}$, each mean annual and annualized cumulative return: $\mu_t\,(b, L, A)\;\; \forall t \in \{2008, ..., 2011, \text{2008-2011}, \text{2009-2011}\}$. For comparison, the table also displays the benchmark annual and annualized cumulative returns.

Recall that $\widetilde{\mu}_t(b, L, A)$ as defined in (5.14) and displayed in the tables and plots of the previous section was the mean of a random $|D_t(b, L, A)|$-elemented subset of the random returns distribution E_t. As described in section 5.4, this $|D_t(b, L, A)|$-elemented subset was used such that $\widetilde{\sigma}_t(b, L, A)$ could be directly compared to $\sigma_t(b, L, A)$. For the following table this is both impossible because we have a variety of (b, L, A) triples, each yielding a different $|D_t(b, L, A)|$, and unnecessary because we are not displaying any standard deviations. Accordingly, each $\widetilde{\sigma}$ value in the following table is the mean of all $\binom{|\mathcal{G}|}{5}$ elements of the appropriate random returns distribution E_t. Specifically, for and only for the following table, we have

$$\widetilde{\mu}_t = \text{mean}\,(E_t) \;\;\; \forall t \in \{2008, ..., 2011\}\,, \tag{5.29}$$

and then, using these annual returns, the annualized cumulative returns $\widetilde{\mu}_{2008-2011}$ and $\widetilde{\mu}_{2009-2011}$ are defined in the usual way (analogous to (5.23), etc.).

The table is as follows.

b	L	2008	2009	2010	2011	2008-2011	2009-2011
0	2	-39.25%	41.04%	16.31%	16.91%	3.9%	24.24%
	3	-33.18%	31.9%	17.33%	19.03%	5.33%	22.58%
	∞	-30.83%	9.66%	17.33%	18.88%	1.42%	15.22%
0.1	2	-44.6%	46.07%	17.46%	21.69%	3.71%	27.81%
	3	-37.06%	34.61%	17.95%	21.71%	5.02%	24.56%
	∞	-35.31%	11.31%	19.52%	10.23%	-1.31%	13.61%
0.25	2	-53.87%	53.77%	19.61%	23.56%	1.19%	31.47%
	3	-44.59%	40.73%	20.27%	23.61%	3.77%	27.9%
	∞	-42.23%	12.41%	21.14%	27.98%	0.17%	20.34%
0.38	2	-41.46%	64.37%	26.36%	29.92%	12.11%	39.22%
	3	-40.04%	47.06%	26.44%	29.5%	9.62%	34.04%
	∞	-38.47%	13.2%	25.95%	25.64%	2.46%	21.45%
0.5	2	-39.92%	-2.22%	-16.95%	4.27%	-15.54%	-5.39%
	3	-43.28%	-0.53%	-17.73%	2.05%	-17.04%	-5.83%
	∞	-44.15%	-4.88%	-18.05%	-1.53%	-19.08%	-8.44%
$\widetilde{\mu}$		-18.62%	16.53%	6.06%	-2.58%	-0.51%	6.38%
$r_{\text{S\&P}}$		-37%	26.46%	15.06%	2.05%	-1.65%	14.09%
r_{CD}		3.48%	0.81%	0.3%	0.15%	1.18%	0.42%

Table 5.1: *Summary of backtesting results for optimal set of investment strategies.*

5.9 Best (b, L, A)

It's clear that the best (b, L) pair in $\mathcal{B} \times \mathcal{L}$ is $(b = 0.38,\ L = 2)$. As per table 5.1, this pair gives the outright highest mean annual returns in 2009 (by a significant margin) and 2011, is only 0.1% behind the highest mean 2010 return, is decidedly middling as mean 2008 returns go, and gives the outright highest annualized cumulative returns for both 2008-2011 and 2009-2011. Furthermore, as per section 5.7, the standard deviations $\sigma_t(b = 0.38, L = 2, A = A_{\max})$ are roughly similar to the standard deviations for the other (b, L, A) triples. In summary, it wins in 2009, 2011, 2008-2011, and 2009-2011, almost wins in 2010, is roughly average in 2008, and is roughly average risk-wise.

Thus, given that $A_{\max}(b = 0.38,\ L = 2) = 0.9$, the single best parameter-choosing strategy is defined by $(b = 0.38,\ L = 2,\ A = 0.9)$. Combining this single best parameter-choosing strategy with the decision-making support tool gives the single best investment strategy.

Though this $(b = 0.38, L = 2, A = A_{\max})$ investment strategy is the best we were able to devise, it is not necessarily good enough: we need to show that it is better than the benchmarks.

From figure 5.1 we see that, in 2009, 2010, and 2011, respectively, our optimal investment strategy outperforms the random benchmark by 48%, 21%, and 44%, it outperforms the S%P 500 by 38%, 11%, and 28%, and it outperforms 1-year CDs by 64%, 26%, and 30%. These are enormous margins and in the case of the random benchmark and the S%P 500 they are significant: in 2009 through 2011, our investment strategy, taking into account all fees and regulatory considerations, significantly outperforms both the stock market as a whole and randomly choosing portfolios of the same stocks that our strategy uses. In dollar terms, assuming a $5000 initial investment as per section 3.3, compared to the S&P 500 the investor would have made an additional $1896, $566, and $1394, respectively, in 2009, 2010, and 2011 using our strategy. In annualized cumulative terms over the period 2009 through 2011, our strategy beats the random benchmark by 38%, the S&P 500 by 25%, and 1-year CDs by 39%. In cumulative terms, if an investor had reinvested in our strategy each year 2009 through 2011, he would have earned a total of $6067 more than if he had reinvested in the S&P 500 each year 2009 through 2011.

In 2008, on the other hand, our strategy is beat by all three benchmarks: the random benchmark performs 23% better, the S&P 500 performs 4% better, and a 1-year CD performs 45% better. While being beat by the random benchmark by 23% is upsetting, the fact that our investment strategy performed only 4% worse than the market as a whole during what was, as previously discussed, perhaps the worst financial crisis since the great depression, is somewhat comforting: even when the market as a whole crashes in a completely anomalous fashion, our portfolio doesn't perform significantly worse. The average investor in the stock market would have done roughly equally poorly through the crisis as an investor using our strategy would have.

Furthermore, when 2008 is taken into account cumulatively, the results are good. Our investment strategy gives a 2008 through 2011 annualized cumulative return that beats the random benchmark by 16%, the S%P 500 by 14%, and 1-year CDs by 11%. So, even if an anomalous stock market crash occurs once every four years, our strategy still performs well.

Finally, the risk in our investment strategy is roughly similar to the risk in the random benchmark strategy: in 2008 and 2011 our strategy was less risky, as measured by standard deviation, by 10% of return and 9% of return, respectively, and in 2009 and 2010 our strategy was more risky, as measured by standard deviation, by 8% of return and 6% of return, respectively.

Overall, our investment strategy outperforms the benchmarks in every metric except for 2008 returns, in which our strategy is outright bested by the benchmarks, and risk, in which our strategy and the benchmarks are roughly equivalent.

6 Conclusions

We successfully accomplished the main goal: devise a profitable and capital non-intensive investment strategy that is practical for an unsophisticated investor to implement. In summary, using the formal mathematical framework of the appropriate pre-Hilbert space we proved in a method, that to the best of our knowledge is original, the well-known financial result that given a fixed set of assets, known expectations, variances, and pairwise covariances of the returns thereof, and a fixed overall return, there always exists a unique portfolio of these assets that is expected to achieve this overall return with minimum variance. Then, we created an investment decision-making support tool by combining the ability to create such fixed-return minimum-variance portfolios with a general investment procedure and an accurate method of calculating real-world returns that accounts for all fees and investment regulations. This tool, upon specifying values for a set of parameters, produces a comprehensively defined investment strategy. Subsequently, we devised several different parameter-choosing strategies, all of which both learn from the past to draw inferences about the future and learn from the past in order to determine exactly how much of the past to learn from. We combined the decision-making support tool with these parameter-choosing strategies to create several different investment strategies that all dynamically learn from historical data. Finally, we backtested these investment strategies to see how well they would have performed in past years and then inspected the results to choose a single best investment strategy. This best investment strategy is capital non-intensive and practical for an unsophisticated investor to implement by design, as well as profitable compared to benchmarks to the tune of outperforming the S&P 500 by up to 38% in average annual returns.

This best investment strategy is practical for an unsophisticated investor to implement because it only requires access to an online brokerage, investing in common stocks, and trading once per year. It is relatively capital non-intensive because it is optimally used with $5000 of initial capital and it can be used, albeit to the slight detriment of the resultant returns, with only $2000 of initial capital. $2000 to $5000 is not a trivial amount of money, but it is certainly a reasonable amount to expect our unsophisticated investor to have and it is significantly less initial capital than is required for various professionally managed alternatives.

Furthermore, we devised an investment decision-making support tool that is a general and powerful platform with the potential to generate investment strategies that are even more profitable than the best investment strategy created here; one must simply devise a parameter-choosing strategy that is better than the optimal ones presented here and combine it with the decision-making tool to obtain a superior investment strategy that is still capital non-intensive and practical for an unsophisticated investor to implement.

But we need not limit ourselves to helping the unsophisticated investor who has only a few thousand dollars to invest. The decision-making support tool can be easily modified to generate capital-intensive strategies for sophisticated investors who are comfortable investing in all asset classes and trading far more frequently than once per year. In fact, for two reasons we expect that such investment strategies have the potential to perform significantly better than even the best unsophisticated capital non-intensive strategy devised here: first, more frequent trading means that the strategy is more rapidly updating its predictions about the future by learning from the newest historical data and second, more asset classes means lower-variance minimum-variance portfolios. Future work would include the development and testing of such potentially enormously profitable investment strategies.

Sources

[1] Luenberger, David G. Optimization by Vector Space Methods. New York: Wiley, 1968. Print.

[2] Merton, Robert C. “An Analytic Derivation of the Efficient Portfolio Frontier.” The Journal of Financial and Quantitative Analysis 7.3 (1972): 1851-1872. JSTOR. Web. 24 Jan. 2012.

[3] ”Transaction Pricing.” E-Trade. Web. 24 Jan. 2012. <https://us.etrade.com/e/t/prospectestation/pricing?id=1206010000#TP>.

[4] ”Pricing.” TD Ameritrade. Web. 24 Jan. 2012. <https://www.tdameritrade.com/pricing.html>.

[5] ”Transaction Prices.” Scottrade. Web. 24 Jan. 2012. <http://www.scottrade.com/online-brokerage/trading-fees-commissions.html>.

[6] ”Trade on Margin.” E-Trade. Web. 24 Jan. 2012. <https://us.etrade.com/e/t/estation/pricing?id=1302000000>.

[7] ”View and Understand Margin Maintenance Requirements.” E-Trade. Web. 24 Jan. 2012. <https://us.etrade.com/e/t/estation/help?id=1302010000>.

[8] ”S&P 500” Wikipedia. Wikimedia Foundation. Web. 24 Jan. 2012. <http://en.wikipedia.org/wiki/Late-2000s_financial_crisis>.

[9] ”Late-2000s Financial Crisis.” Wikipedia. Wikimedia Foundation. Web. 24 Jan. 2012. <http://en.wikipedia.org/wiki/S\%26P_500>.

[10] Global Financial Data. Web. 24 Jan. 2012. <https://www.globalfinancialdata.com>.