

NON-GAUSSIAN ASSET ALLOCATION IN THE FEDERAL THRIFT SAVINGS PLAN

Scott T. Nestler

Robert H. Smith School of Business
University of Maryland
College Park, MD 20742, USA

ABSTRACT

Historical data suggest that returns of stocks and indices are not distributed independent and identically Normal, as is commonly assumed. Instead, returns of financial assets are often skewed and have higher kurtosis. In this study, we investigate how the optimal investment choices in the federal government's Thrift Savings Plan (TSP) change when a non-Gaussian factor model for returns, generated with independent components analysis (ICA) and following the Variance Gamma (VG) process, is used in place of the usual Normally-distributed returns model. Using back-testing and simulation, we hope to show how this method could benefit the more than 3 million TSP participants in achieving their retirement savings objectives.

1 INTRODUCTION

1.1 Motivation and Overview of Approach

Historical data suggest that returns of stocks and indices are not independent and identically distributed Normally, as is commonly assumed. (Praetz 1972; Lo and Mackinlay 1988) Instead, returns of financial assets are often skewed and have higher kurtosis. To account for this, 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 with 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), the factor model for returns is given by $(R - \mu) = D = XB + \varepsilon$ 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, or-

thogonal factors identified using ICA and modeled with the VG model, 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.

1.2 Thrift Savings Plan

1.2.1 Overview

The federal government's Thrift Savings Plan (TSP) is the largest defined contribution retirement savings and investment plan in the world, with over \$173 billion in assets. The TSP was authorized by the United States Congress in 1986; its purpose is to provide Federal workers (both civilian employees and uniformed services members) 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. 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. Also, because there is no minimum vesting period for most TSP assets, savings are portable if a service member or employee leaves government service before reaching retirement eligibility.

1.2.2 Core Investment Funds.

TSP participants can choose from five core 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 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.

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)

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) These four funds are passive in nature; each attempts to track its respective index by holding the same weights of securities represented in the index. The historical correlation between each fund and its tracked index exceeds 0.99, indicating that the fund managers are performing well against their benchmarks.

Table 1: Core TSP fund overview.

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%)

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 one 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.

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. L fund allocations to the five core funds adjust quarterly. The proportion invested in the G fund increases as the retirement horizon nears while allocations to the riskier (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.

1.3 Returns

There are two common definitions of rate of return—log returns and holding-period (or percentage change) returns. 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 analysis, all returns are calculated as being continuously compounded as (natural) log returns,

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

Log returns are generally preferred for a couple of reasons. (Stone 2004) 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.

1.4 Investment Horizon

For this analysis, the focus is on the TSP as a retirement savings vehicle for military personnel as opposed to civilian employees. The point of military retirement, not ultimate retirement (and complete reliance on investment and pension 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 current military retirement system encourages a military career of 20 years. (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 analysis, a representative horizon of 20 years will be used.

1.5 Utility Functions

A utility function maps preferences among alternative choices into measurable utility. Utility functions are order preserving and can be used to rank combinations of risky alternatives. There are a number of commonly-used utility functions to choose from. 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. 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.

1.6 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. 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-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 et al. 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) 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 analysis will use the lower end of the identified range, i.e. a risk aversion parameter of 2.0.

1.7 Organization

Section 2 describes ICA. Section 3 presents the VG process and distribution. Section 4 contains the application of ICA and the VG model to TSP portfolio optimization, and Section 5 concludes.

2 INDEPENDENT COMPONENTS ANALYSIS

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 et al. 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 the knowing the value of one signal provides no information about the value of the others.

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 un-

correlated (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.

2.3 ICA in Finance

Similar to an example on the FTSE 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 (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. (Malaroiu, Kiviluoto et al. 2000; Oja, Kiviluoto et al. 2000)

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 demixing 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 $\mathbf{W}\mathbf{A} = \mathbf{P}\mathbf{D}$ where \mathbf{P} is a permutation matrix and \mathbf{D} is a diagonal matrix. Figure 1 shows a schematic representation of the ICA process. (Back and Weigend 1997)

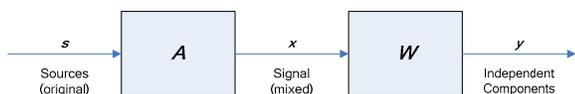


Figure 1: ICA process overview

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 that $\mathbf{y} = \mathbf{W}\mathbf{x} = \mathbf{W}\mathbf{A}\mathbf{s}$. So, \mathbf{s} is a linear function of \mathbf{x} . With appropriate regularity conditions, the Central Limit Theorem tells us that $\mathbf{y} = \mathbf{W}\mathbf{A}\mathbf{s}$ 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) gives. In other words, larger entropy implies more randomness. For continuous RVs, the entropy H is defined as: $H(y) = -\int f(y) \ln f(y) dy$, where $f(y)$ is the density of the random vector 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 $J(y) = H(y_{\text{Gaussian}}) - H(y)$. Thus, maximizing J will result in y being as non-Gaussian as possible. 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 $y = w'x$ at a time, the form of the approximations is $J_G(w) \approx c(E(G(w'x)) - E(G(y_{\text{Gaussian}})))^2$, constrained by $E((w'x)^2) = 1$, 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 et al. 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, $\mathbf{x} \rightarrow \mathbf{x}\mathbf{K}$, where \mathbf{K} is a pre-whitening matrix. Finally, the algorithm estimates a matrix \mathbf{W} such that $\mathbf{x}\mathbf{K}\mathbf{W} = \mathbf{s}$. The function G used in the negentropy approximation is $G(x) = (\ln \cosh(cx)) / c$, with $1 \leq c \leq 2$.

3 VARIANCE GAMMA (VG) MODEL

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) 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. (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, Geman et al. 2002)

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 et al. 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.” From an information perspective, the market does not forget information, so a monotonically non-decreasing process like the family of Gamma distributions is logical here.

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 followed in this chapter follows a recent dissertation (Yen 2004) and papers by authors at the University of Maryland. (Madan and Yen 2004; Madan 2006) However, it is different in that TSP portfolios can only consist of long positions; short sales are not allowed.

3.2 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 $F(x) = P(X \leq x)$ of a RV X is defined as:

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

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. 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.

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

with $\gamma \in R$, $\sigma^2 \geq 0$, and ν is a measure on $R \setminus \{0\}$.

3.3 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 $(W(t), t \geq 0)$ with drift θ and volatility σ : $B(t; \theta, \sigma) = \theta t + \sigma W(t)$. Add an independent Gamma process $(G(t; \nu, \sigma), t \geq 0)$ with mean rate 1 and variance rate ν for the increment $g = G(t+h) - G(t)$ given by the gamma density previously given to get the following VG process: $X(t; \nu, \theta, \sigma) = b(\gamma(t; 1, \nu), \theta, \sigma)$.

Since the VG process is of finite variation, it can be expressed as a difference of two independent Gamma processes. (Geman, Madan et al. 2001) This comes from the fact that the VG process is of finite variation. So, $X_{VG}(t) = G_p(t) - G_n(t)$.

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)

3.4 Simulating From the VG Distribution

There are at least three ways to simulate from the VG distribution. (Fu 2006) 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 $X_{vg} = \theta g + \sigma W(g)$. $W(g) = \sqrt{g} Z$ 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 . $X_{vg} = G_p(t) - G_n(t)$ 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).

3.5 VG Stock Price Model

To construct the VG stock price, replace the Brownian motion in the Black-Scholes model by the VG process. With a continuously 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,

$$\begin{aligned} \exp(-wt) &= E[\exp(X_{vg}(t))] \\ &= \phi_{vg}(t) = \exp\left(-\frac{t}{\nu} \ln\left(1 - \theta\nu - \frac{\sigma^2\nu}{2}\right)\right) \end{aligned}$$

So, $S(t) = S(0) \exp((r + w)t + X_{vg}(t))$ where

$w = \frac{1}{\nu} \ln\left(1 - \theta\nu - \frac{\sigma^2\nu}{2}\right)$, the convexity correction. Considering a single period, say one day ($t=1$), it is then possible to determine $\phi_{\ln S(t)}(u)$ from $\phi_{vg}(t)$ as follows:

$$\begin{aligned} \phi_{\ln S(t)}(u) &= \exp(iu \ln(S(0) + r + \frac{1}{\nu} \ln(1 - \theta\nu - \frac{\sigma^2\nu}{2})) \\ &\quad \cdot \left(1 - iu\theta\nu - \frac{\sigma^2\nu}{2} u^2\right)^{\frac{1}{\nu}} \end{aligned}$$

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

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

with $x = z - m - \frac{1}{\nu} \ln(1 - \theta\nu - \sigma^2\nu/2)$ and where K is the modified Bessel function of the second (third) kind. (Madan, Carr et al. 1998)

4 APPLICATION OF ICA AND VG TO TSP FUND PORTFOLIOS

4.1 Available Returns Data

At the time of this research, data from at least some TSP 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. 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.

Annual and monthly returns data are available from the Thrift Savings Plan website but it does not provide daily historical price or returns data older going back more than five years. As previously mentioned, the correlation between each of the TSP funds and the tracked indices is quite high. Fortunately, daily price and returns data was 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 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.

Shown in Table 2 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 and generally increase in variance, the traditional measure of riskiness, in the order listed.

Table 2: Moments of daily TSP fund returns

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

As expected, each fund other than the G Fund (the risk-free, government bond fund) has 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 histogram of the returns of the risky assets (F, C, S, and I funds); one such example is shown in Figure 2. The solid line represents a smoothed density function of the returns; a Normal density with the same mean and variance is overlaid with a dashed line for comparison purposes.

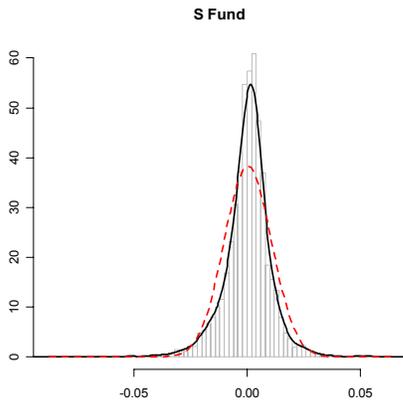


Figure 2: Histogram of daily returns, S fund

4.2 Construction of Independent Components

ICA is then used to generate four independent components (IC), using the $\ln(\cosh)$ as the measure of non-Gaussianity. The first four moments of each of the ICs are shown in Table 3. By design, the ICs have zero mean, unit variance, and are orthogonal.

Table 3: Moments of individual components

IC#	Mean	Variance	Kurtosis	Skewness
1	0	1	14.4	0.589
2	0	1	7.3	-0.054
3	0	1	6.1	0.098
4	0	1	5.5	0.258
5	0	1	2.5	0.144

To estimate the coefficient matrix B and the covariance matrix of the residuals Σ we performed a regression of the core TSP fund returns on the retained factors (ICs). Keeping 5 ICs gives a perfect fit of the data but results in a fully specified model. Retaining 4 ICs still provides a good fit and should lead to good predictive ability. Keeping 3 or less ICs has significant problems with at least one of the core funds. Using the kurtosis level of the Gaussian distribution, three, as a cutoff (rather than the percent variance explained, as in PCA) suggests four ICs be retained in the factor model (IC1-IC4). So, the VG distribution is fit to each of the four retained ICs by maximum likelihood estimation (MLE) using the closed form of the VG density function given earlier. (Madan, Carr et al. 1998) The resulting fitted distribution is shown in Figure 3 with the dotted line; a histogram of the actual returns data is given by the circles; and a Normal(0,1) is overlaid with the dashed line. Clearly, the fitted VG model represents the factor more accurately than the Gaussian distribution.

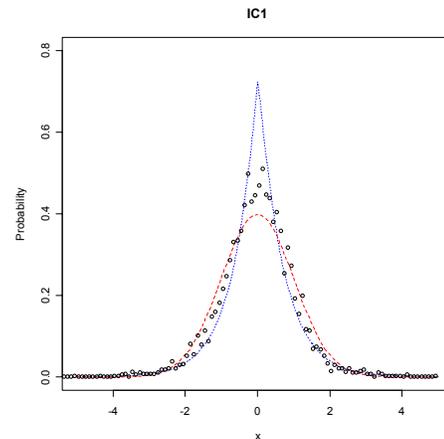


Figure 3: Comparison of returns data, fitted VG, and Normal(0,1)

4.3 Goodness-of-Fit Tests

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) 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 4 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.

Table 4: Fitted VG Parameters and X2 Goodness of Fit Test Statistics

IC#	Fitted VG Parameters			χ^2 Statistic ($\chi^2_{.01,16} = 32.00$)	
	σ	ν	θ	VG(σ, ν, θ)	Norm(0,1)
IC1	0.933	0.969	-0.015	85.51	544.89
IC2	0.980	0.820	-0.001	16.13	291.83
IC3	0.989	0.586	-0.004	33.51	166.32
IC4	0.991	0.468	-0.005	25.00	113.17

4.4 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 $\sigma_d, \nu_d,$ 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 that $X_h \square \sqrt{h}X_d$. So, X_h is VG with $\sigma_h = \sigma_d \sqrt{h}$, $\nu_h = \nu_d$ and $\theta_h = \theta_d \sqrt{h}$. For example, $X_{year} \square \sqrt{252}X_d$ as there are 252 trading days in a year on average.

4.5 Optimal Portfolio Determination

The optimal mix of core TSP funds is determined by maximizing the expected exponential utility in the standard manner of Markowitz. When an investor only cares about mean and variance (i.e. returns are distributed Normally), this is easily determined to be a linear combination of the minimum variance portfolio and the tangency portfolio. For the VG model, the optimal portfolio has also been shown to be deterministic. (Yen 2004) Adjustments are made to preclude short positions in the actual fund positions, although it is possible to have a negative position in one of the factors (ICs).

As shown in Table 5, the optimal portfolios are quite different from one another. The VG portfolio appears considerably more aggressive than the Normal portfolio, even though both used the same risk aversion coefficient. Hopefully, this is the benefit from exploiting the information contained in the higher moments (skewness and kurtosis)—higher returns for the same amount of risk.

Table 5: Optimal portfolios for different returns models.

Returns Model	G Fund	F Fund	C Fund	S Fund	I Fund
Scaled VG	0%	1%	44%	30%	25%
Scaled Normal	42%	54%	0%	4%	0%

5 CONCLUSION

In earlier works of this nature (Madan and Yen 2004, Madan 2006), the application of a non-Gaussian returns model based on ICA and the VG model resulted in improved portfolio returns. It is hoped that once complete, this model will yield similar results in TSP fund portfolios where short positions are not allowed. Results should be available by the time of WinterSim 2007 and will be included in the conference presentation. Ongoing related work includes consideration of multi-period, dynamic models and use of coherent measures of risk.

REFERENCES

Avramidis, A. and P. L'Ecuyer. 2006. Efficient Monte Carlo and Quasi-Monte Carlo option pricing under the variance gamma model. *Management Science* 52: 1930.

Back, A. and A. Weigend. 1997. A first application of independent component analysis to extracting structure from stock returns. *International Journal of Neural Systems* 8: 473-484.

Black, F. and M. Scholes. 1973. The pricing of options and corporate liabilities. *The Journal of Political Economy* 81: 637-654.

Bodie, Z., et al. 2005. *Investments*. Boston, McGraw-Hill.

Carr, P., et al. 2002. The fine structure of asset returns: An empirical investigation. *The Journal of Business* 75: 305.

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

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

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

Federal Thrift Retirement Investment Board. 2006. Financial Statements of the Thrift Savings Fund -- 2005 and 2004.

Friend, I. and M. E. Blume. 1975. The demand for risky assets. *The American Economic Review* 65: 900-922.

Fu, M. 2007. Variance-Gamma and Monte Carlo. In *Advances in Mathematical Finance*, ed Fu, et al. Boston: Birkhauser.

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

- Geman, H., et al. 2001. Time Changes for Levy Processes. *Mathematical Finance* 11: 79-96.
- Grossman, S. J. and R. J. Shiller. 1981. The Determinants of the Variability of Stock Market Prices. *The American Economic Review* 71: 222-227.
- Hyvarinen, A., et al. 2001. *Independent component analysis*. New York, J. Wiley.
- Jennings, W. W. and W. Reichenstein. 2001. The value of retirement income streams: the value of military retirement. *Financial Services Review* 10: 19.
- Law, A. M. and W. D. Kelton. 2000. *Simulation modeling and analysis*. Boston, McGraw-Hill.
- Lo, A. W. and C. A. Mackinlay. 1988. Stock market prices do not follow random walks: evidence from a simple specification test. *The Review of Financial Studies (1986-1998)* 1: 41.
- Madan, D. 2006. Equilibrium asset pricing: with non-Gaussian factors and exponential utilities, University of Maryland.
- Madan, D. and E. Seneta. 1990. The variance gamma (v.g.) model for share market returns. *The Journal of Business* 63: 511-524.
- Madan, D. and J. Yen. 2004. Asset allocation for CARA utility with multivariate Levy returns. *Handbook of Financial Engineering*. J. a. L. Birge, V. Amsterdam, Elsevier.
- Madan, D. B., et al. 1998. The variance gamma process and option pricing. *European Finance Review* 2: 79-105.
- Madan, D. B. and E. Seneta. 1987. Chebyshev polynomial approximations and characteristic function estimation. *Journal of the Royal Statistical Society. Series B (Methodological)* 49: 163-169.
- Malaroiu, S., et al. 2000. *Time series prediction with independent components*. ICA2000, Helsinki, Finland.
- Mandelbrot, B. 1969. Long-run linearity, locally gaussian process, H-spectra and infinite variances. *International Economic Review* 10: 82-111.
- Marchini, J. L., et al. 2006. *FastICA: FastICA algorithms to perform ICA and Projection Pursuit*.
- Merton, R. C. 1973. Theory of rational option pricing. *The Bell Journal of Economics and Management Science* 4: 141-183.
- Merton, R. C. 1976. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics* 3: 125-144.
- Nestler, S. T. 2007. Your retirement savings on cruise control: a critical analysis of the TSP L funds. *PHALANX, The Bulletin of Military Operations Research* 40:10-13.
- Oja, E., et al. 2000. *Independent component analysis for financial time series*. Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000, Lake Louise, Alta., Canada, IEEE.
- Praetz, P. D. 1972. The distribution of share price changes. *The Journal of Business* 45: 49-55.
- Press, S. J. 1967. A compound events model for security prices. *The Journal of Business* 40: 317-335.
- Redding, L. S. 2007. The G fund: portfolio effects of a dominant asset. *Journal of Pension Economics and Finance* (forthcoming).
- Ross, S. A. 1976. The arbitrage theory of capital asset pricing. *Journal of Economic Theory* 13: 341-360.
- Schoutens, W. 2003. *Levy processes in finance: pricing financial derivatives*. Chichester, West Sussex; New York, J. Wiley.
- Stone, J. V. 2004. *Independent component analysis: a tutorial introduction*. Cambridge, Mass., MIT Press.
- Warner, J. T. 2006. Thinking about military retirement. *CRM D0013583.A1/Final*. Alexandria, VA, Center for Naval Analysis.
- Yen, J.-Y. 2004. *Multivariate Levy Processes for Financial Returns*. Ph.D. dissertation, Applied Mathematics, University of Maryland, College Park, Maryland. Available via <https://drum.umd.edu/dspace/bitstream/1903/1985/1/umi-umd-1938.pdf>

AUTHOR BIOGRAPHY

SCOTT T. NESTLER is an assistant professor in the Department of Mathematical Sciences and research analyst in the Operations Research Center, Department of Systems Engineering, at the U.S. Military Academy at West Point. He is also a PhD candidate in the Robert H. Smith School of Business at the University of Maryland. His research interests include the application of discrete-event simulation and simulation-optimization to manpower/personnel and finance problems. His web page can be found via www.wam.umd.edu/~snestler