ABSTRACT

Many businesses make profit yearly and tend to invest some of the profit so that they can cushion their organizations against any future unknown events that can affect their current profit making. Since future happenings in businesses cannot be predicted accurately, estimates are made using experience or past data which are not exact. The probability element (which is normally determined by experience or past data) is important in investment decision making process since it helps address the problem of uncertainty. Many of the investment decision making methods have incorporated the expectation and risk of an event in making investment decisions. Most of those that use risk account for diversifiable risk (non-systematic risk) only thus limiting the predictability element of these investment methods since total risk are not properly accounted for. A few of these methods include the certainty (probability) element. These include value at risk method which uses covariance matrices as total risk and the binning system which always assumes normal distribution and thus does not take care of discrete cases. Moreover, comparison among various entities lacks since the probabilities derived are for individual entities and are just quantile values. Finite investment decision making using real market risk (non-diversifiable risk) was undertaken in this study. Non-diversifiable risk (systematic risk) estimates of a portfolio of stocks determined by a real risk weighted pricing model are used as initial data. The variance of non-diversifiable risk is estimated as a random variable referred to as random error (white noise). The estimator is used to calculate estimates of white noise (wn). A curve estimation of the wn is made using Kernel Density Estimation (KDE). KDE is a non-parametric way to estimate the probability density function of a random variable. KDE is a fundamental data smoothing problem where inferences about the population are made, based on a finite data sample. This is used to derive probability estimates of the non-diversifiable risks of the various stocks. This enables determination of total risk with given probabilities of its occurrence thus facilitating decision making under risky and uncertain situations as well as accentuating comparison among the portfolio of stocks.

1. INTRODUCTION

In the past few years there has been evidence of collapse of well-established business entities. This has been attributed to lack of accurate methods of preventing or measuring risk and uncertainty as opposed to lack of the same methods. Many companies on Wall Street in 2008 went under despite having extensive measures of mitigating risk such as futures and forward. An investigation into some of these methods reveals the lack of a well-estimated market risk measure in the models. It is an obvious fact now that it was the external reactions that brought down the companies on Wall Street. Once the markets got a hint of the internal financial and investment affairs of the companies this spread so rapidly and in a matter of hours these companies had collapsed. A good example is the Lehman Brothers Holdings limited, Merrill Lynch and companies, and American Investment group as explained by Lucchetti et al. [1]. These indicate that market environments are so critical in the existence of business entities such that variables affecting the business entities from the market environments should be estimated with a lot of precision. This paper determines total risk which has both the systematic and non-systematic components. It should be noted that the non-systematic risk is internal in nature, and in most cases well known and relatively less difficult to estimate while systematic is external and in most cases is embodied in market risk.

This paper will determine the risk factor of systematic risk a phenomenon lacking in many risk models. Since we have seen from most examples that market risk is the precursor of most companies downfalls, it is hoped that investors and companies will be able to easily estimate riskiness of the risk measures thus enabling them make informed decisions.

Jorion [2] determines the Value at Risk (VAR) measure as the forecasted volatility, \( \sigma \) multiplied by standard normal deviate, \( \alpha \) for the selected confidence level (e.g. \( \alpha = 2.33 \) for a one-tailed confidence level of 99%). The portfolio variance then becomes \( \sigma_p^2 = w^T \Sigma w \), where \( \Sigma \) is the forecasted covariance matrix for the market risk factors as of the close day \( t \). Hence we have, \( \text{VAR}_t = \alpha \sigma_p^2 \). Although this research takes care of all the other shortcomings of previous researches, the portfolio variance is determined as a covariance which goes against the definition of market risk as
that which cannot be diversified. This has been clearly addressed
in this paper by using non-diversifiable risk which is determined
without covariance.

Dennis et al. [3] in estimating density dependence process noise
and observation error offers a statistical approach for jointly esti-
mating density dependence, process error and observation error.
Although this model is relatively easy for ecologists to use and is
applicable in many population systems, this process noise has a
normal distribution with mean $\mu$ and variance $\sigma^2$ ($E_i \sim N(0, \sigma^2)$).
This paper looks at a case of no assumption of normality for the
noise process. White noise is determined as a random variable on
the precincts of Sklar [4] where he says no common probability
space can be found for a given set of random variables, but such
common probability spaces exists for arbitrary proper subsets of
the given set. In this study the subsets were the portfolios of dif-
ferent companies used giving a common probability space that is
estimated. The results of Wu [5] show that for finite parameters
the consistency of the least squares estimator is equivalent to the
existence of a consistent estimator thus the estimator of white noise
derived in this paper is an unbiased estimator.

2. RISK OF NON-DIVERSIFIABLE RISK

2.1. Determination of White Noise of Non-diversifiable Risk

White noise refers to a purely random process whose random vari-
ables are a sequence of mutually independent, identically distrib-
uted random variables. Thus it describes an event and is a function
with a domain that makes some real number correspond to each
outcome of the experiment. In this paper white noise is taken as the
random error of non-diversifiable risk $N_{\text{err}}$ of an investment $i$. To
add credence to this study it is imperative to show that white noise
is a random variable. Proposition 2.1 below seeks to do so.

Proposition 2.1: Let $V_i$ be the white noise of the non-diversifiable
risk $N_{\text{err}},$ then $V_i(.)$ is a random variable.

Proof: Given $N_{\text{err}}$ and

i. The domain $\Omega \{V_{gi} \} , \{V_{w} \} \}$

ii. The counter domain $r$ is such that $0 \leq r \leq 1.$

iii. The range of returns $i$ is $-\infty \leq i \leq \infty.$

then $V(.)$ is an event. That is $V(.)$ is such that the subset $w_r = \{s:
V(s) \leq r\},$ where $s$ is a subset of the domain. This is true since $0 \leq
V(.) \leq 1.$ If $w_r$ belongs to $W$ for every real number $r,$ where $W$ is the
set of all outcomes of event $V(.)$.

Then the probability of $V(.)$, $P[V(.)]$ is a set function having
domain $V(.)$ and counter domain the interval $[0, 1].$ Therefore $V(.)$
has a probability space $(\Omega, W, P[V(.)])$. Also $W$ consists of four
subsets; $\{V_{gi} \}, \{V_{w} \}$ and $\Omega.$

Such that if:

i. $r < 0,$ then $s; V(s) \leq r = \phi.$

ii. $0 \leq r < 1,$ then $s; V(s) \leq r,$ where $V_{-1 \leq s \leq 1}$

iii. $r \geq 1,$ then $s; V(s) \leq r = \Omega.$

Since $V(.)$ has a probability space, and $W$ consists of the four sub-
sets above. Then for each $r$ the set $\{s; V(s) \leq r\}$ belongs to $W,$ thus
$V(.)$ is a random variable. Since $V(.)$ is a random variable and it is
independent with unique parameters. Therefore the parameters of
white noise for example its mean and variance as well as its unique
probability distribution can be determined. The probability esti-
mates of non-diversifiable risk for investment decisions are then
estimated as shown in the following subsections.

2.1.1. Determination of random error

The non-diversifiable variance estimator

$$N_{\text{err}} = \sum_{i=1}^{n} w_i^2 s_i^2 + \sum_{i=1}^{n} s_i^2$$

derived from the non-diversifiable risk estimated in Anyika et al.
[6] indicates the presence of random error in the risk estimator.
This error is taken to be white noise (wn) thus it can be said to be a
random variable $V_i, V_j, V_{ij}, ... V_{n}$ which is mutually independent
and identically distributed. This is estimated from a sample of data by
first varying the variance of individual return values of $r$ resulting in

$$\hat{w}_i = T \sum_{i=1}^{n} w_i^2 s_i^2$$

where $T = \frac{z-2}{(z-1)^2},$ $z$ being the total number of returns and (2) is
the predicted random error.

From (2) the actual value of $\hat{w}_i$ is given by

$$\hat{w}_i = \sum_{i=1}^{n} w_i^2 s_i^2 C + L$$

where $C$ and $L$ are values representing the location (mean) and
scale (variance) parameters. These parameters are determined such
that the bias and variance of the actual and predicted values of $w_n$ are
minimized as follows;

Let the variance between actual and sample white noise be

$$\text{var}(\hat{w}_n, w_n) = \frac{2}{z-1} \hat{w}_n - \left(\frac{2}{z-1}\right)^2 \hat{w}_n w_n + \frac{2}{z-1}$$

The values of $C$ and $L$ which will minimize variance are given by
the partial derivatives of $C$ and $L,$ $f_c$ and $f_L$ respectively. After several
iterations;

$$f_c = 2w_n - \frac{2}{z-1} \hat{w}_n = L$$

$$f_L = \frac{2}{z-1} \hat{w}_n - w_n = L$$
Thus the value of
\[
\hat{w}_n \left. \frac{4}{3(z-1)} \right| \hat{w}_i,
\]

Proposition 2.2: \( \hat{w}_n \) is an unbiased estimator of \( w_n \).

Proof: From Equation (2)

\[
\hat{w}_n = T \sum_{i=1}^{\infty} \left( \frac{s_i^2 + \hat{w}_n}{z-1} \right)
\]

\[
E(\hat{w}_n) = \frac{z-2}{(z-1)^2} \sum_{i=1}^{\infty} \left( 1 - \frac{1}{z-1} E\left( \frac{\hat{w}_n^2}{z-1} \right) \right)
\]

\[
= \frac{z-2}{(z-1)^2} \sum_{i=1}^{\infty} \left( \frac{1}{z-1} E\left( V_i^2 \right) - E\left( \hat{w}_n^2 \right) \right)
\]

\[
= \frac{z-2}{(z-1)^2} \left( \sum_{i=1}^{\infty} (\mu_i^2 + \hat{w}_n) - \frac{1}{z} \left( \sum_{i=1}^{\infty} (\mu_i^2 + \sum_{j=1}^{\infty} \sum_{j=1}^{\infty} V_j V_k) \right) \right)
\]

\[
= \frac{z-2}{(z-1)^2} \left( z u^2 + z \hat{w}_n - \frac{1}{z} (z u^2 + z \hat{w}_n) - z(z-1) \mu^2 \right)
\]

\[
= \frac{z-2}{(z-1)^2} \left( z u^2 + z \hat{w}_n - u^2 - \hat{w}_n - z(z-1) \mu^2 \right)
\]

\[
= \frac{z-2}{(z-1)^2} \left( z u^2 + z \hat{w}_n - u^2 - \hat{w}_n - z(z-1) \right)
\]

\[
= \frac{z-2}{(z-1)^2} \left( z u^2 + z \hat{w}_n - u^2 - \hat{w}_n - z(z-1) \mu^2 + \mu^2 \right)
\]

\[
= \frac{z-2}{(z-1)^2} \left( z u^2 + z \hat{w}_n - u^2 - \hat{w}_n - z z \mu^2 + \mu^2 \right)
\]

\[
= \frac{z-2}{(z-1)^2} \left( z u^2 + z \hat{w}_n - u^2 - \hat{w}_n - z z \mu^2 + \mu^2 \right)
\]

where \( \hat{w}_n \) and \( \mu \) are the actual variance and mean of non-diversifiable risk respectively.

Dividing Equation (6) by \( z \) results in

\[
\left( \frac{z-2}{z} \right) \hat{w}_n = \left( \frac{z-1}{z} \right) \hat{w}_n
\]

\[
= \left( \frac{1-2}{z} \right) \hat{w}_n
\]

\[
= \left( \frac{1}{z} - \frac{2}{z} \right) \hat{w}_n
\]

\[
= \left( \frac{1}{z} \right) \hat{w}_n
\]

\[
= \hat{w}_n
\]

Thus \( \hat{w}_n \) is an unbiased estimator of \( w_n \).

From the results of Wu [7], Equation (6) and proposition 2.2, \( \hat{w}_n \) is a consistent estimator of \( w_n \).

3. DERIVATION OF A PROBABILITY DENSITY FUNCTION FOR RANDOM ERROR OF NON-DIVERSIFIABLE RISK

3.1. Kernel Density Estimation of White Noise of Non-diversifiable Risk

Let \( V_1, V_2, \ldots, V_n \) denote a sample of size \( n \) from the random variable \( V_i(\cdot) \) with density \( f \). The kernel density estimates of \( f \) at the point \( v \) is given by

\[
\hat{f}_n(v) = \frac{1}{nh} \sum_{i=1}^{n} k \left( \frac{v - V_i}{h} \right)
\]

where the kernel \( k \) satisfies \( \int k(v) dv = 1 \) and the smoothing parameter \( h \) is known as the bandwidth. \( V_i \) is the mean of \( v \).

Proposition 3.1: \( \sum_{i=1}^{n} V_i \) is a minimum sufficient statistics of \( V_i \).

Proof:

Given the function \( f_n = \frac{1}{nh} \sum_{i=1}^{n} 1 - \frac{(v - V_i)}{h} \), the likelihood \( L \) of \( v \) is

\[
L(\theta : V_i) = \frac{1}{nh} \sum_{i=1}^{n} \left( \frac{\sum_{j=1}^{n} V_j - V_i}{h} \right)
\]

\[
L(\theta : V_i) = \frac{\partial}{\partial V_i} \ln \frac{1}{nh} \sum_{i=1}^{n} \left( 1 - \frac{\sum_{j=1}^{n} V_j - V_i}{h} \right)
\]

since, \( \sum_{i=1}^{n} V_i = V_i \), then \( V_i - V_i = 0 \)

Let \( \nu = (v_1, v_2, \ldots, v_n) \) be a point in \( \nu : v = (v_1, v_2, \ldots, v_n) \)

Then

\[
L(\nu : V_i) = \frac{\partial}{\partial V_i} \ln \frac{1}{hn} \sum_{i=1}^{n} \left( 1 - \frac{\sum_{j=1}^{n} V_j - V_i}{h} \right)
\]

and therefore,

\[
L(\nu : V_i) = \frac{1}{hn} \sum_{i=1}^{n} \left( 1 - \frac{\sum_{j=1}^{n} V_j - V_i}{h} \right)
\]

Meaning that it is independent of \( V_i \) and thus, \( \sum_{i=1}^{n} V_i \) is a minimum sufficient statistics of \( V_i \). Lehmann and Scheffé [8] remarks that if the sample space is discrete or a finite dimensional Euclidean space then a minimal sufficient statistic will always exists. Since a minimum sufficient statistic exists then the sample space is discrete and the probability density function exists.
It is generally known that the value of the bandwidth is of critical importance while the shape of the kernel function has little practical impact. Thus we estimate bandwidth and use a given kernel function to get the density estimation of the white noise of non-diversifiable risk.

### 3.2. Bandwidth Selection for Kernel Density Estimation of the \( wn \) of Non-diversifiable Risk

Assuming that the underlying density is sufficiently smooth and that the kernel has fourth moment using the Taylor series

\[
\text{Bias}\{f_h(v)\} = \frac{h^2}{2} \mu_2(k) f''(v) + (h^2)
\]

\[
\text{Var}\{f(v)\} = \frac{1}{nh} R(k) f(v) + 0 - \frac{1}{nh}
\]

where \( R(k) = \int k^2(v) dv \) [9]. Adding the leading variance and squared bias terms produces the Asymptotic Mean Squared bias squared error (AMSE)

\[
\text{AMSE} = \frac{1}{nh} R(k) f(v) + \frac{h^4}{4} \mu_2(k)^2 [f''(v)]^2
\]

The overall measure of the discrepancy between \( f \) and \( f \) is the Mean Integrated Squared Error (MISE) which is given by

\[
\text{MISE}(f_h) = E\{\frac{1}{n} f_h(y) - f(y)\} dy
\]

\[
= \int \text{Bias}(f_h(v))^2 dv + \int \text{Var}(f_h(v)) dv
\]

Under an integrability assumption on \( f \), integrating the expression for AMSE gives the expression for the Asymptotic MISE (AMISE), i.e.,

\[
\text{AMISE}\left( f_h \right) = \frac{1}{nh} R(k) + \frac{h^4}{4} \mu_2(k)^2 R(f'')
\]

where

\[
R(f'') = \int [f''(v)]^2 dv.
\]

The value of the bandwidth that minimizes the AMISE is given by

\[
h_{\text{AMISE}} = \left[ \frac{R(k)}{\mu_2(k)^2 R(f'')} \right]^{\frac{1}{2}} n^{-\frac{1}{2}}
\]

Using the rule of thumb method a global bandwidth \( h \) is based on replacing \( R(f'') \) the unknown part of \( h_{\text{AMISE}} \) by its value for a parametric family expressed as a multiple of a scale parameter, which is then estimated from the data. The method dates back to Deheuvels [10] and Scott [11]. It has been popularized for kernel estimates by Silverman [12].

The plug-in method is used to estimate \( h_{\text{AMISE}} \) in this study. Here the unknown quantity \( R(f'') \) in the expression for \( h_{\text{AMISE}} \) is replaced by an estimate. The “solve - the - equation” plug-in approach developed by Sheather and Jones [13] is based on deriving, the pilot bandwidth for the estimate \( R(f'') \), as a function of \( h \), namely

\[
g(h) = C(k) \left[ \frac{R(f'')}{R(f''')} \right]^{\frac{1}{2}} n^{-\frac{1}{2}}
\]

The unknown functions of \( f \) are estimated using kernel density estimates with bandwidth based on normal rules of thumb resulting in

\[
h_{ij} = \left[ \frac{R(k)}{\mu_2(k)^2 R(f''(g_{ij}))} \right]^{\frac{1}{2}} n^{-\frac{1}{2}}
\]

where \( h_{ij} \) is known as the Sheather–Jones plug-in bandwidth. Under smoothness assumption on the underlying density, \( n^{5/14} (h_{ij}/h_{\text{AMISE}}-1) \) has an asymptotic \( N(0, \sigma_0^2) \) distribution. Thus, the Sheather–Jones plug-in bandwidth has a relative convergence rate of order \( n^{5/14} \), which is much higher than that of biased cross-validation.

The triangle kernel is used for smoothing

This is given by

\[
K_m(t) = \begin{cases} 1 - |t|/c, & |t| \leq 1/c \\ 0, & |t| > 1/c \end{cases}
\]

where \( c \) is the constant used to scale the resulting kernel so that the upper and lower quartiles occur at \( \pm 0.25 \). Substituting the kernel in Equation (16) and the unknowns, \( h \) and \( n \) into the density function (7) gives the function

\[
f_h = \frac{1}{hn} \sum_{i=1}^{n} \left[ 1 - \frac{(y - V_i)}{hc} \right]
\]

This function is used to generate probabilities of non-diversifiable risks of given portfolio thus ensuring that actual systematic risk is determined.

### 4. RESULTS

#### 4.1. Calculating Actual Non-diversifiable Risk

The sample white noise is estimated by varying the variance of individual return values \( r \) as given by Equation (2). The non-diversifiable risk estimates in Table 1 are substituted into Equation (3) to give Equation (5).

Sample \( wn \) estimates are then substituted into Equation (5) to determine the actual white noise.

#### 4.2. Density Estimates of Actual White Noise

R statistical software is used to calculate Sheather–Jones (sj) bandwidth and hence the density estimates of actual \( wn \) as plotted in Figure 1.
A summary of statistics resulting from $\hat{s}$ density estimation in Table 2 enables us to apportion densities of the different quartile ranges. 

$F$-values are calculated as follows:

$$F = \frac{v_i - \bar{v}}{\sigma_v}$$

where $v_i$ is white noise of portfolio $i$, $\bar{v}$ is the mean of white noise of all the portfolios, and $\sigma_v$ is variance of all the portfolios. The probability density estimate of a portfolio $i$ is determined by comparing the $F$-values with the apportioned densities of the different quartile ranges and the maximum value. An $F$-score of positive 0.827722 has its density calculated as follows:

$$0.527674 + (1 - 0.827722)0.130402 = 0.63556$$

where: 0.527674 is the maximum value, 0.130402 is the value apportioned to the first quartile and $1 - 0.827722$ represents the fraction occupied in the first quartile.

Final results of the survey are tabulated in Table 3.

### 4.3. Wilcoxon Signed Rank Test

Wilcoxon signed rank test of hypothesis is used to compare the VaR method of determining risk and Kernel white noise method. Here we test the hypothesis that risks obtained by Kernel white noise are a reflection of actual risks than those obtained by VaR.

$H_0$: The population difference are centered at 0.

$H_a$: The population differences are centered at a value $<0$.

Based on a significance level of $\alpha = 0.01$, the proper test is to reject $H_a$ if $Z < -Z_{\alpha}$. Determining $Z$ and $Z_{\alpha}$

$$Z_{\alpha} = \frac{1}{0} \begin{pmatrix} 1 & 0 \end{pmatrix}$$
CONFLICTS OF INTEREST

The authors declare they have no conflicts of interest.

AUTHORS’ CONTRIBUTION

The original idea and subsequent analysis of the research was undertaken by EAS with the guidance of PW and TA.

ACKNOWLEDGMENTS

I wish to acknowledge the University of Nairobi staff in the school of Mathematics as well as my family members for their encouragement and support during the period of undertaking this research.

REFERENCES

[1] Lucchetti A, Ng S, Craig S, Mollenkamp C. Lehman Files for Bankruptcy, Merrill Sold, AIG Seeks Cash. Wall Str J 2008.
[2] Jorion P. Value at risk: the new benchmark for managing financial risk. New York: McGraw-Hill; 2000.
[3] Dennis B, Ponciano JM, Lele SR, Taper ML, Staples DF. Estimating density dependence, process noise, and observation error. Ecol Monogr 2006;76:323–41.
[4] Sklar A. Random variables, distribution functions, and copulas. A personal look backward and forward. J Math Stat 1996;28:1–14.
[5] Wu CF. Characterizing the consistent directions of least squares estimates. Ann Stat 1980;8:789–801.
[6] Anyika E, Waweru R, Odhiambo R. Non-diversifiable risk in investment portfolio as an aid to investment decision making. J Sci Technol 2005;59:13–27.
[7] Wu CF. Asymptotic theory of nonlinear least squares estimation. Ann Stat 1981;9:501–13.
[8] Lehm ann EL, Scheffé H. Completeness, similar regions, and unbiased estimation. Sankhyā Indian J Stat 1950;10:305–40.
[9] Wand MP, Jones MC. Kernel smoothing. London: Chapman and Hall; 1994.
[10] Deheuvels P. Estimation non paramétrique de la densité par histogrammes généralisés. Rev Stat Appl 1977;25:5–42.
[11] Scott DW. On optimal and data-based histograms. Biometrika 1979;66:605–10.
[12] Silverman BW. Density estimation for statistics and data analysis. London: Chapman and Hall; 1986.
[13] Sheather SJ, Jones MC. A reliable data-based bandwidth selection method for Kernel density estimation. J Roy Statist Soc Ser B 1991;53:683–90.