Beyond 'Bayesian vs. VaR' Dilemma to Empirical Model Risk Management: How to Manage Risk (After Risk Management Has Failed) for Hedge Funds

Yogesh Malhotra, PhD, MSQF

December 04, 2014

Global Risk Management Network, LLC
Cornell Business and Technology Park
Ithaca, NY 14852-4892

Phone: (646) 770-7993 m, (646) 861-3680 v
e-mail: dr.yogesh.malhotra@gmail.com
http://www.linkedin.com/in/yogeshmalhotra

Beyond 'Bayesian vs. VaR' Dilemma to Empirical Model Risk Management: How to Manage Risk (After Risk Management Has Failed) for Hedge Funds

Abstract

In aftermath of the Financial Crisis, some risk management practitioners advocate wider adoption of Bayesian inference to *replace* Value-at-Risk (VaR) models for minimizing risk failures (Borison & Hamm, 2010). They claim reliance of Bayesian inference on *subjective judgment*, the *key limitation* of Bayesian methodology as underscored by statisticians (Kass & Raftery, 1995; Kruschke, 2011; Lynch, 2007), as the most significant *advantage* compared with VaR (Christoffersen, 2012). Despite its well-known limitations, *just like all other quantitative models* (Derman, 1996; Morini, 2011), VaR – [mostly] non-Bayesian *and* [increasingly] Bayesian – continues to be a key methodological foundation of risk management *and* regulation related risk modeling practices in global Finance (Danielsson et al., 2014; Zangari, 1996). Bayesian inference modeling and VaR modeling frameworks are outlined to facilitate *model risk management* (Derman, 1996; Morini, 2011; US Fed & OCC, 2011) for minimizing risk of *any* model – Bayesian, VaR, *or* Bayesian VaR. VaR frameworks are empirically applied for hedge fund risk modeling (Darbyshire & Hampton, 2012, 2014) of a multi-asset fund of funds portfolio of a large Wall Street investment bank. Multiple risk models and measures with transparent assumptions to cross-validate convergent findings across multiple levels of risk analysis are examined for empirical model risk management.

Keywords: Model Risk Management, Risk Modeling, Bayesian Inference, VaR, Portfolio Construction, Portfolio Optimization, Fund of Funds, Hedge Funds.

1. Introduction

In aftermath of the Global Financial Crisis (GFC) of 2008-2009, critical analyses of financial risk management failures and the role of quantitative models such as Value-at-Risk (VaR) continue unabated (Danielsson et al., 2014; US Senate, 2013). *How to Manage Risk (After Risk Management Has Failed)* (Borison & Hamm, 2010) in *Sloan Management Review* is one such article addressed to risk management executives, decision-makers, and modelers. Its authors' *Bayesian vs. VaR* argument advocates for wider adoption of Bayesian inference to *replace* Value-at-Risk (VaR) models. Their central message is that choosing Bayesian *instead of* VaR models would minimize risk management failures *because of* the key role of 'subjective judgment' in the Bayesian methodology. They specifically assert that *if* Bayesian inference had been used in Finance practice *instead of* VaR, *then* risk management failures of GFC would have been minimal. Their basis for choosing Bayesian over VaR is *subjective judgment* which has its advantages, it is however a *key limitation* as recognized by Bayesian statisticians (Kruschke, 2011; Lynch, 2007). Further, since before GFC, *both* non-Bayesian and Bayesian VaR models have been used in Finance practice (Danielsson et al., 2014; Hull & White, 1998; Venkataraman, 1997; Zangari, 1996). Hence, the *Bayesian vs. VaR* dilemma needs to be resolved in order to minimize model *specification* and *estimation* errors in risk modeling (Boucher et al., 2014).

Current research contributes to congruent theme of improving financial risk management practices focused on model risk management. The key problem of model risk in any risk model such as VaR results from the fact that risk cannot be measured, but must be estimated using a statistical model (Boucher et al., 2014; Danielsson et al., 2014) . Hence, model risk occurs because a statistical model is used for risk estimation: *model use entails model risk* (Derman, 1996; Morini, 2011). Using range of different plausible models which can be robustly discriminated between, the disagreement between their range of readings is a succinct measure of model risk (Danielsson et al., 2014). We apply this notion of model risk and model risk management methodology empirically in course of fund-of-funds portfolio construction and optimization for a top Wall Street investment bank which we discuss here.

VaR, originally invented by JP Morgan, introduced quantitative rigor to fathom multi-dimensional complexity of risk with a simple and easy to implement measure (Hull & White, 1998; Jackson et al., 1998). It became the "de facto industry standard" for risk management practices among financial institutions as well as their regulators (Simons, 1996). Despite its well-known limitations (e.g. (Berkowitz et al., 2011; Berkowitz & O'Brien, 2002)) *just like all*

other quantitative models (Derman, 1996; Morini, 2011), VaR – [mostly] non-Bayesian *and* [increasingly] Bayesian – remains the "methodological common root" of Finance risk modeling underlying risk management *and* regulation (Danielsson et al., 2014). It is therefore important to advance beyond the *Bayesian vs. VaR* dilemma to focus on *model risk management* for *all* models – including *Bayesian*, *VaR*, *and*, *Bayesian VaR* – as that is what really matters (Derman, 1996; Morini, 2011; US Fed & OCC, 2011). Hence, the contributions of this paper are as follows.

First, we modulate the 'silver bullet' expectations about 'replacing' VaR with Bayesian models with realities of computational statistical modeling. Specifically, we inform the *Baysian vs. VaR* debate by outlining analytical frameworks of Bayesian inference (based on (Kruschke, 2011)) and VaR (based on (Darbyshire & Hampton, 2012, 2014)). Bayesian statistical inference methodology is anticipated to overcome known limitations of null hypothesis based frequentist significance testing (NHST) statistical inference methodology. Modeling of 'Bayesian priors' – referred by some as 'subjective judgment' – remains a key challenge and limitation of Bayesian methodology. Feasibility as well as precision and accuracy of Bayesian modeling depend on computational statistical algorithms such as Markov Chain Monte Carlo (MCMC) which are themselves reliant upon exponential computing powers (Kruschke, 2011; Lynch, 2007). Such computing power accessibility is becoming available in recent years for mainstream use which explains recent re-emergence of applied interest in Bayesian inference.

Second, we resolve the *Bayesian vs. VaR* dilemma by providing analytical frameworks of Bayesian inference modeling and VaR modeling and advance beyond to empirical model risk management. Related discussion elucidates the central concern of model risk management which is relevant to *every* model – including *Bayesian*, *VaR*, *and*, *Bayesian VaR* – and necessary for minimizing modeling related risk management failures by minimizing model risk. Our current choice of empirical methodology and risk modeling framework is based upon the contextual domain and related current real world practice for risk modeling for multi-asset portfolio hedge funds. Our empirical focus on risk modeling and VaR frameworks for construction and optimization of fund-of-funds portfolio helps fathom the multi-dimensional complexity of financial risk modeling. We empirically examine multiple risk models and measures to cross-validate convergent findings across various levels of risk analysis as one such method of model risk management by applying VaR using classical methodology.

The outline of the paper is as follows. In section 2 we discuss the analytical framework of the Bayesian statistical inference methodology as a viable contender for the classical frequentist

methodologies of statistical inference. In section 3 we discuss the quantitative risk modeling frameworks including VaR that are relevant to the contextual domain and related current real world practice for risk modeling of multi-asset portfolio hedge funds. Section 4 presents the empirical context that applies the risk modeling frameworks including VaR in multi-asset portfolio hedge fund risk modeling for a half-trillion dollar fund-of-funds portfolio comprised of diverse equity, currency, commodity, alternative investments, and, hedge fund asset classes. Empirical findings in section 5 illustrate the use of multiple risk measures and models at various levels of analysis to find convergence across the observations. Section 6 concludes our discussion outlining limitations and directions for future research.

2. Bayesian Modeling

To align expectations of practice with the challenges of computational statistical modeling inherent in Bayesian modeling, we outline the following analytical framework. The proposed framework aims to facilitate Bayesian estimation of parameter values, prediction of data values, and model comparison (based on (Kruschke, 2011)). Our synthesis advances beyond the ambiguity of the *Bayesian vs. VaR* dilemma by clarifying central concerns that characterize Bayesian modeling. First, the role of 'subjective judgment' known more formally as 'Bayesian priors' is recognized as key challenge and limitation of Bayesian inference by its strongest critics and proponents alike (Kruschke, 2011; Lynch, 2007). Second, statistical computational complexity necessary for realizing *more sophisticated* Bayesian inference even when overcome at much expense may not necessarily result in more accurate or precise model. *Hence, regardless of models being used, VaR or Bayesian, model risk management is necessary for minimizing risk management failures.* Readers informed by this framework should be wiser in considering the prescriptive advice about 'replacing VaR models with Bayesian' (Borison & Hamm, 2010). Those new to Bayesian statistical inference probabilistic modeling may find Appendix 1 Bayesian Inference: Probability Background relevant.

Bayes' Rule

Bayes' rule is based on *conditional probability*, the probability of one event given that we know that the other event is true. *Conjoint probability* is the probability of two outcome events occurring together when considering a conjunction of the two events. Given conjoint events x and y, *total* probability of occurrence of a *specific value* of x *regardless of* the probability of *any value* for y is called *marginal probability* of x. M*arginal probability* of a *specific value of x* regardless of any value of y equals sum of all *conjoint probabilities* p(x, y) for the *specific value of x*.

Marginal probability of x = $p(x) = \sum_y \quad p(x,y)$ when x and y are discrete,

$p(x) = \int_y dy \; p(x,y)$ when x and y are continuous.

Above process is called *marginalizing over y* or *integrating out the variable y* (Kruschke, 2011).

Probability of a specific outcome of y given known outcome of x could differ from its probability if outcome of x is not known. Conditional probability of event y is then limited *by* the conjoint probability of x and y for that specific value of y given the specific value of x [for all values of y]. Conditional probability of y given x denoted as $p(y|x)$ equals the conjoint probability of x and y divided by the sum of conjoint probabilities for the specific value of x over all values of y where $p(y,x) = p(x,y)$.

$$p(y|\boldsymbol{x}) = \frac{p(y,x)}{\sum_{\boldsymbol{y}} p(y,x)} = \frac{p(y,x)}{p(\boldsymbol{x})} \quad , \quad p(x|y) = \frac{p(x,y)}{\sum_x p(x,y)} = \frac{p(x,y)}{p(y)}$$ given discrete x and y,

$$p(y|x) = \frac{p(y,x)}{\int_y dy\, p(y,x)} = \frac{p(y,x)}{p(x)}, \quad p(x|y) = \frac{p(x,y)}{\int_x dx \; p(x,y)} = \frac{p(x,y)}{p(y)}$$ given continuous x and y.

In summary, (Conditional Probability = Conjoint Probability / Marginal Probability), which can also be expressed as (Conjoint Probability = Conditional Probability * Marginal Probability). From above expressions, it also follows that: $p(x,y) = p(y|x)p(x) = p(x|y)\, p(y)$.

$p(x|y)$ should *not* be interpreted as denoting temporal order implying that y precedes x. It only implies limiting the calculations of probability to a particular subset of possible events: among all events with value *y*, $p(x|y)$ of them also have value *x* (Kruschke, 2011). When two events x and y have no influence on each other, they are called *independent events*.

When value of y has no influence on value of x, in general, $p(y|x) = p(y) = \frac{p(y,x)}{p(x)}$. Likewise, when value of x has no influence on value of y, in general, $p(x|y) = p(x) = \frac{p(x,y)}{p(y)}$.

Hence for two independent events x and y, conjoint probability $p(x,y)$ equals the product of marginal probabilities $p(x)$ *and* $p(y)$. Symmetrically, when $p(x,y) = p(x)\, p(y)$ for all values of x and y, then $p(x|y) = p(x)$ and $p(y|x) = p(y)$. Both expressions specify *independence of attributes*. The relationship between $p(x|y)$ and $p(y|x)$ called ***Bayes' Rule*** is derived as follows.

From above expressions, $p(y|x)p(x) = p(y,x)$, and, $p(x|y)p(y) = p(x,y)$

=> $p(y|x)p(x) = p(x|y)p(y)$

$$\Rightarrow \boldsymbol{p(y|x)} = \frac{\boldsymbol{p(x\,|\,y)p(y)}}{\boldsymbol{p(x)}} \qquad \textbf{(1)} \qquad \Rightarrow \boldsymbol{p(y|x)} = \frac{\boldsymbol{p(x\,|\,y)p(y)}}{\sum_y \boldsymbol{p(x,y)}} \Leftrightarrow \frac{p(x\,|\,y)p(y)}{\int_y dy\ p(x,y)}$$

As noted earlier, marginal probability of x is $p(x) = \sum_y p(x,y)$ when x and y are discrete, and, $p(x) = \int_y dy\ p(x,y)$ when x and y are continuous. As $\boldsymbol{p(x,y)} = \boldsymbol{p(x|y)p(y)}$, it follows:

$$\Rightarrow \boldsymbol{p(y|x)} = \frac{\boldsymbol{p(x\,|\,y)p(y)}}{\sum_{\boldsymbol{y}} \boldsymbol{p(x\,|\,y)p(y)}} \qquad \textbf{(2)}$$

Above two expressions (1) and (2) called the Bayes' Rule are at the core of Bayesian Inference.

Bayes' Rule holds when x and y are independent as well as when x and y are not independent.

Key Objectives of Bayesian Inference

Pre-existing beliefs about different possible values of a parameter before taking into account some particular set of observations are called *prior beliefs* or simply *priors*. The modified beliefs resulting from taking the particular set of data or observations into account are called *posterior beliefs*. *Even though the terms prior and posterior may seem to suggest historical or temporal ordering, in fact it is not the case* (Kruschke, 2011). Prior simply means the probability distribution of beliefs held *without including* a particular set of data. In contrast, posterior simply means the probability distribution of beliefs held *after including*, i.e., after taking into consideration that particular set of data. Bayesian inference transforms prior beliefs into posterior beliefs. Statistical inference based on data observations typically fulfills one of the following three goals: *estimation of parameter values*, *prediction of data values*, and *model comparison*.

Estimation of parameter values is used to determine the probability distribution of beliefs in different possible values of a parameter. For a random process, underlying true parameter is not known and hence related beliefs are uncertain; therefore posterior beliefs about that parameter are an estimate. Data can result in modification of beliefs given that specific probabilities assigned to different values of the parameter may change. Degree of belief in some possible values of a parameter may increase resulting in corresponding decrease in other possible values. Hence, such reassignment of probabilities across different possible parameter values is called estimation of parameter values as it results in shifting of beliefs across those values (Kruschke, 2011). *Prediction of data values* means inferring values of data other than that we have already considered based upon current beliefs. Again, prediction means inferring

value of data that is not included based on data that has already been included *regardless of the actual temporal relationship* between the two (Kruschke, 2011). Bayesian prediction is based upon taking weighted average of predictions based on respective beliefs, specifically taking a summated weighted average of each possible value of unknown data and the respective [believed] probability of occurrence of the specific value. *Model selection*, also known as *model comparison*, is based upon choosing the model which can generate the observed data with greater likelihood. Bayesian inference can help determine exactly how much more to believe the selected model than those not selected and intrinsically adjusts for model complexity. Complex models, being more flexible, will fit the data better as well as fit random noise better than simpler models (Kruschke, 2011).

Bayes' Rule Applied to Models and Data

In context of application to models and data, a key application of Bayes' Rule is in assessment of conditional probabilities of observed data values and related model parameter values. Its crucial application is in determining the probability of a model when given a set of data. The model itself provides the probability of the data, given specific parameter values and the model structure. Specifically, Bayes' Rule helps to get from the *probability of the data, given the model* to the *probability of the model, given the data*.

Having observed some data, Bayes' Rule is then applied to determine strength of our beliefs across competing parameter values in a model, and, also to determine strength of our beliefs across competing models. Beliefs held prior to the observation of data are called prior beliefs or *priors*. Observed data may modify those priors and result in posterior beliefs or *posteriors*. Again, the notion of "historical data" (Borison & Hamm, 2010) needs to be interpreted carefully especially in the context of Bayesian analysis. *Even though the terms 'prior' and 'posterior' may seem to suggest historical or temporal ordering, in fact it is not the case* (Kruschke, 2011). Prior simply means the probability distribution of beliefs held *without including* a particular set of data. In contrast, posterior simply means the probability distribution of beliefs held *after including*, i.e., after taking into consideration that particular set of data. Bayesian inference transforms prior beliefs into posterior beliefs thus helping us make inference from data to uncertain beliefs. Uncertainty in beliefs results from differing likelihood of diverse possibilities. By helping precisely determine likelihood of diverse possibilities, statistical inference models help precisely define such uncertainty with precise numerical bounds. This is particularly useful with increasing variance in data and increasing uncertainty in beliefs.

Data denotes the observable sample statistic observed for a process to estimate corresponding parameter of the process which cannot be directly observed. The first set of assumptions about the process that generates probabilistic observable data outcomes for the unobservable parameter is the *model of observable events*. The second set of assumptions about our beliefs regarding the likelihood of different levels of the specific process parameter is the *model of our beliefs*. Bayes' rule can be visualized spatially (Kruschke, 2011) in terms of events x listed in i rows R_i and events y listed in intersecting j columns C_j wherein any specific intersection of the two is the conjoint probability $p(R_i, C_j) = p(R_i|C_j)\, p(C_j) = p(C_j|R_i)\, p(R_i)$. Then, normalization of probabilities in row R_i by dividing conjoint probabilities by $p(R_i)$ yields the following.

$$p(C_j \mid R_i) = \frac{p(C_j, R_i)}{p(R_i)} = \frac{p(R_i \mid C_j) p(C_j)}{p(R_i)} \qquad \textbf{(1a)}$$

$$p(R_i|C_j) = \frac{p(R_i, C_j)}{p(C_j)} = \frac{p(C_j \mid R_i)\ p(R_i)}{p(C_j)} = \frac{p(C_j \mid R_i)\ p(R_i)}{\sum_i p(C_j \mid R_i)\ p(R_i)} \qquad \textbf{(2a)}$$

Applying Bayes' Rule in spatial representation to data values D_i in rows and intersecting column parameter values θ_i, we get the following expressions about the Bayesian inference for *model given data*. The following expressions are based upon the earlier observation that conjoint probability equals the product of conditional probability and marginal probability. The first expression is that of the posterior for which we need to avoid the computation of large complex integral in the denominator for ease of computation.

$$\boldsymbol{p(\theta_j \mid D_i)} = \frac{p(\theta_j, D_i)}{p(D_i)} = \frac{\boldsymbol{p(D_i \mid \theta_j)\, p(\theta_j)}}{\boldsymbol{p(D_i)}} \qquad \textbf{(1 b)}$$

$$p(D_i \mid \theta_j) = \frac{p(D_i, \theta_j)}{p(\theta_j)} = \frac{p(\theta_j \mid D_i)\, p(D_i)}{p(\theta_j)} = \frac{p(\theta_j \mid D_i)\, p(D_i)}{\sum_i p(\theta_j \mid D_i)\, p(D_i)} \qquad \textbf{(2 b)}$$

Bayes Rule helps us determine how strongly we believe in the model given the data. It helps us get from the probability of the data given the model $p(D_i \mid \theta_j)$ to probability of the model given the data $p(\theta_j \mid D_i)$ (Kruschke 2011). Writing expression (1b) as follows helps clarify the Bayesian analysis notation.

$$p(\theta_j \mid D_i) = \frac{p(D_i \mid \theta_j)\, p(\theta_j)}{p(D_i)} \quad \text{i.e.} \quad \textit{Posterior} = \frac{\textit{Likelihood} * \textit{Prior}}{\textit{Evidence}} \quad \text{where}$$

Posterior $p(\theta_j|D_i)$ denotes strength of our belief in parameter θ_j when data D_i is considered.
Prior $p(\theta_j)$ denotes strength of our belief in parameter θ_j without considering data D_i.
Likelihood $p(D_i|\theta_j)$ denotes probability that data D_i could be generated by model with parameter θ_j.

Evidence $p(D_i)$ denotes probability of the data according to the model.

For Likelihood $p(D_i|\theta_j)$, θ that maximizes its value is called the Maximum Likelihood Estimate of θ. Evidence is used here as in machine learning and equals the numerical sum across all possible parameter values weighted by respective strength of the belief in those parameter values. Hence,

$$p(D_i) = \sum_\theta p(D_i, \theta_j) = \int_\theta d\theta \; p(D_i|\theta_j) \; p(\theta_j) \quad .$$

Because parameter value θ_i makes sense only in context of the respective model, it helps to make the specific model explicit.

$p(\theta_j \mid D_i, M) = \frac{p(D_i \mid \theta_j, \; M)\, p(\theta_j \mid M)}{p(D_i \mid M)}$. Correspondingly $p(D_i \mid M) = \int_\theta d\theta \, p(D_i \mid \theta_j, \; M) \; p(\theta_j \mid M)$.

Above assessment of the strength of [posterior] beliefs given data for a specific model can be extended to the case of comparison of strength of belief in two different models M_1 and M_2 given observed data.

$p(M_1 \mid D_i) = p(M_1, D_i) \,/\, p(D_i) = p(D_i|M_1)\, p(M_1) \,/\, p(D_i)$ and

$p(M_2 \mid D_i) = p(M_2, D_i) \,/\, p(D_i) = p(D_i|M_2)\, p(M_2) \,/\, p(D_i)$.

Equating the ratios of LHS and RHS above, we get,

$\frac{p(M_1 \mid D_i)}{p(M_2 \mid D_i)} = \frac{p(D_i|M_1)\, p(M_1)}{p(D_i|M_2) \; p(M_2)}$ where the ratio of evidence terms $\frac{p(D_i|M_1)}{p(D_i|M_2)}$ is called the ***Bayes' Factor***.

Hence, for comparison of M_1 and M_2, ratio of posterior beliefs equals Bayes' Factor times the ratio of priors.

What Makes Bayesian Inference Challenging

Beyond estimation of model parameters, Bayesian methodology is far more flexible in evaluating model fit and comparing models, producing parameters samples not directly estimated within the model, handling missing data, while capturing greater uncertainty than the classical approach in prediction and forecasting (Lynch, 2007). *It is however recommended to think of sophistication and complexity of models as a two-edged sword. Simple models are always preferred if they help understanding the assumptions and limits of their scope which helps in managing*

model risk. Complex and sophisticated models may increase the model risk if they obfuscate such understanding and clarity (Kruschke, 2011). Bayesian modeling can help to the extent given that it automatically accounts for model complexity when assessing the strength of belief in any given model. Let's consider the case wherein for estimation of parameter values a simple model with a few parameter values is compared with one containing many parameter values. Given that the same probability is spread out over a larger number of values, the simpler model is favored as it shows greater posterior values for the lesser parameter values. However, the complex model may be favored when the observed data do not fit the simpler model. In any case, the model comparison simply tells about the *relative evidence* for each model and makes sense in the context of *relative comparison*. *Regardless of which model seems relatively superior, it may still not be a good model of the data, but the least worse of the models that are compared* (Kruschke, 2011).

The evidence $p(D_i)$ and $p(D_i \mid M)$ involve a complex integral over a possibly high dimension parameter space $\theta_j \in \Theta$. Such complex integration over high dimension parameter space is the principal inferential operation in Bayesian analysis as compared to optimization in classical inference. *Evaluation of such complex integrals over high dimensional parameter space poses major challenge for actual use of Bayesian analysis. All three goals of Bayesian inference depend upon the solution of the evidence term which is in the denominator of Bayes' formula* (Kruschke, 2011).

Few methods have been typically used to overcome the problems such as those noted above that severely constrain the application of Bayesian analysis. The first method involves using prior and posterior distributions of the same form, i.e., satisfying the condition of *conjugacy*. In other words, the functional forms of distributions $p(D_i|\theta_j)$ and $p(\theta_j)$ combine so that the posterior distribution has same form as prior distribution (e.g. both are normal distributions), then $p(\theta_j)$ is called the *conjugate prior* for $p(D_i|\theta_j)$. Another pure analytical method involves approximation of the actual functions with easier to compute alternatives while demonstrating their reasonableness. Third method involves numerical approximation of the difficult-to-compute integral by approximating the continuous function θ_j as a sum over a fine grid of discrete θ_j values. Such grid approximation is based upon approximating the integral by summation of discrete intervals across the grid. Instead of treating θ_j as a continuous function with associated *probability densities*, it uses discrete finite values of θ_j and aggregates respective *probability masses* as shown below.

$$p(D_i \mid M) = \int_\theta d\theta \, p(D_i \mid \theta_j, M) \, p(\theta_j \mid M) \quad \approx \quad \sum_\theta p(D_i \mid \theta_j, M) \, p(\theta_j \mid M)$$

The above grid approximation method is limited to cases where the number of parameters is relatively very small. *For instance, considering a model which may have say, eight parameters, each having a thousand values, the eight-dimensional parameter space contains (1E3)E8 i.e. 1E24 combinations of parameter values which is a computationally complex problem to solve*. Markov Chain Monte Carlo (MCMC) numerical techniques (Gelfand & Smith, 1990; Malhotra, 2014) that work by simulating a discrete time Markov chain on high dimension parameter space $\theta_j \in \Theta$ by using statistical computing algorithms provide a relatively recent breakthrough for making Bayesian analysis feasible for solving high dimensionality problems. MCMC use Monte Carlo simulations to approximate the true posterior probability density $p(\theta_j|D_i)$ by constructing Markov chains whose steady state distribution matches $p(\theta_j|D_i)$. The samples returned by the MCMC methods of simulation based inference can be assumed as random draws from $p(\theta_j|D_i)$. *Only with availability of MCMC statistical computing algorithms such as Metropolis Hastings algorithm and Gibbs Sampling algorithm and faster inexpensive computing power has Bayesian inference become feasible lately for mainstream use for doing high dimension parameter space analyses* (Gelfand & Smith, 1990; Malhotra, 2014).

'Subjective Judgment' Limitations of Bayesian Inference

A key limitation of Bayesian inference is often attributed to the choice of the appropriate and reasonable prior distribution. For all parameters, proper priors have to be used in order to avoid possible non-integrability of the posterior parameter distribution which would make the Bayesian model selection rather questionable (Kass & Raftery, 1995). Choice of suitable priors is generally a 'contentious issue' (Miazhynskaia et al., 2003): "One wants the priors to reflect one's believes about parameter values and at the same time to use non-informative (flat) priors that does not favor particular values of the parameter over other values." To avoid the "subjectivity" criticism of Bayesian approach as in choice of 'subjective' priors when contrasted from the classical approach, many Bayesian analyses have used uniform, reference, or otherwise 'non-informative' priors (Lynch, 2007). This has lessened the use of priors as a distinguishing characteristic of Bayesian analyses even though most Bayesian analyses specifically attempt to minimize the effect of the prior such as by excluding the 'burn in' period. It may be however argued that explicit priors should be used because prior beliefs influence rational inference from data because new data modifies beliefs from what they were prior to the new data. However, it must be recognized that prior beliefs are ***not*** *capricious and idiosyncratic and unknowable* but based on publicly agreed facts and theories and admissible by a skeptical scientific audience (Kruschke, 2011). Hence, it must be emphasized that Bayesian

analysis doesn't *ipso facto* imply reliance upon *ad hoc and subjective personal judgment* but is rather based upon use of *priors that are agreeable to a skeptical audience* (Kruschke, 2011). In case of disagreement about two sets of priors, either each can be used to conduct separate analysis and then robustness of posterior assessed w.r.t. changes in prior or they can be mixed into joint prior with posterior reflecting the uncertainty in the prior.

The above synthesis of a Bayesian analytical modeling framework is intended to clarify prescriptive advice about 'replacing VaR with Bayesian models' for its actual execution in applied practice. It is important to recognize three key points from the above discussion. First, Bayesian and VaR models cannot and should not be treated as mutually exclusive alternatives in risk modeling for minimizing risk management failures given existence of non-Bayesian VaR, Bayesian VaR, as well as risk models other than VaR. Second, Bayesian approach, even though more sophisticated statistically, comes at much computational expense and does not necessarily ensure more precise or accurate model. Third, but most importantly, regardless of using Bayesian approach with *or* without VaR, model risk management is necessary in all cases for mitigating risk management failures.

3. Value at Risk (VaR) Modeling

The following discussion focus is on VaR and ES models most widely used in hedge fund risk modeling practice (Darbyshire & Hampton, 2012, 2014; J.P. Morgan, 2008). These risk models are used for empirical analysis as described in the next section. Other sophisticated risk management models which share the "methodological common root" of VaR (Danielsson et al., 2014) are reviewed in the concluding discussion of the current section.

Key Concept of Value-at-Risk

For a given portfolio of assets, Value at Risk (VaR) quantifies how much at most can be lost with a given probability over a specific time horizon. Value-at-Risk denotes the worst expected loss over a given time horizon at a given confidence level under normal market conditions (J.P. Morgan, 2008). VaR provides a single number summarizing the firm's exposure to market risk and the likelihood of an unfavorable move in the portfolio's positions. It also provides a predictive tool to prevent portfolio managers from exceeding risk tolerances defined in the portfolio policies. It can be measured at the portfolio, sector, asset class, and security levels. VaR is just an *estimate* and not a uniquely defined value (J.P. Morgan, 2008). Unlike, Expected

Shortfall discussed later, VaR does not provide any information on losses that exceed its value, i.e., VaR is *not* the 'worst case scenario'(J.P. Morgan, 2008).

95% VaR level was defined by the popular Riskmetrics methodology of JP Morgan. 99% VaR level had been the Basel committee's market risk regulatory criterion until 2013 when they proposed replacing it with 97.5% ES in Basel III which is to be implemented sometime until 2019 as of the time of writing. In case of a hedge fund (or a fund of funds), assumptions about hedge fund portfolio returns following a normal distribution and being affected by linear market forces are called the normality and linearity assumptions. The normal distribution of hedge fund returns can be described by just two parameters, mean μ and standard deviation σ. Assuming normal distribution ($X \sim N(\mu, \sigma^2) = N(0, 1)$) for monthly returns and a c% confidence level, where $c = 100(1 - \alpha)$, c% VaR implies that the worst estimated portfolio loss for the next month is no more than $z_\alpha\sigma$, i.e. z_α standard deviations below the mean μ. For c% = 95% and corresponding critical value $z_\alpha = -1.645$, $VaR_c = VaR_{1-\alpha}$ implies 95% probability of portfolio loss not exceeding 1.645σ, i.e., 5% probability of portfolio loss worse than 1.645σ. Similarly, for c% = 99% and corresponding critical value $z_\alpha = -2.2326$, $VaR_c = VaR_{1-\alpha}$ implies 99% probability of portfolio loss not exceeding 2.2326σ, i.e. 1% probability of portfolio loss worse than 2.2326σ. VaR *does not* specify the *amount of loss* expected in excess of VaR for the respective time period, but only specifies that there is only α% probability (i.e., event occurrences out of 100) resulting in loss of at least $z_\alpha\sigma$.

Traditional methods for estimating VaR

Hedge fund industry traditional methods for estimation of VaR for funds-of-funds risk management practices include the following (Darbyshire & Hampton, 2012, 2014; J.P. Morgan, 2008): i. Historical Simulation, ii. Parametric Method, and, iii. Monte Carlo Simulation.

While Historical Simulation is based upon actual data, Parametric Method uses the data only for generating the necessary parameters for specifying the distribution, and Monte Carlo generates data using simulation. Each of the three methods is different in terms of how it defines distribution of losses and has its advantages and limitations as discussed below.

i. Historical Simulation based VaR

Historical simulation relies upon the past data of returns based upon the assumption that historic monthly returns are an accurate representation of future returns with no specific assumptions about the return distribution. The data set of historical monthly % returns needs

to be adequately large to calculate for each historical % return a corresponding simulated P&L value by multiplying the % return with the index AuM. The simulated P&L values are then sorted in order of decreasing losses and increasing profits so that the highest loss is on the top and highest profit on the bottom. For each simulated P&L value, an associated cumulative weight is computed based upon total number of data points starting from the highest profit on the bottom for which the cumulative weight is simply the inverse of the number of data points. That value is incremented for each subsequent lower value of profit (or higher value of loss) with lowest profit (or highest loss) accumulating a final cumulative weight of 100%. The P&L value corresponding to c% confidence level value of the cumulative weights, where c% could be based upon interpolation between the adjacent P&L cumulative weights, is the estimated VaR for the specific confidence interval represented as VaR_c. Its key feature is that it is independent of any assumptions about the underlying statistical distribution or related parameters and is thus non-parametric in nature. Its advantages are the following: it is easy to calculate, easy to understand, does not assume normal distribution, not as data intensive as Monte Carlo, and can be applied to various time periods (J.P. Morgan, 2008). However, historical returns may not be an accurate representation of the future returns. Hence, its disadvantage lie in its assumption that historical correlations will repeat (J.P. Morgan, 2008).

ii. Parametric Method based VaR

For a portfolio of N risky assets, the portfolio variance is given by the expression: $\sigma_p^2 = W^T \sum W$ where W^T is the matrix transpose of W, the vector of individual asset class weights w_i, and, $\sum$, the variance-covariance matrix of the individual assets w_1 thru w_n:

$$\sum = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \vdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \vdots & \sigma_{nn} \end{pmatrix}$$

For portfolio standard deviation $\sigma_p = (W^T \sum W)^{1/2}$, estimated VaR is computed as follows:

$VaR_c = VaR_{1-\alpha} = Pz_\alpha \sigma_p$ where P is the market value of the portfolio. The vector of individual asset weights w_i is derived through the solution of the mean-variance optimization (Markowitz, 1952) for achieving a desired level of portfolio expected return for corresponding level of portfolio risk. The portfolio of weighted assets optimized to yield *minimum variance* (i.e. risk) for a higher expected portfolio return needs to be rebalanced through computation of new weights as market conditions and risk conditions evolve while considering transaction costs involved in such rebalancing. The mean-variance optimization problem can be stated in

terms of a target expected portfolio return r* as: min σ_p^2 = min $W^T \sum W$ s.t. $W^T R = r^*$, $\sum_i^N w_i = 1$ where R is the vector of mean returns of assets in a fully invested portfolio with respective returns elements corresponding to weights in the vector W of individual asset weights w_i. If short selling is not allowed, then additional constraint of $w_i \geq 0$ can be added.

The portfolio variance listed above follows from the following expressions:

$$\sigma_p^2 = \sum_{i=1}^N \sum_{j=1}^N w_i\, w_j \sigma_i \sigma_j \rho_{ij} \quad \Rightarrow \quad \sigma_p^2 = \sum_{i=1}^N \sum_{j=1}^N w_i\, w_j \sigma_{ij} \quad \Rightarrow \quad \sigma_p^2 = W^T \sum W$$

Parametric approach is mathematically simple and intuitive to understand and implement using matrices. Hence, its advantages include the following: it is easy to calculate, easy to understand, has minimal data requirements, and can be applied to various time periods (J.P. Morgan, 2008). Parametric methods suffer from the limitations inherent in the *normality* and *linearity* assumptions about portfolio returns being normally distributed and linear relationships assumed between risk variables. They rely upon strong assumptions about statistical parametric return distributions in terms of mean μ and standard deviation σ such as of the independent and identically distributed (iid) random variables of N (0, 1) normal distribution. Such assumptions of normality are clearly oversimplifications particularly for portfolios of hedge funds and funds of funds for which extensions of traditional VaR methods are discussed later. Furthermore, linear relationships of parametric methods are oversimplifications for portfolios employing sophisticated trading strategies based upon derivatives such as options that have non-linear risk-return characteristics. Therefore, such estimates are not as accurate when asset portfolio consists of non-linear instruments such as in case of specific hedge fund strategies. Hence, disadvantages of parametric VaR include assumption of normality, difficulty of estimating correlations in complex portfolios, and lesser accuracy for non-linear securities such as MBS (J.P. Morgan, 2008).

iii. Monte Carlo Simulation based VaR

Monte Carlo (MC) methods based VaR is based on the premise that the portfolio returns can be characterized by a *stochastic* model typically based upon a *non-deterministic* component. Such a component introduces some degree of uncertainty or randomness in the data generating process (DGP) by use of random number generators. Simulation based upon a specific mathematical stochastic model over thousands or millions of trials generate corresponding time-based *paths* that series of portfolio returns are probabilistically likely to follow over a certain time period. Each of those trials results in a terminal value for the portfolio return (or P&L) at the end of the simulated time period. Just as in the case of

Historical Simulation discussed earlier, VaR_c at a specific confidence interval c is estimated from the simulated P&L distribution by sorting the P&L values and computing P&L value corresponding to c% confidence level value by interpolation between adjacent P&L cumulative weights as needed. Consistent with prior discussion on Bayesian inference, the stochastic model is driven by the μ and σ of the asset returns distribution based upon historical data *as well as* inclusion of a degree of subjective knowledge based upon market experience in the model as necessary. MC methods are robust and probabilistically strong and are excellent for building and understanding non-linearity associated with use of derivatives in multi-asset portfolios. Such subjective knowledge based MC models are "extensively used throughout the financial markets" and are a "much used technique for estimating VaR within the hedge fund community" (Darbyshire & Hampton, 2012, 2014). Hence, incorporation of 'subjective judgments' into the model and flexibility of choosing the appropriate stochastic DGP are the strong points of MC methods based VaR models. MC methods being mathematically complex and challenging are most demanding of computational resources. When dealing with intrinsic complexities of specific multi-asset portfolio strategy with derivatives, they can become mathematically challenging and computationally expensive to implement. Markov Chain Monte Carlo (MCMC) algorithms (Gelfand & Smith, 1990; Malhotra, 2014) are often used in such cases for portfolio modeling especially in case of Bayesian inference models. Advantages of Monte Carlo VaR thus include their ability to use any return distribution or asset correlation and greatest suitability for non-linear assets while disadvantages include requirements of too many assumptions and extensive computing power and time (J.P. Morgan, 2008).

In addition to the traditional VaR methods, portfolio managers also run *stress tests* for testing sensitivity of the models to magnified values of parameters and risk factors to allow for extreme or adverse events that could result in catastrophic losses. From portfolio optimization perspective, stress testing also includes sensitivity analysis that shocks one or several risk factors by a relative small change such as +/- 5 basis points and revalues the portfolio to ascertain the sensitivity of the portfolio to the small change in one or several risk factors (J.P. Morgan, 2008). Similarly, they may also run *scenario analyses* using historical data and associated parameters to test for comparability with high turbulence market events such as the market crash of 1987 and the financial crisis of 2008.

Modified VaR

The normality assumption is the greatest drawback of the above traditional VaR approaches despite use of stress testing and scenario analysis practices. Particularly, hedge fund and fund-

of-funds returns are characterized by negative skew and excess positive kurtosis resulting in asymmetric return distributions with fat tails. Hence, extensions of traditional VaR methods have been proposed to address better estimation and specification of market risk for such portfolios. Contemporary extensions of VaR models are based upon explicit consideration of the standardized third (skew) and fourth (kurtosis) central moments of the returns distribution. In addition, they are also focused on the left tails of the returns distribution wherein most of the extreme losses are concentrated. The concept of Modified VaR is based upon the Modified Sharpe Ratio (MSR) wherein the denominator of Sharpe ratio is modified to account for the higher (third and fourth) moments of the returns distribution. Sharpe ratio which is a measure of risk-free rate per unit of risk, risk being measured in terms of portfolio's standard deviation, is modified by using the Cornish-Fisher expansion (Cornish & Fisher, 1937) to get the MSR. Cornish-Fisher expansion transformation helps transform a standard Gaussian random variable $z_\propto$ into a non-Gaussian z_{cf} random variable as follows:

$$z_\propto \approx N(0,1) \qquad E(z_\propto) = 0 \qquad E(z_\propto^2) = 1 \qquad E(z_\propto^3) = 0 \qquad E(z_\propto^4) = 3$$

$$z_{cf} \approx z_\propto + (z_\propto^2 - 1)\frac{S}{6} + (z_\propto^3 - 3z_\propto)\frac{K}{24} - (2z_\propto^3 - 5z_\propto)\frac{S^2}{36}$$

where sample skew is given by: $S = \frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{s}\right)^3$

and sample excess kurtosis by: $K = \left[\frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{s}\right)^4\right] - \frac{3(n-1)^2}{(n-2)(n-3)}$.

The portfolio Modified VaR is then given by $MVaR_c = MVaR_{1-\alpha} = \mu - z_{cf}\,\sigma_p$

and Modified Sharpe Ratio is given by $MSR = \frac{R_P - R_F}{MVaR_{1-\alpha}}$

where R_P is the annualized return and R_F is the annualized risk-free rate computed using T-bill as a proxy.

The above expression for Modified VaR $MVaR_{1-\alpha}$ represents a more accurate estimate of VaR at a c% confidence level, where c = 100(1 – α), μ = mean of the portfolio returns, and $z_\propto$ = critical value from the normal distribution for the specific confidence interval.

A limitation of the Modified VaR relates to higher confidence intervals (e.g. 99%) leading further into the left tail of the distribution and to inaccurate results. Another limitation is unreliability of MVaR in case of highly skewed and fat-tailed returns or P&L distributions.

Expected Shortfall

In addition to the non-normality and non-linearity related limitations of traditional VaR methodologies, VaR has additional limitation of not being a *coherent risk measure* (Artzner et al., 1999). A risk measure R (such as VaR) that is a *coherent risk measure* should satisfy *all* four following axioms for a random loss L.

- Subadditivity (diversification) $R\ (L_1 + L_2) \leq\ R\ (L_1)\ +\ R(L_2)$
    - Risk of portfolio of two assets should not be greater than the sum of risk of individual assets
- Positive homogeneity (scaling) $R\ (\lambda L) = \lambda R(L)$, for every $\lambda\ > 0$
    - Increasing size of portfolio by λ-times should increase risk by a multiple of $\lambda\ ceteris\ paribus$
- Monotonicity $R\ (L_1) < R(L_2)$ if $L_1 < L_2$
    - Higher risk is associated with higher loss and lesser risk with lesser loss, i.e., more +ve returns
- Transition property $R\ (L + a) < R\ (L) - a$
    - Adding cash or risk-free asset of value a should reduce risk by an equivalent amount a.

As VaR doesn't satisfy the first axiom of subadditivity, an alternative measure called Expected Shortfall was developed (Tasche, 2002). Expected Shortfall (ES) also known as Conditional VaR is the average of all the losses greater than (conditionally to going beyond VaR) VaR specified with the same confidence interval that VaR was estimated (J.P. Morgan, 2008). For example, if VaR is calculated at a 99% confidence level, ES averages the worst 1% losses. As the conditional expectation of loss conditional on its value exceeding VaR, ES is a coherent measure as it is subadditive unlike VaR. ES represents expected value (average) of the severity of losses beyond the VaR confidence threshold as these losses are important to regulators. A risk manager strictly relying upon VaR as the only risk measure may avoid losses within the confidence level while increasing the losses beyond the VaR level which are more severe and thus require the regulators or deposit insurers to backstop such losses. In addition, ES mitigates the disadvantages of VaR that result from the choice of a single confidence level and its impact on risk management decisions particularly as they relate to extreme events.

Mathematically, ES as the conditional expectation of loss conditional on its value exceeding VaR_c is described as: $ES_{1-\alpha}\ = E[L|L > VaR_{1-\alpha}\ ]$ where $ES_{1-\alpha}$ is estimated ES at confidence level

c for a loss distribution continuous in α. ES is the *average* loss in the distribution area beyond VaR in the extreme left-tail i.e. average of all VaRs from level α up to 1.

$$ES_{\alpha} \equiv \frac{1}{1-\alpha}\int_{\alpha}^{1} VaR_c\,(L)dc$$

where, L = a random loss with distribution function F_L, $\propto \epsilon\ (0,1)$ = confidence level close to 1.

It is important to recognize that ES gives only 'Expected' value that is the average value of risk in the left tail if the related VaR confidence level is exceeded. Hence, even though ES is a more conservative estimate than VaR, it is only the average or 'expected' loss in the left tail beyond VaR α. The actual loss (and related risk), however, could be more extreme than the average of the left tail risk. Hence, ES does not provide any information about the severity of loss by which VaR is exceeded. For more precise tail risk analysis of extreme events, Extreme Value Theory techniques (Embrechts et al., 1999; Gumbel, 2004; Pickands III, 1975) such as Block Maxima and Peaks over Threshold represent more sophisticated techniques but [just as in case of Bayesian inference] computationally and mathematically demanding alternatives which are often constrained by lack of adequate representative data for extreme events in case of hedge fund distributions thus leading to broad confidence intervals and weak significance estimates.

Bayesian VaRs beyond 'Bayesian vs. VaR' Dichotomy

Expected Shortfall overcomes classical problems of risk modeling associated with VaR while offering parsimony and transparency for lesser complexity and computational requirements. However, it must be reiterated that VaR and Bayesian modeling cannot be considered a dichotomy. Just like other statistical inference techniques available in both frequentist Null Hypothesis Significance Testing (NHST) and Bayesian statistical inference methodologies, VaR modeling continues to be used with both methodologies. For instance, the quasi-Bayesian and Bayesian versions of VaR have been referenced and applied in Banking & Finance practice since the years preceding the Global Financial Crisis (Hull & White, 1998; Venkataraman, 1997; Zangari, 1996). That being said, it is important to observe that both statistical inference paradigms, NHST *as well as* Bayesian, are moving away from point-estimates toward range based-estimates. In Bayesian VaR approaches, point estimates for parameters are substituted by distributions of parameters reflecting prior knowledge about the various parameter values with posterior distribution of parameters used for further analysis (Aussenegg & Miazhynskaia, 2006; Hoogerheide & van Dijk, 2008). Hence, there are both non-parametric

modeling methods such as historical simulation (discussed earlier), and adjusted historical simulation and parametric modeling methods such as Bayesian, quasi-maximum likelihood (QML) and bootstrap methods for various types of GARCH modeling and analysis. Increasing interest in sophisticated Bayesian VaR models and extensions is evident in research literatures (Aussenegg & Miazhynskaia, 2006; Casarin et al., 2013; Danielsson et al., 2014; Hoogerheide & van Dijk, 2008; Meucci, 2009; Miazhynskaia et al., 2003; Osiewalski & Pajor, 2010).

4. Data and Empirical Research Design

Empirical focus was on quantitative risk modeling of a half-trillion dollar fund-of-funds asset portfolio for a top Wall Street investment bank. Monthly returns over a 21-year period from January 1991 until December 2011 for 12 different asset classes comprising the portfolio were modeled in addition to market benchmark S&P 500 index (SPY). The specific [symbol:]asset classes included: (i) RIY: Developed Large Equity (proxy: Russell Developed Large Cap Index), (ii) RTY: Developed Small Equity (proxy: Russell Developed Small Cap Index), (iii) MXEF: Emerging Market Equity (proxy: MSCI Emerging Markets Index), (iv) LPX50TR: Listed Private Equity (proxy: LPX50 Listed Private Equity Index), (v) DJUBS: Various Commodities (proxy: DJ-UBS Commodity Index), (vi) USTW$: Major Currencies (proxy: Trade Weighted US Dollar Index: Major Currencies), (vii) HFRIEDI: Event Driven Hedge Fund (proxy: HFRI Event - Driven Index), (viii) HFRIEHI: Equity Hedge Fund (proxy: HFRI Equity Hedge (Total) Index), (ix) HFRIMAI: Merger Arbitrage Hedge Fund (proxy: HFRI ED: Merger Arbitrage Index), (x) HFRIMI: Macro Strategy Hedge Fund (proxy: HFRI Macro (Total) Index), (xi) HFRIRVA: Relative Value Hedge Fund (proxy: HFRI Relative Value (Total) Index), (xii) HFRIFOF: Fund of Funds Hedge Fund (proxy: HFRI Fund of Funds Index). All portfolio values, indexes, and returns were measured in US-Dollars (USD).

As noted, the key problem of model risk in any risk model such as VaR results from the fact that risk cannot be measured, but must be estimated using a statistical model (Danielsson et al., 2014) . In other words, model risk occurs because a statistical model is used for estimation of risk: *use of a model in itself entails model risk* (Derman, 1996; Morini, 2011). Consistent with industry practice guidelines, we used a range of different plausible risk models used in hedge fund risk modeling and analysis practice which can be robustly discriminated between, so that the disagreement between their range of readings could help us succinctly assess model risk (Danielsson et al., 2014). Given our focus of quantitative risk modeling on fund-of-funds multi-

asset portfolio construction and optimization, we applied standard practices used in the industry for risk modeling of hedge funds and funds-of-funds.

Statistical analysis of the various asset classes included basic performance plots such as *Value-Added Monthly Index* (*VAMI*) and Histograms; Probability Distributions and Probability Distribution Functions; Normality Tests including Distribution, Normal Q-Q Plot, and Jarque-Bera Normality Test; First Four Moment of Distributions with Skewness and Excess Kurtosis analyzed using both Distributions and Numeric Representations; Regression Plots for finding relationship between each fund asset class and the benchmark market index.

We used Risk-Adjusted Return Metrics applied in standard hedge fund risk modeling practice. These included risk models for Tracking Error, M1/M2 ratio of annualized first and second moments of distributions, Sharpe Ratio, Modified Sharpe Ratio, Sortino Ratio, Drawdown Ratio, Information Ratio, M-Squared Metric, Treynor Ratio, and, Jensen's Alpha. Those new to standard hedge fund risk modeling practice will find the industry specific interpretations and application details in Appendix 2 Hedge Fund Industry Risk-Adjusted Return Metrics relevant.

We used VAMI for tracking the comparative performance of different funds within the fund-of-funds. VAMI is an index of fund performance of a hypothetical $100 or $1000 investment in the specific asset class based on reinvestment of periodic returns. The focus of VAMI is on comparative assessment of risk in terms of draw-down, worst monthly draw-down, worst peak-to-valley-drawdown across different funds and fund managers (National Futures Association, 2013). Details about use of VAMI for risk assessment of funds are available in Appendix 3 Value Added Monthly Index (VAMI) Method.

VaR modeling for portfolio construction and portfolio optimization was done using Historical Simulation, Parametric Method, and Monte Carlo Simulation. Modified VaR was done using Modified Sharpe Ratio. Expected Shortfall was modeled to overcome the known limitations of VaR as a coherent risk measure. In addition to stressing of return to risk ratios for the various asset classes by modifying the assumptions, the portfolio was also stress tested using sensitivity analysis tests including use of equal weights for all asset classes, minimizing variance, maximizing return, and targeting a specific return. Portfolio modeling with the Returns Maximizing portfolio was examined for volatility and chosen for further advanced analysis using VaR, CVAR, ARCH/GARCH, and EVT.

5. Empirical Results

In this section, we discuss the main findings of market risk modeling of a half-trillion dollar fund-of-funds asset portfolio for a top Wall Street investment bank, 21-year monthly returns of 12 different asset classes. The two tracking error measures, quadratic standard deviation (SD) and linear mean absolute deviations (MAD), for each asset class are shown in Table 1. HFRIMAI tracks the S&P index most closely, whereas MXEF tracks S&P index least closely.

Basic performance plots shown in Table 2 for each asset include historical performance in terms of RoR%; VAMI; and, histogram of monthly returns. The ROR% charts show that while returns volatility of RIY, RTY, DJUBS, is comparable to the SPY benchmark; MXEF has more downside risk; LPX50TR has more upside return as well as downside risk; the currencies index USTW$ as well as all the six hedge fund indexes HFRIEDI, HFRIEHI, HFRIMAI, HFRIMI, HFRIRVA, HFRIFOF have *lower upside return as well as downside risk* relative to the benchmark. Relative indicates that the effect of hedges for the various hedge funds is realized consistent with hedging expectations.

VAMI plots for various asset classes show different risk-return behaviors relative to the market index. VAMI for RIY tracks the market VAMI most consistently, whereas RTY VAMI lags the market VAMI for first half but tracks it more consistently for the second half. While USTW$ VAMI remains around the starting value for most of the duration, VAMI for MXEF and DJUBS that lag the market for the first two-thirds period, track the market closely over the last third. *VAMI for all other asset classes show consistent outperformance of market VAMI with VAMI for two of the hedge fund asset classes, HFRIEHI* Equity Hedge Fund *and HFRIEDI* Event Driven Hedge Fund *demonstrating consistently higher highs and higher lows relative to all other asset classes*. Relative performance of VAMI for the hedge fund asset classes is clearly evident in Tables 1 and 2, and, Fig 1 which shows their comparison over the years.

Table 3 shows the comparison of the empirical distributions of the benchmark return. Besides visual analysis of normality and respective Q-Q normality plots, normality of the distributions is also assessed using the Jarque-Barra Test that jointly checks for skewness and excess kurtosis. In addition to the above findings, Table 3 also lists observed values of the first four moments of distribution for all asset class returns: mean, standard deviation (S.D.), skewness and excess kurtosis. Null hypothesis of normality is rejected for all the asset class return distributions. Highest mean value is 10.25 for the MXEF asset class (S.D. 23.94). *Next highest mean values are 10.23 (S.D. 9.46) and 10.09 (S.D. 6.96) for HFRIEHI and HFRIEDI respectively.*

Based on per unit risk analysis for the three highest mean returns, highest mean return per unit risk is delivered by HFRIEDI and HFRIEHI in that order.

The correlation matrix showing relative strength of variability of returns of the asset classes with respect to each other is shown in Table 4. *All asset classes show relatively low correlations with the market index which is a characteristic feature of the hedge funds as active fund managers are compensated for beating the market*. Each of the asset class returns was regressed against the benchmark and Adjusted R-Square for all regressions was found to be insignificant or negligible. *Even though less correlated with the market index, all asset class indices are strongly correlated with each other except for Currencies.* Currencies (USTW$) show low to moderate negative correlation with all other asset classes.

Interestingly, most other asset classes show moderate to strong positive correlations with each other. In particular, three asset classes, HFRIEDI and HFRIEHI besides HFRIFOF, have correlations exceeding 50% with all other asset classes except for commodities with which they have correlations exceeding 40%. As the 'most diversified' portfolio is the market index portfolio, it is expected that the hedge funds will be least correlated with it given the very raison d'être of hedge funds is to beat the market by active investment management. Ergo, it is plausible that the above very high correlations of HFRIEDI (52% to 87%) and HFRIEHI (56% to 87%) besides HFRIFOF (62% to 87%) with most other asset classes relate to hedging characteristics [which as observed above are] uncorrelated with the market index.

Mean-variance optimization was used to compute portfolio asset allocations for minimizing variance and for maximizing returns and compared with portfolio containing equal weights for all asset classes. The covariance matrix created for portfolio mean-variance optimization is shown in Table 5. Table 6 shows the Mean-Variance Optimization Portfolios based upon Equal Weights (Return 6.879%, Variance 0.071%, Sharpe Ratio 2.58); Maximizing Return While Minimizing Variance Portfolio of 9% USTW$, 50% HFRIMAI, 12% HFRIMI, and 29% HFRIRVA (Return 7.305%, Variance 0.008%, Sharpe Ratio 8.00); Minimizing Variance Portfolio of 31% USTW$, 34.6% HFRIMAI, 7.2% HFRIMI, and 27.2% HFRIRVA (Return 5.273%, Variance 0.006%, Sharpe Ratio 6.75); Targeted Return 10% Portfolio of 50% HFRIEDI, 42.50% HFRIEHI and 6.5% HFRIMAI (Return 10.000%, Variance 0.046%, Sharpe Ratio 4.66); *Maximizing Return Portfolio of 50% HFRIEDI and 50% HFRIEHI (Return 10.161%, Variance 0.052%, Sharpe Ratio 4.44)*. As return is maximized at the cost of increasing variance, Sharpe ratio is penalized accordingly.

Risk-adjusted return measures for all asset classes including M1/M2, Sharpe Ratio, MSR, Sortino Ratio, DD Ratio, Information Ratio, M-Squared Ratio, Treynor Ratio, and Jensen Ratio are shown in Table 7. For computation of risk-adjusted return measures, Risk-Free rate was considered as the average T-bill % rate of 3.13 based upon St. Louis Fed data for test duration of Jan '94 - Dec '11. The ranked ordered risk-adjusted return measures show some interesting patterns consistent with observations from prior risk models about the relative risk return characteristics of various asset classes. *Highest risk-adjusted returns in case of each risk-adjusted-measure (RAM) model were demonstrated by the hedge fund asset classes.* Currencies showed the lowest (negative) return in 8 of 9 RAM models. Commodities showed the second lowest (some negative) return in 6 of 9 RAM models. *All three of Sortino ratio, Jensen ratio, and Information Ratio show both HFRIEDI and HFRIEHI as the top two best performing asset classes*. In terms of aggregate RoR% rankings HFRIEDI is ranked 1st by Sortino, Information, and Treynor ratios; 2nd by Jensen ratio; and 3rd by M1/M2, Sharpe, MSR, DD, and M-squared ratios. HFRIEHI is ranked 1st by Jensen ratio and 2nd by Sortino and Information ratios. It is also plausible that the specific fund strategies that exploit known limitations of some ratios and non-stationarity and non-paramaetricity of returns distributions may be influencing finer ranking order of hedge funds relative to each other. *In any case, the broader pattern ranking the two hedge fund asset classes HFRIEDI and HFRIEHI is clearly discernible with this set of risk models and is consistent with prior findings with other risk models.*

Following upon earlier discussion about VaR and its various types as well as ES, Historical Simulation based VaR, Parametric VaR, Modified VaR, and Expected Shortfall were computed for the specific Mean Variance Portfolio Optimizations discussed above and summarized earlier in Table 6. The empirical results of Historical Simulation based VaR, Parametric VaR, Modified VaR, and Expected Shortfall for all asset classes equally weighted portfolio, variance minimizing portfolio, and return maximizing portfolio are presented in Tables 8 (a), 8(b), and 8(c) respectively. Parametric VaR is computed based upon mean-variance optimization. Modified VaR takes into consideration and accounts for non-normality of the returns. Expected Shortfall takes into consideration subadditivity responsible for portfolio diversification of risk with diverse assets, a factor missing from VaR models.

Table 8 (a) VaR and Expected Shortfall: At 95% confidence level, Optimization Portfolios based upon Equal Weights shows monthly Historical Simulation VaR of -$3,466,790 (indicating 5% chance of monthly losses exceeding this figure in any given month, and so on); Parametric VaR of -$4,382,848; Modified VaR of -$7,361,621; and, Expected Shortfall of -$5,258,022 (indicating

average expected loss of -$5,258,022 if the threshold level α of 5% was exceeded without any indication of worst case loss).

Table 8 (b) VaR and Expected Shortfall: At 95% confidence level, Optimization Portfolios based upon Minimum Variance shows Historical Simulation VaR of -$783,190; Parametric VaR of -$1,284,507; Modified VaR of -$1,884,524; and, Expected Shortfall of -$1,681,629.

Table 8 (c) VaR and Expected Shortfall: At 95% confidence level, Optimization Portfolios based upon Maximizing Return shows Historical Simulation VaR of -$2,764,562; Parametric VaR of -$3,766,260; Modified VaR of -$5,733,689; and, Expected Shortfall of -$4,575,377.

The specific values of various types of VaR and ES within the portfolio categories as well as across the categories are consistent with prior observations and discussions. Minimum Variance portfolio is oriented toward minimization of risk and hence demonstrates the lowest values for each type of VaR and ES relative to other portfolio categories. Maximizing Return portfolio is relatively a riskier portfolio and hence shows higher values for each of VaR types as well as for ES relative to Minimum Variance portfolio. The equally weighted portfolio is sub-optimized as apparent from its lowest Sharpe Ratio of 2.58 as compared with all other portfolio categories shown earlier in Table 6: Minimizing Variance portfolio with Sharpe Ratio of 6.75 and Maximizing Return portfolio with Sharpe Ratio of 4.44.

Within each portfolio category, Historical Simulation VaR has the smallest (negative) value (implying least loss), followed by Parametric VaR, Expected Shortfall, and Modified VaR in increasing order of loss. Parametric VaR is limited by its assumption of linear relationships between risk variables (because of exposure to non-linear asset classes such as derivatives) and the assumption of normality about distributions of hedge fund returns. Modified VaR (MVaR) explicitly accounts for the non-normality of hedge fund returns by taking into account skewness and excess kurtosis using the Cornish-Fisher expansion for the z_{α} critical value from the normal distribution for the respective confidence interval c.

Table 9 shows the results of the Portfolio modeling with the Returns Maximizing portfolio chosen for further advanced analysis using VaR, CVAR, ARCH/GARCH, and EVT. Portfolio PDF histogram shows it to have a negative skew with a long left tail and mass of the distribution concentrated on the right. Raw returns and squared returns show positive autocorrelations for short lags which decay to zero as the number of lags increases. Presence of heteroscedasticity in previous analysis indicates GARCH modeling is appropriate and model parameters are first estimated with the default GARCH (1,1) model shown in the table. Based

upon the model fitting with GARCH, generated residuals (innovations) and conditional standard deviations (sigmas) are examined showing volatility clustering. Standardized innovations show existence of autocorrelations. VaR Models examined include conditional VaR, Cornish-Fisher VaR, and EVT with observed findings shown in the table. Portfolio Cornish-Fisher VaR is found to be about 50% of the sum of individual funds VaRs.

6. Conclusion, Limitations, and Future Research

Current empirical model risk management research was motivated by ambiguity in recent research between model risk, modeling method (such as VaR), and statistical inference methodology (such as Bayesian). The ambiguity results from confusing choosing one model over another (or, one inference methodology over another) *ipso facto* as elimination of model risks. Such ambiguity may have serious consequences in further escalating *specification* and *estimation* errors in risk modeling. Ambiguity becomes all the more confusing when it is proposed that replacing a modeling method (such as VaR) with an inference methodology (such as Bayesian) will minimize the problems of [model] risk management (Borison & Hamm, 2010). Consistently, the current research focused on resolving the *Bayesian vs. VaR* dilemma to minimize model *specification* and *estimation* errors in risk modeling (Boucher et al., 2014).

The focus of the current research is on *Bayesian vs. VaR* given the specific contexts. Those contexts included the following: a high visibility journal article recommending replacing VaR models with Bayesian models *and* suggesting that it will minimize model risk; VaR research survey that clearly established its practice in *both* non-Bayesian and Bayesian forms since its beginning; and, Bayesian statistical inference modeling survey that clearly *established as critical a need (if not more)* for model risk management *than* is necessary in VaR modeling. Most importantly, the current focus of financial regulators on model risk management in aftermath of the Global Financial Crisis signifies its critical real world import for risk modeling practice.

The above contexts motivated our delineation of research and practice frameworks for both Bayesian inference as well as VaR modeling. Our primary focus on model risk management guided those delineations as well as related discussions. The same focus also guided the choice of our empirical context of demonstrating how model risk management can be applied in real practice for a top Wall Street investment bank *without replacing* 'VaR with Bayesian.' The choice of the frequentist methodology also underscores that it is *neither easy nor inexpensive* to do Bayesian *right* despite its many advantages over the frequentist methodology.

Further, even if the extra effort and [computational] expense is invested in Bayesian, it still doesn't do away with model risk management. In fact, based on the review of Bayesian VaR methodologies (Aussenegg & Miazhynskaia, 2006; Casarin et al., 2013; Danielsson et al., 2014; Hoogerheide & van Dijk, 2008; Meucci, 2009; Miazhynskaia et al., 2003; Osiewalski & Pajor, 2010), it is apparent that the need for model risk management is probably even more. This is not counterintuitive as often parsimony and transparency of modeling methods and modeling inference methodologies are recommended and preferred for this very reason.

This study has several limitations as choice of any *quantitative* statistical model or methodology entails model risk (Derman, 1996; Morini, 2011). Choosing frequentist inference methodology and VaR models – just like any other methodology and model – results in choosing to 'live with' [*but not at all ignore*] the limitations inherent in *each* such choice. Hence, use of multiple diverse modeling methods and methodologies at various levels of analysis can help cross-check for the various assumptions and boundaries that may not be within scope of one specific methodology or model (Danielsson et al., 2014). Having focused on the specific research for resolving the *Bayesian vs. VaR* dilemma *and* empirically demonstrating its application at a Wall Street bank, subsequent research plans to further address such methodological limitations. Such future research plans to focus on using VaR (and its various extensions including CVAR, ES, and EVT empirically demonstrated herein) as well as other models for analyzing market risk in portfolio construction and optimization. Further, given well-known advantages of Bayesian over frequentist inference (Kruschke, 2011) *as well as* its growing feasibility *with* MCMC (Gelfand & Smith, 1990; Malhotra, 2014), such future research plans to advance on the Bayesian and VaR analytical frameworks proposed herein for empirical analysis of such models.

TABLE 1

Tracking Error	RIY	RTY	MXEF	LPX50TR	DJUBS	USTW$	HFRIEDI	HFRIEHI	HFRIMAI	HFRIMI	HFRIRVA	HFRIFOF
Quadratic (SD)	0.0599	0.0689	0.0814	0.0796	0.0625	0.0476	0.0461	0.0492	0.0432	0.0501	0.0439	0.0461
Linear (MAD)	0.0462	0.0526	0.0627	0.0572	0.0484	0.0366	0.0357	0.0385	0.0334	0.0401	0.0338	0.0358

TABLE 2

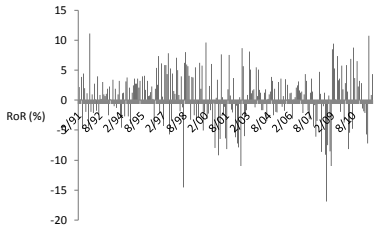
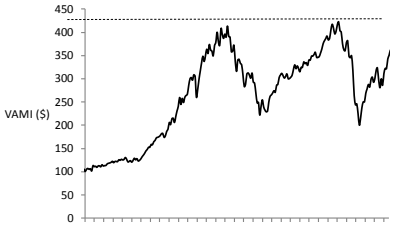
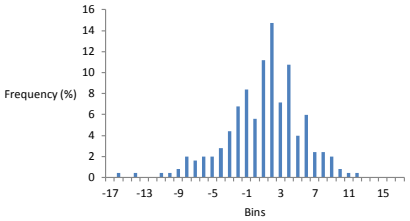
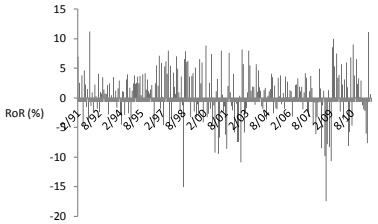
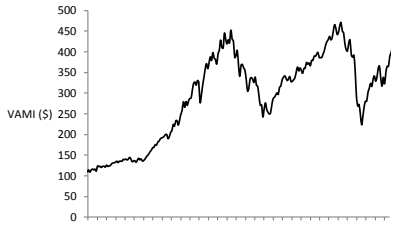
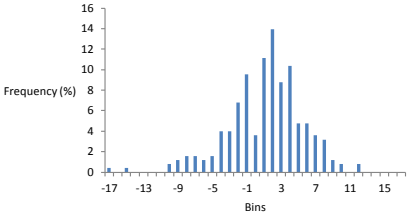
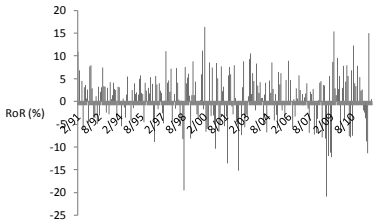
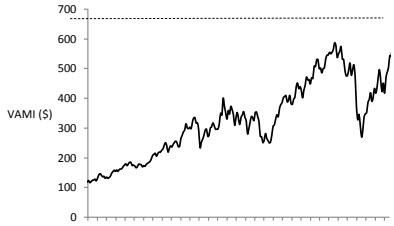
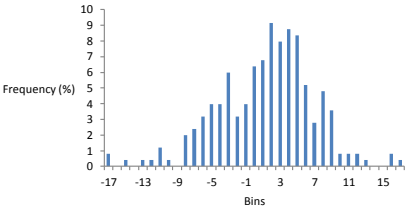
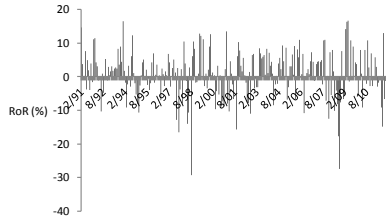
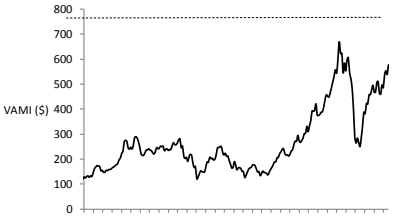
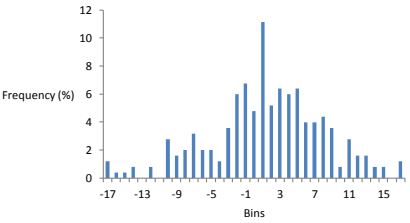
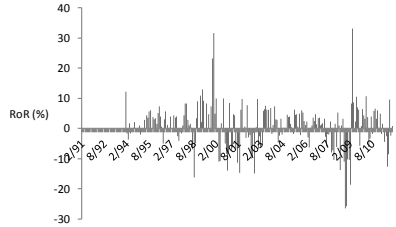
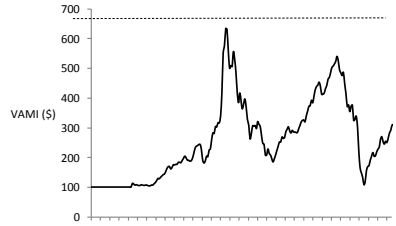
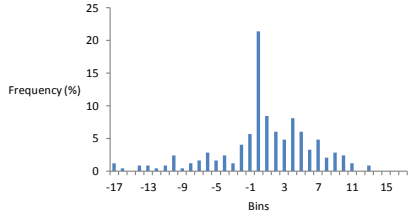
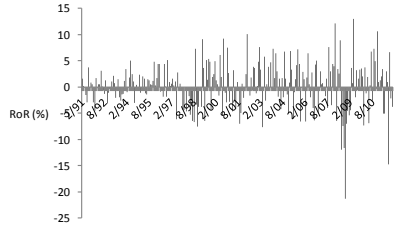
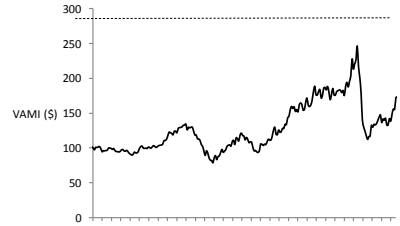
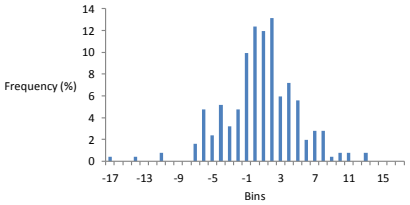
INDEX	RETURNS	VAMI	HISTOGRAM
SPY			
RIY			
RTY			
MXEF			
LPX50TR			
DJUBS			

TABLE 2 (continued)

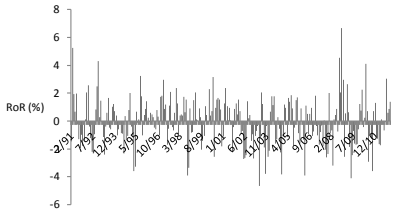
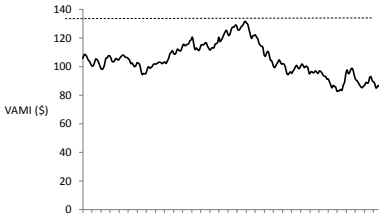
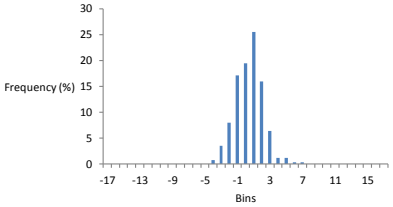
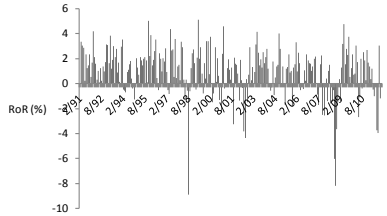
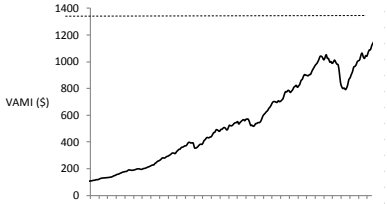
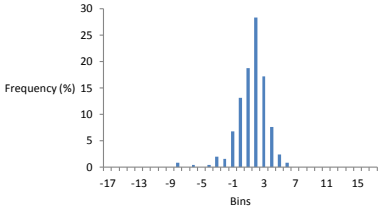
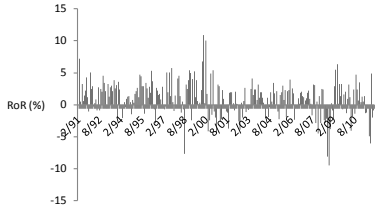
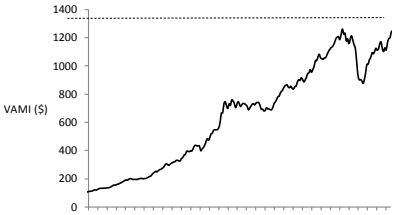
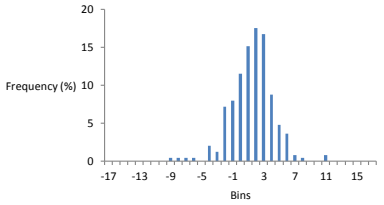
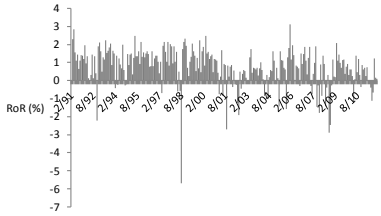
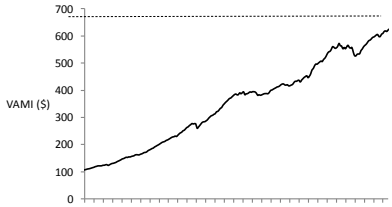
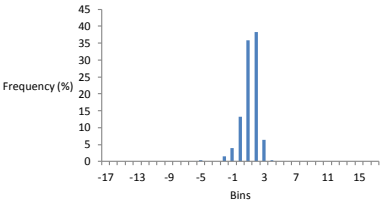
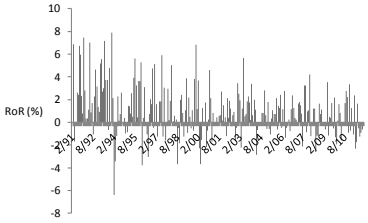
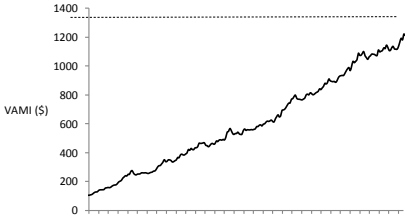
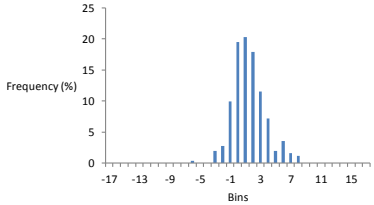
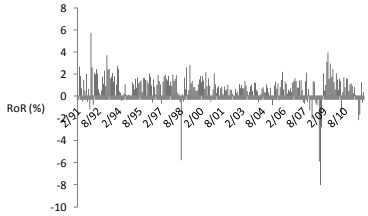
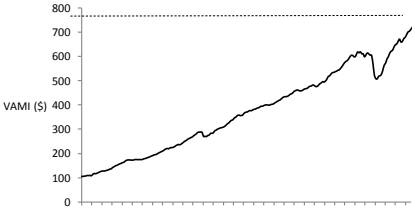
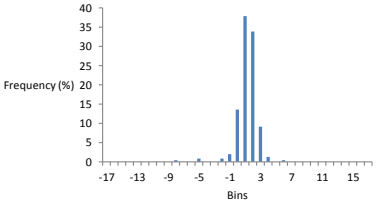
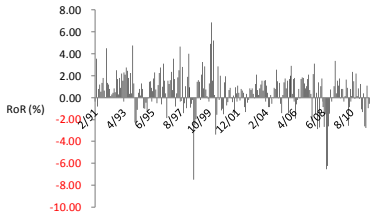
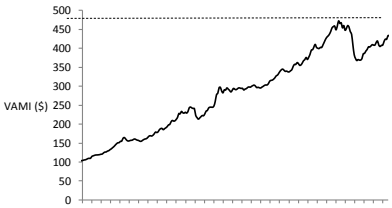
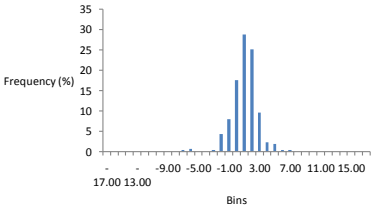
INDEX	RETURNS	VAMI	HISTOGRAM
USTW$			
HFRIEDI			
HFRIEHI			
HFRIMAI			
HFRIMI			
HFRIRVA			
HFRIFOF			

Figure 1: VAMI Values for All Asset Classes in the Multi-Asset Portfolio

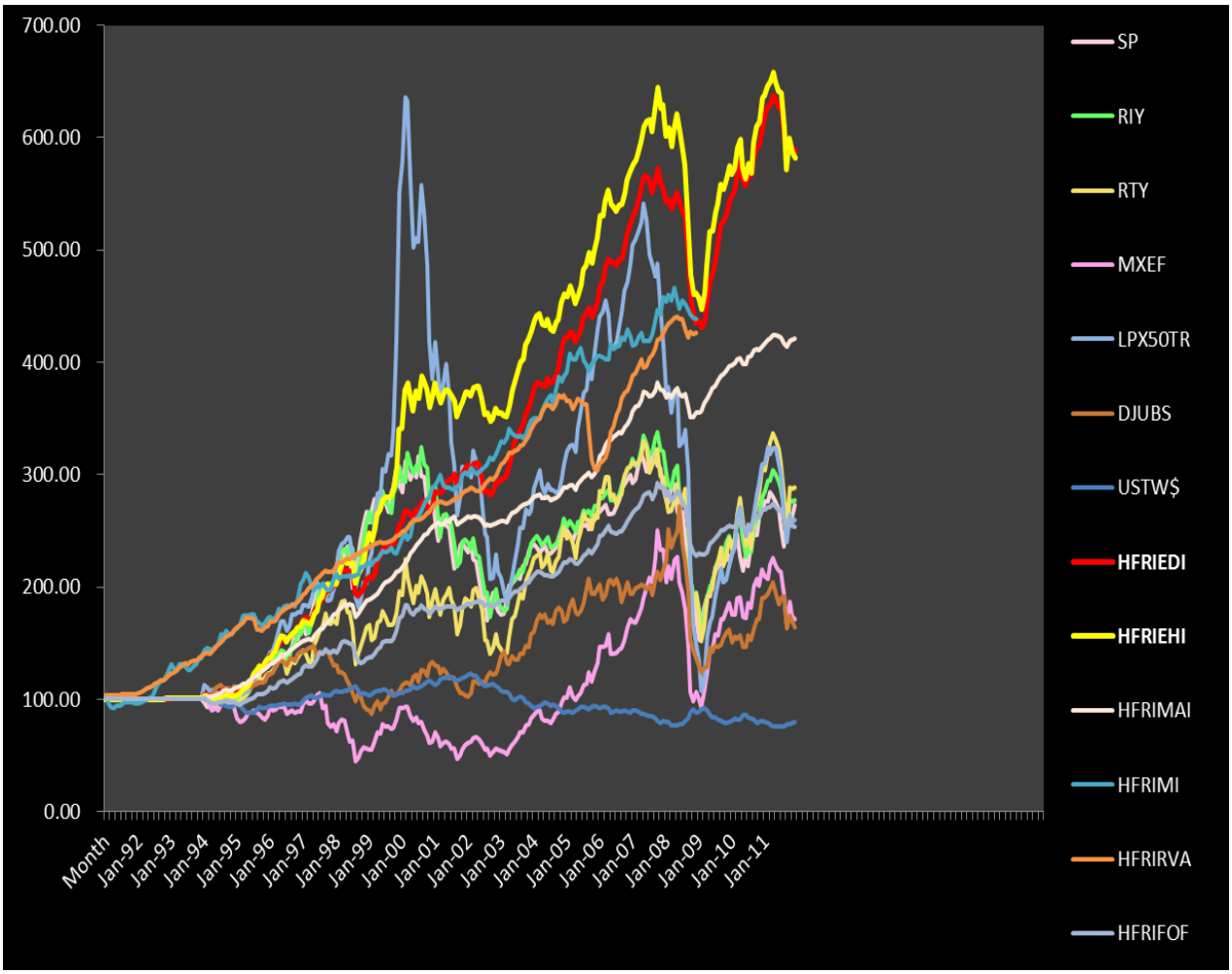


TABLE 3

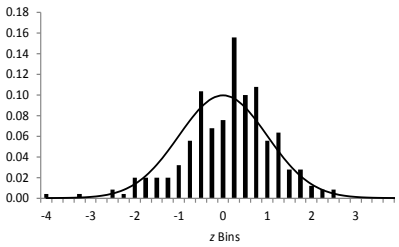
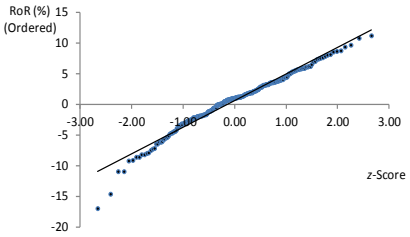
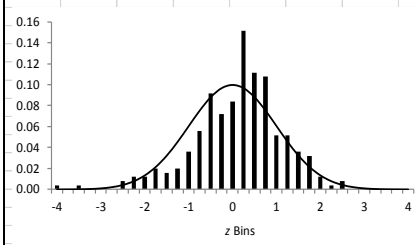
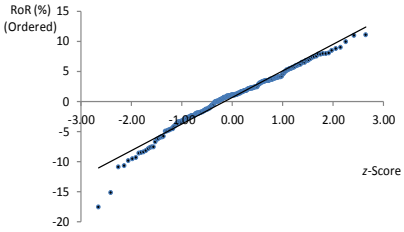
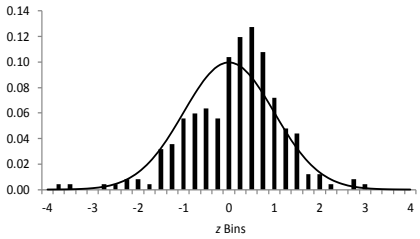
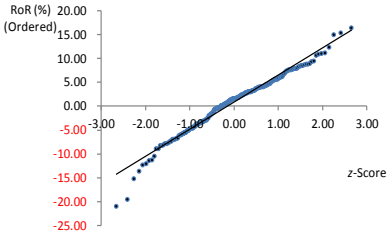
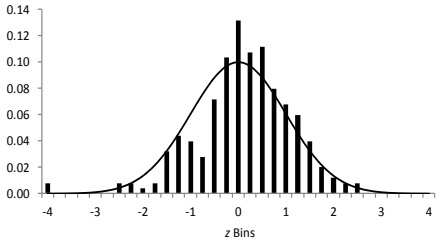
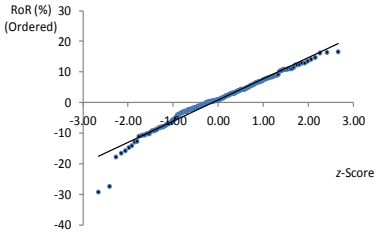
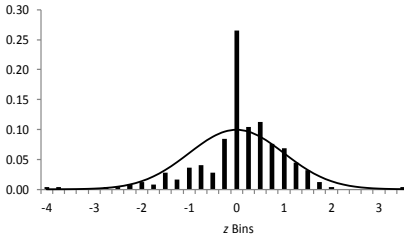
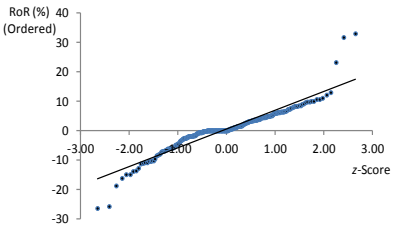
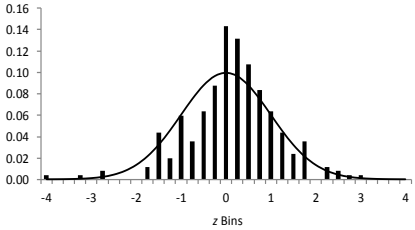
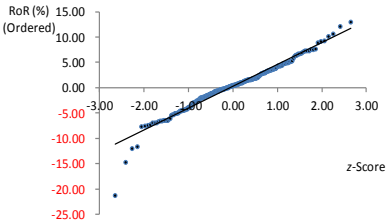
INDEX	EMPRICAL vs. NORMAL DISTBNS.	Q-Q NORMALITY PLOT	MOMENTS & JARQUE-BERA TEST
SPY			**Moments** Mean 7.25 St. Dev. 15.03 Skewness -0.6027 **Excess Kurtosis 1.2005** **Normality Test** Jarque-Bera Value 30.2687 Level of Significance 5% Degrees of Freedom 2 Critical Value 5.9915 *p*-Value 0.0000 **Critical Test** ***p* Test** Normality NO NO **Null Hypothesis of Normality is Rejected**
RIY			**Moments** Mean 6.94 St. Dev. 15.86 Skewness -0.6815 **Excess Kurtosis 1.0725** **Normality Test** Jarque-Bera Value 27.0705 Level of Significance 5% Degrees of Freedom 2 Critical Value 5.9915 *p*-Value 0.0000 **Critical Test** ***p* Test** Normality NO NO **Null Hypothesis of Normality is Rejected**
RTY			**Moments** Mean 9.79 St. Dev. 19.57 Skewness -0.5102 **Excess Kurtosis 1.1073** **Normality Test** Jarque-Bera Value 23.7136 Level of Significance 5% Degrees of Freedom 2 Critical Value 5.9915 *p*-Value 0.0000 **Critical Test** ***p* Test** Normality NO NO **Null Hypothesis of Normality is Rejected**
MXEF			**Moments** Mean 10.25 St. Dev. 23.94 Skewness -0.7002 **Excess Kurtosis 1.9299** **Normality Test** Jarque-Bera Value 59.4637 Level of Significance 5% Degrees of Freedom 2 Critical Value 5.9915 *p*-Value 0.0000 **Critical Test** ***p* Test** Normality NO NO **Null Hypothesis of Normality is Rejected**
LPX50TR			**Moments** Mean 8.36 St. Dev. 24.78 Skewness 0.0725 **Excess Kurtosis 4.5070** **Normality Test** Jarque-Bera Value 183.0062 Level of Significance 5% Degrees of Freedom 2 Critical Value 5.9915 *p*-Value 0.0000 **Critical Test** ***p* Test** Normality NO NO **Null Hypothesis of Normality is Rejected**
DJUBS			**Moments** Mean 3.05 St. Dev. 14.97 Skewness -0.5479 **Excess Kurtosis 2.7376** **Normality Test** Jarque-Bera Value 90.9380 Level of Significance 5% Degrees of Freedom 2 Critical Value 5.9915 *p*-Value 0.0000 **Critical Test** ***p* Test** Normality NO NO **Null Hypothesis of Normality is Rejected**

TABLE 3 (continued)

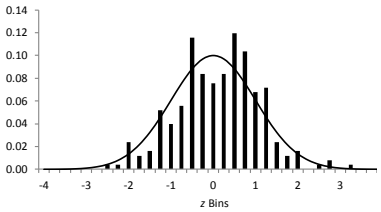
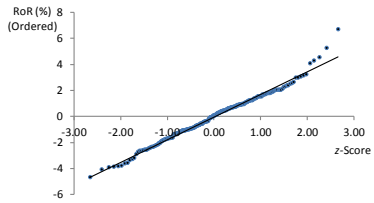
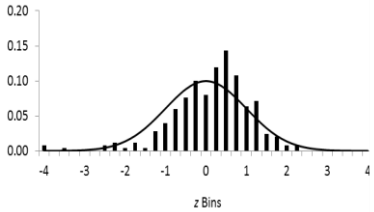
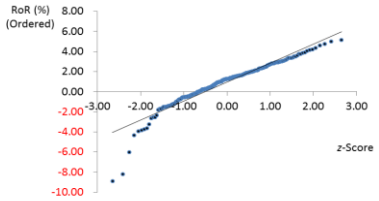
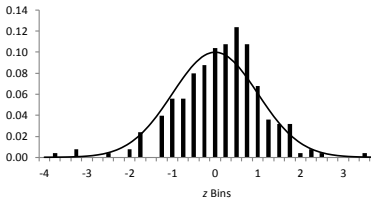
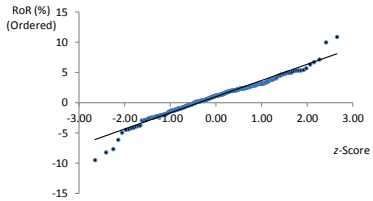
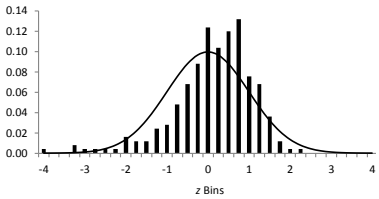
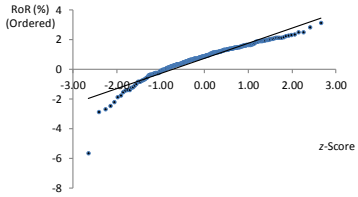
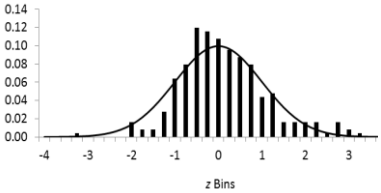
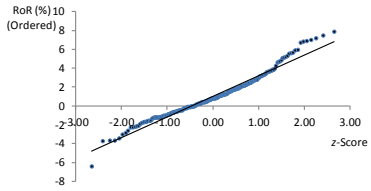
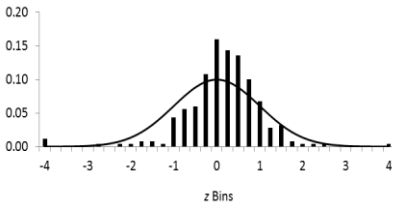
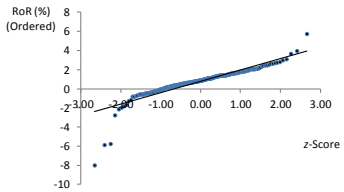
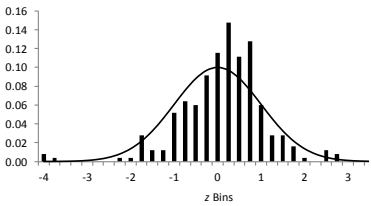
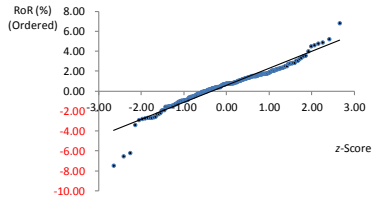
INDEX	EMPRICAL vs. NORMAL DISTBNS.	Q-Q NORMALITY PLOT	MOMENTS & JARQUE-BERA TEST
USTW$			**Moments** Mean -1.07 St. Dev. 5.90 Skewness 0.0491 **Excess Kurtosis 0.7959** **Normality Test** Jarque-Bera Value 7.7128 Level of Significance 5% Degrees of Freedom 2 Critical Value 5.9915 *p*-Value 0.0211 **Critical Test** ***p* Test** Normality NO NO **Null Hypothesis of Normality is Rejected**
HFRIEDI			**Moments** Mean 10.09 St. Dev. 6.96 Skewness -1.2760 **Excess Kurtosis 4.0689** **Normality Test** Jarque-Bera Value 282.5562 Level of Significance 5% Degrees of Freedom 2 Critical Value 5.9915 *p*-Value 0.0000 **Critical Test** ***p* Test** Normality NO NO **Null Hypothesis of Normality is Rejected**
HFRIEHI			**Moments** Mean 10.23 St. Dev. 9.46 Skewness -0.2129 **Excess Kurtosis 1.9992** **Normality Test** Jarque-Bera Value 44.2843 Level of Significance 5% Degrees of Freedom 2 Critical Value 5.9915 *p*-Value 0.0000 **Critical Test** ***p* Test** Normality NO NO **Null Hypothesis of Normality is Rejected**
HFRIMAI			**Moments** Mean 8.05 St. Dev. 3.68 Skewness -1.6106 **Excess Kurtosis 6.1270** **Normality Test** Jarque-Bera Value 458.8732 Level of Significance 5% Degrees of Freedom 2 Critical Value 5.9915 *p*-Value 0.0000 **Critical Test** ***p* Test** Normality NO NO **Null Hypothesis of Normality is Rejected**
HFRIMI			**Moments** Mean 8.40 St. Dev. 6.69 Skewness 0.1786 **Excess Kurtosis 0.9909** **Normality Test** Jarque-Bera Value 21.9863 Level of Significance 5% Degrees of Freedom 2 Critical Value 5.9915 *p*-Value 0.0000 **Critical Test** ***p* Test** Normality NO NO **Null Hypothesis of Normality is Rejected**
HFRIRVA			**Moments** Mean 8.15 St. Dev. 4.37 Skewness -2.8779 **Excess Kurtosis 15.9776** **Normality Test** Jarque-Bera Value 2046.5842 Level of Significance 5% Degrees of Freedom 2 Critical Value 5.9915 *p*-Value 0.0000 **Critical Test** ***p* Test** Normality NO NO **Null Hypothesis of Normality is Rejected**
HFRIFOF			**Moments** Mean 5.33 St. Dev. 6.10 Skewness -0.6651 **Excess Kurtosis 3.6812** **Normality Test** Jarque-Bera Value 161.1758 Level of Significance 5% Degrees of Freedom 2 Critical Value 5.9915 *p*-Value 0.0000 **Critical Test** ***p* Test** Normality NO NO **Null Hypothesis of Normality is Rejected**

TABLE 4 Correlation Matrix

	SP	RIY	RTY	MXEF	LPX50TR	DJUBS	USTW$	HFRIEDI	HFRIEHI	HFRIMAI	HFRIMI	HFRIRVA	HFRIFOF
SP	1.000												
RIY	**0.099**	1.000											
RTY	**0.103**	0.835	1.000										
MXEF	**0.023**	0.751	0.732	1.000									
LPX50TR	**0.130**	0.730	0.745	0.679	1.000								
DJUBS	**(0.033)**	0.331	0.337	0.455	0.292	1.000							
USTW$	**(0.043)**	(0.251)	(0.220)	(0.320)	(0.131)	(0.309)	1.000						
HFRIEDI	**0.109**	0.750	0.802	0.747	0.723	0.415	(0.273)	1.000					
HFRIEHI	**0.115**	0.781	0.843	0.759	0.766	0.462	(0.244)	0.872	1.000				
HFRIMAI	**0.150**	0.596	0.615	0.574	0.501	0.322	(0.143)	0.769	0.679	1.000			
HFRIMI	**(0.063)**	0.345	0.380	0.458	0.315	0.365	(0.106)	0.517	0.563	0.343	1.000		
HFRIRVA	**0.129**	0.596	0.579	0.606	0.609	0.459	(0.295)	0.815	0.721	0.693	0.315	1.000	
HFRIFOF	**0.050**	0.616	0.655	0.732	0.662	0.463	(0.190)	0.850	0.869	0.633	0.715	0.751	1.000

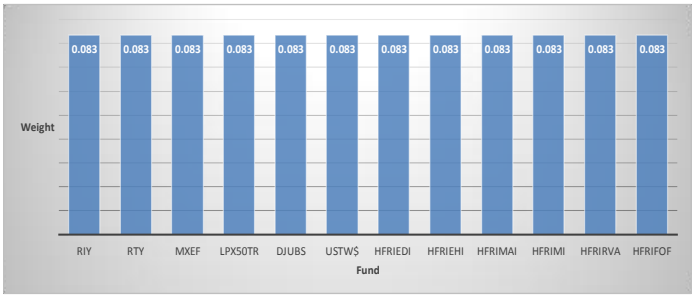
TABLE 5 Covariance Matrix (Σ)

	SP	RIY	RTY	MXEF	LPX50TR	DJUBS	USTW$	HFRIEDI	HFRIEHI	HFRIMAI	HFRIMI	HFRIRVA	HFRIFOF
SP	0.002043	0.000204	0.000272	7.2E-05	0.000419	-6.8E-05	-3.3E-05	9.88E-05	0.000142	7.18E-05	-5.5E-05	7.34E-05	3.94E-05
RIY	0.000204	0.002086	0.00223	0.002423	0.002378	0.000693	-0.00019	0.000687	0.000972	0.000288	0.000303	0.000343	0.000494
RTY	0.000272	0.00223	0.003419	0.003021	0.003109	0.000905	-0.00022	0.00094	0.001344	0.000381	0.000428	0.000426	0.000673
MXEF	7.2E-05	0.002423	0.003021	0.004983	0.003421	0.001473	-0.00038	0.001056	0.001461	0.000429	0.000623	0.000539	0.000907
LPX50TR	0.000419	0.002378	0.003109	0.003421	0.005093	0.000957	-0.00016	0.001034	0.001489	0.000379	0.000433	0.000547	0.000829
DJUBS	-6.8E-05	0.000693	0.000905	0.001473	0.000957	0.002105	-0.00024	0.000381	0.000578	0.000157	0.000323	0.000265	0.000373
USTW$	-3.3E-05	-0.00019	-0.00022	-0.00038	-0.00016	-0.00024	0.000288	-9.3E-05	-0.00011	-2.6E-05	-3.5E-05	-6.3E-05	-5.7E-05
HFRIEDI	9.88E-05	0.000687	0.00094	0.001056	0.001034	0.000381	-9.3E-05	0.000402	0.000476	0.000163	0.000199	0.000206	0.000299
HFRIEHI	0.000142	0.000972	0.001344	0.001461	0.001489	0.000578	-0.00011	0.000476	0.000743	0.000112	7E-05	9.24E-05	0.000118
HFRIMAI	7.18E-05	0.000288	0.000381	0.000429	0.000379	0.000157	-2.6E-05	0.000163	0.000196	0.000112	7E-05	9.24E-05	0.000118
HFRIMI	-5.5E-05	0.000303	0.000428	0.000623	0.000433	0.000323	-3.5E-05	0.000199	0.000295	7E-05	0.000371	7.64E-05	0.000242
HFRIRVA	7.34E-05	0.000343	0.000426	0.000539	0.000547	0.000265	-6.3E-05	0.000206	0.000247	9.24E-05	7.64E-05	0.000159	0.000166
HFRIFOF	3.94E-05	0.000494	0.000673	0.000907	0.000829	0.000373	-5.7E-05	0.000299	0.000416	0.000118	0.000242	0.000166	0.000308

TABLE 6 Mean Variance Portfolio Optimization

Optimization Portfolios based upon Equal Weights

Return Matrix (*R*)		Weight Matrix (*W*)		
Fund	Return (%)	Fund	Weight	Limit
SP	6.83%	SP	0.000	0.000
RIY	6.94%	RIY	0.083	0.500
RTY	7.95%	RTY	0.083	0.500
MXEF	6.06%	MXEF	0.083	0.500
LPX50TR	8.36%	LPX50TR	0.083	0.500
DJUBS	4.04%	DJUBS	0.083	0.500
USTW$	-1.07%	USTW$	0.083	0.500
HFRIEDI	10.09%	HFRIEDI	0.083	0.500
HFRIEHI	10.23%	HFRIEHI	0.083	0.500
HFRIMAI	8.05%	HFRIMAI	0.083	0.500
HFRIMI	8.40%	HFRIMI	0.083	0.500
HFRIRVA	8.15%	HFRIRVA	0.083	0.500
HFRIFOF	5.33%	HFRIFOF	0.083	0.500
			1.00	



Portfolio Return (%) 6.879%

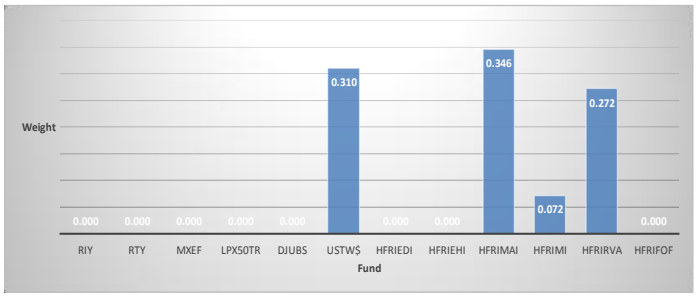
Sharpe Ratio 2.58 Assuming Rf = 0

Target Return (%) 6.879% Maximize

Target Variance (%) 0.071% Minimize

Optimization Portfolios based upon Minimizing Variance

Return Matrix (*R*)		Weight Matrix (*W*)		
Fund	Return (%)	Fund	Weight	Limit
SP	6.83%	SP	0.000	0.000
RIY	6.94%	RIY	0.000	0.500
RTY	7.95%	RTY	0.000	0.500
MXEF	6.06%	MXEF	0.000	0.500
LPX50TR	8.36%	LPX50TR	0.000	0.500
DJUBS	4.04%	DJUBS	0.000	0.500
USTW$	-1.07%	USTW$	0.310	0.500
HFRIEDI	10.09%	HFRIEDI	0.000	0.500
HFRIEHI	10.23%	HFRIEHI	0.000	0.500
HFRIMAI	8.05%	HFRIMAI	0.346	0.500
HFRIMI	8.40%	HFRIMI	0.072	0.500
HFRIRVA	8.15%	HFRIRVA	0.272	0.500
HFRIFOF	5.33%	HFRIFOF	0.000	0.500
			1.00	



Portfolio Return (%) 5.273%

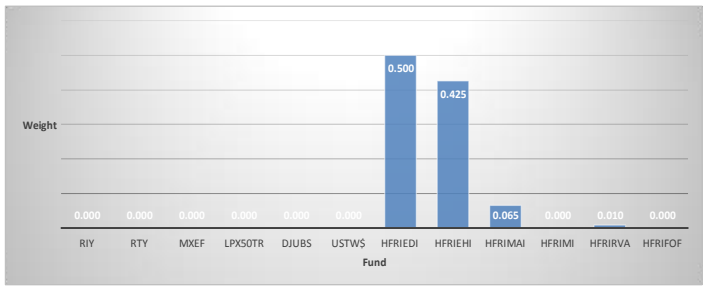
Sharpe Ratio 6.75 Assuming Rf = 0

Target Return (%) 5.273% Maximize

Target Variance (%) 0.006% Minimize

Optimization Portfolios based upon Targeted Return 10%

Return Matrix (*R*)		Weight Matrix (*W*)		
Fund	Return (%)	Fund	Weight	Limit
SP	6.83%	SP	0.000	0.000
RIY	6.94%	RIY	0.000	0.500
RTY	7.95%	RTY	0.000	0.500
MXEF	6.06%	MXEF	0.000	0.500
LPX50TR	8.36%	LPX50TR	0.000	0.500
DJUBS	4.04%	DJUBS	0.000	0.500
USTW$	-1.07%	USTW$	0.000	0.500
HFRIEDI	10.09%	HFRIEDI	0.500	0.500
HFRIEHI	10.23%	HFRIEHI	0.425	0.500
HFRIMAI	8.05%	HFRIMAI	0.065	0.500
HFRIMI	8.40%	HFRIMI	0.000	0.500
HFRIRVA	8.15%	HFRIRVA	0.010	0.500
HFRIFOF	5.33%	HFRIFOF	0.000	0.500
			1.00	



Portfolio Return (%) 10.000%

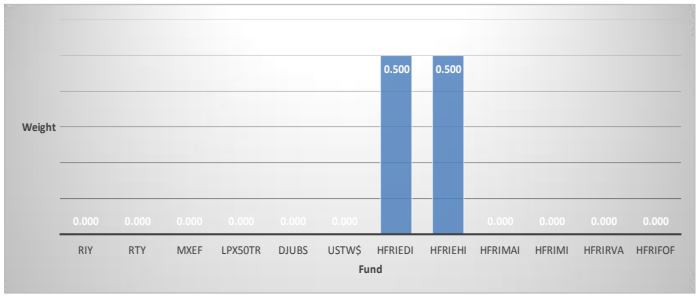
Sharpe Ratio 4.66 Assuming Rf = 0

Target Return (%) 10.000% Optimize

Target Variance (%) 0.046% Minimize

Optimization Portfolios based upon Maximizing Return

Return Matrix (*R*)		Weight Matrix (*W*)		
Fund	Return (%)	Fund	Weight	Limit
SP	6.83%	SP	0.000	0.000
RIY	6.94%	RIY	0.000	0.500
RTY	7.95%	RTY	0.000	0.500
MXEF	6.06%	MXEF	0.000	0.500
LPX50TR	8.36%	LPX50TR	0.000	0.500
DJUBS	4.04%	DJUBS	0.000	0.500
USTW$	-1.07%	USTW$	0.000	0.500
HFRIEDI	10.09%	HFRIEDI	0.500	0.500
HFRIEHI	10.23%	HFRIEHI	0.500	0.500
HFRIMAI	8.05%	HFRIMAI	0.000	0.500
HFRIMI	8.40%	HFRIMI	0.000	0.500
HFRIRVA	8.15%	HFRIRVA	0.000	0.500
HFRIFOF	5.33%	HFRIFOF	0.000	0.500
			1.00	



Portfolio Return (%) 10.161%

Sharpe Ratio 4.44 Assuming Rf = 0

Target Return (%) 10.161% Optimize

Target Variance (%) 0.052% Minimize

TABLE 7 Risk Adjusted Return Measures

	M1/M2	Sharpe	MSR	Sortino	DD Ratio	Information	M squared	Treynor	Jensen
SP	0.436	0.236	0.523	0.607	0.139	-	-	-	-
RIY	**0.439**	0.241	0.538	0.624	0.145	0.063	0.640	38.157	3.400
RTY	**0.393**	0.238	0.516	0.700	0.129	0.554	0.597	36.284	4.233
MXEF	**0.248**	0.120	0.283	0.349	0.070	(0.321)	(1.253)	83.123	2.870
LPX50TR	**0.338**	0.212	0.438	0.695	0.090	0.670	0.185	25.511	4.255
DJUBS	**0.254**	0.057	0.132	0.097	0.027	(1.474)	(2.230)	(27.320)	1.224
USTW$	**(0.182)**	(0.714)	(1.553)	(1.147)	(0.390)	(5.590)	(14.316)	260.688	(3.986)
HFRIEDI	**1.454**	1.003	2.183	1.922	0.501	2.386	12.580	144.022	6.833
HFRIEHI	**1.083**	0.752	1.401	1.723	0.349	2.354	8.642	101.927	6.852
HFRIMAI	**2.194**	1.341	2.814	1.506	0.603	0.943	17.865	140.008	4.860
HFRIMI	**1.260**	0.791	1.341	1.396	0.502	1.086	9.249	(195.247)	5.551
HFRIRVA	**1.870**	1.152	4.487	1.502	0.464	1.012	14.908	139.942	4.960
HFRIFOF	**0.876**	0.362	0.769	0.514	0.164	(1.089)	2.535	114.262	2.227

Ranked Results

M1/M2		Sharpe		MSR		Sortino		DD Ratio	
HFRIMAI	2.194	HFRIMAI	1.341	HFRIRVA	4.487	HFRIEDI	1.922	HFRIMAI	0.603
HFRIRVA	1.870	HFRIRVA	1.152	HFRIMAI	2.814	HFRIEHI	1.723	HFRIMI	0.502
HFRIEDI	1.454	HFRIEDI	1.003	HFRIEDI	2.183	HFRIMAI	1.506	HFRIEDI	0.501
HFRIMI	1.260	HFRIMI	0.791	HFRIEHI	1.401	HFRIRVA	1.502	HFRIRVA	0.464
HFRIEHI	1.083	HFRIEHI	0.752	HFRIMI	1.341	HFRIMI	1.396	HFRIEHI	0.349
HFRIFOF	0.876	HFRIFOF	0.362	HFRIFOF	0.769	RTY	0.700	HFRIFOF	0.164
RIY	0.439	RIY	0.241	RIY	0.538	LPX50TR	0.695	RIY	0.145
SP	0.436	RTY	0.238	SP	0.523	RIY	0.624	SP	0.139
RTY	0.393	SP	0.236	RTY	0.516	SP	0.607	RTY	0.129
LPX50TR	0.338	LPX50TR	0.212	LPX50TR	0.438	HFRIFOF	0.514	LPX50TR	0.090
DJUBS	0.254	MXEF	0.120	MXEF	0.283	MXEF	0.349	MXEF	0.070
MXEF	0.248	DJUBS	0.057	DJUBS	0.132	DJUBS	0.097	DJUBS	0.027
USTW$	(0.182)	USTW$	(0.714)	USTW$	(1.553)	USTW$	(1.147)	USTW$	(0.390)

Ranked Results

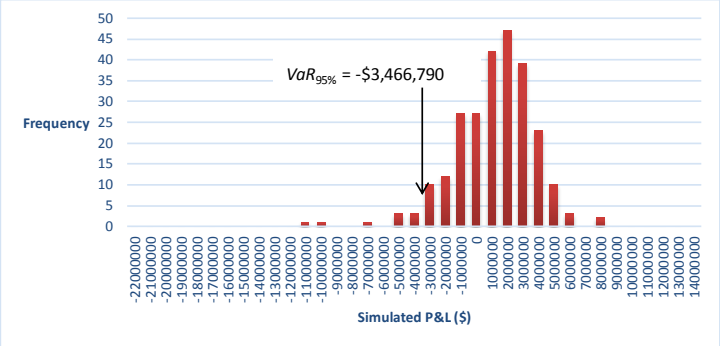
Sharpe		Information		M squared		Treynor		Jensen	
HFRIMAI	1.341	HFRIEDI	2.386	HFRIMAI	17.865	HFRIEDI	144.022	HFRIEHI	6.852
HFRIRVA	1.152	HFRIEHI	2.354	HFRIRVA	14.908	HFRIMAI	140.008	HFRIEDI	6.833
HFRIEDI	1.003	HFRIMI	1.086	HFRIEDI	12.580	HFRIRVA	139.942	HFRIMI	5.551
HFRIMI	0.791	HFRIRVA	1.012	HFRIMI	9.249	HFRIFOF	114.262	HFRIRVA	4.960
HFRIEHI	0.752	HFRIMAI	0.943	HFRIEHI	8.642	HFRIEHI	101.927	HFRIMAI	4.860
HFRIFOF	0.362	LPX50TR	0.670	HFRIFOF	2.535	MXEF	83.123	LPX50TR	4.255
RIY	0.241	RTY	0.554	RIY	0.640	RIY	38.157	RTY	4.233
RTY	0.238	RIY	0.063	RTY	0.597	RTY	36.284	RIY	3.400
SP	0.236	SP	0.000	LPX50TR	0.185	LPX50TR	25.511	MXEF	2.870
LPX50TR	0.212	MXEF	(0.321)	SP	0.000	SP	0.000	HFRIFOF	2.227
MXEF	0.120	HFRIFOF	(1.089)	MXEF	(1.253)	DJUBS	0.000	DJUBS	1.224
DJUBS	0.057	DJUBS	(1.474)	DJUBS	(2.230)	USTW$	0.000	SP	0.000
USTW$	(0.714)	USTW$	(5.590)	USTW$	(14.316)	HFRIMI	(195.247)	USTW$	(3.986)

TABLE 8 (a) VaR and Expected Shortfall: Optimization Portfolios based upon Equal Weights

Historical Simulation VaR

PORT Index AuM ($)	100,000,000
Confidence Level	95%
Critical Value (z_α)	1.645

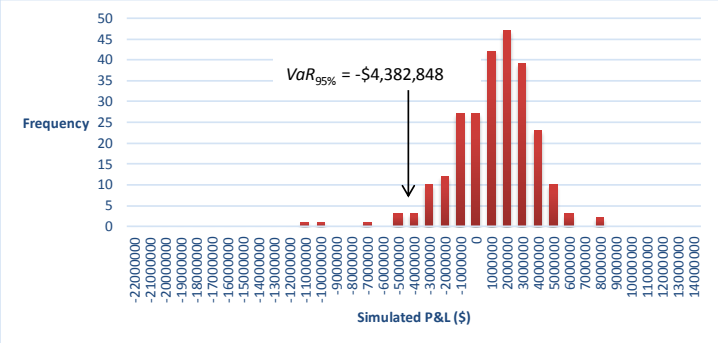
***$VaR_{95\%}$* = -$3,466,790**



Parametric VaR

PORT AuM ($)	100,000,000
Variance (Min.)	7.10
St. Dev.	2.66%
Confidence Level	95%
Critical Value (z_α)	1.645

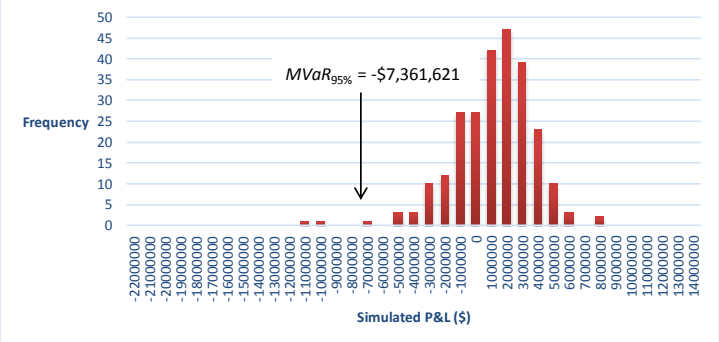
***$VaR_{95\%}$* = -$4,382,848**



Modified VaR

PORT Index AuM ($)	100,000,000
Mean P&L ($)	694,724
St. Dev. P&L ($)	2,549,081
Confidence Level	95%

***$MVaR_{95\%}$* = -$7,361,621**



Expected Shortfall

PORT Index AuM ($)	100,000,000
Mean P&L ($)	694,724
St. Dev. P&L ($)	2,549,081
Confidence Level	95%

***$ES_{95\%}$* = -$5,258,022**

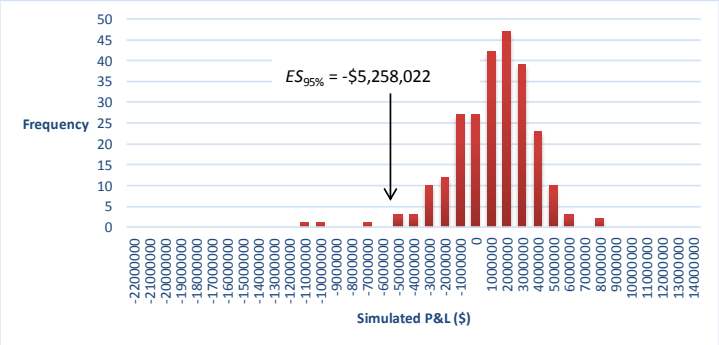
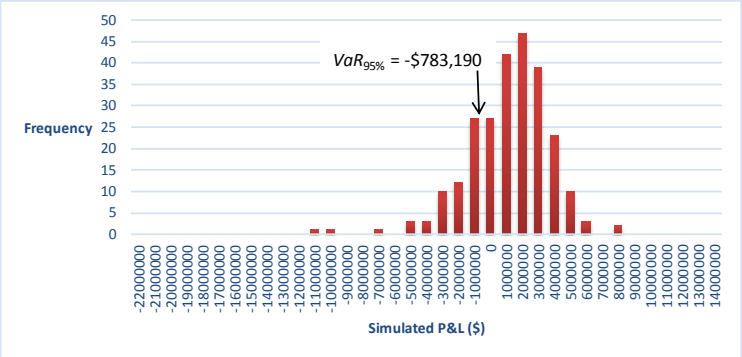


TABLE 8 (b) VaR and Expected Shortfall: Optimization Portfolios based upon Minimum Variance

Historical Simulation VaR

PORT Index AuM ($)	100,000,000
Confidence Level	95%
Critical Value (z_α)	1.645

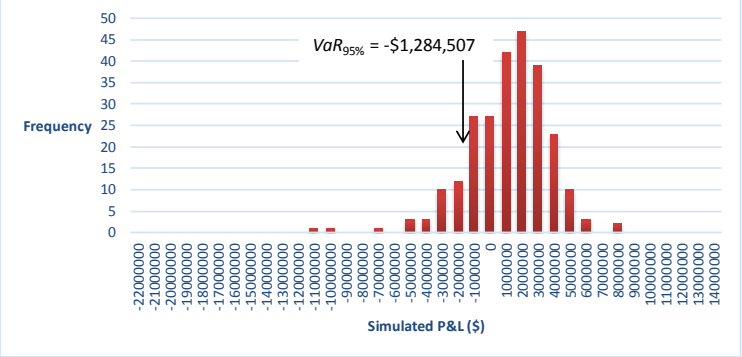
***$VaR_{95\%}$* = -$783,190**



Parametric VaR

PORT AuM ($)	100,000,000
Variance (Min.)	0.61
St. Dev.	0.78%
Confidence Level	95%
Critical Value (z_α)	1.645

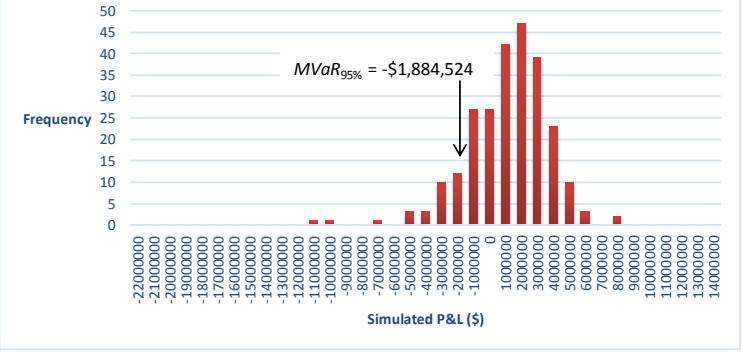
***$VaR_{95\%}$* = -$1,284,507**



Modified VaR

PORT Index AuM ($)	100,000,000
Mean P&L ($)	533,059
St. Dev. P&L ($)	815,251
Confidence Level	95%

***$MVaR_{95\%}$* = -$1,884,524**



Expected Shortfall

PORT Index AuM ($)	100,000,000
Mean P&L ($)	533,059
St. Dev. P&L ($)	815,251
Confidence Level	95%

***$MVaR_{95\%}$* = -$1,681,629**

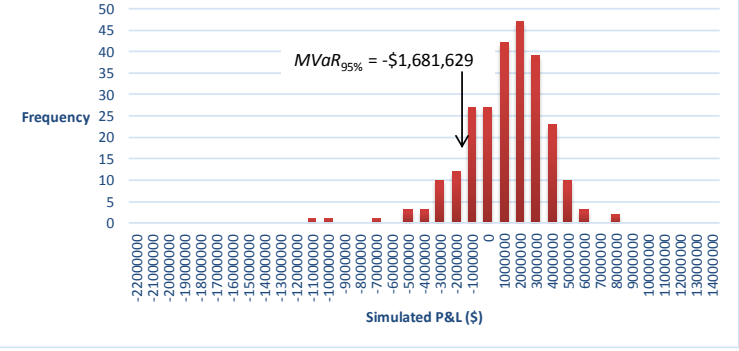
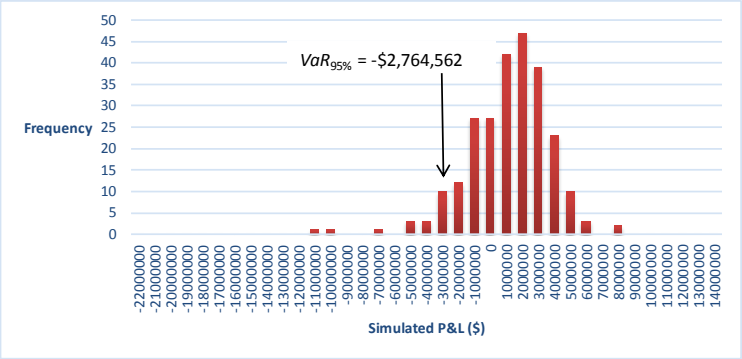


TABLE 8 (c) VaR and Expected Shortfall: Optimization Portfolios based upon Maximizing Return

Historical Simulation VaR

PORT Index AuM ($)	100,000,000
Confidence Level	95%
Critical Value (z_α)	1.645

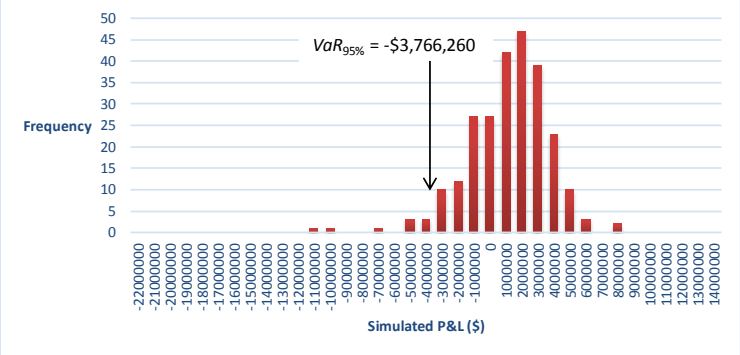
***$VaR_{95\%}$* = -$2,764,562**



Parametric VaR

PORT AuM ($)	100,000,000
Variance (Min.)	5.24
St. Dev.	2.29%
Confidence Level	95%
Critical Value (z_α)	1.645

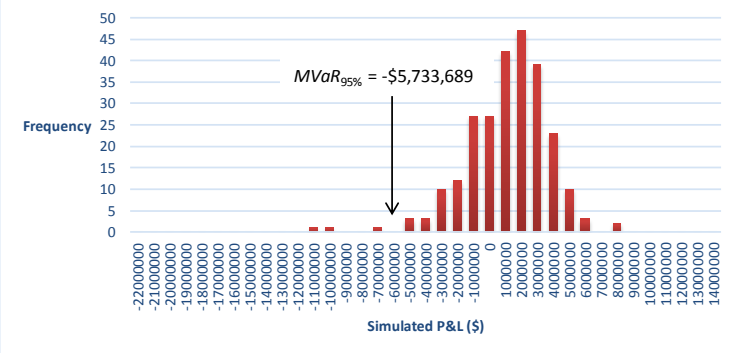
***$VaR_{95\%}$* = -$3,766,260**



Modified VaR

PORT Index AuM ($)	100,000,000
Mean P&L ($)	994,581
St. Dev. P&L ($)	2,218,136
Confidence Level	95%

***$MVaR_{95\%}$* = -$5,733,689**



Expected Shortfall

PORT Index AuM ($)	100,000,000
Mean P&L ($)	994,581
St. Dev. P&L ($)	2,218,136
Confidence Level	95%

***$MVaR_{95\%}$* = -$4,575,377**

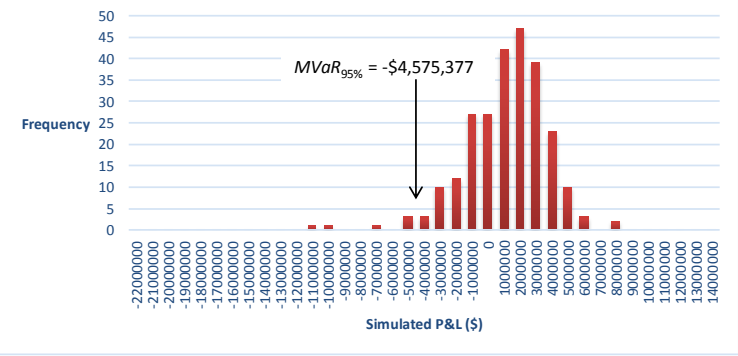
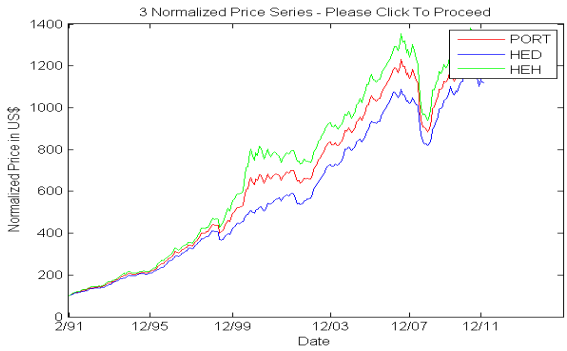
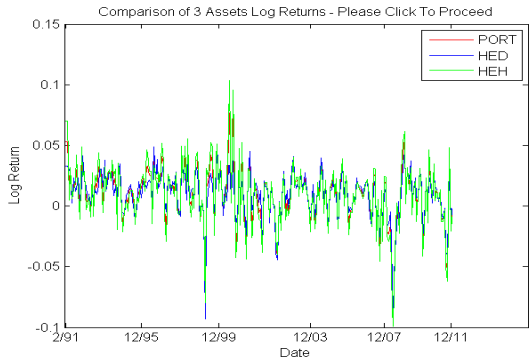
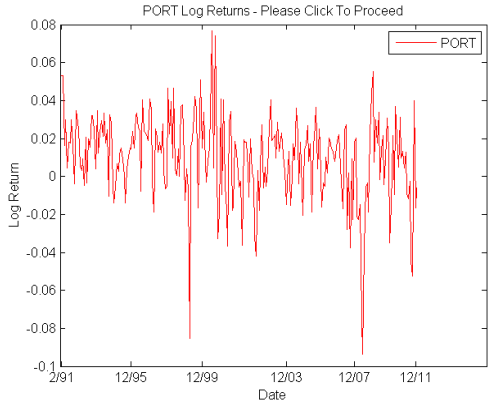
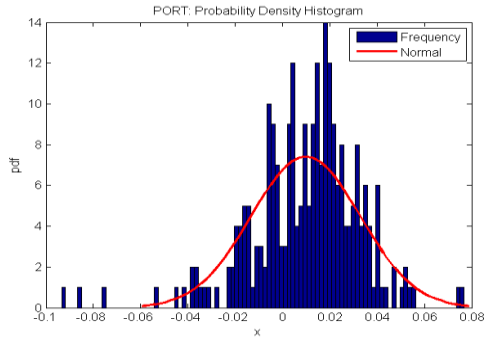


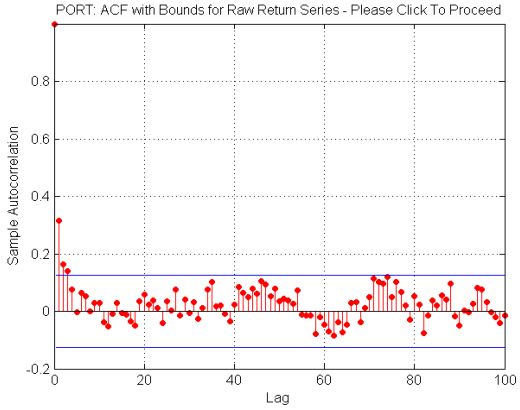
TABLE 9 PORTFOLIO VaR MODELING

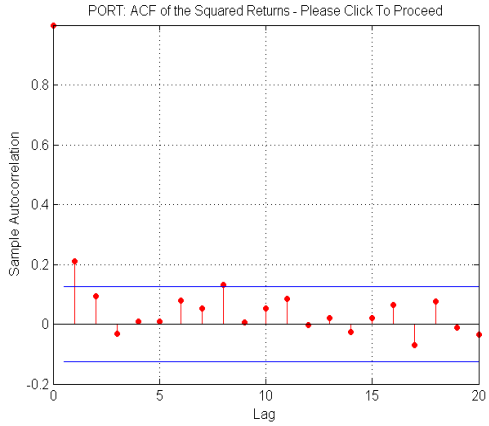


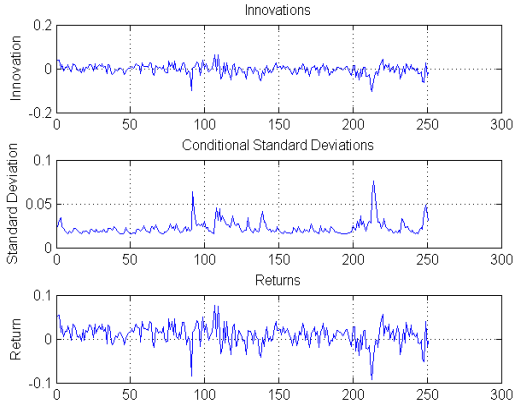












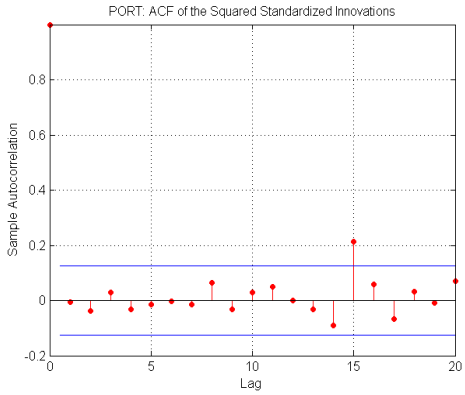
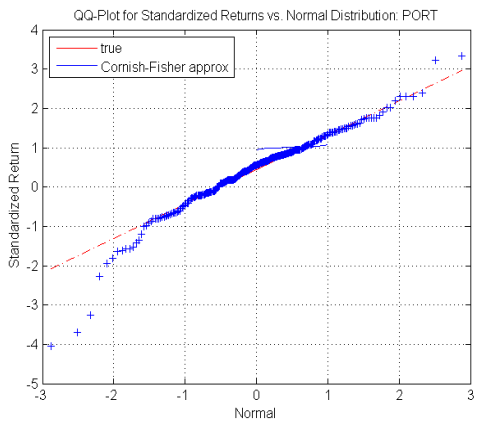
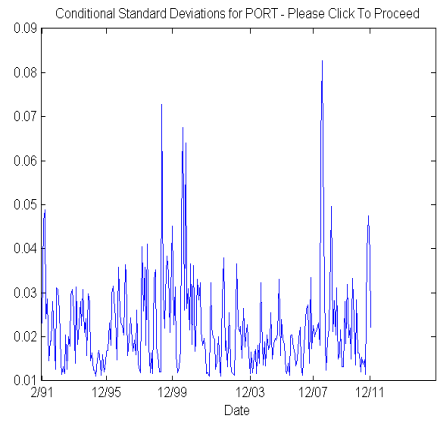
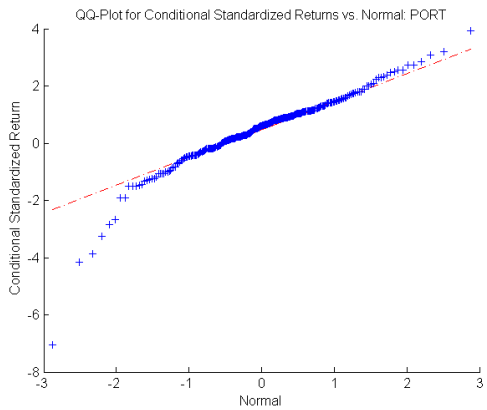
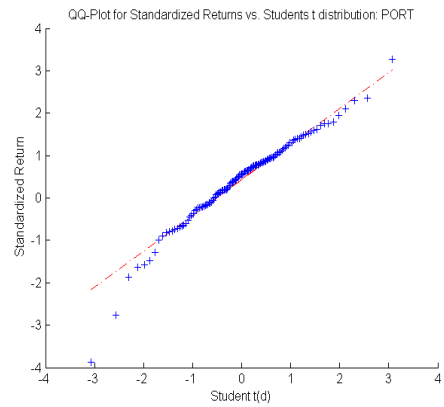


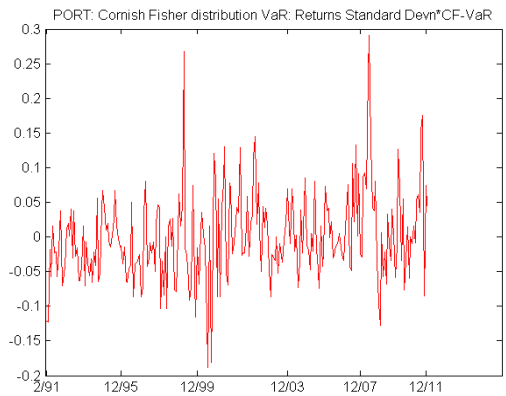
TABLE 9 PORTFOLIO VaR MODELING (continued)

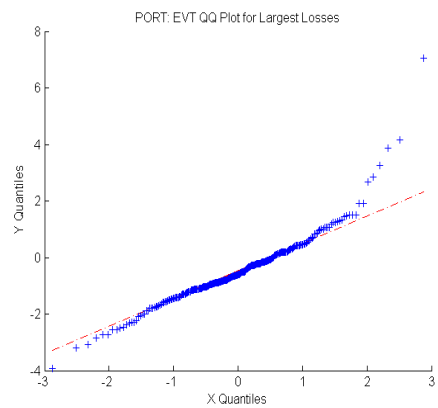












Parameter	Value	Std. Error	T-Statistic
C	0.012609	0.001322	9.5347
K	0.000143	6.50E-05	2.1968
GARCH(1)	0.38558	0.19811	1.9463
ARCH(1)	0.38495	0.12717	3.0269

Appendix 1. Bayesian Inference: Probability Background

How likely an event is, the *likelihood* of a specific outcome, is with respect to the *sample space* which is the set of all mutually exclusive [and] cumulatively exhaustive (MECE) possibilities. Specific *parameter* such as bias (i.e., *probability of a specific outcome*) of a process can be denoted as θ so that the *degree of belief about that parameter value* θ is $p(\theta)$. The *possibilities sample space* or *outcome events sample space* consists of all MECE possibilities or possible outcome events. The *parameter sample space* consists of all values that the specific parameter can have. If the parameter bias can vary from 0% to 100%, respective parameter sample space consists of all continuous data values between 0 and 1. When a specific process is sampled, it is sampled from the parameter sample space. For the specific parameter sampled, the outcome events are then sampled from the outcome events sample space. For specific outcome events that can be observed, *probability* of occurrence of any specific event is its *long-run relative frequency*. Such long-run relative frequency can be observed by actually sampling from the sample space and tracking counts of different outcomes. Sampling can be done using *computerized simulation* in which the computer generates the outcomes randomly. A long run, being a finite random sample, can only approximate the probability by long-run relative frequency. Or, it can be calculated with greater precision by *deriving it mathematically* based on known properties of the process.

Probabilities are non-negative numbers assigned to the set of MECE possibilities. The probabilities should sum to 1.0 for all MECE possibilities. For two mutually exclusive, i.e., independent events, the probability that one *or* the other occurs equals the sum of respective individual probabilities. *Probability distribution* is the list of all possible MECE outcomes *and* their corresponding probabilities. The probability of discrete outcome value is called *probability mass* to distinguish it from the probability of continuous outcome value which is called *probability density*. If a continuous distribution is discretized then the amount of the probability in a specific interval is given by its probability mass. *Probability density* of an interval is the probability mass of that interval divided by the interval width. For a continuous distribution, since the probability of any specific discrete *exact* infinitesimal point is zero, probability is denoted as *probability density* which is the ratio of the probability to the respective interval width. Hence for a *uniform* scale that is divided into N intervals, the probability of any infinitesimal interval converges to zero in the limit as N grows to infinity. However, its probability density which is the ratio of probability mass (1/N) to its width (1/N) always remains 1 = ((1/N)/ (1/N)).

Probability mass cannot exceed 1, however probability density being a ratio of probability mass to respective interval width can be lesser or greater than 1. If the uniform interval scale is changed from 0–1 to 0-0.5, then the amount of probability per unit interval width doubles, hence probability density becomes 2 everywhere (((1/N)/ (0.5/N))). In case of a logarithmic scale, every additional unit interval width contains lesser and lesser probability in a smaller interval width thus having exponentially smaller probability density. For instance a log-10 circular scale will contain 1 to 10 (10^0 to 10^1) within the first half, i.e., 0.5 probability mass, and 10 to 100 (10^1 to 10^2) in the second half.

Let continuous variable be denoted as x and interval width by Δx. Let interval index be denoted as i, and the interval between x_i and x_i+Δx be denoted as [x_i, x_i+Δx]. Then,

$Probability\ mass$ of the ith interval: $p([x_i, x_i + \Delta x])$ and

Sum of probability masses for all the intervals: $\sum_i \ p([x_i, x_i + \Delta x]) = 1.0$

Dividing and multiplying by the interval width Δx: $\sum_i \ \Delta x \ \frac{p([x_i, x_i+\Delta x])}{\Delta x} = 1.0$

As $\Delta x \to 0$, the above equation becomes: $\int \ dx \ \ p(x) \ = 1.0$ where $p(x)$ is the *probability density*.

Appendix 2. Hedge Fund Industry Risk-Adjusted Return Metrics

Tracking Error, or Standard Deviation of Excess Return, is a statistical measure of dispersion measuring volatility of excess returns over a given period (J.P. Morgan, 2008). For each asset class modeled, tracking error was measured in terms of quadratic standard deviation (SD) and linear mean absolute deviations (MAD). The tracking error measures how closely the fund follows the index to which it is benchmarked: lower the error, more closely the fund follows risk-and-return characteristics of the benchmark. While SD being the quadratic form may be more difficult to interpret, its linear alternative MAD seems more intuitive for hedge fund managers who may prefer seeing it in linear terms.

Quadratic Tracking Error

$$SD = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(r_{N-1,N} - r_{N-1,N}^{bm})^2}$$

Linear Tracking Error

$$MAD = \frac{1}{N-1}\sum_{i=1}^{N}|r_{N-1,N} - r_{N-1,N}^{bm}|$$

N = No. of sample data points, $r_{N-1,N}$= Fund return, $r_{N-1,N}^{bm}$ = Benchmark return

Basic performance plots for each asset included historical performance of RoR%, VAMI, and, histogram of monthly returns. Tests for normality of each asset's returns statistical distribution included charts of empirical versus normal distribution and normal Q-Q plots as well as the Jarque-Bera Normality Test which is a joint test of skewnesss and excess kurtosis. Descriptive statistics included the first four moments of distribution. The correlation matrix was computed to show relative strength of variability of returns of the asset classes with respect to each other. Mean-variance optimization (Markowitz, 1952) was used to compute the portfolio asset allocations for minimizing variance and for maximizing returns and then compared with the portfolio with equally weighted asset classes.

Relative risk and returns behavior of different asset classes and robustness in consistency of their behavior was monitored in course of risk modeling using different models. Different risk measures based on varying risk estimation assumptions facilitated stress testing and sensitivity analysis. As risk and returns may not vary proportionally for all indexes or portfolios, their relative performance can be more accurately measured by using *risk-adjusted return measures*. *Risk Adjusted Return* is an ex-post risk measure in which the portfolio return is adjusted by the standard deviation or beta of the portfolio (J.P. Morgan, 2008). Most commonly used risk-adjusted return measures in the hedge fund industry are based upon the ratio of risk free returns to risk:

$$Risk\ adjusted\ returns = \frac{R_P - R_F}{Risk}.$$

The risk-adjusted return measures help assess the true performance of the hedge fund managers delivering real alpha (reflecting real skill) compared to others delivering sophisticated alternative beta (available at a lower cost) or traditional market beta (available free). Alpha is a measure of performance on a risk-adjusted basis as it takes into consideration the risk-free rate. In current context, it refers to the excess return of the portfolio relative to the return of the benchmark. Beta is a measure of the volatility, or systematic risk, of a fund or portfolio in relation to the overall market. Beta of 1 indicates moment in same direction and by same percentage as the overall market. Beta greater (lesser) than 1 indicates that the fund is expected to move more (less) than the market and hence is more (less) risky. Portfolio Beta is the weighted average of the Betas of the various assets held in the portfolio.

The ratio of annualized first and second moments of distributions is another such measure:

$$M1/M2 = \frac{Return}{Volatility}.$$

In the above computation, R_P is the annualized return while R_F is the annualized risk-free rate (such as for a US treasury bill). *Sharpe ratio* (Sharpe, 1994) uses the volatility of returns σ_P as the measure of risk:

$$Sharpe\ Ratio = \frac{R_P - R_F}{\sigma_P} = \frac{Return\ - Risk\ Free\ Return}{Standard\ Deviation\ Of\ Returns}.$$

Also known as the "reward to variability ratio, it relates the reward to the portfolio's risk, as measured by the portfolio's standard deviation (J.P. Morgan, 2008). By using the standard deviation, Sharpe Ratio measures the total risk of the portfolio, not just risk in relation to the market. As compared with prior measure M1/M2, Sharpe Ratio introduces a static benchmark to the numerator by subtracting the risk-free rate from the return. Sharpe ratio thus penalizes the fund manager whose return is lower than risk-free rate and shows negative Sharpe ratio for managers delivering returns lower than the risk-free rate.

The *Modified Sharpe Ratio* introduced earlier in the discussion on MVaR accounts for the third and fourth moments of the returns (and P&L) distribution, skewnes and excess kurtosis, and is given by:

$$\mathrm{MSR} = \frac{R_P - R_F}{\mathrm{MVaR}_{1-\alpha}} = \frac{Return\ - Risk\ Free\ Return}{Modified\ VaR}.$$

The *Sortino Ratio* (Sortino & Forsey, 1996) modifies Sharpe Ratio so that the fund manager is penalized only for downside risk (volatitlity) but not for upside volatility which enhances returns. It uses the concept of the minimum acceptable return (MAR). It divides the returns into those that are greater than MAR and those that are less than MAR. Higher Sortino ratio implies that the manager is better at controlling downside risk and is not penalized for producing high upside returns.

$$Sortino\ Ratio = \frac{R_P - MAR}{\sqrt{\frac{1}{T}\sum_{t,R_P<MAR}^{T}\left(R_{P,t} - MAR\right)^2}} = \frac{Return\ - Risk\ Free\ Return}{Negative\ Semi - Deviation\ of\ Returns}.$$

The *Drawdown Ratio*, another variant of Sharpe Ratio, uses maximum historical drawdown as the risk measure.

$$DD\ Ratio = \frac{R_P - R_F}{|\max DD|}.$$

Maximum drawdown is defined as maximum loss in VAMI or NAV terms from the preceding highest high to the lowest low during the period that the fund has not recovered its value to the last highest high. Variants of Drawdown Ratio include the *Sterling Ratio* which uses an average of the most significant drawdowns and the *Burke Ratio* which uses the square root of the sum of the squares of each drawdown. The key idea in both the variations is about penalizing significant long-term drawdowns relative to several milder drawdowns.

The *Information Ratio* (Goodwin, 1998) measures a portfolio's performance against risk and return relative to a benchmark or alternative measure. The higher the Information Ratio, the greater the added value for a given level of risk, relative to the benchmark. Information Ratio uses a market reference benchmark instead of the risk-free rate. Thus, greater added value for a given level of risk, relative to the benchmark, i.e. excess returns on a benchmark portfolio B in period t, can be described as:

$\Delta_t = R_{P,t} - R_{B,t}$ and their arithmetic average from t = 1 to T is given by: $\bar{\Delta} = \frac{1}{T}\sum_{t=1}^{T}\Delta_t$. Then, standard deviation of the excess returns from the benchmark is given by $\sigma_\Delta = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(\Delta_t - \bar{\Delta})^2}$. Then,

$$Information\ Ratio = \frac{\bar{\Delta}}{\sigma_\Delta} = \frac{Excess\ Return}{Std.\ Devn.\ of\ Tracking\ Error}\ .$$

The *M-Squared Metric* helps see how the hedge fund outperforms the benchmark return to which it has had its risk profile matched. It does so by interpreting the fund's return as the return that would have been produced had the fund's volatility been equal to that of the market benchmark.

$$M^2 = \frac{\sigma_M}{\sigma_P}(R_P - R_F) - R_F\ .$$

The *Treynor Ratio*, also known as the "reward to volatility ratio," measures the excess return achieved by a fund manager per unit of risk incurred (J.P. Morgan, 2008). Based on systematic risk, it uses the beta of the fund relative to a benchmark as the risk measure in the denominator:

$$Treynor\ Ratio = \frac{R_P - R_F}{\beta_P}.$$

Treynor Ratio, just like Information Ratio, is more commonly used for active traditional equity portfolios.

Jensen's Alpha (Jensen, 1967) is used to determine the Excess Return over the required rate of return as predicted by the Capital Asset Pricing Model (CAPM) given the portfolio's beta and the average market return (J.P. Morgan, 2008). It is the sum of risk-free rate and beta adjusted market excess returns subtracted from fund's net return:

$\alpha_P = R_P - [R_F + \beta_P (R_M - R_F)]$ based upon CAPM: $(R_P - R_F) = \alpha_P + \beta_P (R_M - R_F)$.

The above expression highlights the three parts that make up a hedge fund return: alpha (measurable skill), beta continuum (from skill to no skill) (Anson 2008) and the risk-free rate (no skill).

Jensen's Alpha Ratio (J.P. Morgan, 2008) is a risk-adjusted performance measure that represents the average return on a portfolio over and above that predicted by the Capital Asset Pricing Model (CAPM), given the portfolio's beta and the average market return (Jensen's Alpha).

$$\textit{Jensen's Alpha Ratio} = \frac{Average\ Jensen's\ Alpha}{Std.\ Devn.\ of\ Jensen's\ Alpha} \ .$$

Appendix 3. Value Added Monthly Index (VAMI) Method

The VAMI method generally assumes an initial investment of $100 or $1,000 and shows how such an investment would have fared over a certain period of time. In order to calculate annual ROR using VAMI, you must first calculate the value of the investment at the end of each subperiod or month based upon the monthly RORs computed in accordance with one of the above mentioned methods. The following calculation assuming initial investment of $1,000 (National Futures Association, 2013):

Annual and Year-to-Date Rates of Return

In the first month of the period: VAMI for month = (1 + ROR for month) x 1000 For all subsequent months: VAMI for month = (1 + ROR for month) x VAMI for prior month

Annual ROR would then be calculated as follows:

Annual ROR = (year-end VAMI - $1,000) divided by $1,000. When calculating the annual RORs for subsequent years, the value of the initial investment should be the prior year-end VAMI.

Computing Monthly and Peak-to-Valley Draw-Downs

Draw-down means losses experienced by a pool or trading program over a specified period.

Worst monthly draw-down is simply the trading program's worst monthly percentage ROR.

Worst peak-to-valley draw-down is the greatest cumulative percentage decline in month-end net asset value (NAV) due to losses sustained by the accounts during any period in which the initial month-end NAV is not equaled or exceeded by a subsequent month-end NAV. In order to calculate this amount, the firm should calculate a continuous VAMI for the time period presented. Using this method the firm should determine the first month in which the VAMI is not followed by a VAMI that is greater than or equal to that month's VAMI. This would be the first peak. The next peak would be the next month in which the VAMI is greater than the previous peak's VAMI and is followed by a lower VAMI. Once all the peaks have been identified, determine all the months that have the lowest VAMIs during a period between two peaks. These would be the valleys. Then determine the percentage change between each peak and valley using the following calculation:

(Valley VAMI – Peak VAMI) divided by Peak VAMI

The worst peak-to-valley draw-down will be the largest percentage change from a peak to a valley. The peak month and the valley month should be reported in the capsule. A peak-to-valley draw-down that began prior to the beginning of the most recent five calendar years is deemed to have occurred during such five calendar year period.

Bibliography

Artzner, P., Delbaen, F., Eber, J. M., & Heath, D. (1999). Coherent Measures of Risk. *Mathematical Finance, 9*(3), 203-228.

Aussenegg, W., & Miazhynskaia, T. (2006). Uncertainty in Value-at-Risk Estimates Under Parametric and Non-Parametric Modeling. *Financial Markets and Portfolio Management, 20*(3), 243-264.

Berkowitz, J., Christoffersen, P. F., & Pelletier, D. (2011). Evaluating Value-at-Risk Models with Desk-Level Data. *Management Science, 57*(12), 2213-2227.

Berkowitz, J., & O'Brien, J. (2002). How Accurate Are Value-at-Risk Models at Commercial Banks? *Journal of Finance, 57*(3), 1093–1111.

Borison, A., & Hamm, G. (2010). How to Manage Risk (After Risk Management Has Failed). *MIT Sloan Management Review, Fall*, 51–57.

Boucher, C. M., Daníelsson, J., Kouontchou, P. S., & Maillet, B. B. (2014). Risk models-at-risk. *Journal of Banking & Finance*.

Casarin, R., Chang, C.-L., Jimenez-Martin, J.-A., McAleer, M., & Pérez-Amarald, T. (2013). Risk management of risk under the Basel Accord: A Bayesian approach to forecasting Value-at-Risk of VIX futures. *Mathematics and Computers in Simulation, 94*, 183–204.

Christoffersen, P. F. (2012). *Elements of financial risk management* (2nd ed.). Waltham, MA: Academic Press.

Cornish, E. A., & Fisher, R. A. (1937). Moments and Cumulants in the Specification of Distributions. *Extrait de la Revue de l'Institute International de Statistique, 4*, 1-14.

Danielsson, J., James, K., Valenzuela, M., & Zer, I. (2014). *Model Risk of Risk Models*. Finance and Economics Discussion Series, Divisions of Research & Statistics and Monetary Affairs. Working Paper. Federal Reserve Board, Washington, D.C.

Darbyshire, P., & Hampton, D. (2012). *Hedge fund modelling and analysis using Excel and VBA*. Chichester, West Sussex ; Hoboken, N.J.: Wiley.

Darbyshire, P., & Hampton, D. (2014). *Hedge fund modelling and analysis using MATLAB®*. Hoboken: Wiley.

Derman, E. (1996). Model Risk *Quantitative Strategies Research Notes* (April ed.): Goldman, Sachs & Co.

Embrechts, P., Resnick, S. I., & Samorodnitsky, G. (1999). Extreme Value Theory as a Risk Management Tool. *North American Actuarial Journal, 3*(2).

Gelfand, A. E., & Smith, A. F. M. (1990). Sampling-Based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association, 85*(410), 398-409.

Goodwin, T. H. (1998). The Information Ratio. *Financial Analysts Journal, 54*(4).

Gumbel, E. J. (2004). *Statistics of Extremes*: Courier Dover Publications.

Hoogerheide, L. F., & van Dijk, H. K. (2008, September 30). Bayesian Forecasting of Value at Risk and Expected Shortfall using Adaptive Importance Sampling. *Tinbergen Institute Discussion Paper.* TI 2008-092/4.

Hull, J., & White, A. (1998). Value At Risk When Daily Changes In Market Variables Are Not Normally Distributed. *Journal of Derivatives, 5*(3), 9-19.

J.P. Morgan. (2008). Investment Performance, Analytics, and Risk: Glossary of Terms: J.P. Morgan Investment Analytics and Consulting.

Jackson, P., Maude, D. J., & Perraudin, W. (1998). Bank capital and Value at Risk: Bank of England.

Jensen, M. C. (1967). The Performance of Mutual Funds in the Period 1945-1964. *Journal of Finance, 23*(2), 389-416.

Kass, R., & Raftery, A. (1995). Bayes factor. *Journal of American Statistical Association, 90*, 773–792.

Kruschke, J. K. (2011). *Doing bayesian data analysis : a tutorial with R and BUGS*. Burlington, MA: Academic Press.

Lynch, S. M. (2007). *Introduction to Applied Bayesian Statistics and Estimation for Social Scientists*. New York, NY: Springer New York.

Malhotra, Y. (2014). *Markov Chain Monte Carlo Models, Gibbs Sampling, & Metropolis Algorithm for High-Dimensionality Complex Stochastic Problems: Applications in Network and Computer Security*. Working Paper. Computational Quantitative Finance-Risk Management. Global Risk Management Network, LLC. Cornell Business & Technology Park, Ithaca.

Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance, 7*(1), 77–91.

Meucci, A. (2009). *Risk and Asset Allocation* (Corr. 3rd printing 2009 edition ed.): Springer.

Miazhynskaia, T., Fruhwirth-Schnatter, S., & Dorffner, G. (2003). A Comparison of Bayesian Model Selection based on MCMC with an application to GARCH-Type Models. *Adaptive Information Systems and Modelling in Economics and Management Science*: Vienna University of Economics & University of Vienna.

Morini, M. (2011). *Understanding and managing model risk a practical guide for quants, traders and validators* (1. ed.). Hoboken: Wiley.

National Futures Association. (2013). Disclosure Documents: A Guide for CPOs and CTAs.

Osiewalski, J., & Pajor, A. (2010). Bayesian Value-at-Risk for a Portfolio: Multi- and Univariate Approaches Using MSF-SBEKK Models. *Central European Journal of Economic Modelling and Econometrics, 2*, 253-277.

Pickands III, J. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics, 3*(1), 119-131.

Sharpe, W. F. (1994). The Sharpe Ratio. *The Journal of Portfolio Management, 21*(1), 49–58.

Simons, K. (1996). Value at Risk - New Approaches to Risk Management. *New England Economic Review*: Federal Reserve Bank of Boston.

Sortino, F., & Forsey, H. (1996). On the Use and Misuse of Downside Risk. *The Journal of Portfolio Management*.

Tasche, D. (2002). Expected shortfall and beyond. *Journal of Banking and Finance, 26*(7), 1519–1533.

US Fed, & OCC. (2011). *Supervisory Guidance On Model Risk Management*. (SR Letter 11-7 / OCC 2011-12).

US Senate. (2013). JPMorgan Chase Whale Trades: A Case History Of Derivatives Risks And Abuses: U.S. Senate Committee on Homeland Security & Governmental Affairs Permanent Subcommittee on Investigations Hearing.

Venkataraman, S. (1997). Value at risk for a mixture of normal distributions: the use of quasi-Bayesian estimation techniques. *Economic Perspectives* (Vol. March/April, pp. 2-13): Federal Reserve Bank of Chicago.

Zangari, P. (1996). An improved methodology for measuring VAR *RiskMetrics Monitor*: Reuters/JP Morgan.