ABSTRACT

Title of Dissertation: EMPIRICAL ANALYSES ON FEDERAL THRIFT SAVINGS PLAN PORTFOLIO OPTIMIZATION

Scott T. Nestler, Ph.D., 2007

Co-Directed By: Professor Michael Fu
Department of Decision and Information Technologies
Robert H. Smith School of Business

Professor Dilip Madan
Department of Finance
Robert H. Smith School of Business

There is ample historical data to suggest that log returns of stocks and indices are not independent and identically distributed Normally, as is commonly assumed. Instead, the log returns of financial assets are skewed and have higher kurtosis. To account for skewness and excess kurtosis, it is necessary to have a distribution that is more flexible than the Gaussian distribution and accounts for additional information that may be present in higher moments.

The federal government's Thrift Savings Plan (TSP) is the largest defined contribution retirement savings and investment plan in the world, with nearly 3.6 million participants and over $173 billion in assets. The TSP offers five assets

(government bond fund, fixed income fund, large-cap stock fund, small-cap stock fund, and international stock fund) to U.S. government civilian employees and uniformed service members. The limited choice of investments, in comparison to most 401(k) plans, may be disappointing from a participant's perspective; however, it provides an attractive framework for empirical study.

In this study, we investigate how the optimal choice of TSP assets changes when traditional portfolio optimization methods are replaced with newer techniques. Specifically, the following research questions are posed and answered:

(1) Does use of a non-Gaussian factor model for returns, generated with independent components analysis (ICA) and following the Variance Gamma (VG) process, provide any advantage over conventional methods with returns assumed to be Normally-distributed, in constructing optimal TSP portfolios?

(2) Can excess TSP portfolio returns be generated through rebalancing to an optimal mix? If so, does changing the frequency of rebalancing from annual to monthly or even daily provide any further benefit that offsets the increased computational complexity and administrative burden?

(3) How does the use of coherent measures of risk, with corresponding portfolio performance measures, in place of variance (or standard deviation) as the traditional the measure for risk and Sharpe Ratio as the usual portfolio performance measure affect the selection of optimal TSP portfolios?

We show through simulation that some of the newer schemes should produce excess returns over conventional (mean-variance optimization with Normally-distributed returns) portfolio choice models. The use of some or all of these methods could benefit the nearly 4 million TSP participants in achieving their retirement savings and investment objectives. Furthermore, we propose two new portfolio performance measures based on recent developments in coherent measures of risk.

EMPIRICAL ANALYSES ON FEDERAL THRIFT SAVINGS PLAN PORTFOLIO OPTIMIZATION

by

Scott T. Nestler

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
PhD in Management Science
2007

Advisory Committee:

Professor Michael Fu, Co-Chair
Professor Dilip Madan, Co-Chair
Professor George Quester
Dr. Paul Thornton
Professor Russ Wermers

Acknowledgements

I am grateful to my advisors, Michael Fu and Dilip Madan, for all their guidance and support over the past several years. The Mathematical Finance Research Interaction Team that they jointly run at Maryland is a hidden gem. I thank my committee members Paul Thornton and Russ Wermers for helpful comments and suggestions, and George Quester for being willing to serve as the Dean's Representative, an important but under-recognized service to academia. My office mates Stacy Cuffe, Andy Hall, and William Mennell, helped to push me along when I was in need of such encouragement. I must acknowledge the Department of Mathematical Sciences, United States Military Academy at West Point, for having confidence in me and providing me with the necessary support to be full-time graduate student again. I also thank Jerry Hoberg, Mark Loewenstein, and Paul Zantek for their advice at critical moments during my research. Any remaining errors are my own. I thank my wife, Kristin, for creating a warm, supportive environment where it was possible for me to devote the effort required for this undertaking. Finally, I ask Anna and Sophia to forgive me for the time and energy I spent on this effort that would otherwise have been theirs.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Overview

There is ample historical data to suggest that log returns of stocks and indices are not independent and identically distributed Normal, as is commonly assumed. (Cootner 1964; Fama 1965; Lo and Mackinlay 1988; Mandelbrot 1963; Mitchell 1915; Osborne 1959; Praetz 1972) Instead, the log returns of financial assets are skewed and have higher kurtosis. To account for skewness and excess kurtosis, it is necessary to have a distribution that is more flexible than the Gaussian distribution and accounts for additional information that may be present in higher moments. (Harvey *et al.* 2004)

The federal government's Thrift Savings Plan (TSP) is the largest defined contribution retirement savings and investment plan in the world, with nearly 3.6 million participants and over $173 billion in assets as of Dec 31, 2005. (Federal Thrift Retirement Investment Board 2006) The TSP offers five assets (a government bond fund, a fixed income fund, a large-cap stock fund, a small-cap stock fund, and an international stock fund) to U.S. government civilian employees and uniformed[1] service members. The limited choice of investments, in comparison to most 401(k)

[1] Uniformed, not uninformed. Uniformed services include the five military services (Army, Navy, Air Force, Marine Corps, and Coast Guard) plus the National Oceanic and Atmospheric Administration Commissioned Corps and Public Health Service Commissioned Corps.

plans, may be disappointing from an investor's perspective; however, it provides an attractive framework for empirical study in an academic setting.

In this study, we investigate how the optimal choice of TSP assets changes when traditional portfolio optimization methods are replaced with newer techniques. Specifically, the following research questions are posed and answered:

(1) Does use of a non-Gaussian factor model for returns, generated with independent components analysis (ICA) and following the Variance Gamma (VG) process, provide any advantage over conventional methods with returns assumed to be Normally-distributed, in constructing optimal TSP portfolios?

(2) Can excess TSP portfolio returns be generated through rebalancing to an optimal mix? If so, does changing the frequency of rebalancing from annual to monthly or even daily provide any further benefit that offsets the increased computational complexity and administrative burden?

(3) How does the use of coherent measures of risk, with corresponding portfolio performance measures, in place of variance (or standard deviation) as the traditional measure for risk and Sharpe Ratio as the usual portfolio performance measure affect the selection of optimal TSP portfolios?

We show through simulation that some of the newer schemes should produce excess returns over conventional (mean-variance optimization with Normally-distributed returns) portfolio choice models. The use of some or all of these methods could benefit the nearly 4 million TSP participants in achieving their retirement savings and investment objectives. Furthermore, we propose two new portfolio performance measures based on recent developments in coherent measures of risk.

1.2 Thrift Savings Plan

1.2.1 General

The TSP was authorized by the United States Congress in the Federal employees' Retirement System Act of 1986 and is administered by the Federal Retirement Thrift Investment Board (FRTIB), an independent Government agency. It was first offered to civilian employees in 1988 as an integral part of the Federal Employee Retirement System (FERS). Later, the plan was opened to members of the uniformed services and civilians covered by the older Civil Service Retirement System (CSRS) that FERS replaced. The purpose of the TSP is to provide Federal employees a tax advantaged savings and investment plan similar to those provided by many private companies with 401(k) plans. The TSP is primarily a defined contribution plan funded by voluntary contributions by the participant but can also include matching contributions from the government. This is a significant difference from defined benefit plans like military retirement, CSRS, or the basic benefit portion of FERS, where contributions are borne entirely by the government. The TSP is tax advantaged in two ways— employee contributions are made from pre-tax pay and taxes on contributions and earnings are deferred until withdrawal. In addition to the tax benefits, because there is no minimum vesting period for most TSP assets, they are portable if a service member or employee leaves government service before reaching retirement eligibility.

1.2.2 Core Investment Funds.

TSP participants are offered five investment funds: the Government Securities Investment Fund (G Fund), the Fixed Income Investment Fund (F Fund), the Common Stock Index Investment Fund (C Fund), the Small Capitalization Stock Index Investment Fund (S Fund), and the International Stock Index Investment Fund (I Fund). The G, F, and C funds formed the totality of investment options when the TSP began in 1988 while the S and I funds were added in May 2001.

The G Fund invests in short-term non-marketable U.S. Treasury securities that are issued only to the TSP. By law, the G Fund earns an interest rate that is equal to the average market rates of return on outstanding U.S. Treasury marketable securities with four or more years to maturity. (5 USC §8438 (a)) The implication of this is that, unless the yield curve is inverted, the G Fund is a riskless asset with an above-market rate of return. (Redding 2007) However, although it is riskless in one respect (guaranteed to not have a negative return), it is not riskless with regards to having no variance, the usual definition of "riskless" in portfolio theory.

Fund	Description	Invests In or Tracks	Assets
G	Government	Short-term U.S. Treasury Securities	$66.6B (39.2%)
F	Fixed-Income	Lehman Brothers U.S. Aggregate Index	$10.2B (6.0%)
C	Common Stock	Standard & Poor's 500 Stock Index	$66.7B (39.3%)
S	Small Cap Stock	Dow Jones Wilshire 4500 Completion Index	$13.7B (8.1%)
I	International Stock	Morgan Stanley Capital International (MSCI) Europe, Asia, Far East (EAFE) Index	$12.6B (7.4%)

Table 1. Core TSP Funds.

The F, C, S, and I funds are managed by Barclays Global Investors. Each of these four funds tracks a popular index, as shown in Table 1. (Federal Thrift Retirement Investment Board 2006) As of Dec. 31, 1985, the F Fund approximately consisted of the following asset mix: 39% mortgage-backed securities (primarily guaranteed by Government National Mortgage Association, Fannie Mae, and Freddie Mac), 23% investment-grade corporate securities (both U.S. and non-U.S.), 25% Treasury securities, 11% Federal agency securities, and 2% asset-backed securities and taxable municipals. The C, S and I funds are passive in nature; each attempts to track its respective index by holding the same weights of stocks of the companies represented in the index. In order to cover withdrawals and loans from the TSP, each of these funds maintains a liquidity reserve that is invested in instruments similar to G Fund holdings. The historical correlation between each fund and its tracked index exceeds 0.99, indicating that the fund managers are performing well against their benchmarks.

Note the percentages shown in the ‘Assets’ column of Table 1. These represent what participants are doing in aggregate; we will call this mix (39% G fund, 6% F fund, 39% C fund, 8% S fund, and 8% I fund) the TSP “Market Portfolio” (TSP MP). Although no individual investor may actually be investing their funds in this manner, this is the overall distribution of monies in the TSP funds.

Employees can make payroll contributions to any of the TSP investment funds in whole percentage increments. Also, they can make daily inter-fund transfers to redistribute existing account balances at no direct costs. Any rebalancing costs are covered internally by the fund manager and spread across all investors in the fund as lowered returns. An investor who rebalances daily incurs no greater cost than an investor who makes no trades. This is one instance where the actuality of real-world investing actually mirrors the idealistic “no transaction costs” assumption of academia. This will become important for some of our analysis concerning dynamic, multi-period portfolio models with rebalancing at various intervals.

1.2.3 Lifecycle Funds

In 2005, the TSP added five additional investment options, called the Lifecycle (or L) Funds. The intent of the L Funds is to provide the highest possible rate of return for the risk taken, given an individual’s retirement time horizon. The L fund managers use mean-variance optimization to allocate assets to the five core TSP funds described above by seeking the maximum expected returns for a specified level of risk. An individual’s retirement time horizon is also considered, as each L fund corresponds to a date range of ten years. Investors are encouraged to pick the L fund that corresponds to their planned retirement date, as shown in Table 2.

Choose	If your time horizon is:
L 2040	2035 or later
L 2030	2025 through 2034
L 2020	2015 through 2024
L 2010	2008 through 2014
L Income	Sooner than 2008

Table 2. L Fund Selection Criteria.

Figure 1 shows the initial allocation of the L Funds to the five core TSP funds. (Federal Thrift Retirement Investment Board 2005) The L fund allocations to the five core funds adjust quarterly. Notice how the proportion invested in the G fund increases as the retirement horizon nears (moving from left to right), while allocations to the more risky (but potentially more rewarding) F, C, S, and I funds decrease as time passes to reflect the change in investment objective from growth to preservation of assets. Once the target retirement date is reached the percentages do not change since rebalancing among the asset classes (funds) no longer occurs. A new dated L fund will be introduced every 10 years as an existing, dated L fund becomes the L Income fund. For example, the L2050 fund should be introduced in 2015 for investors with a retirement time horizon of 2045 or later, once the L2010 allocations match the L Income percentages.

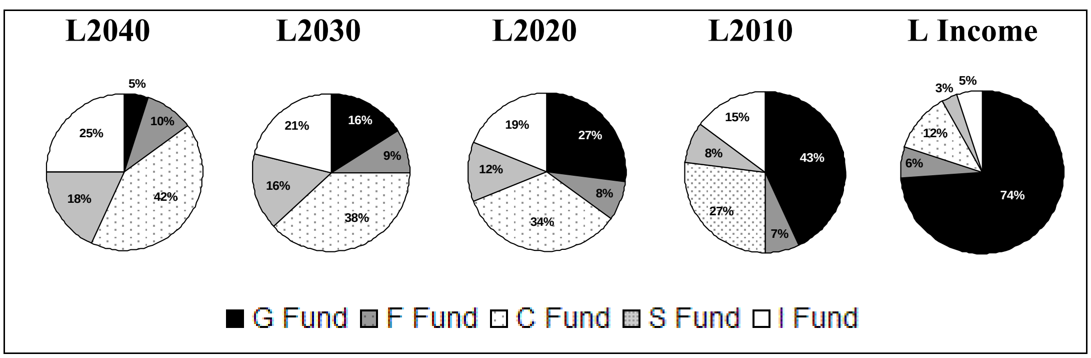

Figure 1. L Fund Allocations to Core TSP Funds.

Although there are only five actual L Funds, they can be seen as a continuum of portfolios, because of the quarterly rebalancing described above. Using linear interpolation between the five portfolios results in the view given in Figure 2.

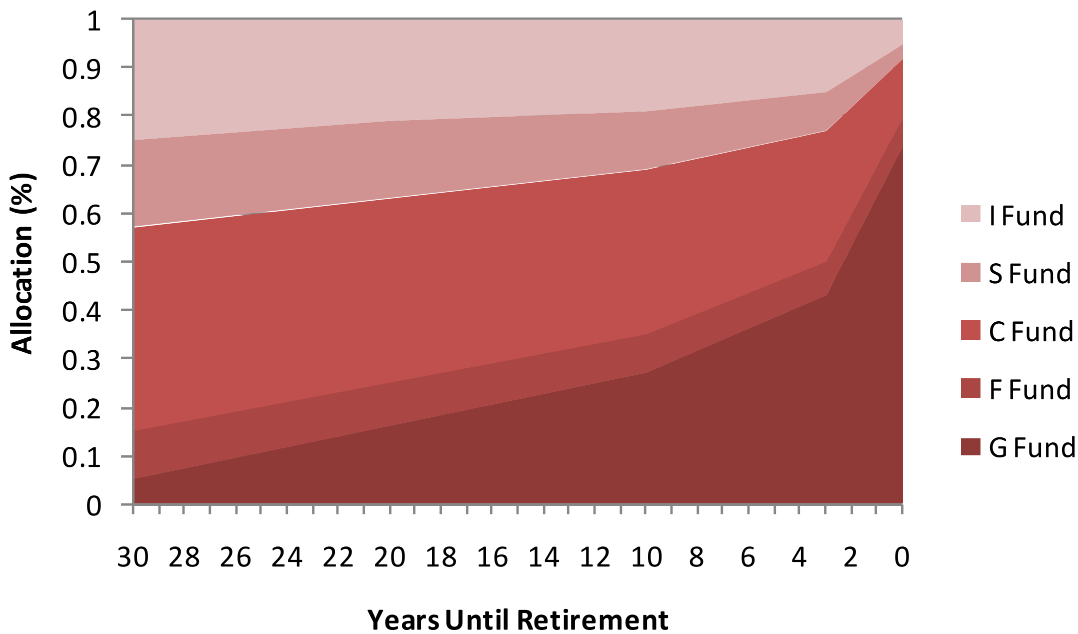

Figure 2. Continuous View of L Funds Over Time.

1.3 Description of Returns Data

1.3.1 General

There are two common definitions of rate of return. Let $r_{i,t}$ denote the return of the *i*-th asset at time *t* and $S_{i,t}$ be the price of the *i*-th asset at time *t*. In our analyses, all returns are calculated as being continuously compounded as log[2] returns,

$$r_{i,t} = \ln(S_{i,t}) - \ln(S_{i,t-1}) = \ln\left(\frac{S_{i,t}}{S_{i,t-1}}\right), \tag{1}$$

instead of the holding-period returns,

$$r_{i,t} = \left(\frac{S_{i,t} - S_{i,t-1}}{S_{i,t-1}}\right) \tag{2}$$

For the purposes of our analyses, each method is more appropriate in at least one case. However, in an effort to combine the three major themes of this research, one standard technique is necessary. Log returns are generally preferred for a couple of reasons. First, log returns over a period of length *ks* are just the sum of the log returns of *k* periods of length *s*. Additionally, most continuous-time models for the stock price $S_{i,t}$ include an exponential of some stochastic process, so continuously-compounded log returns are the obvious choice. Furthermore, returns calculated using log differences are approximately the same as those computed as a percentage

[2] All references to logarithms in this work refer to the natural logarithm, i.e. $\ln(e^x) = x$, regardless of the use of the symbols 'ln' or 'log'.

change; it can be shown by a Taylor series expansion that $\ln(1+x) \approx x$ for x near zero. So, even for those analyses where the holding period return may be more suitable, the form chosen is a sufficiently close estimate.

At the time of this research, data from at least some funds or indices were available from January 1988 until August 2006. In order to maintain commonality between annual, monthly, and daily returns, the 17-year period from January 1, 1989 to December 31, 2005 was selected. Data from the start of the TSP in January 1988 through the end of 1988 were excluded because daily data is not available for the Lehman Brothers Aggregate (LBA) in this time period. Also, data from 2006 were not used, because annual returns were not yet available when this research was begun. For dates prior to 2001, since the S and I funds did not exist, the returns data for their tracked indices is used, as their tracking since their inception has show correlation exceeding 0.99. Finally, because the MSCI EAFE index includes equities from multiple exchanges located in different time zones and with different trading days and holidays, some daily observations were deleted. After matching, we have 4235 daily, 204 monthly, and 17 annual returns.

1.3.2 Annual and Monthly Returns

Annual and monthly returns data are available from the Thrift Savings Plan website. Figure 3 shows the compound annual returns for each of the core TSP funds/indices over the period 1988-2005. As shown by the horizontal dashed line, all five funds outpaced inflation (2.9%) but some performed better than others.

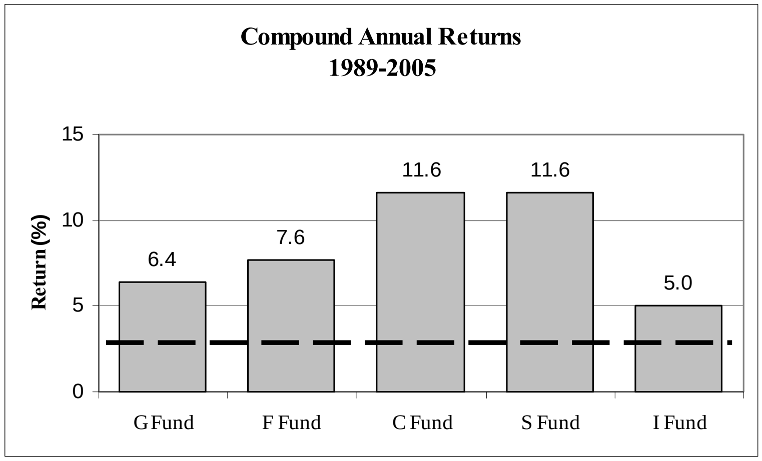

Figure 3. Annual Returns of TSP Core Funds. Horizontal line shows inflation rate during this time period.

1.3.3 Daily Returns

The Thrift Savings Plan website does not provide daily historical price or returns data. As previously mentioned, the correlation between each of the TSP funds and the tracked indices is quite high. Fortunately, daily price and returns data were available for each of the tracked indices. Historical daily data for the S&P 500, DJ Wilshire 4500, and MSCI EAFE indices were obtained via Bloomberg, while the LBA data was provided to the author by the Lehman Brothers Family of Funds.

The G Fund daily returns were computed by starting from the monthly returns and assuming a constant return for each day in the month. Let r_d be the daily rate and r_m the monthly rate, and n days in a month,

$$r_m = (1+r_d)^n - 1, \tag{3}$$

or equivalently

$$r_d = (1+r_m)^{\frac{1}{n}} - 1. \tag{4}$$

For example, if the monthly return was 2% and there are 21 trading days in the month, the daily return would be

$$r_d = (1+0.02)^{\frac{1}{21}} - 1 = .0009434 \text{ or } .09434\% \tag{5}$$

<u>1.4 Investment Horizon</u>

For our analyses, the point of military retirement, not ultimate retirement (and complete reliance on investment income) will be used for two reasons. First, upon military retirement, the investor's opportunities greatly increase because the retiree can move assets from the TSP into other qualified plans, e.g. IRA or 401(k) plans, which have a far broader range of investment options. Also, post-military retirement employment options vary greatly by individual. Some will have no employment after military retirement; others will choose a lower paying but otherwise rewarding second career (e.g. teacher); and some will have a high-paying position (e.g. government contractor). The investment problem faced at that time, although similar in general, is significantly more complex and beyond the scope of this research.

The current military retirement system encourages a military career of 20 years. (U.S. Department of Defense 2006; Warner 2006) As a result, most TSP investors face an investment horizon of 20 years or less; a minority have a planning horizon that is 30 years or longer. For the purposes of our analyses, a representative horizon of 20 years will be used.

1.5 Utility and Risk Aversion

1.5.1 Utility Functions

A utility function maps preferences among alternative choices into measurable utility. Utility functions have two properties—they are order preserving and can be used to rank combinations of risky alternatives. (Copeland and Weston 1988) A concave utility function for a risk averse-investor is shown in Figure 4. In summary, a risk-averse investor with less wealth will obtain a greater amount of utility from the same amount of wealth as a risk-averse investor starting with greater wealth.

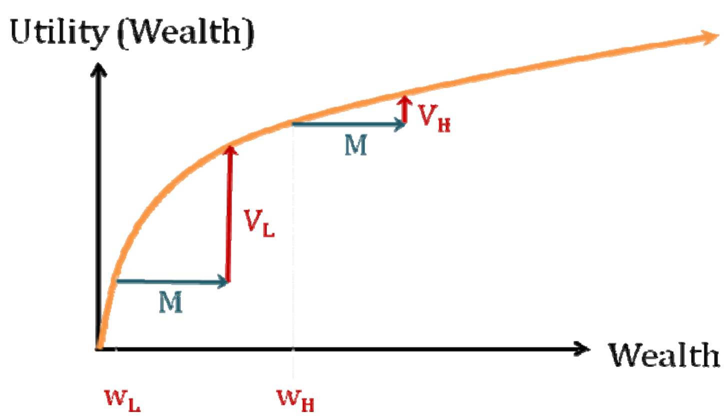

Figure 4. Utility Function for Risk-Averse Investors. Investor with low initial wealth w_L gets greater utility V_L from same amount of additional wealth (M) than investor with higher initial wealth w_H.

There are a number of commonly-used utility functions to choose from, as shown in Table 3. We will use the negative exponential utility function, which has constant absolute risk aversion, for ease of computation. Exponential utility is sometimes bypassed in favor of other utility functions, like logarithmic or power, because it has constant absolute risk aversion or no "wealth effect." As a result, all investors with the same degree of risk aversion who maximize exponential utility will hold the same portfolio, regardless of their initial wealth.

Name	**U[w]**	**A[w]**	**R[w]**
Risk Neutral	w	0	0
Negative Exponential	$-e^{-\gamma w}$	γ (constant)	γw
Quadratic	$w-\frac{\gamma w^2}{2}, \forall w \in [0, \frac{1}{\gamma}]$	$\frac{\gamma}{1-\gamma w}$	$\frac{\gamma w}{1-\gamma w}$
Power	$\frac{w^\gamma}{\gamma}, \gamma < 1 \; and \; \gamma \neq 0$	$\frac{1-\gamma}{w}$	$1-\gamma$
Logarithmic	$\ln(w)$	$\frac{1}{w}$	1 (constant)

Table 3. Examples of Utility Functions, A[w] = absolute risk aversion, R[w] = relative risk aversion, w = wealth, γ = risk aversion coefficient.

However, the variation in wealth among TSP investors is relatively limited. After all, they are all government employees earning somewhat modest salaries with limited bonus potential. The vast majority of TSP participants are neither below the poverty line nor multi-millionaires. Considering this, together with the computational tractability advantage, makes exponential utility a reasonable choice in this situation.

1.5.2 Risk Aversion Parameter

Risk aversion is a measure of the risk premium (additional expected return to compensate for risk) required by an investor for choosing a riskier investment over a guaranteed return. Risk aversion may be measured in absolute or relative terms. For a utility function $U(w)$, the absolute risk aversion is defined as $A(w) = -\frac{U''(w)}{U'(w)}$ and the relative risk aversion is given by $R(w) = wA(w) = -\dfrac{wU''(w)}{U'(w)}$; instances for specific utility functions are shown in Table 3. An estimated 74% of household net worth is estimated to be invested in a broad array of risky assets. Assuming that this portfolio has similar reward-to-risk characteristics as the S&P 500 since 1926 (a risk premium of 8.2% and standard deviation of 20.8%), the implied coefficient of risk aversion is 2.6. (Bodie, Kane and Marcus 2005) Other studies taking into account a wide range of available assets estimate the degree of risk aversion for representative investors at 2.0 to 4.0. (Friend and Blume 1975; Grossman and Shiller 1981) As many in the military will also have retirement income from a defined benefit pension, it is reasonable to conclude that they will fall at the lower (riskier) end of this range. The reason for this is that the value of their military retirement may be considered "pseudo-bonds," as they are backed by the U.S. government and indexed for inflation. (Jennings and Reichenstein 2001; Nestler 2007) This idea is shown graphically in Figure 5.

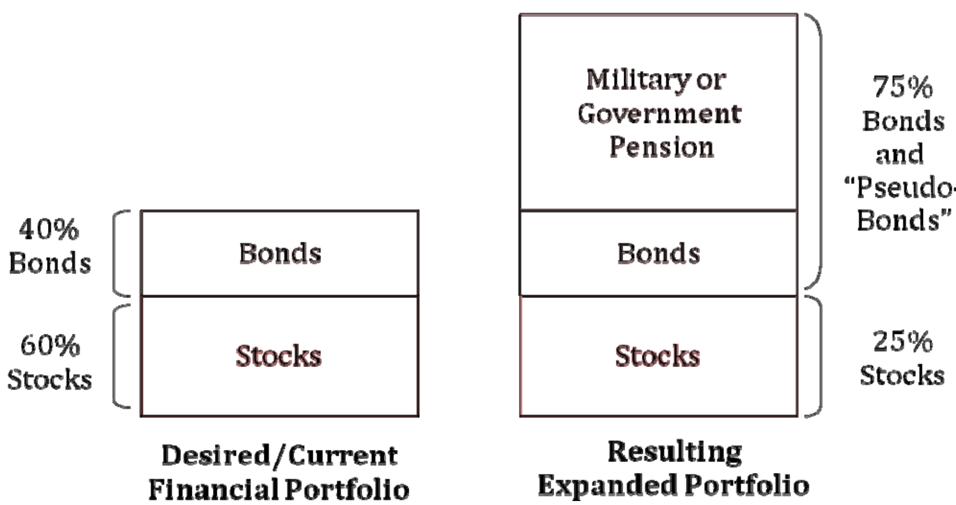

Figure 5. Impact of Considering Defined Benefit Pension as "Pseudo-Bonds" in Expanded Portfolio.

With discounting at Treasury Inflation Protected Securities (TIPS) rates, the net present value of a military retirement with 20 years of service ranges between \$450K and \$900K. As a result of considering this asset as part of their total investment portfolio, those expecting a military retirement can afford to be less risk averse in their TSP and other investments. (Nestler 2007) Therefore, our analyses will use the lower end of the identified range, i.e. a risk aversion parameter of 2.0.

1.6 Organization

Chapter 2 develops a non-Gaussian factor model (called VG-ICA) for returns using ICA with components assumed to follow the VG distribution. In Chapter 3, different measures of risk and portfolio performance are considered, including two new measures. Chapter 4 presents the results of applying the techniques outlined in the earlier chapters to the TSP portfolio optimization problem, while Chapter 5 concludes and discusses future work.

Chapter 2: The VG-ICA Model

2.1 Motivation and Overview of Approach

Historical data suggest that log returns of stocks and indices are not independent and identically distributed Normally, as is commonly assumed. (Lo and Mackinlay 1988; Praetz 1972) Instead, the log returns of financial assets are skewed and have higher kurtosis. To account for skewness and excess kurtosis, it is necessary to have a distribution that is more flexible than the Gaussian distribution (i.e. not completely described by the first two moments), but is still based on a stochastic process that has independent and stationary increments, like Brownian motion. Processes with these characteristics are known as Lévy processes. Distributions which meet the criteria described above include Variance Gamma (VG), Normal Inverse Gaussian (NIG), and Generalized Hyperbolic Model. (Schoutens 2003)

In the manner of Arbitrage Pricing Theory (Ross 1976), a factor model for returns is given by:

$$\begin{aligned} &(R-\mu) = XB + \varepsilon \\ &D = XB + \varepsilon \end{aligned}$$

where R is a vector of asset returns with means μ, D is a vector of de-meaned asset returns, X is a subset of the zero mean, unit variance, orthogonal factors identified using ICA, and ε is the noise component. The only use of the average returns is to de-mean the data. The noise component will be modeled as having multivariate Gaussian density, as this is the finite variance, zero mean RV with the maximum uncertainty in terms of entropy.

2.2 Independent Component Analysis

2.2.1 General

Independent Component Analysis (ICA) is a statistical technique for extracting useful information from a complex dataset through decomposition into independent components. The intent is to find the underlying factors behind a set of observed data. ICA is a linear transformation method that originated in the signal processing field. (Hyvarinen, Karhunen and Oja 2001) The basic idea behind the technique is to find a representation of data that is suitable for some type of analysis, like pattern recognition, visualization, removal of noise, or data compression with the goal of extracting useful information. ICA can be considered a type of blind source separation (BSS), with the word blind indicating that the method can separate data into source signals even if very little is known about their nature. ICA works by exploiting the fact that the source signals are independent from one another. If two (or more) signals are statistically independent, then knowing the value of one signal provides no information about the value of the others. Numerous overviews of ICA are available. (Cardoso 1998; Hyvarinen 1999; Hyvarinen, Karhunen and Oja 2001; Stone 2004)

2.2.2 Comparison with Other Methods

ICA is closely related to the better known, classical, linear transformations like principal component analysis (PCA), factor analysis (FA), and projection pursuit. However, it differs from each of these methods in at least one assumption. PCA and FA both rely on data that is assumed to be Gaussian, but ICA does not. These second order methods (PCA and FA) make this assumption so that the higher moments are not considered (since the Normal distribution is completely described by its first two moments, the mean and variance). To account for higher moments, ICA requires (or assumes) complete statistical independence, while PCA and FA rely on the weaker assumption of uncorrelated (or linearly independent) signals. Projection pursuit and ICA both allow for non-Gaussian data and use information beyond second order, but projection pursuit does not permit a noise term in the model; ICA does.

2.2.3 ICA in Finance

Similar to an example on the FTSE[3] 100 (Stone 2004), consider the prices of the 500 stocks comprising the S&P 500 index to be a set of time-varying measurements. Each of these depends on some relatively small number of factors

---

[3] The letters F-T-S-E are no longer an acronym that stands for anything. Originally, it represented that FTSE was a joint venture between the Financial Times (F-T) and London Stock Exchange (S-E).

(e.g. unemployment, retail sales, weather, etc.) with each series of stock prices viewed as some mixture of these factors. If the factors can be extracted using ICA, they can then be used to predict the future movement of these stock prices. More details on the use of ICA in finance are available in the literature. (Back and Weigend 1997; Malaroiu, Kiviluoto and Oja 2000; Oja, Kiviluoto and Malaroiu 2000)

2.2.4 Theory and Mechanics of ICA

Consider the matrix equation $\mathbf{x} = \mathbf{A}\mathbf{s}$, where $\mathbf{x}$ is the data matrix that is believed to be a linear combination of non-Gaussian, independent components $\mathbf{s}$ and $\mathbf{A}$ is the unknown mixing matrix. The goal of ICA is to find a de-mixing matrix $\mathbf{W}$ such that $\mathbf{y} = \mathbf{W}\mathbf{x}$. If $\mathbf{W} = \mathbf{A}^{-1}$, then $\mathbf{y} = \mathbf{s}$, or the original source signals have been perfectly recovered in the independent components $\mathbf{y}$. Usually, this is not the case and it is only possible to find $\mathbf{W}$ such that $\mathbf{WA} = \mathbf{PD}$ where $\mathbf{P}$ is a permutation matrix and $\mathbf{D}$ is a diagonal matrix. Figure 6 shows a schematic representation of the ICA process. (Back and Weigend 1997)

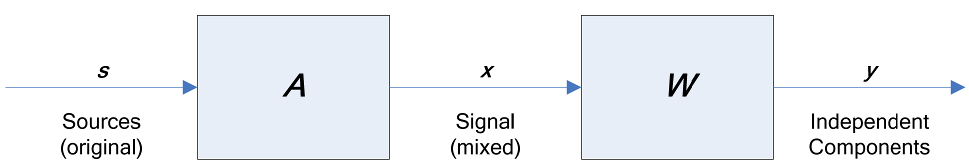

Figure 6. ICA Process Overview.

As previously mentioned, ICA assumes that the independent components are non-Gaussian. So, the rotation chosen should be the one that maximizes non-Gaussianity. Substituting in for $\mathbf{x}$ above, we have $\mathbf{y} = \mathbf{Wx} = \mathbf{WAs}$. So, $\mathbf{s}$ is a linear function of $\mathbf{x}$. With appropriate regularity conditions, the Central Limit Theorem tells us that $\mathbf{y} = \mathbf{WAs}$ is more Gaussian than $\mathbf{s}$, unless $\mathbf{W} = \mathbf{A}^{-1}$, and $\mathbf{y} = \mathbf{s}$ as previously mentioned.

In information theory, entropy is the degree of information that observing a random variable (RV) provides. In other words, larger entropy implies more randomness. For continuous RVs, the differential entropy H is defined as:

$$H(\mathbf{y}) = -\int f_Y(\tilde{\mathbf{y}}) \ln f_Y(\tilde{\mathbf{y}}) d\tilde{\mathbf{y}} \tag{6}$$

where $f_Y(\cdot)$ is the density function of Y. The Gaussian RV has the largest entropy among all RVs of equal variance. (Cover and Thomas 1991) The negentropy J of a RV Y (or random vector y) is defined as the difference between the entropy of a Gaussian RV and the entropy of y, or $J(\mathbf{y}) = H(\mathbf{y}_{Gaussian}) - H(\mathbf{y})$. Thus, maximizing J will result in $\boldsymbol{y}$ being as non-Gaussian as possible. Note that the only case where negentropy equals zero is when y is Gaussian. As actually computing entropy (and therefore negentropy) can be difficult, several approximations have been developed. For ICA, where the objective is to find one independent component $\mathbf{y} = \mathbf{w}'\mathbf{x}$ at a time, the form of the approximations is:

$$J_G(\mathbf{w}) \approx c(E(G(\mathbf{w}'\mathbf{x})) - E(G(\mathbf{y}_{Gaussian})))^2 \tag{7}$$

constrained by $E((\mathbf{w}'\mathbf{x})^2) = 1$, with G a non-quadratic function and c a constant.

All ICA computations in this paper were performed using the FastICA package for R; FastICA is a fixed point iteration algorithm that is also available for Matlab. (Marchini, Heaton and Ripley 2006) The first step in the process is to center the data by subtracting the mean of each column of the data matrix $\mathbf{x}$. Next, it is whitened by projecting the data onto its principal component directions, $\mathbf{x} \rightarrow \mathbf{xK}$, where $\mathbf{K}$ is a pre-whitening matrix. Finally, the algorithm estimates a matrix $\mathbf{W}$ such that $\mathbf{xKW = s}$. The function G used in the negentropy approximation is

$$G(x) = \frac{1}{c} \ln \cosh(cx), \text{ with } 1 \le c \le 2. \tag{8}$$

2.2.5 ICA Examples

Figure 7 gives an example of how ICA can be used to un-mix two independent uniform distributions that were mixed with a deterministic mixing matrix. (Marchini, Heaton and Ripley 2006) The first panel shows the mixed distributions; the second panel shows the limitations of PCA (rotation by maximizing explained variance); the third panel shows good separation by ICA.

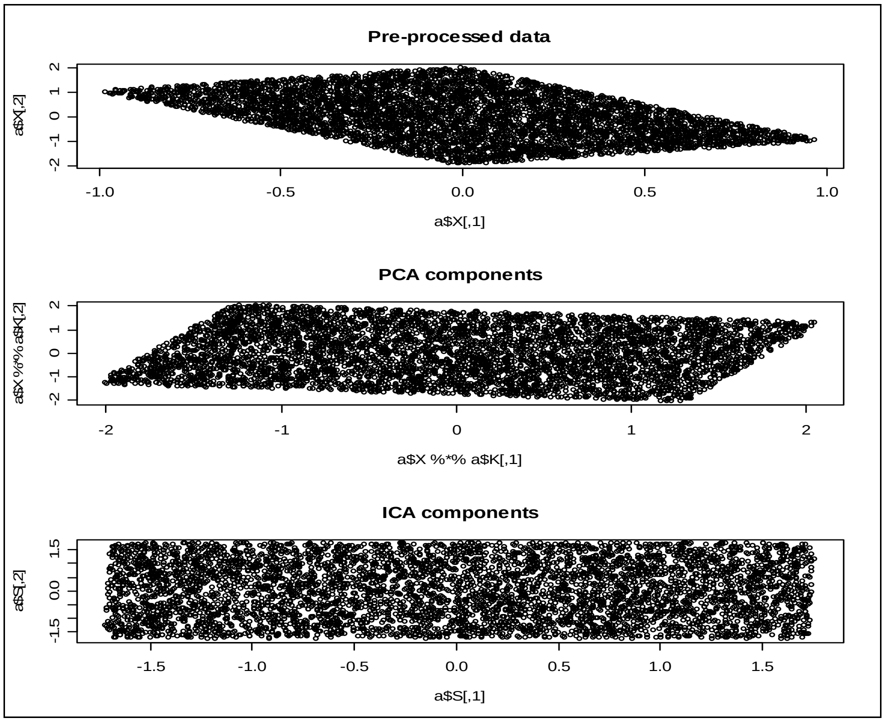

Figure 7. Example of un-mixing two independent uniform distributions using ICA.

Figure 8 shows an example of ICA as blind source separation. (Marchini, Heaton and Ripley 2006) The two leftmost figures show the original source signals, a sine wave and a saw tooth wave. The two center figures show the mixed signals that were provided to the ICA algorithm; the two figures on the right show the resulting un-mixed estimates of the original signals.

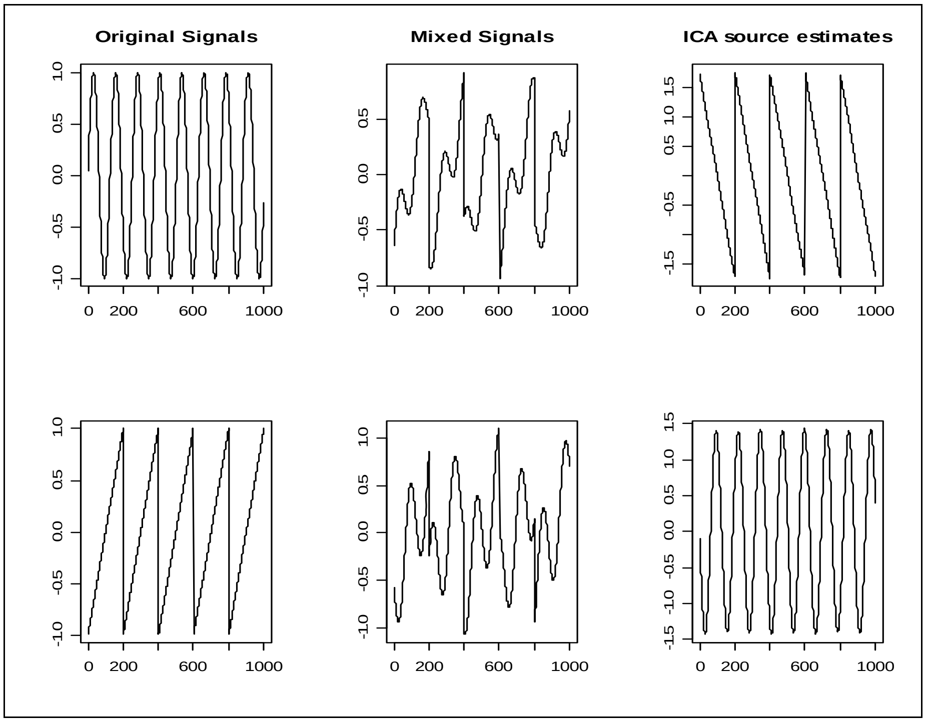

Figure 8. Example of un-mixing two source signals using ICA.

2.3 The Variance Gamma (VG) Process and Distribution

2.3.1 Motivation and History

Looking at stock price movements, it becomes obvious that they do not look very much like a Brownian motion. Instead of moving continuously, they make lots of finite up and down movements. It has long been observed that empirical data shows that, in comparison to the Normal distribution, returns have fatter tails and a higher center that is more peaked. (Fama 1965; Praetz 1969) Early efforts provided the use or development of several distributions to account for this, including the scaled t-distribution (Praetz 1972), the stable Paretian family of distributions (Mandelbrot 1969), and the compound events model (Press 1967). However, these efforts lacked the development of underlying continuous-time stochastic processes that are required for option pricing. (Madan and Seneta 1987) Some models do rely on established processes, like the Black-Scholes and Merton diffusion (Black and Scholes 1973; Merton 1973), the pure jump process of Cox and Ross (Cox and Ross 1976) and the jump diffusion of Merton (Merton 1976). These models do provide continuous paths, except at jump events, but are of infinite variation. However, it appears that index returns tend to be pure jump processes of infinite activity and finite variation, as the index return processes seem to have diversified away diffusion risk that may be present in individual stock returns. (Carr *et al.* 2002) Jump components are important in stock price modeling because pure diffusion models suffer from problems with volatility smiles in short-dated options. (Bakshi, Cao and Chen 1997)

The VG model endeavors to tackle this shortcoming by modeling the evolution of stock prices by considering “experienced time” as a RV. The original model (Madan and Seneta 1990) did not allow for skewness, but an extension (Madan, Carr and Chang 1998) does provide this additional control. The VG formulation came from considering the distribution of the reciprocal of variance of a zero-mean Normal to be gamma distributed (Praetz 1972) but with the modification that the variance itself is gamma distributed; hence the name “Variance Gamma.”

There is evidence that estimated independent components (ICs) produced from financial time series fall into two categories: (i) infrequent, large shocks that are responsible for major stock price moves, and (ii) frequent, small changes that contribute very little to the changes in stock prices. (Back and Weigend 1997) For this reason, use of the VG distribution to the estimated ICs seems ideal. In many respects, the approach used in this chapter follows a recent dissertation (Yen 2004) and papers by authors at the University of Maryland. (Madan 2006; Madan and Yen 2004) However, it is different in that TSP portfolios can only consist of long positions; short sales are not allowed. A similar model was used in another recent work on pricing multi-asset products. (Xia 2006)

2.3.2 Gamma Random Variables

From an information perspective, the market does not forget information, so the process used should be monotonically non-decreasing. The family of Gamma distributions, with two parameters, the mean, μ, and the variance, ν, and density function:

$$f_h(x) = \left(\frac{\mu}{\nu}\right)^{\frac{\mu^2 h}{\nu}} \frac{x^{\frac{\mu^2 h}{\nu}-1} \exp(-\frac{\mu}{\nu}x)}{\Gamma(\frac{\mu^2 h}{\nu})}, \quad \text{for } x > 0 . \tag{9}$$

is one such model that is useful.

2.3.3 Lévy Processes

A stochastic process $X = \{X(t) : t \geq 0\}$ is a Lévy process if it: (i) X has independent increments, (ii) $X(0) = 0$ almost surely, (iii) X has stationary increments, (iv) X is stochastically continuous, and (v) X is right continuous with left limits almost surely. (Schoutens 2003)

The characteristic function $\phi_X(u)$ of the distribution $F(x) = P(X \le x)$ of a RV X is defined as:

$$\phi_X(u) = E[\exp(iux)] = \int_{-\infty}^{\infty} \exp(iux) dF(x). \tag{10}$$

A probability distribution with characteristic function $\phi_X(u)$ is infinitely divisible if, for any positive integer *n*, $\phi_n(u) = \phi(u)^{1/n}$ is also a characteristic function. The Lévy-Khintchine formula gives the unique characteristic function of any infinitely divisible function. Shown here is the log characteristic function for such a distribution.

$$\psi(u) = \ln \phi(u) = i\gamma u - \frac{1}{2}\sigma^2 u^2 + \int_{-\infty}^{\infty} (\exp(iux) - 1 - iux 1_{\{|x|<1\}})\nu(dx), \tag{11}$$

with $\gamma \in R$, $\sigma^2 \ge 0$, and ν is a measure on $R \setminus \{0\}$. Notice that the Lévy-Khintchine formula has three constituents: a deterministic part with drift coefficient γ, a Brownian piece with diffusion coefficient σ, and a pure jump component (the last term).

2.3.4 Variance Gamma Process

The VG process is a pure jump Lévy process; it contains no deterministic or Brownian motion components. There are two representations of the VG process: as a time-changed Brownian motion with a Gamma subordinator, and as a difference of two Gamma processes. For the first representation, start with a Brownian motion $(W(t), t \ge 0)$ with drift θ and volatility σ. That is, $B(t; \theta, \sigma) = \theta t + \sigma W(t)$. Using an independent Gamma process $(G(t;\nu,\sigma), t \geq 0)$ with mean rate 1 and variance rate ν for the increment $G(t+h)-G(t)$ given by the gamma density previously given to get the following VG process: $X(t;\nu,\theta,\sigma) = B(G(t;1,\nu),\theta,\sigma)$.

Since the VG process is of finite variation, it can be expressed as a difference of two independent Gamma processes. (Geman, Madan and Yor 2001) This comes from the fact that the VG process is of finite variation. In other words, $X_{VG}(t;\theta,\nu,\sigma) = G_p(t;\nu_p,\mu_p) - G_n(t;\nu_n,\mu_n)$, where $G_p(t)$ is for the positive changes and $G_n(t)$ is for the negative move, where $\mu_p = \frac{1}{2}\sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} + \frac{\theta}{2}$,

$\mu_n = \frac{1}{2}\sqrt{\theta^2 + \frac{2\sigma^2}{\nu}} - \frac{\theta}{2}$, $\nu_p = \mu_p^2 \nu$ and $\nu_n = \mu_n^2 \nu$, The characteristic functions of the two independent Gamma processes are:

$$\phi_{G_p}(u) = \left(\frac{1}{1 - iu\nu_p/\mu_p}\right)^{\frac{\mu_p^2 t}{\nu_p}} \text{ and } \phi_{G_n}(u) = \left(\frac{1}{1 - iu\nu_n/\mu_n}\right)^{\frac{\mu_n^2 t}{\nu_n}}. \qquad (12)$$

Then, the characteristic function of the difference of the two Gamma processes G_p and G_p is:

$$\phi_{G_p - G_n}(u) = \left(\frac{1}{1 - iu\left(\frac{\nu_p}{\mu_p} - \frac{\nu_n}{\mu_n}\right) + u^2 \frac{\nu_p}{\mu_p}\frac{\nu_n}{\mu_n}}\right)^{\frac{t}{\nu}}. \qquad (13)$$

Each representation has an advantage over the other. Viewing the VG process as a time-changed Brownian motion is useful for determining the characteristic functions and density, as shown above. This is useful for fitting the VG distribution to data. However, considering the VG process to be a difference of two gamma processes provides a basis for more efficient simulation of observations from a fitted or hypothesized VG distribution. (Avramidis and L'Ecuyer 2006)

The VG distribution is closely related to other, better known distributions. For example, the Laplace distribution is a special case of the symmetric (i.e. $\theta = 0$) VG distribution. Similarly, the t- distribution may be considered to be a generalization of the Cauchy distribution. As a result of similarity in structure of the Laplace and Cauchy distributions, the VG and t- distributions are virtually indistinguishable from one another in terms of tail structure. (Fung and Seneta 2006) The parameter σ affects the spread of the distribution in usual way; the effect of the parameters ν and θ can be seen in Figure 9 and Figure 10.

A useful economic interpretation of the parameters can be obtained by re-parameterizing the model in terms of realized quadratic variation (or volatility), a directional premium, and a size premium. (Madan and Yen 2004) They also note that the Gaussian model is a special case that results from allowing the variance of the Gamma process approach zero (or letting the kurtosis level approach 3).

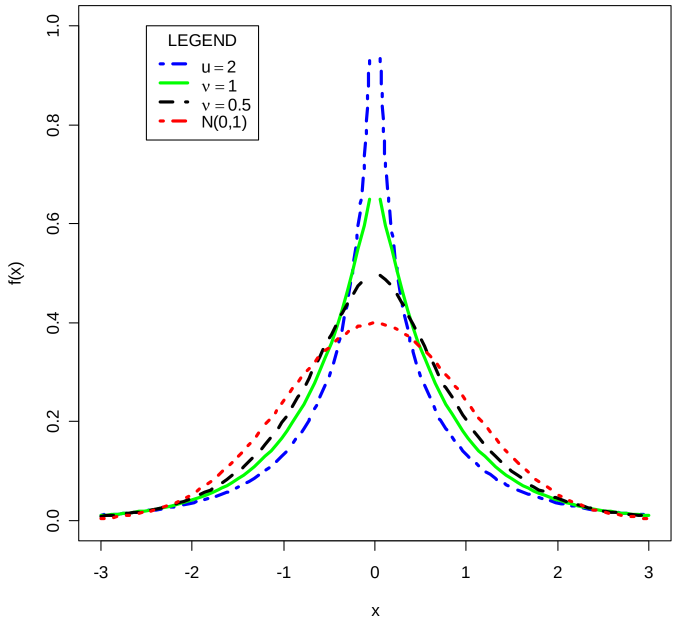

Figure 9. Effect of Parameter ν on VG Distribution.

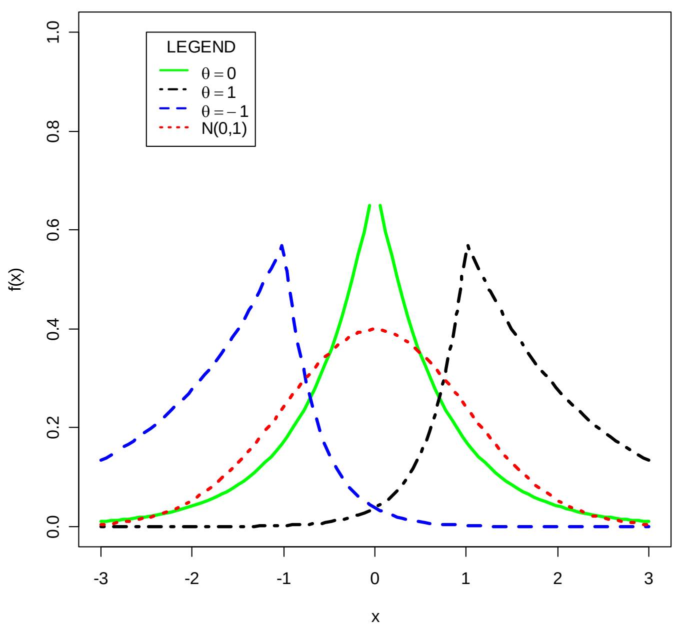

Figure 10. Effect of Parameter θ on VG Distribution.

2.3.5 Simulating From the VG Distribution

There are at least three ways to simulate from the VG distribution. (Fu 2007) The first two methods are exact and come from the representations of the VG process—a time-changed Brownian motion with a Gamma subordinator and as a difference of two non-decreasing (Gamma) processes.

To simulate VG as Brownian motion with a Gamma time-change, let X_{VG} be a VG process with parameters σ, ν, and θ, or $X_{VG} = \theta g + \sigma W(g)$. $W(g) \overset{d}{=} \sqrt{g} Z$ and Z is a standard Normal RV that is independent of g. To obtain a sample path of X_{VG}, first simulate a gamma process g with shape parameter $1/\nu$ and scale parameter ν. Then, independently simulate a standard Brownian motion, or random numbers with zero mean and variance $\sigma^2 g$. Combine these as shown above and you have X_{VG}, the desired VG process.

To simulate VG as the difference of two independent Gamma processes, let X_{VG} be a VG process with parameters σ, ν, and θ. $X_{VG}(t) = G_p(t) - G_n(t)$ where $G_p(t)$ and $G_n(t)$ are two independent Gamma processes with mean rates μ_p, μ_n, and variance rates ν_p, ν_n, respectively. To obtain a sample path of X_{VG}, simulate G_p with shape parameter μ_p^2 / ν_p and scale parameter ν_p / μ_p; and G_n with shape parameter μ_n^2 / ν_n and scale parameter ν_n / μ_n. Take the difference as described above and the result is X_{VG}, the desired VG process.

The remaining method is an approximation based on a compound Poisson process. Newer methods including bridge sampling (starting from the end and filling in as needed) have recently been introduced. (Avramidis and L'Ecuyer 2006) Also, general variance reduction techniques used in simulation are useful in some situations. (Fu 2007)

2.4 VG Stock Price Model

To construct the VG price for stocks (or factors), replace the Brownian motion in the Black-Scholes model by the VG process. With a continuously compounded risk free rate *r*, we then have: $S(t) = S(0)\frac{\exp(rt + X_{VG}(t))}{E[\exp(X_{VG}(t))]}$. Therefore, $E[S(t)] = S(0)\exp(rt)$. More specifically, $\exp(-wt) = E[\exp(X_{VG}(t))] = \phi_{VG}(t) = \exp(-\frac{t}{\nu}\ln(1-\theta\nu - \frac{\sigma^2\nu}{2}))$. So, $S(t) = S(0)\exp((r+w)t + X_{VG}(t))$ where $w = \frac{1}{\nu}\ln(1-\theta\nu-\frac{\sigma^2\nu}{2})$, the convexity correction. It is then possible to determine $\phi_{\ln S(t)}(u)$ from $\phi_{VG}(t)$ as follows:

$$\begin{aligned}
\phi_{\ln S(t)}(u) &= E(\exp(iu\ln S(t))] \\
&= \exp(iu\ln(S(0) + rt + \frac{t}{\nu}\ln(1-\theta\nu-\frac{\sigma^2\nu}{2}))\phi_{VG}(u) \qquad (14)\\
&= \exp(iu\ln(S(0) + rt + \frac{t}{\nu}\ln(1-\theta\nu-\frac{\sigma^2\nu}{2}))(1-iu\theta\nu-\frac{\sigma^2\nu}{2}u^2)^{\frac{t}{\nu}}
\end{aligned}$$

The density function for VG is obtained by using Fourier inversion on this characteristic function, resulting in:

$$h(z) = \frac{2\exp(\theta x/\sigma^2)}{\nu^{t/\nu}\sqrt{2\pi}\sigma\,\Gamma(\frac{t}{\nu})}\left(\frac{x^2}{2\sigma^2/\nu+\theta^2}\right)^{\frac{t}{2\nu}-\frac{1}{4}} K_{\frac{t}{\nu}-\frac{1}{2}}\left(\frac{1}{\sigma^2}\sqrt{x^2(2\sigma^2/\nu+\theta^2)}\right) \quad (15)$$

with $x = z - mt - \frac{t}{\nu}\ln(1-\theta\nu-\sigma^2\nu/2)$ and where $K_{\frac{t}{\nu}-\frac{1}{2}}$ is the modified Bessel function of the second (third) kind with the indicated number of degrees of freedom. (Madan, Carr and Chang 1998) For a single period, say one day (t=1), these simplify to:

$$\phi_{\ln S(1)}(u) = \exp(iu\ln(S(0)+r+\frac{1}{\nu}\ln(1-\theta\nu-\frac{\sigma^2\nu}{2})))(1-iu\theta\nu-\frac{\sigma^2\nu}{2}u^2)^{\frac{1}{\nu}} \quad (16)$$

$$h(z) = \frac{2\exp(\theta x/\sigma^2)}{\nu^{1/\nu}\sqrt{2\pi}\sigma\,\Gamma(\frac{1}{\nu})}\left(\frac{x^2}{2\sigma^2/\nu+\theta^2}\right)^{\frac{1}{2\nu}-\frac{1}{4}} K_{\frac{1}{\nu}-\frac{1}{2}}\left(\frac{1}{\sigma^2}\sqrt{x^2(2\sigma^2/\nu+\theta^2)}\right) \quad (17)$$

with $x = z - m - \frac{1}{\nu}\ln(1-\theta\nu-\sigma^2\nu/2)$.

Chapter 3: Measures of Risk and Portfolio Performance

3.1 Overview

When considering various investment options and selecting portfolios under uncertainty, there are two common approaches. (De Georgi 2002; Stoyanov, Rachev and Fabozzi 2005) One method for comparison is to use the concept of stochastic dominance, which is closely linked to utility theory. Unfortunately, this approach often results in optimization problems that are difficult to solve, so it may not result in useful information. More commonly seen is the use of a portfolio performance measure in the form of a reward-to-risk ratio that evaluates the balance between expected reward and risk. One portfolio is preferred to another if it has higher expected reward and lower risk. Of course, there is usually a tradeoff involved; generally, with higher returns come higher risk. One must first settle on an appropriate measure of risk, which, judging by the number of measures proposed to date is clearly not an easy task. If possible, use of stochastic dominance is desirable because risk measures (which assign a single number to a random wealth) have difficulty summarizing all distribution information, whereas stochastic orders compare cumulative distribution functions. (Ortobelli *et al.* 2005)

3.2 Stochastic Dominance

Stochastic dominance (SD) generalizes utility theory by eliminating the need to explicitly specify a utility function. Instead, general statements about wealth preference, risk aversion, etc. are used to decide between investment alternatives. (Heyer 2001) SD is one tool that can be used to address the balance between risk and reward when considering different investment alternatives. It is an analytical tool that is intuitive, easy to implement when presented with empirical output from simulation models. One of the challenges when using utility theory is that there is no way to assess the overall acceptability of competing options as there is no objective, absolute scale for utility. It is completely dependent on the specification of the utility function. This makes utility theory notionally elegant but largely ineffective in practice. Most investors do not have the willingness or means to select and parameterize their own utility function. SD allows us to use features from utility theory (like increasing wealth preference, risk aversion, etc.) without using a specific utility function.

3.2.1 First-Order Stochastic Dominance (FOSD)

FOSD assumes only monotonicity (that investors like more money rather than less money and are non-satiated); as it has the weakest assumptions, it is the strongest result among the various orders of SD. When comparing two return distributions, It is easy to see when FOSD occurs by comparing the cumulative distribution functions (CDFs).. Given two random variables A and B, with CDFs $F_A(x)$ and $F_B(x)$, respectively, A dominates B, if and only if, $F_B(x) \geq F_A(x), \forall x$, with at least one strict inequality. This can be seen graphically in Figure 11 since $F_B(x)$ is above $F_A(x)$

everywhere (or equivalently, $F_A(x)$ is to the right of $F_B(x)$ everywhere), the probability of getting at least x is higher under $F_A(x)$ than $F_B(x)$.

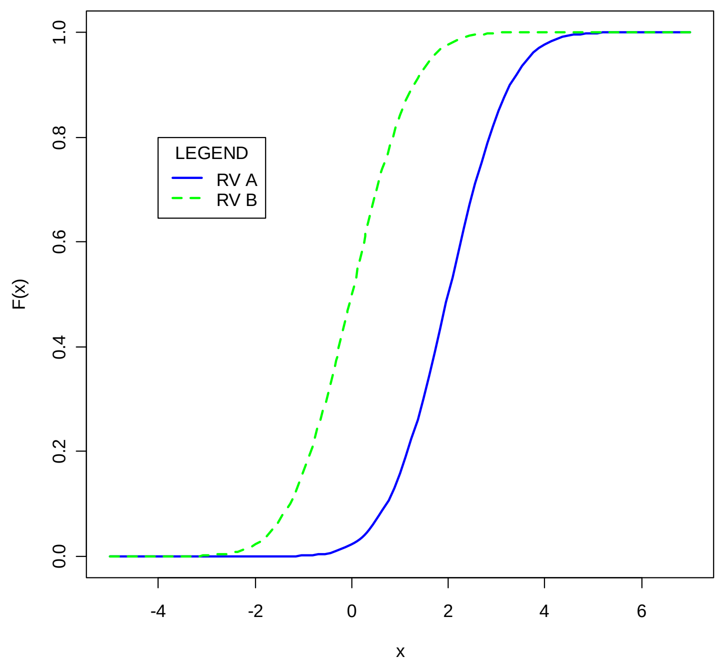

Figure 11. First-Order Stochastic Dominance (FOSD).

3.2.2 Second-Order Stochastic Dominance (SOSD)

SOSD adds risk aversion to the lone assumption of FOSD. This implies that expected utility is less than or equal to the utility of expected returns. Since Jensen's Inequality holds, a risk-averse investor will not play a fair game and will be willing to pay for insurance.

Graphically, if the curves of $F_A(x)$ and $F_B(x)$ cross on the CDF plot, then it is obvious that FOSD does not apply and neither investment option results in higher wealth at every level of probability. However, is unclear whether either distribution dominates the other in other ways. A stochastically dominates B in a second order sense, if and only if, $\int_{-\infty}^{x}[F_B(u)-F_A(u)]du \geq 0, \forall x$, with at least one strict inequality. Once again, this can be easily seen by comparing a graph of the CDFs of A and B. As shown in Figure 12, although neither distribution has FOSD over the other (since the curves cross), A does have SOSD over B because the area indicated by I is larger than the area labeled II, so the constraint integral in the definition of SOSD is satisfied. This can be interpreted as option A providing a uniformly higher partial expected value at every wealth limit. Alternatively, A has "uniformly less downside risk at every level of probability". (Heyer 2001)

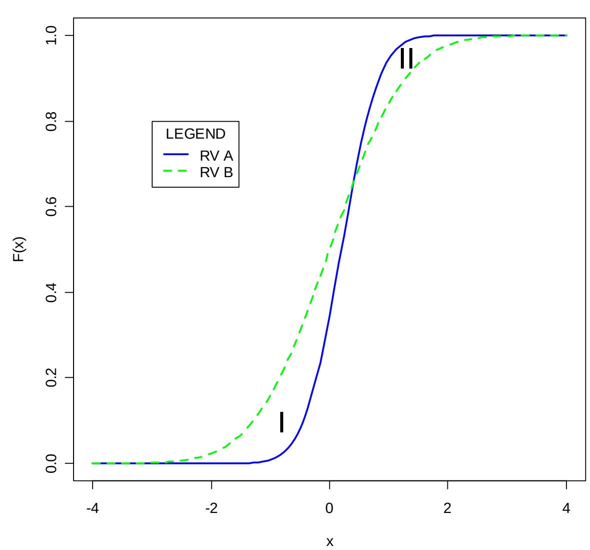

Figure 12. Second-Order Stochastic Dominance (SOSD).

3.2.3 Implications Between Orders

As lower (numbered) orders of SD have fewer assumptions than higher orders, they are stronger results and, therefore, imply the higher orders. So, FOSD implies SOSD. However, the implications do not go the other way; SOSD does not imply FOSD. Dominance is also transitive, meaning that if A dominates B and B dominates C, then A also dominates C.

3.2.4 Empirical Application

One of the advantages of using SD when comparing alternatives is the ease with which it can be applied to empirical data. The following is a straightforward procedure for testing the presence of each of the orders of SD. (Heyer 2001) Given n terminal wealth outcomes or cash flows produced by simulation under two investment alternatives:

(1) Sort the outcomes for each in ascending order to produce empirical estimates of the CDFs.

(2) Test for FOSD by computing the difference between each percentile of option A and option B, placing the differences in vector S_1. If every element of S_1 is positive, then FOSD and SOSD both apply.

(3) Test for SOSD but computing the running sums of S_1. (i.e., for each element of $S_{1,}$ compute the sum of that element and every prior element of

S_1) and place it in a vector S_2. If every element of S_2 is positive, then SOSD applies.

3.3 Cash Flow, Gains, and Losses

For the purposes of this study, and the following definitions of various risk measures, we shall define X as the discounted loss obtained from a particular realization of an investment strategy. That is, X is the negative of the difference between the final portfolio value at the end of the investment horizon and the value of the corresponding riskless investment at the end of the same period. A positive (negative) value of X means that the strategy under- (out-) performed the riskless portfolio over the same horizon. Much of our analysis will consider various statistics and measures on the distribution of the portfolio loss (or gain) over a large number of simulated sample paths, as opposed to looking at the returns.

3.4 Risk Measures

There are two common types of risk measures—dispersion risk measures, or measures of variability, and safety risk measures, which focus on potential losses. (Ortobelli *et al.* 2005)

3.4.1 Dispersion Risk Measures

Traditionally, variance (or really, standard deviation) has been used as the measure of risk. However, it has numerous flaws; foremost among these is that it equally weights positive and negative deviation from the mean. Markowitz actually acknowledged this and proposed semi-variance as a way to account for this deficiency. Others have proposed other measures that were variations on this theme. (Fishburn 1977; Konno and Yamazaki 1991; Markowitz 1959) However, for a variety of reasons, none really ever caught on as a standard.

3.4.2 Safety Risk Measure: Value at Risk

First introduced by J.P. Morgan (now J.P. Morgan Chase) under the proprietary name of RiskMetrics ® in 1994, Value at Risk (VaR) became the standard risk measurement throughout the financial industry over the past two decades. (Krause 2002) It was written into the Basel II Accords, which govern the capital reserve requirements for banks and other financial institutions. (Federal Reserve Board 2007) VaR has three parameters: the time horizon of interest, the confidence level λ, and the appropriate currency unit. Common levels of λ are .95 and .99. VaR is an attractive measure because it is easy to understand. It can be interpreted as the expected maximum loss over a fixed horizon for a given confidence level.

With the discounted portfolio loss X as described above (and not as often used, with losses defined as positive), VaR is defined as follows:

$$P(X \geq VaR_{1-\lambda}(X)) = \lambda \tag{18}$$

So, if we choose λ to be .05 or 5%, then $VaR_{95\%}$ would represent the magnitude of the loss that could be expected 5% of the time. However, VaR itself has several shortcomings. When applied to non-Gaussian returns, VaR is not sub-additive, as shown by examples where the VaR of a portfolio is not less than the sum of the VaRs of the individual assets . The economic interpretation is that VaR does not, in general, reward diversification, which has long been known to reduce risk. (Warnung 2007) Also, as shown in Figure 13, (Cherny and Madan 2006) it is possible for two distributions to have the same $VaR_{1-\lambda}$, but yet be quite different, even in the left tail. Clearly, the distribution on the right is "better" than the distribution on the left, even though their λ-quantiles (q_λ) and therefore $VaR_{1-\lambda}$ are identical. This is because VaR only accounts for the size (and not the shape) of the tail.

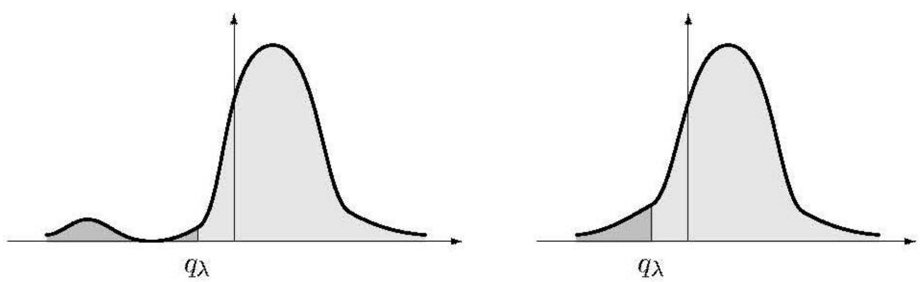

Figure 13. Different Distributions With Same VaR.

From a mathematical perspective, standard deviation and VaR have significant weaknesses. (Artzner *et al.* 1999) To get past these shortcomings requires the use of what are termed "coherent" measures of risk.

3.4.3 Coherent Measures of Risk

Coherent measures of risk satisfy four axioms: translation invariance (adding riskless wealth causes a decline in wealth at risk), monotonicity (more wealth is preferred to less wealth), subadditivity (aggregated risk of investments is lower than the sum of the individual risks), and positive homogeneity (multiplying wealth at risk by a positive factor causes risk to grow proportionally). (Artzner *et al.* 1999)

3.4.4. Conditional Value at Risk

Conditional Value at Risk (CVaR), is defined as:

$$CVaR_{1-\lambda}(X) = E[X \mid X > VaR_{1-\lambda}(X)]. \tag{19}$$

In words, CVaR is the expected value of all losses, given that they exceed the VaR level for a specified λ. So, $CVaR_{95\%}(X)$ would be the average of all losses greater than $VaR_{95\%}(X)$. CVaR is known by many names, including Tail VaR, expected shortfall (ES), expected Tail Loss (ETL) and Average VaR. The origins of CVaR are relatively recent; it appears to have appeared in the literature simultaneously in several sources. (Acerbi, Nordio and CSirtori 2001; Rockafellar and Uryasev 2001) One sign of its acceptance is that it is part of the Solvency II framework, the replacement for the Basel II Capital Adequacy Accord. Another way to consider CVaR is as follows:

$$CVaR_{1-\lambda}(X) = \frac{1}{1-\lambda}\int_{\lambda}^{1} VaR_u(X)du. \qquad .(20)$$

What this means is that CVaR considers the shape of the tail into account by putting all VaR measurements in the tail of the distribution into one number. (Warnung 2007) As defined here, CVaR will always be greater than VaR for a fixed value of λ. It has been shown that CVaR is the largest alternative coherent risk measure to VaR. (Delbaen 1999)

To estimate CVaR using Monte Carlo simulation, the following procedure can be used. Given n (e.g. 1000) samples of the cash flow X at the investment horizon:

(1) Compute VaR for the desired level of λ.

(2) Find all instances of X greater than $VaR_{1-\lambda}$.

(3) Compute their average. This is $CVaR_{1-\lambda}$..

The disadvantage of this method is that it requires finding $VaR_{1-\lambda}$, which involves calculating the quantiles of X. An alternative estimation technique that does not require determining $VaR_{1-\lambda}$ exists; this would be useful if one only wished to consider CVaR and not VaR. (Warnung 2007) CVaR also has favorable properties that allow linearization of what is ordinarily a nonlinear optimization problem. (Krokhmal, Palmquist and Uryasev 2001; Rockafellar and Uryasev 2001)

Although CVaR does address both the size and shape of the tail, it does not consider the rest of the distribution. Figure 14 shows two distributions with not only the same λ-quantiles (q_λ) and VaR_λ, but also identical $CVaR_\lambda$. However, the rest of the distributions are very different. The distribution on the right has a higher expected value and longer upper tail than the one on the left; if this were representative of cash flows, the distribution on the right is plainly preferable to the one on the left.

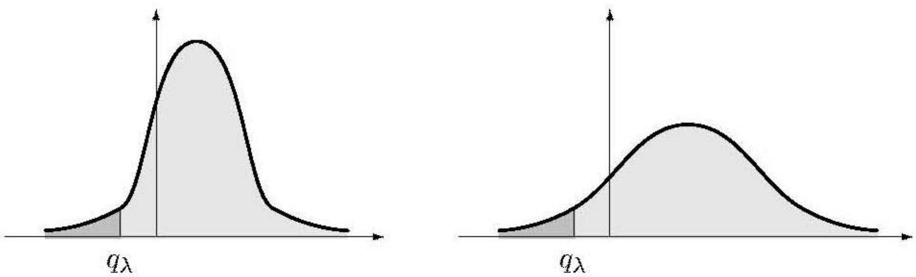

Figure 14. Different Distributions with Same CVaR.

In some cases, as in the TSP problem at hand, the rest of the distribution is of interest. The potential upside should certainly be considered when making portfolio allocation decisions, not just the loss tail. It is necessary to look at the complete distribution to consider higher-order statistics like skewness and kurtosis, as preferences for these vary among individual investors. In general, there is a preference for skewness, or at least a dislike for negative skewness. (Harvey *et al.* 2004) Also, it is noted that kurtosis comes from two sources, peakedness and tailweightedness, which typically have opposite effects on preferences. (Eberlein and Madan 2007) Although this realization is not new, attempts to address it by including information about the entire distribution have only recently been introduced.

3.4.5 Weighted VaR

Weighted VaR is a coherent risk measure, defined as follows:

$$WVaR_\mu(X) = \int_{[0,1]} CVaR_\lambda(X)\mu(d\lambda) \tag{21}$$

where μ is a probability measure on [0,1] and CVaR is as previously defined. The function μ serves to "distort" CVaR in a way that emphasizes the portions of interest. WVaR has some nice properties that CVaR does not; primarily, it is "smoother" than CVaR. (Cherny 2006)

A particular instance of Weighted VaR is Beta VaR (BVaR). (Cherny and Madan 2006) BVaR measures risk by the expectation of the average of the β smallest of α independent copies of the random cash flow. For example, if $\alpha = 50$ and $\beta = 5$, then BVaR(50,5) is the average of the 5 smallest of 50 independent copies of the random cash flow. The advantage of BVaR over other risk measures is that it depends on the entire distribution of X and not just on the tail, as is the case for CVaR. The effect of varying α and β can be seen in Figure 15 and Figure 16.

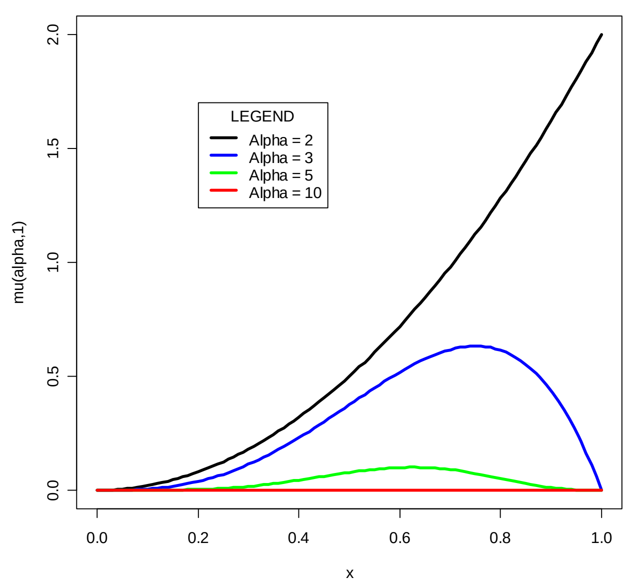

Figure 15. Effect of Varying α With β = 1.

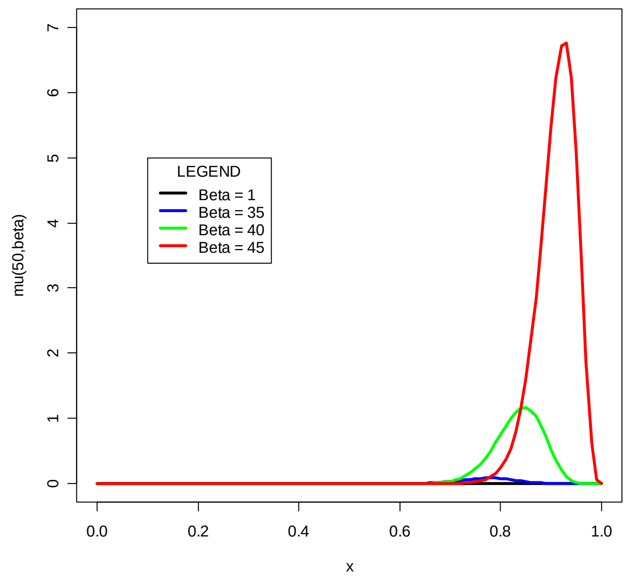

Figure 16. Effect of Varying β with α = 50.

In addition to these desirable properties, BVaR is faster to estimate than CVaR, as it does not require the complete ordering of the cash flow realizations. It can be applied to a wide variety of models, as it uses no assumptions on the structure of the cash flow evolution. The procedure to estimate BVaR is straightforward and simple to implement, as given here. (Cherny and Madan 2006; Warnung 2007)

(1) Simulate α cash flows, $X_1, \ldots, X_\alpha$.
(2) Sort the sample and pick the β smallest cash flows $X_{(1:\alpha)}, \ldots, X_{(\beta:\alpha)}$.
(3) Calculate their average $BVaR_k = \frac{1}{\beta}\sum_{i=1}^{\beta} X_{(i:\alpha)}$.
(4) Estimate Beta VaR by $B\hat{V}aR = \frac{1}{n}\sum_{k=1}^{n} BVaR_k$

A special case of BVaR (with $\beta = 1$) is Alpha VaR (AVaR), which measures risk by the expectation of the smallest of α independent copies of the random cash flow. So, AVaR(50) is the expectation of the smallest of 50, independent copies of the random cash flow.

Four recently introduced (Cherny and Madan 2007) distortions which could be considered are: MINVaR, MAXVaR, MAX MINVaR, and MINMAXVaR. The distortion function for each is shown below. In each case, $x \in \mathbb{R}_+ \; and \; y \in [0,1]$.

$$\begin{array}{ll} MINVaR & 1-(1-y)^{x+1} \\ MAXVaR & y^{\frac{1}{x+1}} \\ MAXMINVaR & (1-(1-y)^{x+1})^{\frac{1}{x+1}} \\ MINMAXVaR & 1-(1-y^{\frac{1}{x+1}})^{x+1} \end{array} \tag{22}$$

MINVaR considers the expectation of the minimum of (1+x) independent draws from the distribution. MAXVaR looks at finding the distribution G(x) from which (1+x) draws are made and taking the best outcome to get the distribution F(x). The other two combine the first two methods to construct worst case outcomes and stress cash flows prior to taking the expectations.

To summarize the recent developments in risk measures, the more features of a distribution that can be put into one number (the risk measure), the better it can capture risk preferences in a convincing manner. (Warnung 2007)

3.5 Portfolio Performance Measures

Portfolio performance measures that take the form of a Reward-to-Risk ratios usually have some measure of reward, typically expected return (or cash flow), in the numerator and a measure of risk in the denominator. Thus, a higher ratio is considered better.

3.5.1 Sharpe Ratio

The Sharpe Ratio, or reward-to-variability ratio, been used for over 40 years. (Sharpe 1966) It is fully compatible with Normally-distributed returns (actually all elliptical distributions), but it can lead to incorrect decisions when returns exhibit skewness and kurtosis. (Biglova *et al.* 2004) Notice that it is essentially the inverse of the coefficient of variation from statistics.

$$SR = \frac{E(r_p) - r_f}{\sigma} \text{ or } SR = \frac{E(X)}{\sigma_X} \tag{23}$$

3.5.2 STARR Ratio

The STARR (Stable Tail Adjusted Risk Ratio) is the ratio of the expected excess return and the CVaR. (Martin, Rachev and Siboulet 2003) It penalizes downside risk but does not take into account the upside potential. Essentially, it replaces the symmetric, non-coherent standard deviation as the risk measure with the coherent downside Conditional VaR.

$$STARR = \frac{E(r_p) - r}{CVaR_\gamma(r_p)} \; or \; STARR = \frac{E(X)}{CVaR_\gamma(X)} \tag{24}$$

3.5.3 Rachev's R- Ratio

The Rachev Ratio (R-Ratio) can be interpreted as the ratio of the expected tail return above a certain threshold level λ_1, divided by the expected tail loss beyond some threshold level λ_2. The part of the distribution between λ_1 and λ_2 is not considered. (Rachev *et al.* 2005) It does reward extreme rewards adjusted for extreme losses, so it enables modeling different levels of an investor's risk aversion or tolerance.

$$R = \frac{CVaR_{\lambda_1}(r_f - r_p)}{CVaR_{\lambda_2}(r_p - r_f)} \text{ or } R = \frac{CVaR_{\lambda_1}(-X)}{CVaR_{\lambda_2}(X)} \tag{25}$$

The idea behind the R-ratio is to try to simultaneously maximize the level of return and get insurance for the maximum loss. (Biglova *et al.* 2004)

One thing to note is that the STARR and R-Ratio only assume finite mean of the return distribution and require no assumption on the second moment. So, they can evaluate return distributions of assets with heavy tails. In comparison, the Sharpe ratio is defined for returns having a finite second moment, which limits its usefulness. The STARR and R-Ratios exhibit better risk-adjusted performance because they are compliant with the ability of their respective coherent risk measures to capture distributional features of data, including the part of risk due to heavy tails; the R-Ratio also adds the ability to adjust for upside reward and downside risk simultaneously. (Rachev *et al.* 2005)

3.5.4 Alpha VaR Ratio and Beta VaR Ratio

As Cherny & Madan suggested Alpha VaR as a substitute for CVaR, it is logical to develop a reward-to-risk measure using Alpha VaR. Following in the manner of Rachev's R-Ratio (by using Alpha VaR on both the loss and gain tails), we therefore propose a new portfolio performance measure, the AVaR-Ratio (AVR), defined as follows:

$$AVR = \frac{AVaR_{\alpha_1}(r_f - r_p)}{AVaR_{\alpha_2}(r_p - r_f)} \text{ or } AVR = \frac{AVaR_{\alpha_1}(-X)}{AVaR_{\alpha_2}(X)} \tag{26}$$

Going one step further, using Beta VaR results in another proposed measure, BVaR-Ratio (BVR), as shown here:

$$BVR = \frac{BVaR_{\alpha_1,\beta_1}(r_f - r_p)}{BVaR_{\alpha_2,\beta_1}(r_p - r_f)} \; or \; BVR = \frac{BVaR_{\alpha_1,\beta_1}(-X)}{BVaR_{\alpha_2,\beta_1}(X)} \tag{27}$$

.They too are based on coherent measures of risk, but they have advantages over CVaR, as previously outlined in the section on risk measures. By careful choice of parameters, nearly any investor's risk preferences for both gains and losses can be attained. These performance measures will be considered, along with the others discussed, using the TSP portfolio optimization problem.

Chapter 4: Results

4.1 Overview of Empirical Work

To provide an answer to research question 1, two portfolios are constructed using models where returns are assumed to be Gaussian. First, the TSP "Market Portfolio" described earlier is developed, as this is what TSP participants as a whole are actually doing with their money. Admittedly, this includes a wide range of investors, from new employees to those already retired. Also, the large allocation to the G Fund may reflect the fact that this is the default fund for new TSP participants. What if they instead were to invest as the FRTIB suggests, using the appropriate L Fund? For someone entering the service and planning on a 20-year career, the L 2030 Fund would be most appropriate. This is the second Gaussian portfolio used for comparison with the proposed VG-ICA portfolio, whose construction is described below.

4.2 VG-ICA Model for TSP Funds

4.2.1 Daily Return Data

Daily data (as opposed to monthly or yearly) is used because they can better capture the distributional properties . (Rachev *et al.* 2005) Figure 17 shows the time series of 4235 daily returns for each of the five core TSP funds for the period 1989-

2005. The G fund appears as a flat line because all five time series have been created on the same vertical scale to show scaling relative to one another. A rescaled version of the G fund returns are shown in Figure 18.

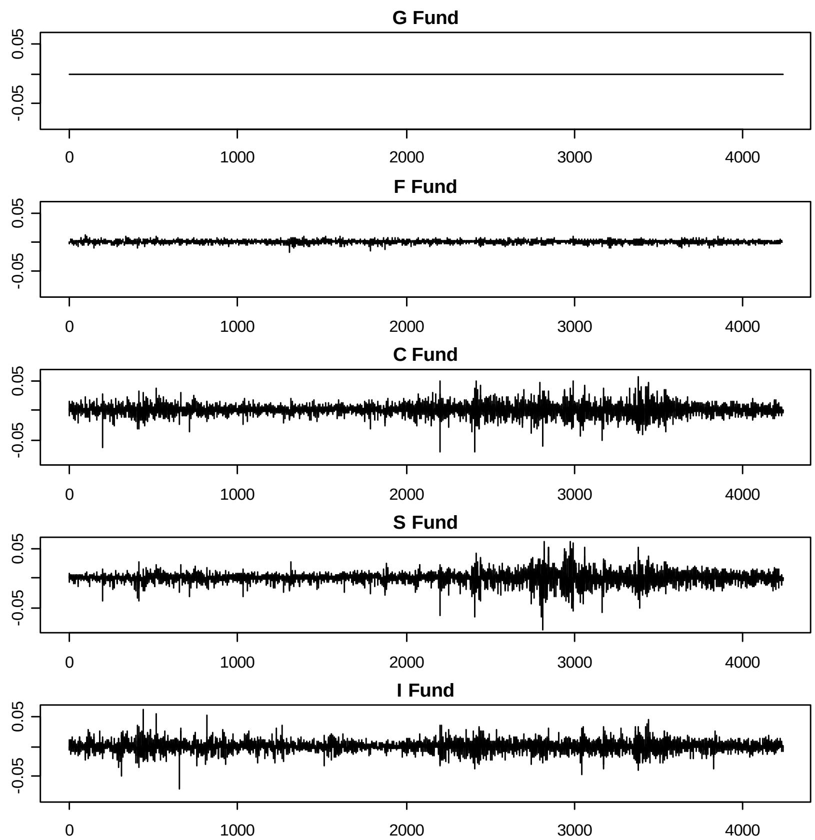

Figure 17. Time series of core TSP fund returns, 1989-2005.
Note: Horizontal axis units are days; vertical axis units are daily returns.

Even with appropriate scaling, the returns for the G fund appear much different than for the other funds. Notice that they are all positive and more stable than the returns of the other four funds. Due to the riskless nature of the G fund, this is somewhat to be expected. The difference is exaggerated by the fact that only monthly returns data were available for the G fund. These were converted to daily returns as described earlier. From a macro level, the G fund returns illustrate the decline of riskless rate over the time period considered.

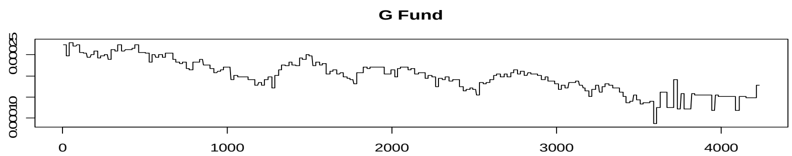

Figure 18. Time series of G fund returns, 1989-2005.
Note: Horizontal axis units are days; vertical axis units are daily returns.

Shown in Table 4 are the first four moments of the daily returns for each of the five core TSP funds. Notice that all funds have positive means over the time period of interest. Also, the funds generally increase in variance, the traditional measure of riskiness, in the order listed (with the exception of the I fund).

Fund	Mean	Variance	Skewness	Kurtosis
G Fund	0.00020	0.00000	-0.14865	2.54154
F Fund	0.00031	0.00001	-0.26061	5.14093
C Fund	0.00033	0.00010	-0.15973	7.08139
S Fund	0.00035	0.00011	-0.41086	9.12962
I Fund	0.00009	0.00009	-0.13723	6.24030

Table 4. Moments of Daily TSP Fund Returns, 1989-2005.

As expected, the returns of the G Fund (the risk-free, government bond fund) are approximately Normal, as can be seen in Figure 19, where a histogram and kernel density estimate (KDE) of the daily returns are shown, with a Normal distribution having the same mean and variance provided for comparison purposes.

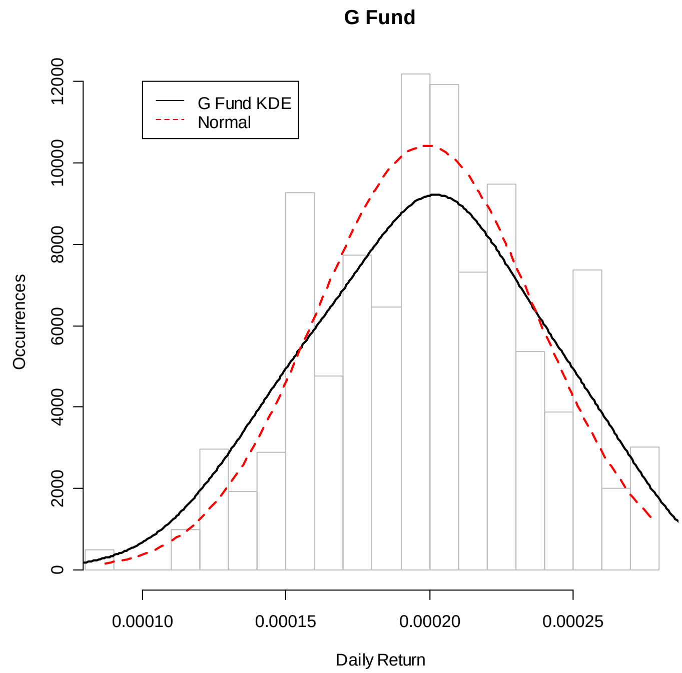

Figure 19. Distribution of Daily Returns for G Fund.

As expected, each of the risky funds has slight negative skewness and excess kurtosis relative to the Normal distribution (i.e. greater than 3). This can be seen in the peaked modes and heavier tails of histograms of the returns of the risky assets (F, C, S, and I funds), as shown in Figure 20 through Figure 23. As before, the solid line represents a smoothed kernel density function of the returns; a Normal density with the same mean and variance is overlaid with a dashed line for comparison purposes.

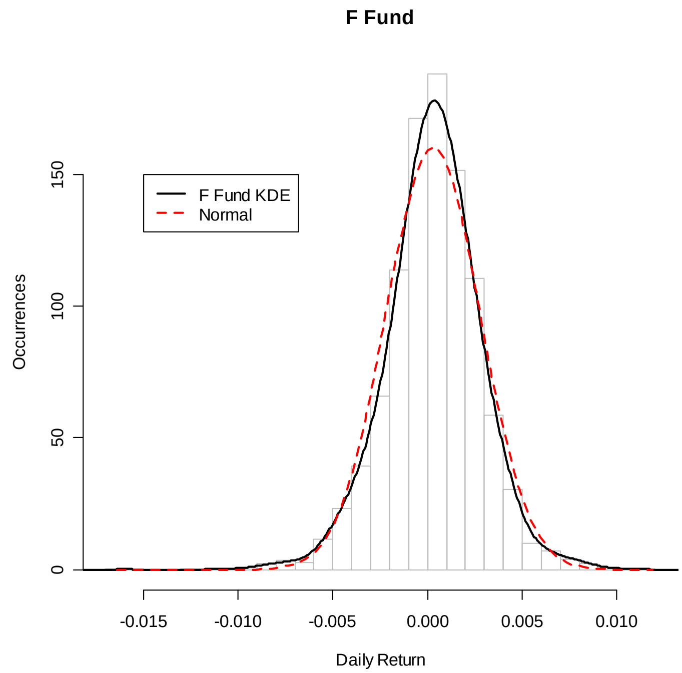

Figure 20. Distribution of Daily Returns for F Fund.

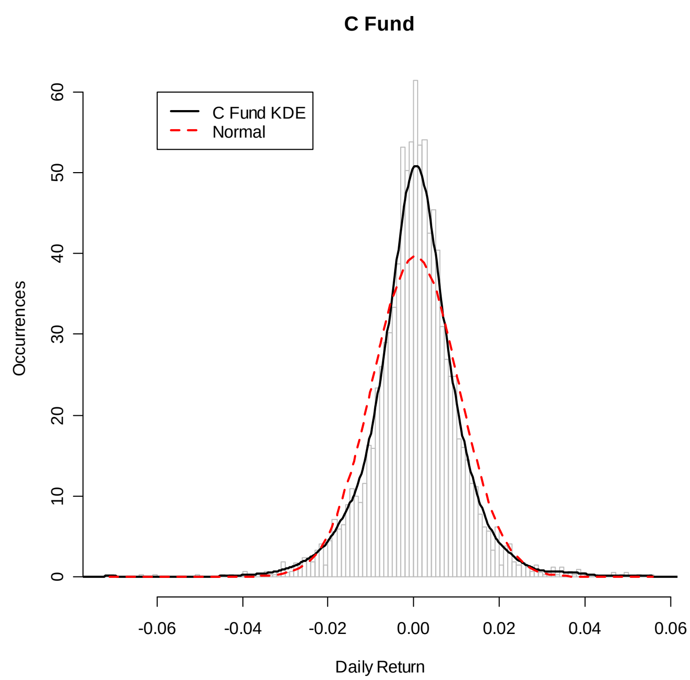

Figure 21. Distribution of Daily Returns for C Fund.

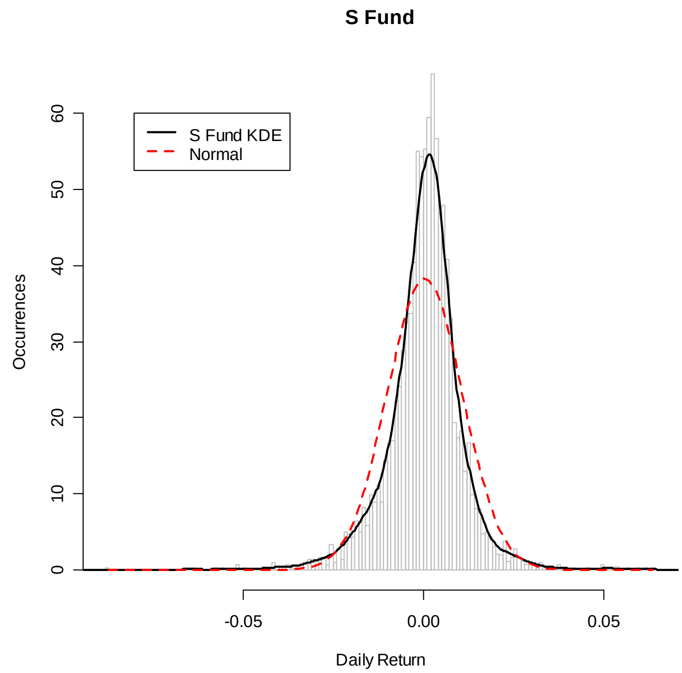

Figure 22. Distribution of Daily Returns for S Fund.

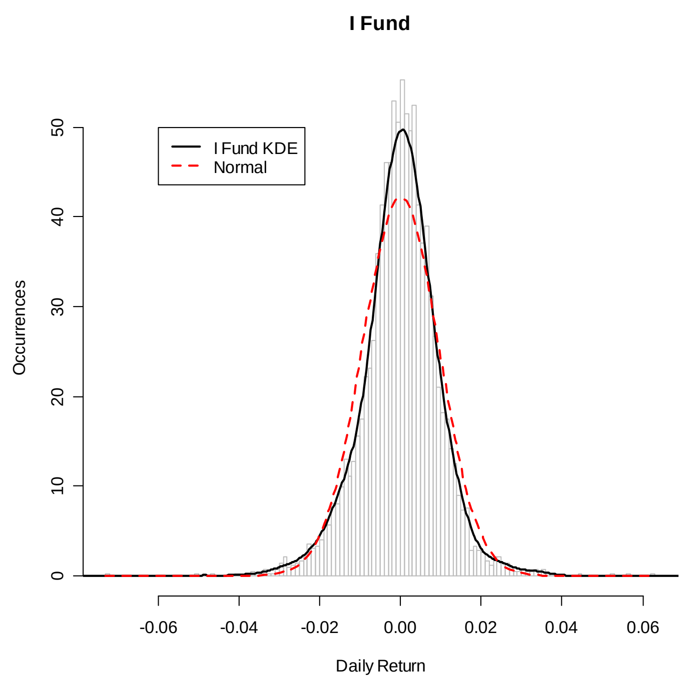

Figure 23. Distribution of Daily Returns for I Fund.

4.2.2 Construction of Independent Components

ICA is then used to generate five independent components (IC), using the *ln(cosh)* function described in Equation (8) as the measure of non-Gaussianity. The first four moments of each of the ICs are shown in Table 5. By design, the ICs have zero mean, unit variance, and are orthogonal.

IC#	Mean	Variance	Kurtosis	Skewness
1	0	1	14.4	-0.589
2	0	1	7.3	-0.055
3	0	1	6.1	-0.098
4	0	1	5.5	-0.258
5	0	1	2.54	-0.144

Table 5. First Four Moments of Individual Components.

To estimate the coefficient matrix B and the covariance matrix of the residuals Σ, we performed a regression of the four risky assets on the retained factors (ICs).

# ICs Kept	F Fund	C Fund	S Fund	I Fund
5	1	1	1	1
4	.9980	.9994	.9998	.9949
3	.0611	.8983	.9652	.9869
2	.0611	.8962	.9629	.1276

Table 6. Adjusted R^2 values from regression of de-meaned returns on IC1-IC4.

As seen in Table 6, keeping all 5 ICs gives a perfect fit of the data and should lead to good predictive ability. However, it results in a completely specified model, which for other applications (e.g. with a significantly greater number of assets) would not be feasible. In this case, there is no need for dimension reduction, but it appears that keeping 4 ICs should result in adequate results. Keeping 3 or less ICs has significant problems with the F Fund; dropping to 2 ICs or fewer results in a poor fit of the I Fund. Here is the resulting coefficient matrix:

$$W = \begin{bmatrix} & \text{G fund} & \text{F fund} & \text{C fund} & \text{S fund} & \text{I fund} \\ \text{Intercept} & 0 & 0 & 0 & 0 & 0 \\ \text{IC}1 & .000002 & .0000087 & -.003369 & -.008446 & -.001584 \\ \text{IC}2 & .000003 & .000613 & -.008905 & -.005761 & -.002988 \\ \text{IC}3 & -.000003 & -.000008 & .000462 & -.000492 & -.008761 \\ \text{IC}4 & 0 & -.002411 & -.003196 & -.001938 & -.000846 \end{bmatrix}$$

Figure 24. Matrix of Coefficients From Regressing Core Funds on ICs.

Figure 25 is the ICA analogy to a scree plot[4] in PCA, except that kurtosis instead of percent variance explained (or eigenvalues) is shown on the vertical axis. Using the kurtosis level of the Gaussian distribution, three, as a cutoff (shown as the dashed, horizontal line) suggests four ICs be retained in the factor model (IC1-IC4). This means that we will not have a Gaussian noise term in our model as originally hypothesized. This is fine, as the number of factors is small and no dimension reduction is required for computational tractability.

[4] Scree is the little rocks and pieces of soil that accumulate at the bottom of a steep hill. In PCA, one of the standard ways to decide how many components to keep is to look for the bend in the curve.

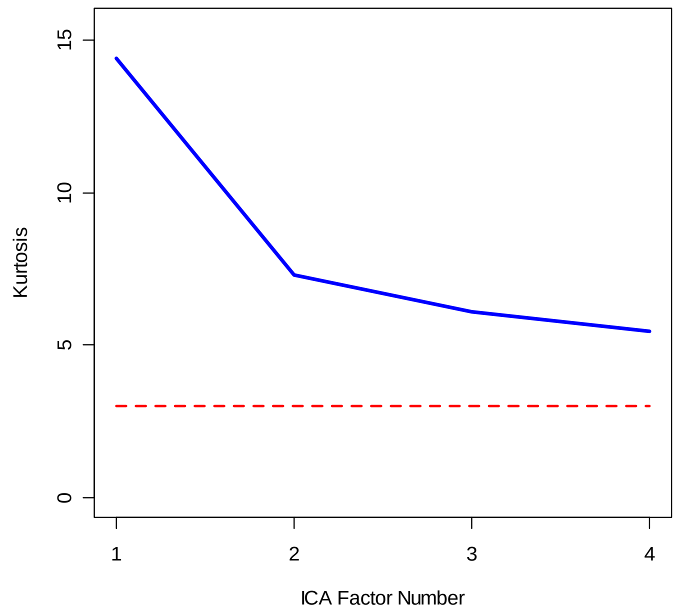

Figure 25. Kurtosis of Independent Components.

Confirmation of the decision to keep four ICs is provided by the fact that all entries in the covariance matrix of the residuals are relatively small in magnitude, as shown in Figure 26

$$Cov(\mathbf{E}) = \mathbf{\Sigma} = \begin{bmatrix} 6.1\times10^{-5} & 3.1\times10^{-5} & 5.8\times10^{-5} & 6.2\times10^{-6} & -2.3\times10^{-5} \\ 3.1\times10^{-5} & 5.2\times10^{-5} & 6.6\times10^{-5} & 5.8\times10^{-5} & -2.1\times10^{-5} \\ 5.8\times10^{-5} & 6.6\times10^{-5} & 2.7\times10^{-4} & 9.6\times10^{-6} & -6.2\times10^{-4} \\ 6.2\times10^{-6} & 5.8\times10^{-5} & 9.6\times10^{-5} & 1.0\times10^{-4} & 1.5\times10^{-4} \\ -2.3\times10^{-5} & -2.1\times10^{-5} & -6.2\times10^{-4} & 1.5\times10^{-4} & 1.9\times10^{-3} \end{bmatrix}$$

Figure 26. Covariance Matrix of Residuals.

4.2.3 Fitting VG Distribution to Components

So, with four ICs retained, the VG distribution is fit to each of them by maximum likelihood estimation (MLE) using the closed form of the VG density function. (Madan, Carr and Chang 1998) For each of the components, Figure 27 shows the data (small, black circles) overlaid with the fitted VG distribution (dashed blue line) and a Normal(0,1) distribution (dotted red line) for comparison.

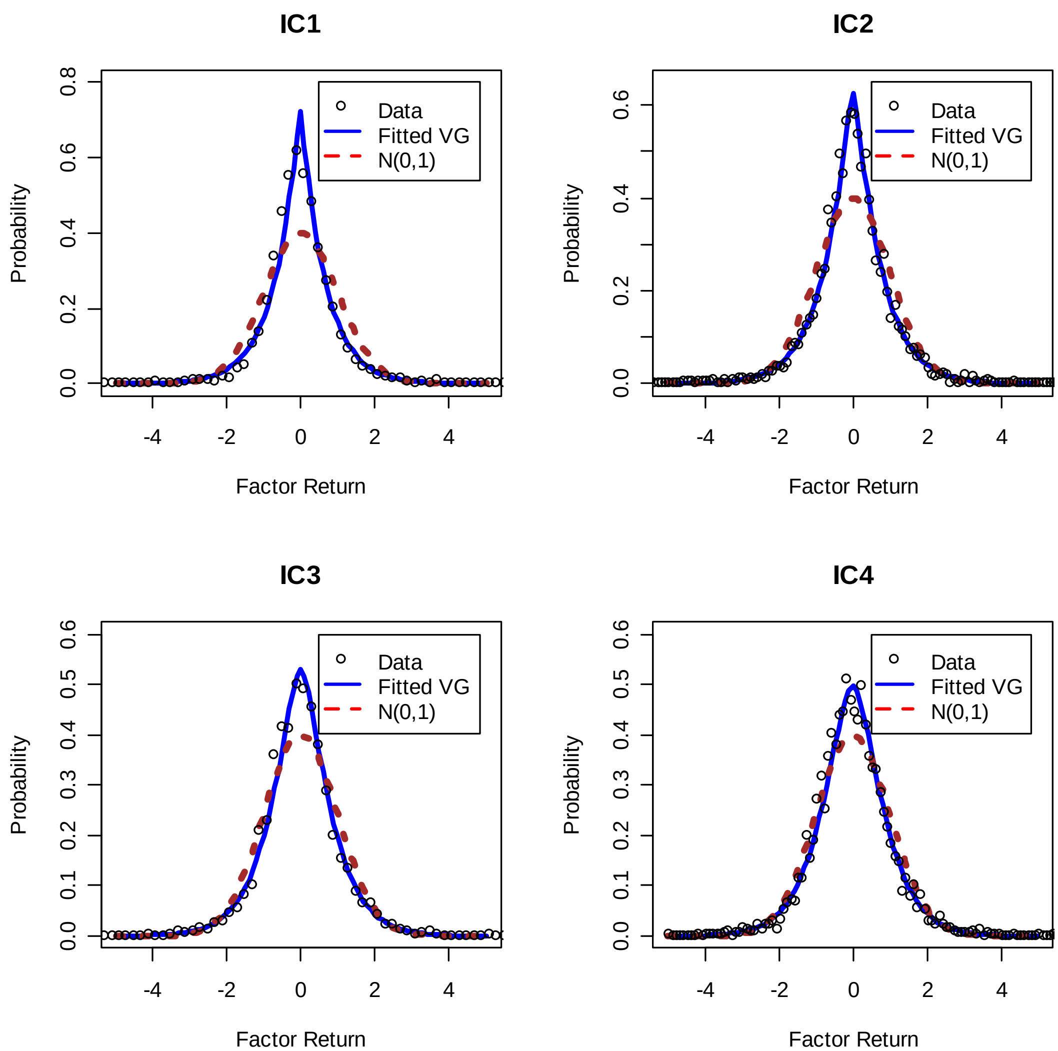

Figure 27. Independent Component Histograms Overlaid with Fitted VG and N(0,1).

While visually comparing the fit of the VG and Gaussian distributions to the IC data using their densities or a Q-Q plot is useful, the use of a goodness-of-fit test provides a formal assessment of whether the data are an independent sample from a specific distribution. One of the more commonly used goodness-of-fit tests is the chi-square (χ^2) test of Pearson; as we have a large number of data points, this test may be more appropriate than other tests, like the Kolmogorov-Smirnov or Anderson-Darling tests. (Law and Kelton 2000) The first step in the procedure is to bin the data into k adjacent intervals. Let N_j be the number of observed data points in the j-th interval. Also required is the expected proportion p_j of the data points that would fall in the j-th interval if sampling from the hypothesized distribution. Then, with n sample data points, np_j is the number of points that would be expected to fall in the j-th interval. The k intervals should be chosen so that $k > 3$, $np_j > 5$, and they are equi-probable rather than equally spaced. The chi-square test statistic is then calculated as:

$$\chi^2 = \sum_{j=1}^{k} \frac{(N_j - np_j)^2}{np_j} \tag{28}$$

Since we needed to estimate the three parameters of the VG distribution, the number of degrees of freedom is approximately k-m-1, where m is the number of parameters estimated. Table 7 gives the results of the chi-square test on the ICs for both the Normal(0,1) and VG with estimated parameters. The test was conducted with $k = 20$ intervals, so the degrees of freedom for the chi-square RV is 20 – 3 – 1 = 16. Clearly, the VG distribution does a much better job of fitting the IC data than the Normal(0,1) distribution.

IC#	Fitted VG Parameters (Annualized)			χ^2 Statistic ($\chi^2_{.01,17} = 33.41$)	
	σ	ν	Θ	VG(σ,ν,θ)	Normal(0,1)
IC1	14.814	0.00385	-3.774	99.01	546.89
IC2	15.558	0.00326	-0.222	13.37	295.73
IC3	15.704	0.00232	-1.019	34.94	171.70
IC4	15.739	0.00186	-1.149	25.29	118.07

Table 7. Fitted VG Parameters and X^2 Goodness of Fit Test Statistics.

4.2.4 Scaling Law

Construction of independent components and fitting of VG distributions has all been done with daily return data. Let X_d be the VG distribution for daily returns with parameters σ_d, ν_d, and θ_d. In some cases, we wish to examine other periods, like months, or years. In these cases, we use a scaling law that says the distribution for any time horizon *h* is $X_h \sim \sqrt{h} X_d$. So, X_d is VG with $\sigma_h = \sigma_d \sqrt{h}$, $\nu_h = \nu_d$, and $\theta_h = \theta_d \sqrt{h}$. For example, $X_{year} \sim \sqrt{252} X_d$, as there are 252 trading days in a year on average. Similar scaling relationships can be derived for a month using 21 trading days.

4.2.5 Choice of Risk Aversion Parameter

Confirmation of our assumed risk aversion parameter was obtained as follows. By performing mean-variance optimization on the annual returns data, the expected return and standard deviation of each of the five L funds portfolios were calculated. Then, given the investment horizon of 20-30 years (represented by the L2020, L2030, and L2040 portfolios), the implied risk aversion coefficient range of 2 to 3 is obtained, as indicated by the arrows in Figure 28.

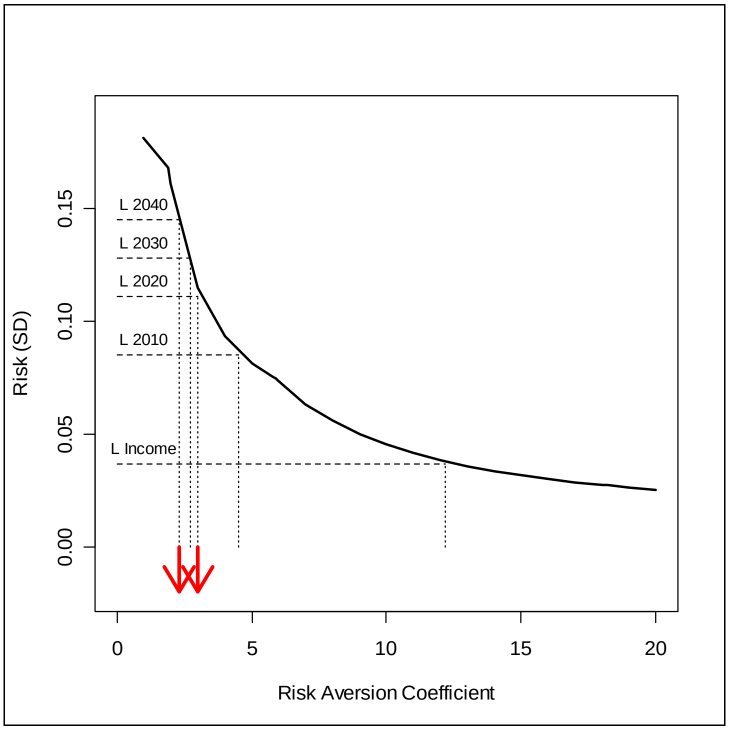

Figure 28. Implied Risk Aversion Parameter from L Funds Standard Deviation.

4.2.6 Optimal VG-ICA Portfolios

The optimal VG-ICA portfolio is computed by maximizing expected utility of a negative exponential utility function with risk aversion parameter 2, as previously outlined in section 1.5. The procedure used is provided as Theorem 6.2 in Yen. (Yen 2004) The choice of the optimal position, y_i, in each asset *i* is found by maximizing the expected exponential utility with risk aversion coefficient η, or the function:

$$1 - e^{-\eta y'(\mu - r)} E[e^{-\eta y' x}] = 1 - e^{-\eta y'(u-r)} E[e^{=\eta y' As}], \tag{29}$$

where r is the riskless rate, s are the independent components and matrix A is as described in Section 2.2.4, the introduction to ICA. The resulting optimal position for each independent component i is given by:

$$\tilde{y}_i = \frac{\theta_i}{\nu_i} - \frac{1}{\eta \zeta_i \nu_i} \pm \sqrt{\left(\frac{\theta_i}{\sigma_i^2} - \frac{1}{\eta \zeta_i \nu_i} \right)^2 + 2 \frac{\zeta_i + \frac{\theta_i}{\eta}}{\zeta_i \sigma_i^2 \nu_i}}, \tag{30}$$

with $\zeta = A^{-1} \frac{\mu - r}{\eta} - \frac{\theta}{\eta}$. The only modification was to limit any assets from optimal short positions by constraining them to zero (one at a time, starting with the most negative) and re-optimizing until all positions were long. The resulting position is then normalized (by dividing by the sum) so that allocations in terms of percentages were identified. Although the optimal positions $\tilde{y}_i$ varied with changes in the risk aversion parameter η, the optimal portfolio allocations (in terms of percentages) did not.

Table 8 shows: optimal VG-ICA portfolios, including the one computed from daily data and those scaled to annual and 20-year horizons; the riskless portfolio; the TSP “Market Portfolio”; the L 2030 portfolio (the recommended L Fund for those with this investment horizon); and the L 2040 portfolio (the L Fund that is closest in composition to the VG-ICA portfolio). Due to the similarity of the scaled VG-ICA portfolios and the computational tractability of the daily model, the scaled portfolios were not used.

Model	G Fund	F Fund	C Fund	S Fund	I Fund
VG-ICA (Daily)	0%	1%	43%	30%	26%
VG-ICA (Scaled to Annual)	0%	1%	44%	40%	25%
VG-ICA (Scaled to 20 years)	0%	0%	46%	31%	23%
Riskless	100%	0%	0%	0%	0%
TSP "Market Portfolio"	39%	9%	49%	8%	8%
L 2030	16%	9%	38%	16%	21%
L 2040	5%	10%	42%	18%	25%

Table 8. All Portfolios for Comparison.

4.2.7 Scenarios for Simulation

Three different scenarios are used to compare the portfolios under consideration, as outlined below:

(1) Scenario 1: $10,000 initial investment; no further contributions; no rebalancing during investment period.

(2) Scenario 2: Same as Scenario 1, but rebalancing the VG-ICA portfolio is performed at yearly, monthly, and daily intervals for comparison with the un-rebalanced VG-ICA portfolio.

(3) Scenario 3: No initial investment, but monthly contributions based on a 10% of an officer's base pay savings rate, increasing with promotion and

longevity raises (see Figure 29); no rebalancing during investment period. Over the 20-year horizon, total contributions would be approximately $170K.

Although Scenario 1 may be unrealistic (as most TSP participants do not have a lump sum to invest at the beginning of their career), it was chosen as a simple-to-understand example. Also, it serves as a basis for comparison in Scenario 2, where rebalancing is considered. Scenario 3 is the way that most investors contribute to the TSP, by allocating a percentage, say 10% (the actual median savings rate of TSP participants), of their base pay to be deducted from their monthly paycheck.

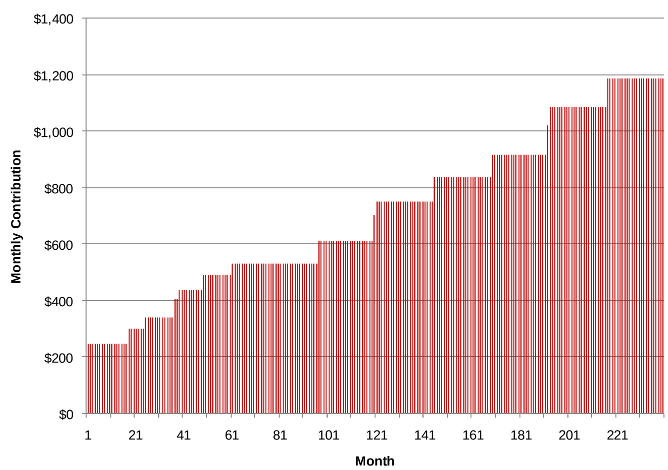

Figure 29. Monthly contributions over 20 years for realistic scenario. Contributions total nearly $170K for typical military officer.

4.3 Results

4.3.1 Scenario 1 Results

As shown in Figure 30, at the 20-year horizon (5040 days), the VG-ICA model results in an expected discounted cash flow of $47,362, which is 228% higher than the TSP "Market Portfolio" and 258% higher than the L2030 portfolio. Also, the potential upside is significantly higher than for the other portfolios (182% and 117%, respectively). However, the downside for the VG-ICA model is slightly worse, which means it is potentially riskier; this will require further examination.

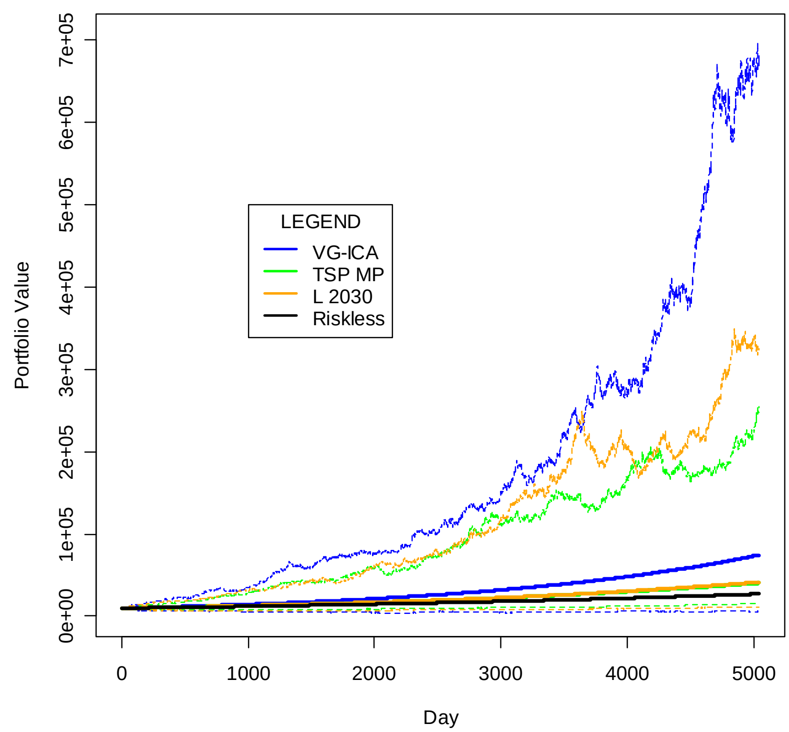

Figure 30. Portfolio Values Over Time, Scenario 1. Heavy solid lines indicate expected portfolio values, while lighter dotted lines above and below indicate both up- and down-side potential of portfolios.

Figure 31 are the empirical distribution functions of each of the portfolios at end of the 20-year simulation. The dashed vertical lines represent the mean of each distribution. Notice how the right tail of the VG-ICA cash flow is significantly longer than the other two. Also, the red vertical line at 0 represents the performance of the riskless asset.

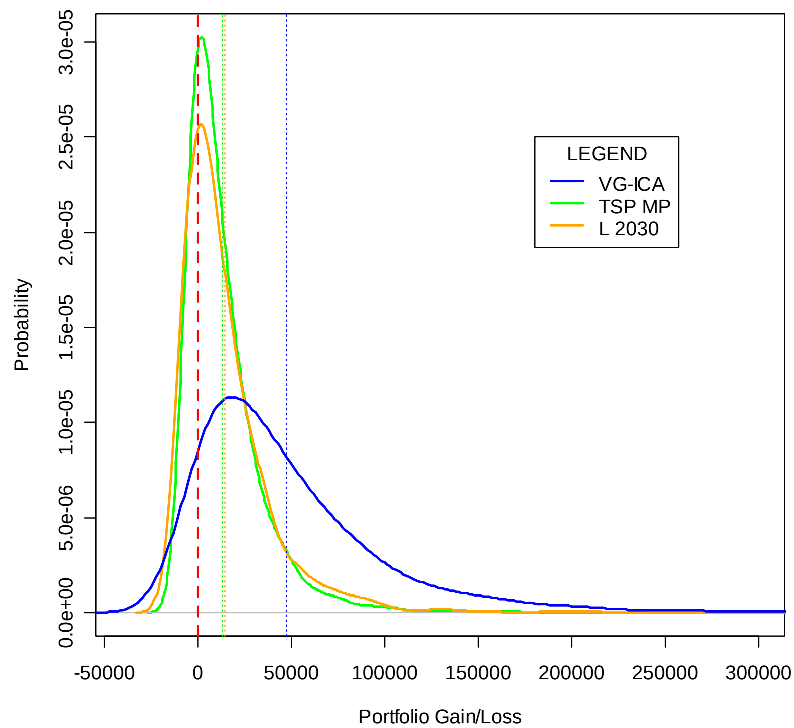

Figure 31. PDF Comparison, Scenario 1. Red, vertical line at 0 is the value of the riskless portfolio.

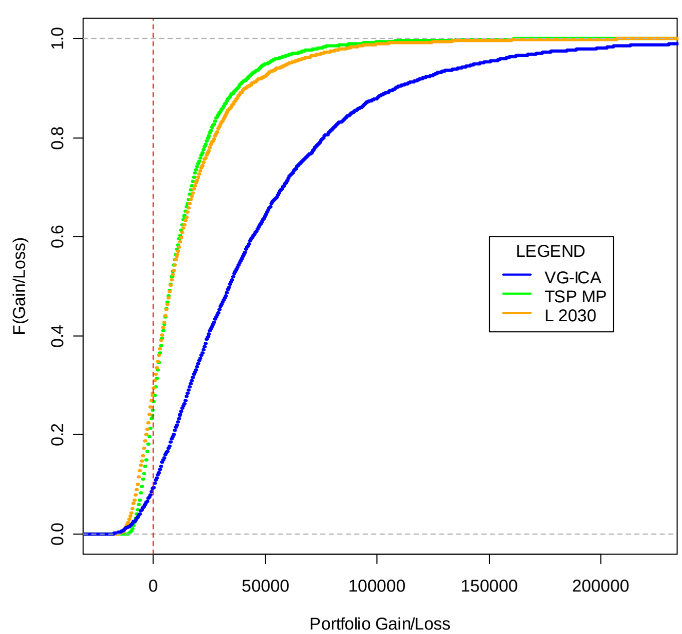

Figure 32. CDF Comparison, Scenario 1. Red, vertical line at 0 is value of the riskless portfolio.

Although this is interesting to look at to see the general character of each of the discounted cash flow distributions, one cannot easily glean all desired information from the empirical PDF. Additional insights can be gained by examining the empirical CDFs, including some direct interpretations of risk measures. For example, in Figure 32 it is possible to see that the VG-ICA portfolio (the blue curve) has a lower probability of losing money (relative to the riskless portfolio) than the other two investment alternatives. It also appears, at first glance, that perhaps the VG-ICA portfolio stochastically dominates the other distributions. However, zooming in on the lower tail of the distributions, as is done in **Error! Reference source not found.**, reveals that this is not the case. Obviously, as all three CDFs cross one another, FOSD does not apply. More details on other stochastic dominance orders will follow in a later section.

It is easy to observe that the VG-ICA portfolio has approximately a 10% probability of "losing" money (relative to the riskless portfolio), whereas the probability for the other two portfolios are significantly higher, at 26% for the TSP "Market Portfolio" and 28% for the L 2030 portfolio. It is also easy to see, by imagining a horizontal line at 0.05), that the 95% VaR for the VG-ICA portfolio will be greater (i.e. less negative) than for the other two portfolios.

This graphical analysis, although enlightening, will now be augmented with the more formal techniques for comparing portfolio outcomes as described in Chapter 3.

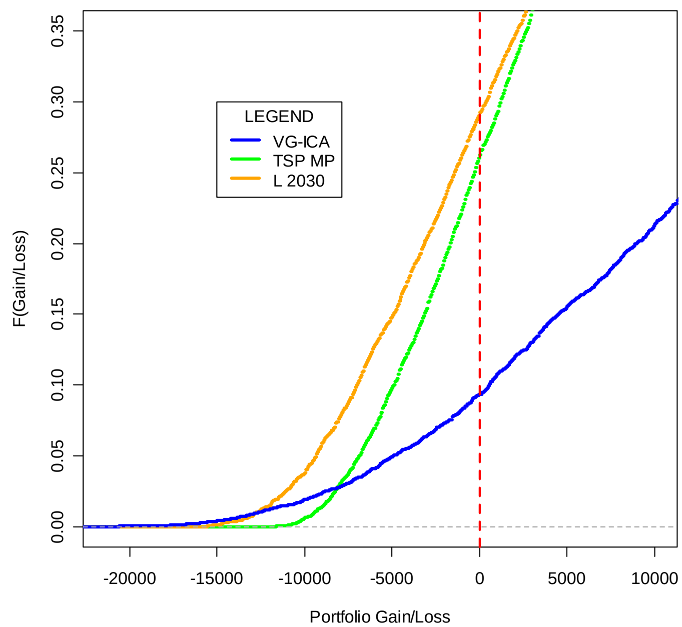

Figure 33. Zoomed CDF Comparison, Scenario 3. Red, vertical line at 0 is value of the riskless portfolio.

4.3.2 Scenario 2 Results

As the TSP allows investors to rebalance at no cost, Scenario 2 was designed to test if there was an advantage to be gained by periodically rebalancing the VG-ICA portfolio to an optimal ratio. Periods tested included: yearly, monthly, and daily. As shown in Figure 34, it appears that there is actually a disadvantage to doing so. The expected values of the rebalanced portfolios are 24%-26% lower than the portfolio that is not rebalanced. Also, the potential upside is lowered by 41%-42% relative to the portfolio that is not rebalanced.

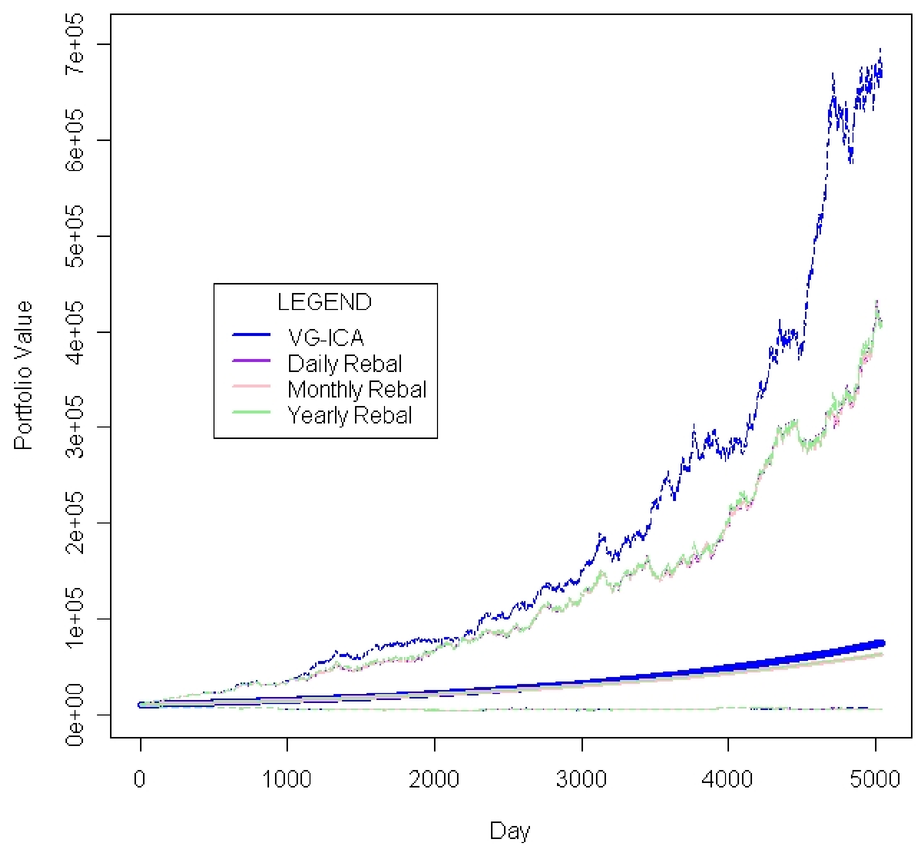

Figure 34. Simulated Portfolio Values Over Time, Scenario 2. Heavy solid lines indicate expected portfolio values, while lighter dotted lines above and below indicate both up- and down-side potential of portfolios.

Table 9 shows the change in the VG-ICA portfolio after 5040 days (20 years) if not rebalanced. Notice the move away from the I fund and toward the C and S funds.

Portfolio	G Fund	F Fund	C Fund	S Fund	I Fund
Original VG-ICA	0%	1%	43%	30%	24%
After 20 years	0%	1%	53%	43%	3%

Table 9. Change in VG-ICA Portfolio Over Time.

To see what would happen, the process was allowed to continue for another 20 years; this time the shift was from the C, with the S fund being the beneficiary, as seen in Table 10. The consistency of this test was verified with different random number streams.

Portfolio	G Fund	F Fund	C Fund	S Fund	I Fund
After 20 years	0%	1%	53%	43%	3%
After 40 years	0%	1%	38%	54%	3%

Table 10. Continued Evolution of VG-ICA Portfolio Over Time.

The CDF comparisons in Figure 35 and Figure 36 show that each of the three rebalanced portfolios are nearly identical, but have a larger probability (12%) of "losing" money (versus the riskless asset) than the VG-ICA portfolio that is no rebalanced (9%). As it appears to be disadvantageous and increases the computational requirement, rebalancing is not recommended and is, therefore, not considered in Scenario 3.

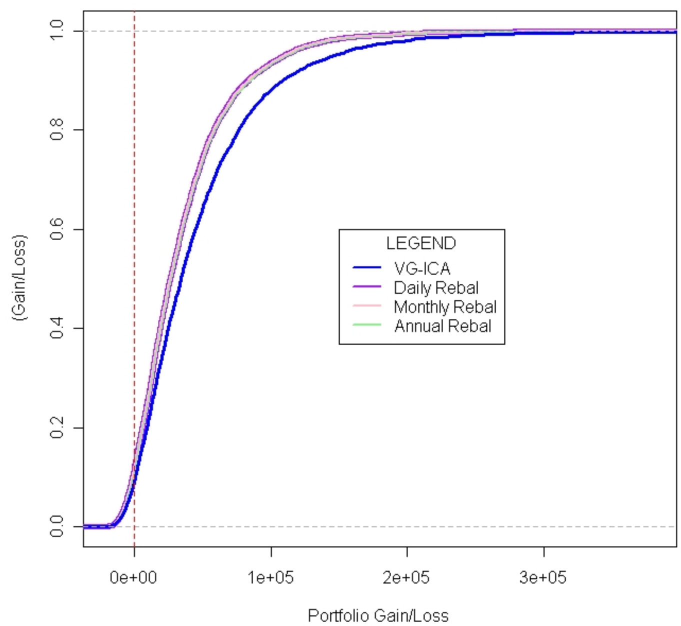

Figure 35. CDF Comparison, Scenario 2. Red, vertical line at 0 is value of riskless portfolio.

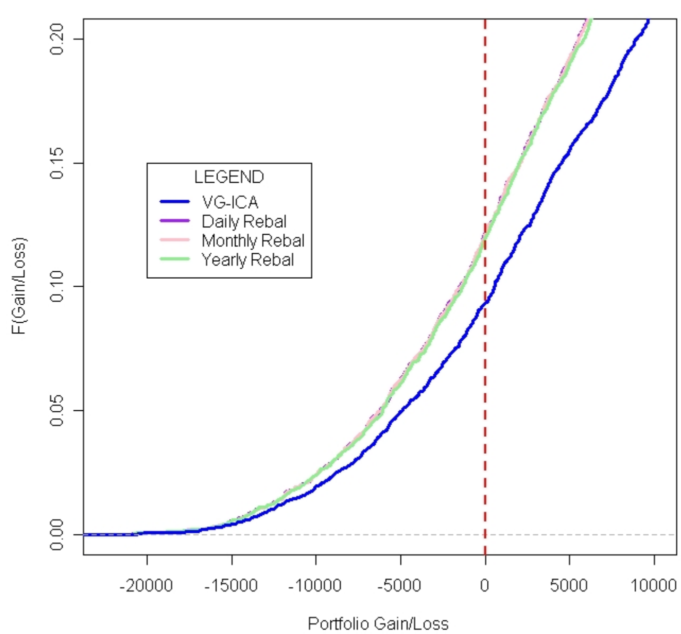

Figure 36. Zoomed CDF Comparison, Scenario 2. Red, vertical line at 0 is value of riskless portfolio.

4.3.3 Scenario 3 Results

As shown in Figure 37, the VG-ICA once again has the highest expected value at the end of the 20 year horizon. In this case, it exceeds the expected values of both the TSP "Market Portfolio" and L 2030 and L2040 portfolio by approximately 33%. As before, the upside is also significantly higher, with the VG-ICA besting the TSP "Market Portfolio" by 82% and the L Fund portfolios by 69%. Also as seen in Scenario 1, the downside for the VG-ICA is slightly worse than for the other portfolios. This will be examined both graphically and with reward-to-risk measures.

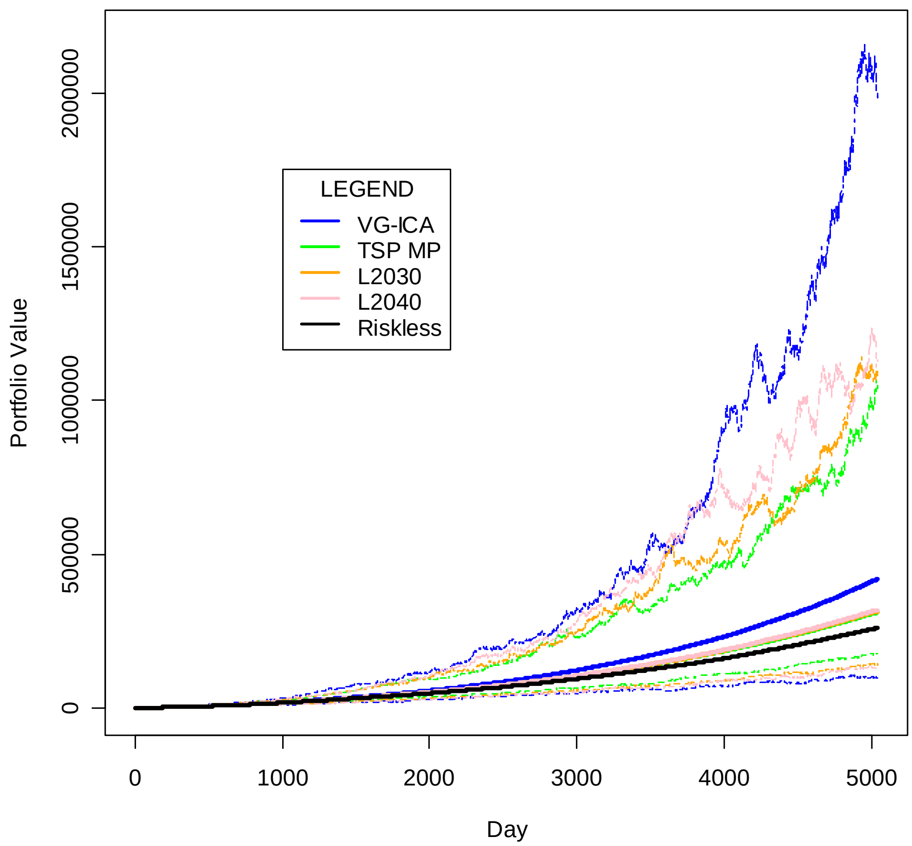

Figure 25. Simulated Portfolio Values Over Time, Scenario 3. Heavy solid lines indicate expected portfolio values, while lighter dotted lines above and below indicate both up- and down-side potential of portfolios.

Figure 37 presents are the empirical distribution functions (PDFs) of each of the discounted portfolio values at end of the 20-year simulation. As before, the dashed vertical lines represent the mean of each distribution and the red vertical line at 0 represents the performance of the riskless asset. Once again, the right (gain) tail of the VG-ICA cash flow is significantly longer than the other two.

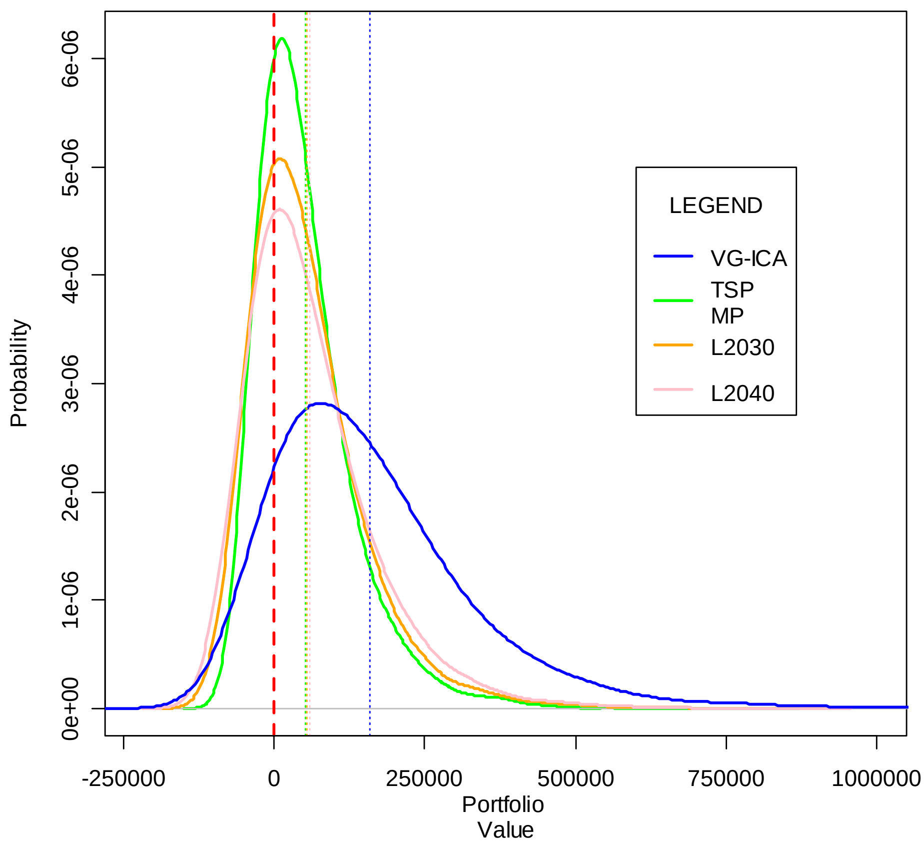

Figure 37. PDF Comparison, Scenario 2. Red, vertical line at 0 is value of riskless portfolio.

Considering the empirical CDFs in this scenario again provides additional information. In Figure 38, it is clear that the VG-ICA portfolio (the blue curve) has a lower probability of “losing” money (relative to the riskless portfolio) than the other two investment alternatives. As was the case with Scenario 1, it is again difficult to tell what is going on down in the tails of the distribution, so Figure 39 provides a closer look at that area.

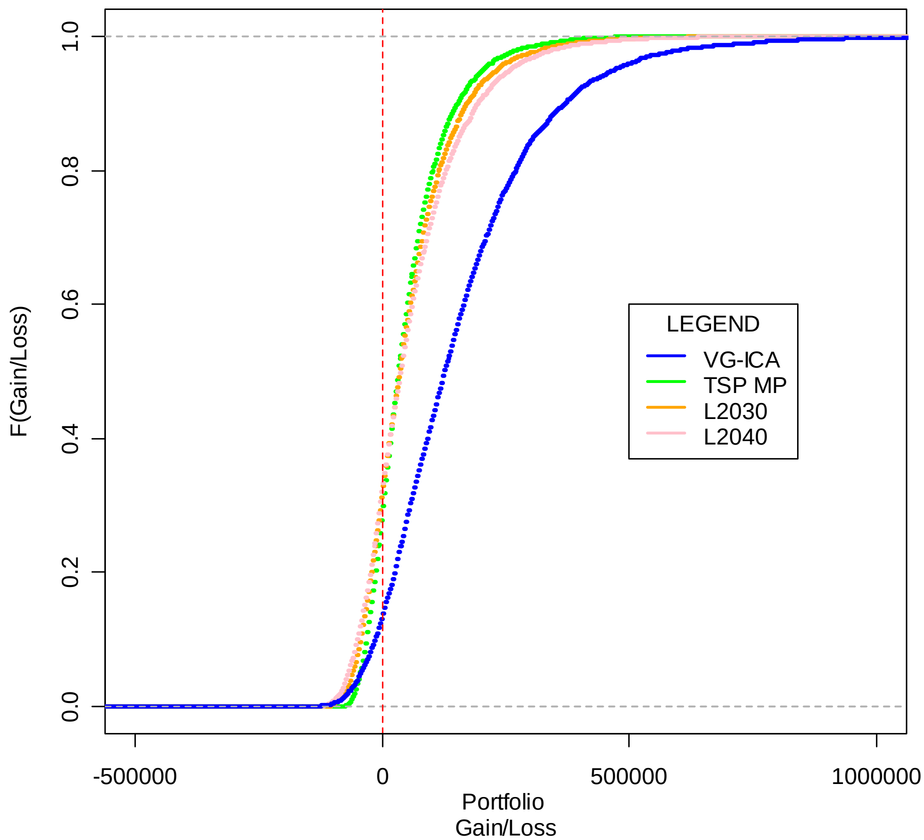

Figure 38. CDF Comparison, Scenario 3. Red, vertical line at 0 is value of riskless portfolio.

In this case, the VG-ICA portfolio has approximately a 13% probability of underperforming the riskless portfolio, while this likelihood for the TSP “Market Portfolio” and the L 2030 and L2040 portfolio are between 28% and 32%. Using the technique previously described, the 95% VaR for the VG-ICA portfolio will be approximately the same as for the TSP “Market Portfolio” and greater (i.e. less negative) than for L 2030 and L 2040 portfolios.

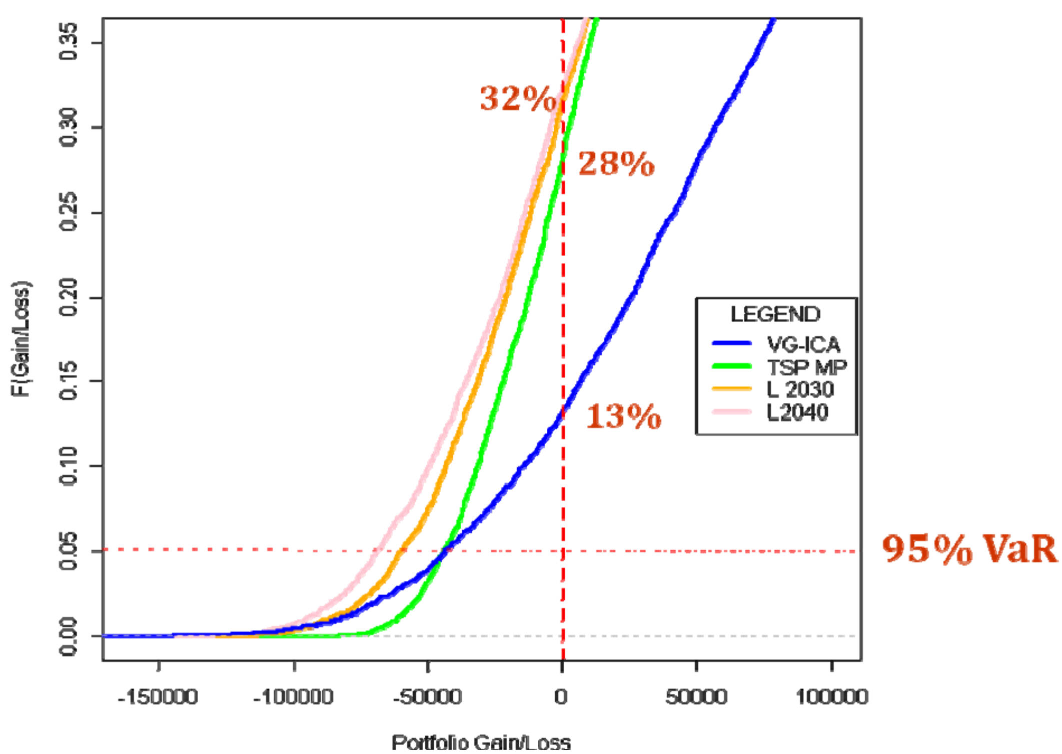

Figure 39. Zoomed CDF Comparison, Scenario 3. Red, vertical line at 0 is value of riskless portfolio.

<u>4.4 Comparison Using Stochastic Dominance</u>

Comparison the portfolios’ performance using stochastic dominance first makes sense, as it relies on few assumptions and, therefore, any results obtained are strong. However, it turns out that there are no conclusive results from comparing these portfolios from a stochastic dominance perspective. In Table 11, each pair of letters represents whether the option on the left dominates the alternative in the column header in a first-, second-, and third-order sense, respectively. As shown, none of the portfolios considered (VG-ICA, TSP Market Portfolio, L2030, and L2040) stochastically dominate each other in a first- or second-order sense.

	VG-ICA	**TSP "Market Portfolio"**	**L 2030**	**L 2040**
VG-ICA		FF	FF	FF
TSP Market Portfolio	FF		FF	FF
L 2030	FF	FF		FF
L 2040	FF	FF	FF	

Table 11. Stochastic Dominance Relationships, Scenarios 1 and 3.

4.5 Comparison Using Reward-to-Risk Ratios

Turning our attention to the various risk and portfolio performance measures described in Chapter 3, the results shown in Table 12 and Table 13 are obtained for the portfolios under consideration in Scenario 1. Although the .05 level of significance was used in these examples; others (.01 and .10) were also tested; although the numbers are different, the qualitative results are similar.

Risk Measure (↓ better)	VG-ICA	TSP "Market Portfolio"	L 2030	L 2040
Std Dev	$ 52,626	$ 20,647	$ 24,421	$ 26,637
95% VaR	$ 4,906	$ 6,875	$ 9,529	$ 10,575
95% CVaR	$ 9,506	$ 8,456	$ 11,579	$ 12,916
AVaR(50)	$ 11,904	$ 8,956	$ 12,121	$ 13,753
BVaR(50,5)	$ 5,012	$ 6,840	$ 9,359	$ 10,462

Table 12. Comparison of Risk Measures, Scenario 1.

Performance Measure (↑ better)	VG-ICA	TSP "Market Portfolio"	L 2030	L 2040
Sharpe Ratio	0.90	0.64	0.59	0.60
STARR Ratio	4.98	5.60	1.25	1.22
R-Ratio(.05,.05)	20.56	5.29	7.23	7.04
AVR(50,50)	21.23	9.98	8.83	8.11
BVR((50,5),(50,5))	32.21	8.52	7.23	7.08

Table 13. Comparison of Portfolio Risk Measures, Scenario 1.

The shaded cells represent the portfolio that had the lowest risk measure (in Table 12) or highest portfolio performance measure (in Table 13). With some of the less informed (i.e., rely only on a dispersion measure or only use one tail, as opposed to the entire distribution) and non-coherent risk and portfolio performance measures, the TSP "Market Portfolio" has the highest ranking. However, once measures that are coherent and consider information about the entire distribution are examined, the VG-

ICA portfolio dominates the other portfolios, as seen with the R-Ratio and the newly proposed AVR and BVR measures. Of note is that in no case do either of the L Fund portfolios (developed using Normally-distributed returns and recommended by the Federal Thrift Retirement Investment Board) turn in the best performance. Similar results were not computed for Scenario 2, given the analysis previously conducted with the graphs of the distributions. In the same manner as for Scenario 1, the following results obtain in Scenario 3, where there is no initial wealth, but monthly contributions over the entire, 20-year period.

The same risk and portfolio performance measures for Scenario 3 (monthly contributions) are given in Table 14 and Table 15. Once again, the L Fund portfolios fail to have the best measure in any instance. Although the "TSP Market Portfolio" seems to have the lowest risk with nearly all measures, once the risk-adjusted rewards are considered, the VG-ICA model again outperforms the other models. The only case where this is not true is in the VaR measure, which only accounts for the size of the loss tail; measures that consider the potential gains all favor the VG-ICA model.

Risk Measure (↓ better)	**VG-ICA**	**TSP "Market Portfolio"**	**L 2030**	**L 2040**
Std Dev	$ 168,885	$ 80,890	$ 94,515	$ 105,525
95% VaR	$ 43,382	$ 44,146	$ 60,378	$ 66,910
95% CVaR	$ 68,056	$ 54,789	$ 74,684	$ 82,783
AVaR(50)	$ 77,575	$ 59,250	$ 81,352	$ 87,754
BVaR(50,5)	$ 43,938	$ 44,010	$ 60,203	$ 67,757

Table 14. Comparison of Risk Measures, Scenario 3.

Performance Measure (↑ better)	VG-ICA	TSP “Market Portfolio”	L 2030	L 2040
Sharpe Ratio	0.94	0.63	0.57	0.54
STARR Ratio	2.33	2.90	0.72	0.69
R-Ratio(.05,.05)	8.87	2.82	4.10	4.07
AVR(50,50)	9.35	5.41	4.62	4.67
BVR((50,5),(50,5))	11.60	5.05	4.24	4.12

Table 15. Comparison of Portfolio Performance Measures, Scenario 3.

When considered in the traditional reward-risk framework of mean-variance (or mean-standard deviation), the VG-ICA portfolio lies where it should be expected to appear, given that it is predominantly a linear combination of the C, S, and I Funds. Also, the TSP “Market Portfolio” falls roughly halfway between the L 2010 and L 2020 portfolios, which is no surprise, given its composition.

However, when considered in one of the new proposed reward-risk frameworks, specifically, using a version of AlphaVaR as both the measure of risk and the measure of reward, the picture changes dramatically, as seen in Figure 41. Here, the VG-ICA and F Fund portfolios potentially lie on the hypothesized efficient Alpha VaR frontier. This indicates the fact that different portions of the returns distribution have been (de-)emphasized.

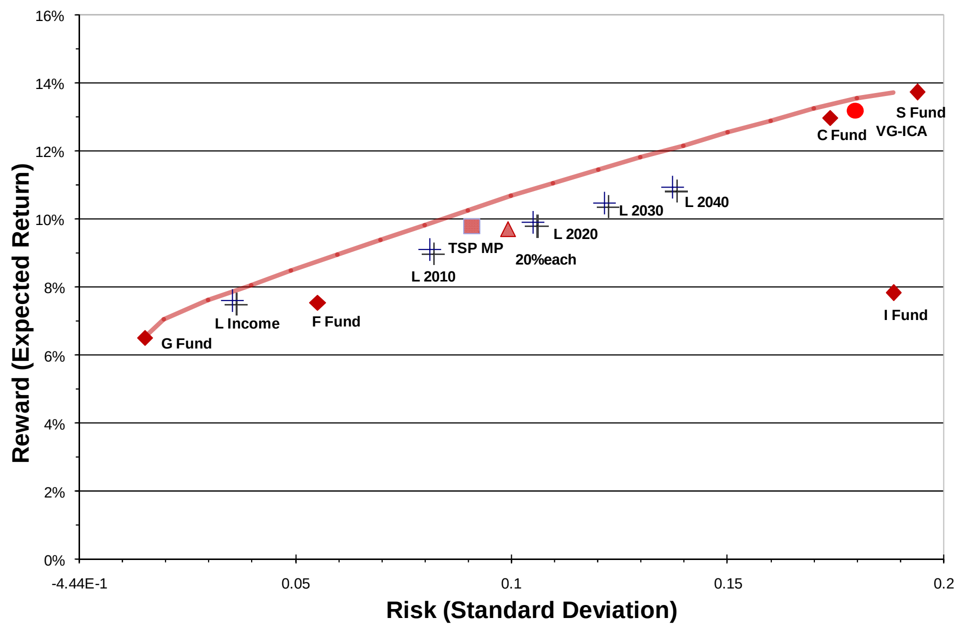

Figure 40. All Portfolios Viewed in Traditional Reward-Risk Space.

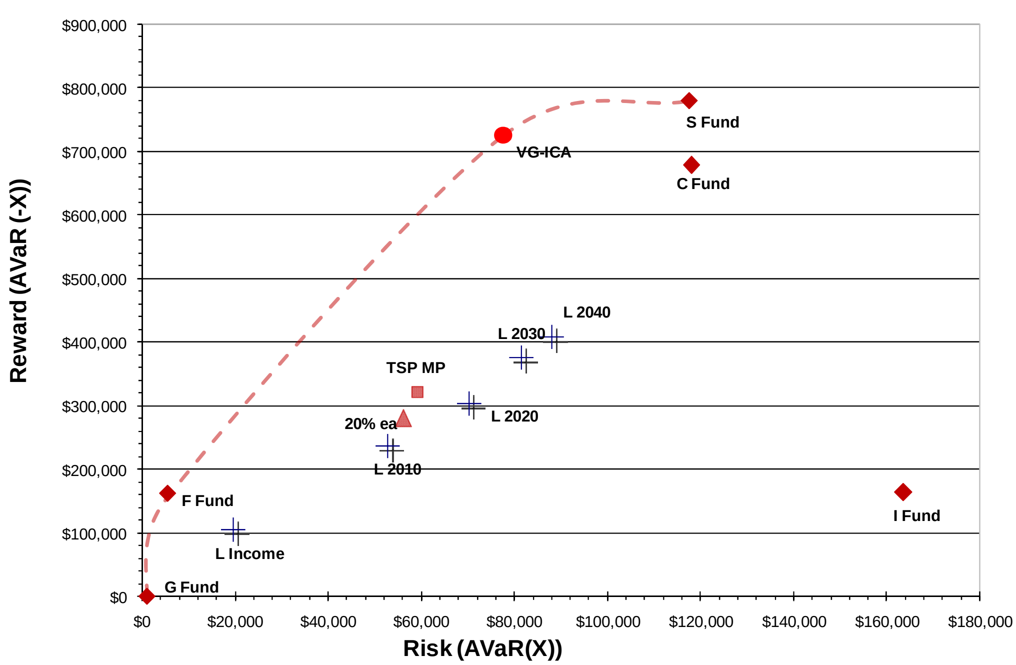

Figure 41. All Portfolios Viewed in New (Alpha VaR) Reward-Risk Space.

Chapter 5: Conclusions and Future Work

5.1 Conclusions

Using the Federal TSP as a framework, this empirical analysis set out to answer three questions. First, "does use of a non-Gaussian factor model for returns, generated with independent components analysis (ICA) and following the Variance Gamma (VG) process, provide any advantage over conventional methods with returns assumed to be Normally-distributed?" Based on the portfolio optimization and simulation performed, "yes, it does." Not only is the expected excess cash flow for the VG-ICA model noticeably higher, the slightly worse downside is more than offset by the significantly greater upside, as demonstrated with the various portfolio performance measures.

The second question considered looked at whether excess returns could be generated through rebalancing to an optimal mix at annual, monthly, and daily intervals, considering the increased computational complexity and administrative burden involved in such an effort. The answer to this question is no; for the VG-ICA model constructed for the TSP funds, it appears that rebalancing is not advantageous.

Finally, as regards the third question, "how does the use of coherent measures of risk, with corresponding portfolio performance measures … affect the selection of optimal TSP portfolios?" Although we did not optimize directly against the coherent measures of risk and portfolio performance measures (but instead computed optimal portfolios with maximization of expected utility), models for returns with heavy tails (specifically the VG-ICA model) performed well against other models when compared with these newer, coherent measures of risk and performance.

In addition to the answers to these three questions, this effort also includes the following two contributions. This was the first known application of the VG-ICA model of Madan & Yen to a portfolio with no short positions. By constraining (one at a time) any short positions developed in the portfolio to zero, a portfolio that included only long (or zero) positions was attained. Based on the results of Scenarios 1 and 3, the VG-ICA model seems to perform as well in this environment as it does in its natural long-short state.

Finally, to reference the second question in "The Only Three Questions That Count: Investing by Knowing What Others Don't" (Fisher 2007), knowing this information about the behavior of the Federal TSP funds may help those interested enough to fathom what others cannot. Keeping in mind the saying that "all models are wrong, but some are useful," I hope that this one falls into the latter category.

5.2 Future Work

Several potential avenues for future research were identified in the course of this work. First, rather than determining the optimal VG-ICA portfolio by maximizing expected utility, optimization by maximization of expected return subject to coherent risk measure constraints AVR and BVR should be considered. This is along the line of work using CVaR for portfolio optimization, either in the objective function or as a constraint. (Rockafellar and Uryasev 2001) The TSP portfolio problem would also make a for examining the newer classes of weighted risk, like MINVaR, MAXVaR, MINMAXVaR, and MAXMINVaR. Additionally, shortcomings in some existing reward-to-risk ratios were identified when they were applied to extremely conservative portfolios which had very small probabilities of loss. This too deserves further investigation.

Were the VG-ICA model to be used in practice, re-estimation of the VG parameters and re-computation of the optimal portfolio of factors (and rebalancing of actual fund allocations) should be done periodically using a rolling horizon, rather than assuming that they are stationary over the entire 20-year horizon, as was done in this work.

References

Acerbi, D., C. Nordio, and C. CSirtori. "Expected Shortfall as a Tool for Financial Risk Management." In *Working paper, Available online at www.gloriamundi.org* (2001).

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath. "Coherent Measures of Risk." *Mathematical Finance* 9 (1999), 203-228.

Avramidis, A., and P. L'Ecuyer. "Efficient Monte Carlo and Quasi-Monte Carlo Option Pricing Under the Variance Gamma Model." *Management Science* 52 (2006), 1930-1944.

Back, A., and A. Weigend. "A First Application of Independent Component Analysis to Extracting Structure from Stock Returns." *International Journal of Neural Systems* 8 (1997), 473-484.

Bakshi, G., C. Cao, and Z. Chen. "Empirical Performance of Alternative Option Pricing Models." *The Journal of Finance* 52 (1997), 2003-2049.

Biglova, A., S. Ortobelli, S. Rachev, and S. Stoyanov. "Different Approaches to Risk Estimation in Portfolio Theory." *Journal of Portfolio Management* (2004), 103-112.

Black, F., and M. Scholes. "The Pricing of Options and Corporate Liabilities." *The Journal of Political Economy* 81 (1973), 637-654.

Bodie, Z., A. Kane, and A. J. Marcus. *Investments*. Boston: McGraw-Hill (2005).

Cardoso, J.-F. "Blind Signal Separation: Statistical Principals." *Proceedings of the IEEE* 86 (1998), 2009-2025.

Carr, P., H. Geman, D. B. Madan, and M. Yor. "The fine structure of asset returns: An empirical investigation." *The Journal of Business* 75 (2002), 305-332.

Cherny, A. S. "Weighted V@R And Its Properties." In *Working Paper* (2006).

Cherny, A. S., and D. B. Madan. "Coherent Measurement of Factor Risks." In *Available at SSRN: http://ssrn.com/abstract=904543* (2006).

—. "On Measuring the Degree of Market Efficiency." In *Available at SSRN: http://ssrn.com/abstract=955472* (2007).

Cootner. *The Random Character of Stock Market Prices*. Cambridge, MA: MIT Press (1964).

Copeland, T. E., and J. F. Weston. *Financial theory and corporate policy*. Reading, Mass.: Addison-Wesley Pub. Co. (1988).

Cover, T. M., and J. A. Thomas. *Elements of information theory*. New York: Wiley (1991).

Cox, J. C., and S. A. Ross. "The valuation of options for alternative stochastic processes." *Journal of Financial Economics* 3 (1976), 145-166.

De Georgi, E. "Reward-Risk Portfolio Selection and Stochastic Dominance." In *Working Paper Series*. Zurich: Institute for Empirical Research in Economics, University of Zurich (2002).

Delbaen, F. "Coherent Risk Measures on General Probability Spaces." In *Working Paper*. Zurich: ETH (1999).

Eberlein, E., and D. B. Madan. "Hedge Fund Replication Beyond Alphas and Betas." In *Available online at https://www.stochastik.uni-freiburg.de/~eberlein/papers/rrtors2jfe.pdf* (2007).

Fama, E. F. "The Behavior of Stock-Market Prices." *The Journal of Business* 38 (1965), 34-105.

Federal Reserve Board. "Basel II Capital Accord, Basel I Initiatives, and Other Basel-Related Matters." http://www.federalreserve.gov/generalinfo/Basel2/ (2007).

Federal Thrift Retirement Investment Board. "L Funds Information Sheet." (2005).

—. "Financial Statements of the Thrift Savings Fund -- 2005 and 2004." (2006).

Fishburn, P. C. "Mean-Risk Analysis with Risk Associated with Below-Target Returns." *American Economic Review* 66 (1977), 116-126.

Fisher, K. *The Only Three Questions That Count: investing by Knowing What Others Don't*. Hoboken, NJ: John Wiley & Sons (2007).

Friend, I., and M. E. Blume. "The Demand for Risky Assets." *The American Economic Review* 65 (1975), 900-922.

Fu, M. "Variance-Gamma and Monte Carlo." In *Advances in Mathematical Finance*, M. Fu, R. A. Jarrow, J.-Y. J. Yen, and R. A. Elliott, eds. Boston: Birkhauser (2007), 21-36.

Fung, T., and E. Seneta. "Tailweight, quantiles, and kurtosis: A study of competing distributions." *Operations Research Letters* Forthcoming (2006).

Geman, H., D. B. Madan, and M. Yor. "Time Changes for Levy Processes." *Mathematical Finance* 11 (2001), 79-96.

Grossman, S. J., and R. J. Shiller. "The Determinants of the Variability of Stock Market Prices." *The American Economic Review* 71 (1981), 222-227.

Harvey, C. R., J. C. Liechty, M. W. Liechty, and P. Muller. "Portfolio Selection With Higher Moments." *Available at SSRN: http://ssrn.com/abstract=634141* (2004).

Heyer, D. D. "Stochastic Dominance: A Tool for Evaluating Reinsurance Alternatives." *Casualty Actuarial Society Forum* (2001), 95-118.

Hyvarinen, A. "Survey on Independent Component Analysis." *Neural Computing Surveys* 2 (1999), 94-128.

Hyvarinen, A., J. Karhunen, and E. Oja. *Independent component analysis*. New York: J. Wiley (2001).

Jennings, W. W., and W. Reichenstein. "The value of retirement income streams: the value of military retirement." *Financial Services Review* 10 (2001), 19.

Konno, H., and H. Yamazaki. "Mean-Absolute Deviation Portfolio Optimization Model and its Applications to Tokyo Stock Market." *Management Science* 37 (1991), 519-531.

Krause, A. "Coherent risk measurement: an introduction." *Balance Sheet* 10 (2002), 13-17.

Krokhmal, P., J. Palmquist, and S. Uryasev. "Portfolio Optimization With Conditional Value-at-Risk Objective and Constraints." In *Working Paper*. Gainesville, FL: University of Florida (2001).

Law, A. M., and W. D. Kelton. *Simulation modeling and analysis*. Boston: McGraw-Hill (2000).

Lo, A. W., and C. A. Mackinlay. "Stock Market Prices Do Not Follow Random Walks: Evidence From a Simple Specificatoin Test." *The Review of Financial Studies (1986-1998)* 1 (1988), 41.

Madan, D. "Equilibrium Asset Pricing: With Non-Gaussian Factors and Exponential Utilities." *Quantitative Finance* 6 (2006), 455-463.

Madan, D., and E. Seneta. "The Variance Gamma (V.G.) Model for Share Market Returns." *The Journal of Business* 63 (1990), 511-524.

Madan, D., and J. Yen. "Asset allocation for CARA utility with multivariate Levy returns." In *Handbook of Financial Engineering*,15, J. a. L. Birge, V., ed. Amsterdam: Elsevier (2004).

Madan, D. B., P. Carr, and E. Chang. "The Variance Gamma Process and Option Pricing." *European Finance Review* 2 (1998), 79-105.

Madan, D. B., and E. Seneta. "Chebyshev Polynomial Approximations and Characteristic Function Estimation." *Journal of the Royal Statistical Society. Series B (Methodological)* 49 (1987), 163-169.

Malaroiu, S., K. Kiviluoto, and E. Oja. "Time Series Prediction With Independent Components." In *ICA2000*. Helsinki, Finland (2000).

Mandelbrot, B. "The Variation of Certain Speculative Prices." *Journal of Business* 36 (1963), 394-419.

—. "Long-Run Linearity, Locally Gaussian Process, H-Spectra and Infinite Variances." *International Economic Review* 10 (1969), 82-111.

Marchini, J. L., C. Heaton, and B. Ripley. "fastICA: FastICA algorithms to perform ICA and Projection Pursuit." (2006).

Markowitz, H. *Portfolio Selection--Efficient Diversification of Investment*. New York: Wiley (1959).

Martin, D., S. Rachev, and F. Siboulet. "Phi-Alpha Optimal Portfolios and Extreme Risk Management." In *Wilmott Magazine of Finance* (2003).

Merton, R. C. "Theory of Rational Option Pricing." *The Bell Journal of Economics and Management Science* 4 (1973), 141-183.

—. "Option pricing when underlying stock returns are discontinuous." *Journal of Financial Economics* 3 (1976), 125-144.

Mitchell, W. C. "The Making and Use of Index Numbers." U. S. B. o. L. Statistics, ed. (1915).

Nestler, S. T. "Your Retirement Savings on Cruise Control: A Critical Analysis of the TSP L Funds." *PHALANX, The Bulletin of Military Operations Research* 40 (2007), 10-13.

Oja, E., K. Kiviluoto, and S. Malaroiu. "Independent Component Analysis for Financial Time Series." In *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000*. Lake Louise, Alta., Canada: IEEE (2000).

Ortobelli, S., S. Rachev, S. Stoyanov, F. J. Fabozzi, and A. Biglova. "The Proper Use of Risk Measures in Portfolio Theory." In *Working Paper*. Karlsruhe (2005).

Osborne, M. F. M. "Brownian Motion in the Stock Market." *Operations Research* 7 (1959), 145-173.

Praetz, P. D. "Australian Share Prices and the Random Walk Hypothesis." *Australian Journal of Statistics* 11 (1969), 123-39.

—. "The Distribution of Share Price Changes." *The Journal of Business* 45 (1972), 49-55.

Press, S. J. "A Compound Events Model for Security Prices." *The Journal of Business* 40 (1967), 317-335.

Rachev, S., T. Jasic, S. Stoyanov, and F. J. Fabozzi. "Momentum Strategies Based on Reward-Risk Stock Selection Criteria." In *Working Paper*. Santa Barbara, CA: Univeristy of California, Santa Barbara (2005).

Redding, L. S. "The G Fund: Portfolio Effects of a Dominant Asset." *Journal of Pension Economics and Finance* (2007).

Rockafellar, R. T., and S. Uryasev. "Conditional value-at-risk for general loss distributions." *Journal of Banking and Finance* 26 (2001), 1443-1471.

Ross, S. A. "The arbitrage theory of capital asset pricing." *Journal of Economic Theory* 13 (1976), 341-360.

Schoutens, W. *Levy processes in finance : pricing financial derivatives*. Chichester, West Sussex ; New York: J. Wiley (2003).

Sharpe, W. F. "Mutual Fund Performance." *Journal of Business* 39 (1966), 119-138.

Stone, J. V. *Independent component analysis : a tutorial introduction*. Cambridge, Mass.: MIT Press (2004).

Stoyanov, S., S. Rachev, and F. J. Fabozzi. "Optimal Financial Portfolios." In *Technical Report*: University of California, Santa Barbara (2005).

U.S. Department of Defense. "Completing the Transition to an All-Volunteer Force: Report of the Defense Advisory Committee on Military Compensation." In *Office of the Secretary of Defense*. Alexandria, VA (2006).

Warner, J. T. "Thinking About Military Retirement." In *CRM D0013583.A1/Final*. Alexandria, VA: Center for Naval Analysis (2006).

Warnung, R. "Beyond VaR: Manageable Alternatives." *GARP Risk Review* (2007), 38-40.

Xia, Q. "Extending the Levy Processes to Multiasset Products Pricing." In *PhD Dissertation in Applied Mathematics and Scientific Computation*. College Park: University of Maryland (2006).

Yen, J.-Y. "Multivariate Levy Processes for Financial Returns." In *PhD Dissertation in Applied Mathematics and Scientific Computation*. College Park: University of Maryland (2004).