State-value weighted entropy as a measure of investment risk

DAVID N. NAWROCKI and WILLIAM H. HARDING

Villanova University, Villanova, PA 19085, USA

This paper explores a problem with the use of the entropy measure as a measure of investment performance. The commonly used entropy measure ignores the dispersion of security frequency classes used in the calculation of entropy. Therefore, state-value weighting of the entropy is proposed and tested using a portfolio selection heuristic algorithm. The results indicate that weighting the entropy value will increase the investment performance of the entropy risk measure.

I. INTRODUCTION

Entropy and information theory analysis experienced a brief surge of popularity in the finance and economics literature during the early 1970s. This popularity was questioned by Horowitz and Horowitz (1976). They noted numerous articles that implied that entropy analysis measures meaningful information that is not available to standard statistical techniques such as variance or correlation analysis. This is incorrect since there does not seem to be any statistical measure that tells whether information is meaningful in an economic sense. While Philippatos and Wilson (1972, 1974) argue that entropy is a better statistical measure of risk than the variance measure (because entropy is a nonparametric measure), entropy has not exhibited a clear superiority in published studies. Therefore, the main thrust of the Horowitz and Horowitz article is well founded. An additional argument was made by White (1974) who states that, since entropy analysis is not integrated into economic theory including the theory of choice under uncertainty, it should not be used in economic and financial studies.

This early misuse is unfortunate since entropy and information theory has a lot to contribute to economics and finance because of its integration into dynamic models of economic behaviour. Shackle (1952) was the earliest researcher to apply the concepts of information theory to economics in the field of business decisions. He breaks a decision into two components: a representative gain (focus gain) and a representative loss (focus loss). He places a subjective probability on each and measures the potential surprise to the investor. An occurrence with a low probability contains more surprise than an occurrence with a high probability with the amount of surprise measuring the risk of the investment.

Murphy (1965) uses entropy models to model an uncertain environment where adaptive
control processes provide optimal economic decisions. Georgescu-Roegen (1971) models the production of goods and services as an entropic process. Cozzolino and Zahner (1973) provide a proof where a maximum entropy model is equivalent to the random walk model of financial market behaviour. Philippatos and Gressis (1976) provide conditions where mean-variance, mean-entropy and second degree stochastic dominance are equivalent. This last study answers White's criticism by integrating entropy into the theory of choice under uncertainty.

Nicolis and Prigogine (1977) provide a framework to use entropy in dynamic disequilibrium models. Under these conditions, we cannot predict the future. We can only predict scenarios and assign probabilities to each scenario. Nicolis and Prigogine use entropy and bifurcation theory to describe disequilibrium processes in chemical systems. Nawrocki (1984) proposes the use of entropy and bifurcation theory to explain the behaviour of financial markets. He notes the growing literature that discusses disequilibrium models of financial market behaviour and argues the use of entropy because of its integration into dynamic disequilibrium models. Entropy provides a superior analogue – not necessarily the superior statistical measure.

Because of the potential theoretical benefits of using entropy analysis, this paper is interested in the statistical use of entropy in finance and economics. Philippatos and Wilson (1972) suggest entropy as a measure of portfolio (investment) risk because it does not make assumptions concerning the underlying probability distribution. Unfortunately, as we tested the entropy measure with security portfolios, we found that entropy in its discrete form is not well defined for use in finance and economics. Specifically, the generally accepted calculation for entropy is not a good measure of security risk because it ignores the state values of the different frequency classes used in the calculation.

A survey of the information theory literature indicated that a weighted entropy scheme has merit. As a result, the focus of this study is to suggest alternative forms of entropy as a measure of risk and to check the performance of each entropy measure in an empirical context by exploring the possible use of the entropy measure in a simple portfolio heuristic.

The organization of the paper is as follows: First, a simple example illustrating the potential problem with the current entropy measure is presented. Next, an overview of entropy and the derivation of a state-weighted entropy is discussed. Finally, an empirical test of the different entropy measures is described along with a discussion of the results.

II. AN ILLUSTRATION OF THE PROBLEM

The entropy formula, \( H = -\Sigma p_i \log p_i \), that is generally used can cause problems when using entropy as a risk measure.\(^1\) Given the probabilities for security A and B below, it is obvious that security A is less risky than security B even though both securities will have the same entropy value.\(^2\)

\(^1\)See Horowitz and Horowitz (1976) for a general survey of entropy applications in the finance and economics literature.
\(^2\)This example assumes the same return state values for both securities.
Entropy measure and investment performance

<table>
<thead>
<tr>
<th>State i</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( p_i )</td>
<td>( p_i )</td>
</tr>
<tr>
<td>Low return</td>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.2</td>
</tr>
<tr>
<td>High return</td>
<td>5</td>
<td>0.1</td>
</tr>
</tbody>
</table>

The entropy of both securities is 1.4708 using natural logarithms \((2(0.1)\ln 0.1 + 2(0.2)\ln 0.2 + 0.4\ln 0.4)\). The problem is that the entropy ignores the structure of the dispersion contained in the frequency classes. The next section provides an overview of entropy and suggests that a weighted entropy that accounts for the dispersion structure is consistent with the literature in entropy theory.

### III. A General Overview of Entropy

The use of entropy in economics and finance harks back to the development of information theory by Shannon and Weaver (1949). The basic concepts are that the information, \( I \), associated with a message is a function of the probability, \( f_i \), of the message being sent. The lower the value of \( f_i \), the greater the informational value, if the message were received such that:

\[
I = I(f_i), \text{ and} \\
I(f_i) \rightarrow \text{larger as} \\
1/f_i \rightarrow \text{larger and} \\
0 < f_i < 1 
\]

For two independent messages:

\[
f_{ij} = f_i f_j
\]

and the information value of both is:

\[
I(f_{ij}) = I(f_i) + I(f_j)
\]

A form obeying these attributes is:

\[
I(f_i) = -C \log f_i
\]

and the information gain for \( n \) messages is:

\[
IG = \sum_{i=1}^{n} I(f_i) = -C \sum_{i=1}^{n} \log f_i
\]

Having obtained the total information gain, \( IG \), for \( n \) messages, the usual procedure is to consider the expected information gain, \( IG \). This expected value is generally obtained as the
weighted mean where the weights used are the probabilities, \( f_i \), associated with message as in Morse (1962).

\[ IG = -C \sum_{i=1}^{n} f_i \log f_i \]  

(6)

Within a constant factor this is the entropy, \( S \), i.e.

\[ S = C'IG = -C'C \sum_{i=1}^{n} f_i \log f_i = C^* \sum_{i=1}^{n} f_i \log f_i \]  

(7)

It has been true in general that the economics literature uses the following convention.³

\[ H = S \]  

(8)

\[ p_i = f_i \]  

(8a)

\[ C^* = 1.0 \]  

(8b)

yielding the entropy form:

\[ H = - \sum_{i=1}^{n} p_i \log p_i \]  

(9)

In accordance with the above formalism, we wish to consider the previous problem analysis where two securities can have the same entropy, but have different levels of risk.

Philippatos and Wilson (1974) are aware of this potential problem. They state that entropy analysis will differ from variance analysis because entropy analysis depends on the number of potential states in a probability distribution, whereas variance analysis depends on the weight given to each state, i.e. the state value. This difference can be seen in the formulas used in the two measures.

\[ H_i = S_i = - \sum_{j=1}^{n} p_{ij} \log p_{ij} \]  

(10)

\[ V_i = \sum_{j=1}^{n} p_{ij} (s_{ij} - U_i)^2 \]  

(11)

where \( n \) is the number of states in the distribution, \( p_{ij} \) is the probability of security \( i \) being in state \( j \), \( s_{ij} \) is the state value for security \( i \) in state \( j \) and \( U_i \) is the mean value for security \( i \). Note that both \( H_i \) and \( S_i \) are generally accepted symbols for entropy.

The problem with the entropy measure exists since the entropy measure may not provide properly for the weight associated with state values. The variance weights each state on the basis of the squared deviation from the mean.

³See Philippatos and Wilson (1972) and Horowitz and Horowitz (1976).
IV. A WEIGHTED ENTROPY CONCEPT

We note in the previous section leading to Equation 9 that the weighting approach used involved the factor $p_i$. It seems reasonable that a somewhat different weighting scheme may be in order. Recall that the original development considered $C^* = 1.0$, in order to arrive at:

$$H = -C^* \sum_{i=1}^{n} p_i \log p_i = - \sum_{i=1}^{n} p_i \log p_i$$

(12)

for a particular group of messages i.e. return values on a given security. However, the more general form would be:

$$H = -C^* \sum_{i=1}^{n} p_i \log p_i \text{ or}$$

(13)

$$H = - \sum_{i=1}^{n} C^* p_i \log p_i$$

(13a)

where the constant value property of $C^*$ allows it to move inside the summation.

For our particular system, there appears to be no a priori reason why $C^*$ should be constant over all states. That is, this principle of equal weight is not appropriate whether the weight is one (1.0) or some other constant value. We therefore let $C^* \to C^*_i$, i.e. let $C^*$ become dispersion dependent (not constant) and consider the consequences of:

$$H = - \sum_{i=1}^{n} C^*_i p_i \log p_i$$

(14)

Since the above entropy is associated with a single stock and the entropy for a portfolio of $m$ stocks ($i = 1, 2, 3, \ldots, m$) having $n$ rate of return states ($j = 1, 2, 3, \ldots, n$) would be given as (assuming perfect independence between securities):

$$H_p = \sum_{i=1}^{m} H_i = \sum_{i=1}^{m} \sum_{j=1}^{n} C^*_i p_{ij} \log p_{ij}$$

(15)

It is interesting to note that a variable transformation formally relates the above portfolio entropy (Philippatos and Wilson, 1972, Equation 10) with Equation 6 that is provided by Dinkel and Kochenberger (1979) for:

$$K_{ij} = 1$$

(16)

$$t_{ij} = C^*_i p_{ij}$$

(16a)

$$C^*_i = 1/K_{ij}$$

(16b)

This transformation is important since it relates our entropy Equation 15 to an area in mathematical programming known as constrained entropy models. This is a model where the objective function is a weighted entropy equation as in our Equation 15. Therefore, we have converted Philippatos and Wilson's entropy model of portfolio analysis into a constrained entropy model.

We are concerned here with exploring reasonable forms for the state-dependent weighting
factor $C^*_j$. Possible forms for the scheme are:

$$C^*_j = s_{ij}$$ (17)

where $s_{ij}$ is the state value or the mean value of the frequency class $j$. Other alternatives using the global mean, $U_i$, could be

$$C^*_j = (s_{ij} - U_i)^2$$ (18)

and

$$C^*_j = |s_{ij} - U_i|$$ (19)

When we wish to compare individual security entropies, this weighted entropy could be calculated as follows:

$$H_i = - \sum_{j=1}^{n} C^*_j p_{ij} \log p_{ij}$$ (20)

We recognize that though Equations 18 and 19 may not represent a great amount of original thought, they do represent two of the most commonly used weighting techniques used in statistics: squared deviations and absolute deviations. We consider these techniques to be a reasonable starting point for consideration of weighted values in the constrained entropy calculation.

V. METHODOLOGY

The methodology used in the earlier study by Nawrocki (1983) is replicated for this study. This study also uses 72-month historical periods, 60-month holding periods, 1% transaction fees, and the second degree stochastic dominance (SSD) to evaluate portfolio performance. Three portfolio strategies, the Elton, Gruber and Padberg (1976) optimal diagonal matrix algorithm, an equal weight market index and a buy-and-hold (BH) with an initial equal weight in each security, are used as benchmarks for comparing portfolio performance. (It should be noted that the market index is constantly rebalanced each month back to the original equal proportions while the buy-and-hold is never rebalanced.)

The security sample includes 62 securities that are randomly selected from the CRSP tape universe that had complete data for the 1950–75 period. As such, this sample is a subset of the 150 securities used by the Nawrocki (1983) study.

The return-to-risk ratio portfolio heuristic that was developed in the earlier Nawrocki study is used in this paper to test the mean absolute deviation ($MAD$), the standard deviation ($\sigma$), the target semivariance ($SV_T$), the unweighted entropy ($H$), the entropy weighted by squared deviations from the global mean ($H_{sq}$) and the entropy weighted by absolute deviations ($H_{abs}$) as alternative risk measures.

Portfolio sizes of 5, 10, 15, 20, 25 and 30 securities are tested in this study with the entropy values calculated using the following procedure. All calculations use natural logarithms. Six

frequency classes are used with three classes above the mean and three classes below the mean. The width of each class is set to 0.05 (5.0% return) since 0.05 is a close approximation to the standard deviation of the market index. The frequency classes are shown below:

<table>
<thead>
<tr>
<th>Class</th>
<th>Appropriate returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( R_i &gt; U_i + 0.10 )</td>
</tr>
<tr>
<td>2</td>
<td>( U_i + 0.05 &lt; R_i &lt; U_i + 0.10 )</td>
</tr>
<tr>
<td>3</td>
<td>( U_i &lt; R_i &lt; U_i + 0.05 )</td>
</tr>
<tr>
<td>4</td>
<td>( U_i - 0.10 &lt; R_i &lt; U_i - 0.05 )</td>
</tr>
<tr>
<td>5</td>
<td>( R_i &lt; U_i - 0.10 )</td>
</tr>
</tbody>
</table>

where \( U_i \) is the expected mean security return and \( R_i \) is the periodic security return for period \( t \).

VI. EMPIRICAL RESULTS

The empirical results are provided in Tables 1 and 2. Table 1 provides information concerning the risk–return performance of the different portfolios. \( R/V \) ratios are computed using the following formula:

\[
R/V = (E_p - R_f)/\sigma_p
\]

The market index provides the best risk–return performance. The Elton, Gruber and Padberg (ELTON) optimal algorithm and the target semivariance heuristic offer the best performance of the portfolio techniques. Next comes the entropy weighted by squared deviations, the entropy weighted by absolute deviations and the standard deviation. The unweighted entropy generally provides the worst results with the exception of the 25 and 30 security portfolios where the entropy is comparable to the \( MAD \) and the standard deviation \( (V) \). The unweighted entropy

<table>
<thead>
<tr>
<th>Table 1. Reward to variability ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S )</td>
</tr>
<tr>
<td>( MAD )</td>
</tr>
<tr>
<td>( H )</td>
</tr>
<tr>
<td>( Hsq )</td>
</tr>
<tr>
<td>( Habs )</td>
</tr>
<tr>
<td>( V )</td>
</tr>
<tr>
<td>( SV_e )</td>
</tr>
<tr>
<td>( ELTON )</td>
</tr>
<tr>
<td>Market</td>
</tr>
<tr>
<td>BH</td>
</tr>
</tbody>
</table>

\(^5\)The monthly mean return of the market index is 1.009093 and the standard deviation is 0.043943 for the 1950–75 period.
performs poorly with the 5 and 10 security portfolios where having a good risk measure is more crucial than with larger size portfolios. It is clear that the entropy weighted with squared deviations provides an improvement over the unweighted entropy.

The results in Table 1 are not very strong because of the distribution and utility assumptions of the standard deviation measure that is used in the calculation of the $R/V$ ratios (Equation 21). In order to provide stronger results the second degree stochastic dominance (SSD) analysis is performed on the portfolio returns.

These SSD results are presented in Table 2. The market index was undominated by other portfolios while the BH strategy was dominated by a number of portfolios. With 5 of 6 portfolios undominated, the ELTON optimal algorithm was the best portfolio technique followed by the target semivariance (4 of 6) and the entropy weighted by squared deviations (3 of 6). Of the other risk measures, only $MAD$ provides an undominated portfolio.

These results indicate that the weighted entropy scheme performs well in comparison to some optimum approaches and thus the technique deserves consideration. The unweighted entropy performs poorly in comparison to these same techniques.

The $MAD$ results are surprising because of the poor performance of the $MAD$ measure as compared to earlier results published in Nawrocki (1983). The $MAD$ performance indicates that $MAD$ is more sensitive to the size of the security sample than is the target semivariance ($SV_t$).

**VII. SUMMARY**

This paper explores the use of entropy as a risk measure in the security analysis. A problem with the entropy calculation as generally used in the finance literature was noted. The problem results from the entropy calculation that ignores the dispersion of the security frequency classes used in the entropy calculation. Our analysis of the information theory literature indicates that this problem does not have to exist. There is precedence in the literature for adding a constant to the entropy calculation. In fact, the field of constrained entropy optimization models uses weighted entropy schemes. This constant is used to introduce dispersion related state-values into the entropy calculation. Given this result, we suggest two alternative calculations for entropy; entropy weighted by squared deviations from the mean and entropy weighted by absolute deviations from the mean.
An empirical test using a portfolio selection heuristic previously developed in Nawrocki (1983) is used to provide an example of the relative investment performance of the various entropy measures. Using second degree stochastic dominance to evaluate portfolio performance, portfolios derived from an entropy measure weighted with squared deviations from the mean dominated portfolios derived with unweighted entropy values and entropy values weighted with absolute deviations from the mean. Other undominated portfolios include the target semivariance measure and the Elton, Gruber and Padberg (1976) optimal algorithm.

ACKNOWLEDGEMENTS

The authors wish to thank the referee of this journal for helpful comments. Support for this study was provided by a summer research grant from Villanova University.

REFERENCES