ISSN 2090-3359 (Print)
ISSN 2090-3367 (Online)

Advances in Decision Sciences

Volume 22(A)
22nd Anniversary Special Issue
December 2018

Michael McAleer

Editor-in-Chief
University Chair Professor
Asia University, Taiwan

Published by Asia University, Taiwan

ADS@ASIAUNIVERSITY

Joint Distribution of Forecasts and Outcomes: Impact of Non-Normality on the Measurement of Forecasting Skill, with Applications to Analysts' Target Prices *

S.A. Bond

University of Cincinnati

Q. Chang

Office of the Comptroller of the Currency

J. Knight

University of Western Ontario

S.E. Satchell **
Birkbeck College and University of Sydney

Revised: December 2018

* The authors would like to thank Michael McAleer and a reviewer for invaluable comments and suggestions.

** Corresponding author: ses999gb@yahoo.co.uk

Abstract

The purpose of this paper is to provide a detailed analysis of the joint distribution of forecasts and outcomes in the context of financial forecasting. We use Edgeworth expansions to model this joint distribution. In turn this allows us to assess the impact of non-normality, either in forecasts, or outcomes or both. This leads to multiple results; we can deduce the distribution of the forecast error; we can analyse the properties of the hit rate (*H*), as a statistical concept in its own right, and also in its relation to the information coefficient (*IC*); two tools that are used in the assessment of forecasting ability by active fund managers and financial analysts. Our paper contributes to the recent econometric literature on directional forecasting as well as the empirical literature that examines analyst performance. We find that the close link between *H* and *IC* under normality breaks down in the more general case. We provide further evidence on the richness of this approach by looking at simulation and empirical evidence.

Key Words: Hit rate, Information Coefficient, Edgeworth expansions, financial forecasting, Analysts' Target Prices.

JEL: G10, G12.

1. Introduction

The ability of financial market participants to make accurate forecasts related to future market prices or other variables of interest has resulted in an extensive amount of research in the finance and economics literature. This literature includes assessing the performance of active fund managers, and earnings and target price forecasts made by analysts. These examples can be considered as categories or applications within the general field of economic forecasting. For instance, a particular view of active fund management is that active management equals passive management plus asset returns forecasts. Two key tools that have often been used to assess the forecasting ability among fund managers or analysts are the hit rate and the information coefficient.

In this paper we will provide a general framework to assess these statistics of forecasting ability by providing a general bivariate framework to address such questions. We use Edgeworth expansions to set up a bivariate model of forecasts and outcomes and specialise this to look at the distributions of forecast errors, hit rates and information coefficients.

Specifically, we investigate the relationship between the hit rate, defined as the probability that the forecaster correctly forecasts the sign of the future asset return, and the Information Coefficient, defined as the correlation between the forecasted return and the actual return. This work has been motivated by research on directional forecasts by financial econometricians, see Christoffersen et al (2007). However the broader issue of assessing financial forecasters by directional forecasts has been studied for some time, see Cowles (1933) for an early reference, as well as Henriksson and Merton (1981)[1].

The hit rate has been used in the academic finance literature to assess the performance of analysts' price targets. For example, Asquith, Mikhail and Au (2005) report hit rates of 54.28% for the 818 cases of analysts' price targets they considered from their sample of Institutional Investors All-American analyst-reports. In that paper, Asquith et al define a hit as the future price exceeding the target at some point in the next 12 months, if the target exceeds the current price, or going below the target, if the target is below the current price. Further analysis of price targets using similar approaches have been completed by Bradshaw and Brown (2007), and Dechow and You (2013).

While our interest in this paper is primarily on furthering the development of the hit rate in a financial context, we do note our connection to a broader literature on directional forecasting following from Pesaran and Timmermann (1992). Anatolyev (2009) provides an overview of this area. The hit rate is also used extensively in economic forecasting. For instance, in a macroeconomic context Garratt et al (2009) discuss the hit rate when assessing the forecasting performance of monetary aggregates on output growth and inflation.

In this paper we also highlight the potential consequences of the behavioural aspects of a forecaster's actions by explicitly considering non-normality in the joint distribution of returns and outcomes. Such behaviour has been well discussed in a long literature on the way in which

---

[1] These calculations have been advocated in the practitioner literature as well, for example, by Grinold and Kahn (2000).

professional forecasters construct their forecasts (see for instance Keane and Runkle 1990, Ehrbeck and Waldmann 1996, Graham 1999, and Chen and Jiang 2006).

The structure of our paper is as follows. In section 2, we consider some general definitions and express the hit rate as a function of the bivariate distribution of returns and forecasts, we also look at some preliminary results. In section 3, we define our joint distribution of actuals and outcomes in terms of Edgeworth expansions and show how higher moments will affect the relationship between the information coefficient and the hit rate. In section 4, we use our model to investigate the distribution of the forecast error in terms of fundamental parameters. In section 5 we calculate some simulations based on the true model using a cubic market model, a structure that will exhibit non-normality under appropriate circumstances. In section 6, we carry out an empirical study based on forecasts of Target Prices by analysts for US equity; this illustrates some of the empirical challenges involved in applying such procedures.

2. Background and Some Preliminary Results

We now present some basic definitions of the quantities we wish to study in this paper. We denote the actual return which we wish to forecast by Y. The forecast of Y is denoted by X. We define the joint (outcome) distribution function of X and Y by F(x,y); in particular"

$$Prob(X \leq x, Y \leq y) = F(x, y)$$

$$Prob(X \leq x) = F_1(x)$$

$$Prob(Y \leq y) = F_2(y).$$

The hit rate, *H*, defined as the probability that the forecaster correctly forecasts the sign of the outcome, is given by the following expression:

$$H = 1 - F_1(0) - F_2(0) + 2F(0,0). \qquad (1)$$

An alternative measure, the Hansen-Kuiper score is the hit rate minus the probability of getting the score wrong. Since this is just *2H-1*, we shall not investigate this further. The Information Coefficient, *IC*, is defined as the correlation between *X* and *Y*. In the case that *X* and *Y* are bivariate standardised normal with correlation coefficient ρ, then $IC = \rho$; however *H* requires further calculation. By recourse to polar coordinates and a number of other transformations, we find that $H = \frac{1}{2} + \frac{Sin^{-1}(\rho)}{\pi}$. This result is usually attributed to Sheppard (1900). The inverse of this relationship can easily be derived; it is $\rho = \sin(\pi(H - \frac{1}{2}))$. We see immediately that $H > \frac{1}{2}$ is required for the IC to be positive.

In practice, most forecasting problems involve non-zero means and non-unit variances. Moreover, standardising financial data will typically involve estimating volatility processes which can be quite involved. We therefore present our analysis in terms of non–standardised variables. We now assume that $\binom{X}{Y} \sim N\left(\begin{pmatrix}\mu_x\\ \mu_y\end{pmatrix}, \begin{pmatrix}\sigma_x^2 & \rho\sigma_y\sigma_x\\ \rho\sigma_y\sigma_x & \sigma_y^2\end{pmatrix}\right)$; we denote the standardised bivariate normal distribution function by $\Phi(x, y; \rho)$ and the standardised univariate normal distribution by $\Phi(x)$. The relevant normal densities (pdf's) will be denoted by $\varphi(x)$. We define the Sharpe ratio, against a zero benchmark by $SR_x = \frac{\mu_x}{\sigma_x}$. Then the hit rate can be expressed as, using equation (1):

$$H = 1 - \Phi(-SR_x) - \Phi(-SR_y) + 2\Phi(-SR_x, -SR_y; \rho) \quad (2)$$

It is an attractive feature that the hit rate is a function of only the Sharpe Ratios of the actual and the forecast and the *IC*. However, when we come to differentiate *H* with respect to either Sharpe Ratio, we find the following result.

Proposition 1

Given the assumptions of bivariate normality for the actual and the forecast, and the formula for H given by equation (2), the derivative of H with respect to the Sharpe ratio of X is given by

$\frac{\partial H}{\partial SR_x}$=$2\varphi(SR_x)(1 - \Phi\left(\frac{\rho SR_x - SR_y}{\sqrt{1-\rho^2}}\right)) + \varphi(SR_x)$, which is obviously positive. By symmetry, an analogous result exists for the Sharpe ratio of Y.

Proof: See Appendix.

The monotonicity of H with respect to both Sharpe ratios seems rather counter-intuitive. Taken from a point where both ratios are equal, this says that the incremental gains are equal. This seems to be unlikely to be true in practice. This is one motivation for exploring more complex distributions for F(X,Y).

Another reason is that it is common practice for forecasters to move their forecasts when compared in cross-section. Moving it to the consensus, avoids the odium of being the one person who got it wrong. This has the advantage that, if reality exhibits an extreme value, one may be regarded as the best forecaster relative to the others; whilst we do not address this cross-sectional problem directly, we may want to be able to bring higher moments into our model, both in forecasts and in actual.

It should also be noted that the actual events for which the hit rate is computed require modification at times. Returning to Asquith et al (2005), their definition requires that the actual exceeds the target (forecast) over the 12 month period if the forecast is positive with a corresponding definition if the forecast is negative. They note on para1, page 208, that;

> "Price targets that project a change of 0-10% and 10-20% are achieved 74.4% and 59.6% of the time, respectively. In contrast, price targets that project a change in price of 70% or more are realised in fewer than 25% of the time."

This suggests a monotonic reduction in the hit rate as we condition on the size of the forecast, given their definition of hitting. We therefore define the AMA (Asquith, Mikhail and Au) conditional hit rate as:

$$\text{CHRAMA}=\int_{s}^{\infty} pdf(x|s)dx \text{ if } y = s>0 \text{ or} \quad (3)$$

$$=\int_{-\infty}^{s} pdf(x|s)dx \text{ if } y = s<0 \ .$$

Proposition 2

If the conditional value of *s* is positive, and the beta of the actual on the forecast is less than l then CHRAMA, as given by equation (3), is decreasing in *s* under the assumptions of bivariate normality.

Proof. The conditional pdf(x|s) is $N(\mu_x + \beta(s - \mu_y), \sigma_x^2(1 - \rho^2))$ where *N(a,b)* means normally distributed with mean *a* and variance *b*. Assuming *s*>0 and integrating between *s* and infinity, we see that:

$$\text{CHRAMA}=\Phi\left(\frac{SR_x}{\sqrt{1-\rho^2}} + (\beta - 1)\frac{s}{\sigma_x\sqrt{1-\rho^2}} - \beta\frac{SR_y}{\sqrt{1-\rho^2}}\right).$$

If we differentiate this the sign of the derivative will be determined by the relative magnitudes of beta and 1. QED.

We note that we would expect beta to be less than 1 in most financial applications, certainly given the context of target prices. Proposition 2 gives a theoretical justification for the results discussed above in Asquith et al. Furthermore, if we work with the standardised model, we see that CHRAMA becomes $\Phi\left(-\sqrt{\frac{1-\rho}{1+\rho}}\, s\right)$. This is necessarily less than 50% when the IC is positive, so this may well explain the low numbers that this measure achieves. An alternative definition of the conditional hit rate could be based on hitting times and continuous stochastic processes; this is a topic for future research.

3. Bivariate Non-Normality; Edgeworth Expansions

A simple way to capture non-normality of an unknown distribution is to use Edgeworth expansions, thereby expressing the distribution as an approximation given by a normal with correction terms for skewness and kurtosis. Whilst there are a number of legitimate criticisms one can make of this technique, it does allow for a fairly general form of analysis. Furthermore, it is widely used in financial econometrics, see Zhang et al (2009), for a recent application. We provide full details of the Edgeworth expansion of the three distribution functions involved in our definition of H, i.e., $F_1(x)$, $F_2(y)$ and F(x, y) in the appendix.

For the univariate distribution functions $F_1(x)$ and $F_2(y)$ the expansions are quite straightforward and can be found in most standard statistics books. The expansion can be expressed generically by the following:

$$F(z) = F_0(z) - \frac{1}{3!}K_3F_3 + \frac{1}{4!}K_4F_4$$

Where:

$$F_0(z) = \int_{-\infty}^{z} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{w-\mu}{\sigma}\right)^2\right) dw$$

With:

$$F_n = \frac{\partial^n}{\partial z^n} F_0(z)$$

and K_3 and K_4 are the third and fourth order cumulants, given by:

$K_3 = E(z - \mu)^3$

$K_4 = E(z - \mu)^4 - 3\sigma^4$

with $\mu = E(z)$ and $\sigma^2 = Var(z)$.

As:

$$F_0(z) = \int_{-\infty}^{\left(\frac{z-\mu}{\sigma}\right)} \phi(w)\,dw$$

$$= \Phi\left(\frac{z-\mu}{\sigma}\right)$$

The expansion is usually given in the standardised form:

$$F(z) = \Phi\left(\frac{z-\mu}{\sigma}\right) - \left[\frac{1}{3!}\frac{K_3}{\sigma^3}H_2\left(\frac{z-\mu}{\sigma}\right) + \frac{1}{4!}\frac{K_4}{\sigma^4}H_3\left(\frac{z-\mu}{\sigma}\right)\right]\varphi\left(\frac{z-\mu}{\sigma}\right)$$

where $H_r(\cdot)$ is the r^{th} Hermite polynomial defined by the identity:

$$(-1)^r \frac{d^r}{dx^r}\varphi(x) = H_r(x)\varphi(x) .$$

For the bivariate case things get a lot more complicated due to the bivariate cumulants and all the third and fourth order derivatives of the joint distribution, that is:

$$F(x,y) = F_0(x,y) - \frac{1}{3!}\left[K^{1,1,1}\frac{\partial^3 F}{\partial x^3} + K^{2,2,2}\frac{\partial^3 F}{\partial y^3} + 3K^{1,1,2}\frac{\partial^3 F}{\partial x^2 \partial y} + 3K^{1,2,2}\frac{\partial^3 F}{\partial x \partial y^2}\right]$$

$$+\frac{1}{4!}\left[K^{1,1,1,1}\frac{\partial^4 F}{\partial x^4} + K^{2,2,2,2}\frac{\partial^4 F}{\partial y^4} + 4K^{1,2,2,2}\frac{\partial^4 F}{\partial x \partial y^3} + 4K^{2,1,1,1}\frac{\partial^4 F}{\partial x^3 \partial y} + 6K^{1,1,2,2}\frac{\partial^4 F}{\partial x^2 \partial y^2}\right]$$

where $K^{i,j,k}$ etc. are the bivariate cumulants and the derivatives are taken with respect to $F_0(x,y)$, defined below. In particular, from McCullagh (1987, Ch. 2), we have:

$$K^{i,j,k} = \frac{\partial^3}{\partial t_i \partial t_j \partial t_k} \ln M(t_1,t_2)\Big|_{\substack{t_1=0 \\ t_2=0}}$$

$$K^{i,j,k,l} = \frac{\partial^4}{\partial t_i \partial t_j \partial t_k \partial t_l} \ln M(t_1,t_2)\Big|_{\substack{t_1=0 \\ t_2=0}}$$

with:

$$M(t_1,t_2) = E[\exp(t_1 x + t_2 y)]$$

and:

$$F_0(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y}(2\pi\sigma_x\sigma_y\sqrt{1-\rho^2})^{-1}\exp\left(-\frac{1}{2(1-\rho^2)}\left[\left(\frac{w_1-\mu_x}{\sigma_x}\right)^2+\left(\frac{w_2-\mu_y}{\sigma_y}\right)^2-2\rho\left(\frac{w_1-\mu_x}{\sigma_x}\right)\left(\frac{w_2-\mu_y}{\sigma_y}\right)\right]\right)dw_2\,dw_1 .$$

As in the univariate case, we can standardise the expansion by noting that:

$$F_0(x,y)=\int_{-\infty}^{\frac{x-\mu_x}{\sigma_x}}\int_{-\infty}^{\frac{y-\mu_y}{\sigma_y}}\phi(w,u;\rho)\,du\,dw$$

$$=\Phi\left(\frac{x-\mu_x}{\sigma_x},\frac{y-\mu_y}{\sigma_y};\rho\right).$$

Consequently:

$$\frac{\partial^{i+j}F_0(x,y)}{\partial x^i\partial y^j}=\frac{1}{\sigma_x^i\sigma_y^j}\frac{\partial^{i+j}}{\partial x_1^i\partial x_2^j}\Phi(x_1,x_2;\rho)\Bigg|_{\substack{x_1=\frac{x-\mu_x}{\sigma_x}\\ x_2=\frac{y-\mu_y}{\sigma_y}}}$$

Further, noting that:

$$\Phi(x_1,x_2;\rho)=\int_{-\infty}^{x_1}\int_{-\infty}^{x_2}(2\pi\sqrt{1-\rho^2})^{-1}\exp\left(-\frac{1}{2(1-\rho^2)}(w_1^2+w_2^2-2\rho w_1w_2)\right)dw_2\,dw_1$$

$$=\int_{-\infty}^{x_1}\frac{1}{\sqrt{2\pi}}e^{-w_1^2/2}\int_{-\infty}^{x_2}(2\pi\sqrt{1-\rho^2})^{-1}\exp\left(-\frac{1}{2(1-\rho^2)}(w_2-\rho w_1)^2\right)dw_2\,dw_1$$

$$=\int_{-\infty}^{x_2}\frac{1}{\sqrt{2\pi}}e^{-w_2^2/2}\int_{-\infty}^{x_1}(2\pi\sqrt{1-\rho^2})^{-1}\exp\left(-\frac{1}{2(1-\rho^2)}(w_1-\rho w_2)^2\right)dw_1\,dw_2$$

all the derivatives are readily calculated and are given in the Appendix.

We now re-evaluate the relationship between H, the hit rate, and IC(ρ).

<u>Proposition 3</u>

Using Edgeworth expansions and in the special case of $\mu_x = \mu_y = 0$, $\sigma_x = \sigma_y = 1$, evaluated at $x = y = 0$ the hit rate can be expressed as:

$$H = \frac{1}{2} + \frac{\sin^{-1}(\rho)}{\pi} + \frac{1}{24\pi\sqrt{(1-\rho^2)}}\left[(E(x^4)+E(y^4)-6)\frac{\rho(3-2\rho^2)}{1-\rho^2}\right.$$

$$-\frac{4}{1-\rho^2}(E(xy^3)+E(x^3y)-6\rho)$$

$$\left.+\frac{6\rho}{1-\rho^2}(E(x^2y^2)-2\rho^2-1)\right]. \qquad (4)$$

Proof. See Appendix.

The contributions of the three terms inside the square brackets on the right-hand side of equation (4) are interesting; in each case the term in the brackets represents the additional contribution from non-normality. They can be thought of as the contributions due to univariate kurtosis, odd-powered cross-kurtosis (sometimes called co-kurtosis) and even-powered cross kurtosis, respectively. We note if skill is positive, on inspection of equation (4), that: increases in univariate kurtosis increase the hit rate; that increases in cross-kurtosis involving odd terms, decrease the hit rate; and, that increases in cross-kurtosis involving even terms increase the hit rate. The bracketed terms will be zero if the data are normal but will be positive for fat-tailed distributions. For example if both marginals are leptokurtotic, then $(E(x^4)+E(y^4)-6)$ will be positive. Similar comments apply to the other two bracketed expressions. Furthermore from the symmetries of the problem, no cubic terms enter into the hit rate, which may be unrealistic in practice. When $\rho = 0$, there is considerable simplification for the approximation. The result is given in the following corollary.

Corollary

If $\rho = 0$, equation (4) simplifies further and, we find:

$$H = \frac{1}{2} - \frac{1}{6\pi}\left(E(xy^3) + E\,(x^3 y)\right). \quad (5)$$

This shows that the forecaster could have no skill as measured by the IC but still have a hit rate that is greater, or less than, 50%. This is a striking result as the impact of non-normality as measured by odd cross-moment kurtosis is to increase/decrease the hit rate when the forecaster has an IC of zero.

4. Distribution of the Forecast Error

In addition to using an Edgewoth expansion to develop an expression for the hit rate we now use it to examine the distribution of the forecast error. With X being the forecast of the variable Y we define the forecast error, ε, by ε = X – Y. If the joint distribution of X and Y is given by a Bivariate Edgeworth expansion we can readily show that ε will have a univariate Edgewoth expansion with moments and cumulants associated with those of X and Y. Consequently, the pdf and cdf of the forecast error can be derived. Using sample estimates of the relevant cumulants we can easily develop an approximation and use it to calculate various probabilities, for example P(|ε| > η), and compare them with the sample data.

It is clear that if the vector Z has a multivariate Edgeworth distribution then a one to one transformation of Z will also have a multivariate Edgeworth distribution. For example, if $Z \to V = AZ$ where the elements in V are given by $V^r = a_i^r Z^i = \sum_{i=1}^{m} a_i^r Z^i$. That is:

$$\begin{pmatrix} V^1 \\ V^2 \\ . \\ . \\ V^m \end{pmatrix} = \begin{pmatrix} a^1 \\ a^2 \\ . \\ . \\ a^m \end{pmatrix} \begin{pmatrix} Z^1 \\ Z^2 \\ . \\ . \\ Z^m \end{pmatrix}.$$

The cumulants of V will consequently change as follows (see McCullagh, 1987, Sect. 2.4): $a_i^r K^i$, $a_i^r a_j^s K^{i,j}$, $a_i^r a_j^s a_k^t K^{i,j,k}$ and $a_i^r a_j^s a_k^t a_l^u K^{i,j,k,l}$, where the $K^i, K^{i,j}$ etc are the cumulants associated with the vector Z.

If interest only centres on one of the V^r say, $V^m = a_i^m Z^i$ then the cumulants of this scalar variable are given by $a_i^m K^i$, $a_i^m a_j^m K^{i,j}$, and so on.

For our purposes, we only need to deal with a bivariate Edgeworth distribution corresponding to the vector $Z = (X, Y)'$ and the transformation from $Z \to V$ where:

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} z_1 \\ z_1 - z_2 \end{pmatrix} = \begin{pmatrix} X \\ X - Y \end{pmatrix}.$$

We now prove the following lemma.

<u>Lemma</u>

If the two-dimensional vector Z has a bivariate Edgeworth distribution then the transformed vector V given by $v = Az$ where $A = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}$ will also have a bivariate Edgeworth distribution with the individual elements in v having univariate Edgeworth distributions.

Proof

The bivariate Edgeworth distribution of the vector is given by:

$$f(z) = f_0(z) + \frac{K_z^{i,j,k}}{3!} f_{ijk}(z) + \frac{K_z^{i,j,k,l}}{4!} f_{ijkl}(z),$$

where $K_z^{i,j,k}$ and $K_z^{i,j,k,l}$ are the bivariate third and fourth order cumulants and $f_{ijk}(z)$ and $f_{ijkl}(z)$ are derivatives of $f_0(z)$ given by: $f_{ijk}(z) = \frac{\partial^3 f_0(z)}{\partial z^i \partial z^j \partial z^k}$, and so on.

Transforming from $z \to v$ we have the corresponding distribution for v given by:

$f(v) = f_0(v) + \frac{K_v^{i,j,k}}{3!} f_{ijk}(v) + \frac{K_v^{i,j,k,l}}{4!} f_{ijkl}(v)$, with its associated moment generating function given by:

$M_v(s) = M_0(s)[1 + \frac{s_i s_j s_k}{3!} K_v^{i,j,k} + \frac{s_i s_j s_k s_l}{4!} K_v^{i,j,k,l}]$ and by setting $s_1 = 0$, we have the marginal mgf for v_2:

$$M_{v_2}(s_2) = M_0(0, s_2)[1 + \frac{s_2^3}{3!} K_{v_2}^{2,2,2} + \frac{s_2^4}{4!} K_{v_2}^{2,2,2,2}] .$$

As $M_0(s) = \exp(s_i K_v^i + \frac{1}{2} s_i s_j K_v^{i,j})$, we have $M_0(0, s_2) = \exp(s_2 K_{v_2}^2 + \frac{1}{2} s_2^2 K_{v_2}^{2,2})$ which we recognize as the mgf of a univariate Normal random variable with mean $K_{v_2}^2$ and variance $K_{v_2}^{2,2}$.

Higher-order cumulants of v_2 can be readily found, and hence inverting the mgf term-by-term, we obtain the univariate Edgeworth distribution given by:

$$f(v_2) = f_0(v_2)[1 + \frac{\lambda_3(v_2)}{3!} h_3(w) + \frac{\lambda_4(v_2)}{4!} h_4(w)],$$

where $f_0(v_2) = (K_{v_2}^{2,2})^{-0.5}\varphi(w)$, with $w = \frac{v_2 - K_{v_2}^{2}}{\sqrt{K_{v_2}^{2,2}}}$ and $\varphi(.)$ is the standard Normal pdf. Also:

$$\lambda_3(v_2) = \frac{K_{v_2}^{2,2,2}}{(K_{v_2}^{2,2})^{1.5}} \text{ and } \lambda_4(v_2) = \frac{K_{v_2}^{2,2,2,2}}{(K_{v_2}^{2,2})^{2}}$$

and finally $h_3(w)$ and $h_4(w)$ are the standard Hermite polynomials: $h_3(w) = w^3 - 3w$ and $h_4(w) = w^4 - 6w^2 + 3$.

We can now use the above theorem to develop an approximation to the pdf of the forecast error and thereby characterize its distribution rather than just examining its moments. To simplify notation we now let the Edgeworth distribution for the forecast error be given by:

$$f(\varepsilon) = \sigma_\varepsilon^{-1}\varphi(w)[1 + \frac{\lambda_3}{3!} h_3(w) + \frac{\lambda_4}{4!} h_4(w)]$$

where ε has a mean of μ_ε and variance of σ_ε^2 and the third and fourth order cumulants are given by K_3 and K_4. Thus $h_3(w)$ and $h_4(w)$ are as defined earlier and λ_3 and λ_4 the standardized cumulants also defined as earlier and finally, $w = \frac{\varepsilon - \mu_\varepsilon}{\sigma_\varepsilon}$.

By replacing the cumulants by unbiased estimators, we can easily develop the approximation for any set of data. Further, by developing a corresponding expansion for the distribution function, which can be shown to be given by:

$$F(\varepsilon) = \Phi(w) - \varphi(w)[\frac{\lambda_3}{3!} h_2(w) + \frac{\lambda_4}{4!} h_3(w)]$$

where $\Phi(.)$ is the standard Normal cdf and $h_2(w) = w^2 - 1$ the second degree Hermite polynomial, we can then calculate various probabilities associated with the forecast error. For example, we have:

$$P(\varepsilon \geq \omega) = 1 - F(\omega)$$

and

$$P(|\varepsilon| \geq \omega) = 2(1 - \Phi(w)) + 2\varphi(w)\frac{\lambda_4}{4!}h_3(w), \text{ where } w = \frac{\omega - \mu_\varepsilon}{\sigma_\varepsilon}.$$

In the next section, we will estimate these approximations using data on individual stocks and their forecast errors.

4.1 Simulation

To assess whether the approach outlined above gives us any additional insights, we build a bivariate process used in the skill /performance literature. We suppose that the data-generating process describes the conditional distribution pdf(y | x) by the following equation:

$$y_t = \beta_1 x_t + \beta_2(x_t^2 - 1) + \beta_3 x_t^3 + v_t \ . \qquad (6)$$

Such a model is called a cubic market model and variations of this specification have been suggested by Kraus and Litzenburger (1976) and Barone-Adesi (1985) and many authors subsequently. Here we assume that x_t, the forecast, is N(0,1) and v_t, the error, is N(0, σ_v^2); and that the forecast and error are independent of each other. Equation (6) has been structured to have a mean of zero. To calibrate the population variance of the return to be 1, we re-scale our initial parameters with the constant defined by

$c=(\beta_1^2 + 2\beta_2^2 + 15\beta_3^2 + 6\beta_1\beta_3 + \sigma_v^2)^{1/2}$.

Thus our two variables will both be standardised and their correlation will be $\beta_1 + 3\beta_3$. If we wish to make the correlation zero, we set this to zero and get the data-generating process:

$$y_t = \beta_2(x_t^2 - 1) + \beta_3(x_t^3 - 3x_t) + v_t \qquad (7)$$

Interestingly, equation (7) is written in terms of orthogonal Hermite polynomials, of degrees 2 and 3 respectively. Coming back to the general case, it is now possible, although very tedious, to calculate all the terms that occur in Proposition 3. For the uncorrelated case, we have:

$$E(y_t x_t^3) = 6\beta_3$$

$$E(x_t y_t^3) = 144\beta_2^2\beta_3 + 324\beta_3^3.$$

It is apparent from the above that the behaviour of H for the case of corollary 3 is that a hit rate in excess of 50% requires β_3 to be negative when the IC of the forecaster is zero.

To carry out our simulation, we start with estimated cubic market model values from Hwang and Satchell (2001, page 110) based on the use of the cubic model in emerging markets, and take the example of Columbia with parameters suitably re-calibrated as discussed in section 2. The interpretation here is that we are using the US market to forecast the emerging market return although the data are contemporaneous. We set σ_v^2 so that the squared coefficient of multiple determination is 27.33%. We chose this example because the relative magnitude of the initial β_3 to β_1 is large and both are positive so that some non-normal effects enter into the data. The final values are $\beta_1 = 0.2724$, $\beta_2 = 0.1885$, $\beta_3 = 0.0528$ *and* $\sigma_v = 0.8524$. This leads to an IC=ρ=.4308. We see that there is clear evidence of additional benefit from the non-linear terms as the coefficient of multiple determination is 0.5231.

To calculate the approximation for H, as given in (4), we need expressions for the moments of x_t and y_t along with the specified cross moments. The required moments are given in the following proposition.

<u>Proposition 4</u>

For the model given in (6) with x_t, the forecast, N(0,1) and v_t, the error, N(0, σ_v^2) along with independence between x_t and v_t, we have the following moments in terms of the underlying parameters:

$$E(x_t^4)+E(y_t^4)-6 = 3\beta_1^4 + 60\beta_2^4 + 10395\beta_3^4 + 60\beta_1^3\beta_3 + 60\beta_1^2\beta_2^2 + 630\beta_1^2\beta_3^2 + 3780\beta_1\beta_3^3 + 936\beta_1\beta_2^2\beta_3$$

$$+4500\beta_2^2\beta_3^2 + 6\sigma_v^2(\beta_1^2 + 2\beta_2^2 + 15\beta_3^2 + 6\beta_1\beta_3) + 3(\sigma_v^4 - 1)$$

$$E(x_t y_t^3)+E(x_t^3 y_t) - 6\rho = 3\beta_1^3 + 945\beta_3^3 + 45\beta_1^2\beta_3 + 30\beta_1\beta_2^2 + 234\beta_2^2\beta_3 + 315\beta_1\beta_3^2$$

$$+3(\beta_1 + 3\beta_3)(\sigma_v^2 - 2) + 3(\beta_1 + 5\beta_3)$$

$$E(x_t^2 y_t^2) - 2\rho^2 - 1 = \beta_1^2 + 10\beta_2^2 + 87\beta_3^2 + 18\beta_1\beta_3 + \sigma_v^2 - 1.$$

Proof: See Appendix.

Using our approximation, we find a hit rate, H, of 0.6061 while our simulation based on equation (6) and the above values with a million replications gives a true hit rate, H, of 0.6196, indicating that, for this example, our approximations works extremely well. These hit rate values when compared with that under joint normality of 0.6417 indicate that the non-normality reduces the hit rate. Further analysis reveals that the individual contributions for each of the three terms are given by: 0.0281254, -0.0943828 and 0.03065759, respectively. That is, the contribution from the marginal kurtosis of x and y is positive and equals 0.0281258, that from the odd cross moment kurtosis is negative and equals -0.0943828 and that from the even cross moment kurtosis is positive and equals 0.03065759. A comparison of moments between the true and simulated is given in the following table and illustrates the accuracy of the simulation.

Table 1

Moments of Distribution

	True	**Simulated**
Mean	0.0	0.00012
Standard Deviation	1.0	1.00045
Skewness	0.4745	0.49536
Kurtosis	4.3765	4.58743

4.2 Empirical application

In this section we consider two empirical applications. The first is an application of the model used in the simulation where the betas are estimated by regression and the hit rate and information coefficient estimated accordingly. The second application examines the forecast error and the accuracy of the Edgeworth expansion when used to calculate percentiles of the empirical distribution.

We obtain analyst report data from I/B/E/S, and stock price and return data from CRSP. Our target price sample is obtained from I/B/E/S's detailed (stock-split adjusted) price target file. "Price Target" represents the projected price level forecast by the analyst within a specific time horizon. We use the "cumulative factor to adjust stock prices" routine on CRSP to convert realized stock prices to the same-split adjusted basis so all stock prices used in our analysis are based on the same-split adjusted basis. We retain target prices issued for our sample firms by identifiable analysts spanning the calendar years 1999-2010 that meet certain criteria. First, they must be '12-month-ahead' target prices. Analysts occasionally provide target prices for other time-horizons but the vast majority are for 12 months ahead, and we seek to evaluate all target price forecasts over the same forecast horizon. Second, to determine target price performance, we require the share price 12 months after the target price release date. Third, we require the firm's closing share price three trading days prior to the release date of target price forecast.

Calculation process:

Changes in target price for a covered stock are calculated as *(TP/P)-1*, where *TP* is the target price forecast, and P is the closing price three trading days prior to the target price release date. Changes in realized price for a covered stock are calculated as *(P12/P)-1*, where *P12* is the stock price 12-months following the target price release date. Changes in market consensus for a covered stock are calculated as *(CONS/P)-1*, where the market consensus is the average target price for the same stock in the same month by different analysts.

We use a dummy indicator variable, *hit*, equal to one if the changes in target price (or changes in market consensus) have the same sign (+ or –) with the changes in realized closing price as of the end of the 12-month forecast horizon. So the hit rate is calculated as *hit rate = number of hits/ number of forecasts*.

The betas are estimated from the following equation:

$$y_t = \beta_1 x_t + \beta_2 (x_t^2 - 1) + \beta_3 x_t^3 + v_t$$

where y_t represents changes in realized price, and x_t is changes in target price or changes in market consensus. Changes are log changes and thus interpretable as returns. Both y_t and x_t are normalized using the sample mean and standard deviation of each stock.

Table 2

Market Model Estimates, Hit Rate and Information Coefficient

Company Ticker	Beta 1	Beta 2	Beta 3	Sigma	Hit Rate	Hit Rate (Approx)	IC (Approx)
AA	1.022	0.086	-0.032	0.347	0.7121	0.8316	0.7485
AXP	2.237	-0.763	0.097	0.207	0.9545	1.0431	0.8482
DD	1.091	0.107	-0.094	0.332	0.6543	0.8649	0.7061
GE	0.683	-0.112	-0.023	0.280	0.9556	0.7823	0.6686
IBM	1.076	-0.021	-0.015	0.042	0.8381	0.9425	0.9110
JPM	0.322	-0.160	0.136	0.584	0.7075	0.7066	0.5846
MCD	0.681	0.349	0.154	0.602	0.8940	0.8593	0.6750
PFE	0.873	-0.241	-0.074	0.440	0.9750	0.7964	0.5845
T	0.654	0.368	0.053	0.340	0.9180	0.7939	0.6406
XOM	0.893	-0.203	-0.084	0.502	0.8515	0.7867	0.5716

Note: Columns two to five show estimates from the non-linear market model in equation (6). Column six shows the sample hit rate and column seven is the hit rate calculated from equation (4). The approximated IC from equation (7) is shown in column (8).

In Table 2, we report some calculations based on forecasts and actual returns for 10 US firms chosen from components of the Dow Jones Index for the purposes of illustration. Columns 2 to 4 refer to the estimated betas from equation (6) in their unnormalised form. Column (5) reports the estimated residual standard deviation of equation (6). Armed with these estimates, we can deduce the hit rate from equation (4) and proposition (4). The result of these computations are presented in column (6) and can be compared with the sample hit rates in column (5). The comparison shows a close match in some cases but also, shows, as can happen when using Edgeworth models, an approximate hit rate greater than one for the case of AXP. The approximate IC corresponds to an estimated value of $\beta_1 + 3\beta_3$ using normalised values as discussed above equation (7).

Table 3

Sample IC and Number of Targets

Company Ticker	IC	No. of Targets
AA	.9143	609
AXP	.8879	617
DD	.9194	476
GE	.9934	603
IBM	.9748	586
JPM	.8742	566
MCD	.9963	480
PFE	.9943	573
T	.9873	1746
XOM	.9477	566

It should be noted that the results are not especially interesting as forecasting ICs here seems extremely and implausibly high,

For the empirical analysis of the forecast error we use the same stock data and define our forecast error as FE_{t+1} = ($\ln P_{F,t+1}$ – $\ln P_t$) - ($\ln P_{t+1}$ – $\ln P_t$) = $\ln P_{F,t+1}$ - $\ln P_{t+1}$ =lnX – lnY where X is the forecast target price and Y the actual. The cumulants for FE_{t+1} are calculated for each stock and using these sample values along with the formulas developed in Section 4 we can develop an approximate cumulative distribution function for each stock. The stocks considered in the Table below are only those whose skewness and kurtosis ensure a positive density. Then using the empirically calculated percentiles for each stock we can easily compare the accuracy of our Edgeworth approximation by comparing areas under our approximate distribution with those from the data. The results of such a calculation are given in the following table.

Table 4

Empirical and Approximated Density Function

Dow Jones Index Components

Company			Cumulative Density				
Ticker	0.010	0.025	0.050	0.950	0.975	0.990	Distribution
AXP	0.00461	0.01289	0.03027	0.96175	0.98822	0.99450	N
	0.01170	0.02209	0.03787	0.96635	0.98672	0.99211	E
BA	0.02486	0.05211	0.08856	0.96546	0.98032	0.98491	N
	0.01142	0.04120	0.08444	0.95405	0.97196	0.97789	E
BAC	0.05991	0.07565	0.09519	0.96216	0.99195	0.99578	N
	0.04524	0.06306	0.08581	0.94625	0.98322	0.98982	E
CSCO	0.02366	0.03587	0.06521	0.98199	0.99413	0.99829	N
	0.01017	0.01860	0.04332	0.96173	0.97987	0.99080	E
CVX	0.00336	0.00929	0.01580	0.94098	0.96053	0.97670	N
	0.01366	0.02587	0.03542	0.96260	0.98004	0.99213	E
GE	0.00578	0.01593	0.02867	0.93991	0.95545	0.97449	N
	0.01784	0.03341	0.04741	0.95918	0.97345	0.98878	E
HPQ	0.02422	0.04469	0.07286	0.96470	0.97244	0.98334	N
	0.01545	0.03892	0.07255	0.95954	0.96882	0.98208	E
IBM	0.01885	0.04985	0.07015	0.95415	0.98476	0.99349	N
	0.01156	0.04204	0.06349	0.94605	0.97879	0.98965	E
INTC	0.01109	0.02742	0.04890	0.95587	0.98591	0.99351	N
	0.01933	0.03348	0.04955	0.95908	0.98166	0.98862	E
KFT	0.01456	0.03623	0.07238	0.96547	0.97795	0.99140	N
	0.00778	0.02875	0.06742	0.95861	0.97227	0.98811	E
KO	0.03364	0.05596	0.07802	0.96657	0.98013	0.98948	N
	0.02413	0.04826	0.07303	0.95825	0.97372	0.98520	E
MCD	0.01056	0.06489	0.11192	0.97991	0.99558	0.99957	N
	0.00749	0.02718	0.06347	0.94482	0.97137	0.99306	E
MRK	0.01267	0.02312	0.04255	0.96084	0.97500	0.98162	N
	0.01665	0.02835	0.04871	0.96703	0.98108	0.98730	E
T	0.01898	0.02428	0.03338	0.95961	0.99073	0.99655	N
	0.01942	0.02411	0.03215	0.95732	0.98663	0.99356	E
UTX	0.01620	0.03359	0.04948	0.94925	0.96829	0.97851	N
	0.01636	0.03752	0.05645	0.95298	0.97559	0.98711	E
VZ	0.02174	0.05206	0.07038	0.95936	0.97527	0.99663	N
	0.01282	0.04190	0.06098	0.94914	0.96556	0.99267	E
XOM	0.00244	0.00687	0.05548	0.92550	0.95995	0.97367	N
	0.00849	0.01660	0.06276	0.93956	0.96967	0.98049	E

From the above table we note that the rows denoted with N at the far right give the areas calculated using just a normal distribution. Those with an E are the Edgeworth approximations, which take into account corrections for skewness and kurtosis. It is hard to draw general conclusions but what we do note is that for many stocks the Edgeworth gives a closer approximation.

5. Conclusions and Future Research

In this paper we derive a general expression, based on Edgeworth expansions involving skewness and kurtosis terms, for the joint pdf of forecasts and outcomes. We apply our approach to calculate properties of measures of skill such as the IC and the hit rate; we also investigate the distribution of the forecast error. Our analysis uses data on analysts' forecasts of target prices for large US companies.

Our results show the intuitively attractive link between the hit rate and IC, which holds under normality whereby a hit rate in excess of 50% occurs if and only if IC is positive, breaks down when we allow our distribution to be of a more general kind.

Higher moments impact on the hit rate in different ways. In particular, increases in higher power even cross-moments, $\mathrm{E}(x^2y^2)$, increase the hit rate, whilst univariate kurtosis and higher-power odd cross-moments decrease the hit rate.

Our empirical and simulation results do not provide strong evidence that our approach will automatically lead to more accurate forecasts. More research is needed to create practical forecasting tools based on this model. However, it does give us additional insights into the properties of forecasts and certainly suggests that the assumption of bivariate normality is overly simplistic.

References

Anatolyev, S. (2009) “Non Parametric Retrospection and Monitoring of the Predictability of Financial Returns”, *Journal of Business and Economic Statistics*, 27, pp 149-160.

Asquith, P, Mikhail, M. B, and A.S. Au (2005) “Information Content of Equity Analyst *Reports”, Journal of Financial Economics, 75, pp245-282.*

Barone-Adesi, G. (1985)”Arbitrage Equilibrium with Skewed Asset Returns”, *Journal of Financial and Quantitative Analysis”20,299-313.*

Bradshaw, M.T. and L.D. Brown (2007) “*Do Sell-Side Analysts Exhibit Differential Target Price Forecasting Ability*”, Working paper, Harvard Business School.

Dechow, P.M. and H. You (2013), “Understanding and Predicting Target Price Valuation Errors”, Working paper, University of California, Berkeley.

Diebold, F, X, P. Christoffersen, R. Mariano, A. Tay and Y.K. Tse (2007) “Direction-of-Change Forecasts Based on Conditional Variance, Skewness and Kurtosis Dynamics: International Evidence," *Journal of Financial Forecasting*, 1(2), pp 3-24.

Cowles, A. (1933)“Can Stock Markets Forecasters Forecast?”,*Econometrica,*1,pp309-324.

Ehrbeck, T. and Waldmann, R. (1996) “Why are Professional Forecasters Biased? Agency Versus Behavioural Explanations”, *Quarterly Journal of Economics*, 111, pp 21-40.

Garratt, A., Koop, G., Mise, E. and S.P. Vahey (2009) “Real-time Prediction with U.K. Monetary Aggregates in the Presence of Model Uncertainty”, *Journal of Business and Economic Statistics*, 27, pp 480-491.

Grinold, R.C. and R.N. Kahn (2000) *Active Portfolio Management*, McGraw-Hill, New York.

Henriksson, R.D. and R.C. Merton (1981) "On Market Timing and Investment Performance. II. Statistical Procedures for Evaluating Forecasting Skills", *Journal of Business*, 54, pp 513-533.

Hwang,S. And S.E.Satchell (2001)"Modelling Emerging Markets Risk Premia using Higher Moments" in *Return Distributions in Finance*(editors J.Knight and S.E.Satchell)Butterworth and Heinemann, Oxford.

Keane, M.P. and D.E. Runkle (1990) "Testing the Rationality of Price Forecasts: New Evidence from Panel Data", *American Economic Review*, 80, pp 714-735.

Kraus,A. And R.H.Litzenberger (1976) "Skewness Preference and the Valuation of Risky Assets"," *Journal of Finance"*,31,1085-1100.*l*

Pesaran, M.H. and A. Timmermann. (1992) 'A Simple Nonparametric Test of Predictive Performance", *Journal of Business and Economic Statistics*, 10, pp 461-465.

Sheppard W.F.(1900) On the calculation of the double integral expressing normal correlation. *Transactions of the Cambridge Philosophical Society*, 19, 23-66

Zhang L, P.A.Mykland, and Yacine A¨ıt-Sahalia (2009) "Edgeworth Expansions for Realized Volatility and Related Estimators", mimeo.

Appendix

Proof of Proposition 3

For the univariate distribution functions $F_1(x)$ and $F_2(y)$ the expansions are quite straightforward and can be found in most standard statistics books. The expansion can be expressed generically by the following:

$$F(z) = F_0(z) - \frac{1}{3!}K_3F_3 + \frac{1}{4!}K_4F_4$$

where

$$F_0(z) = \int_{-\infty}^{z} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{w-\mu}{\sigma}\right)^2\right) dw \ ,$$

With:

$$F_n = \frac{\partial^n}{\partial z^n} F_0(z)$$

and K_3 and K_4 are the third and fourth order cumulants given by:

$K_3 = E(z - \mu)^3$

$K_4 = E(z - \mu)^4 - 3\sigma^4$

with $\mu = E(z)$ and $\sigma^2 = Var(z)$, as:

$$F_0(z) = \int_{-\infty}^{\left(\frac{z-\mu}{\sigma}\right)} \phi(w)\,dw$$

$$= \Phi\left(\frac{z-\mu}{\sigma}\right).$$

The expansion is usually given in the standardised form:

$$F(z) = \Phi\left(\frac{z-\mu}{\sigma}\right) - \left[\frac{1}{3!}\frac{K_3}{\sigma^3}H_2\left(\frac{z-\mu}{\sigma}\right) + \frac{1}{4!}\frac{K_4}{\sigma^4}H_3\left(\frac{z-\mu}{\sigma}\right)\right]\varphi\left(\frac{z-\mu}{\sigma}\right)$$

where $H_r(\cdot)$ is the r^{th} Hermite polynomial defined by the identity:

$$(-1)^r \frac{d^r}{dx^r}\varphi(x) = H_r(x)\varphi(x)\,.$$

For the bivariate case, the analysis becomes more complicated due to the bivariate cumulants and all the third- and fourth-order derivatives of the joint distribution, that is:

$$F(x,y) = F_0(x,y) - \frac{1}{3!}\left[K^{1,1,1}\frac{\partial^3 F}{\partial x^3} + K^{2,2,2}\frac{\partial^3 F}{\partial y^3} + 3K^{1,1,2}\frac{\partial^3 F}{\partial x^2 \partial y} + 3K^{1,2,2}\frac{\partial^3 F}{\partial x \partial y^2}\right]$$

$$+\frac{1}{4!}\left[K^{1,1,1,1}\frac{\partial^4 F}{\partial x^4} + K^{2,2,2,2}\frac{\partial^4 F}{\partial y^4} + 4K^{1,2,2,2}\frac{\partial^4 F}{\partial x \partial y^3} + 4K^{2,1,1,1}\frac{\partial^4 F}{\partial x^3 \partial y} + 6K^{1,1,2,2}\frac{\partial^4 F}{\partial x^2 \partial y^2}\right]$$

where $K^{i,j,k}$ etc. are the bivariate cumulants and the derivatives are taken with respect to $F_0(x,y)$, defined below. In particular, from McCullagh (1987, Ch. 2) we have:

$$K^{i,j,k} = \frac{\partial^3}{\partial t_i \partial t_j \partial t_k}\ln M(t_1,t_2)\Big|_{\substack{t_1=0\\ t_2=0}}$$

$$K^{i,j,k,l} = \frac{\partial^4}{\partial t_i \partial t_j \partial t_k \partial t_l}\ln M(t_1,t_2)\Big|_{\substack{t_1=0\\ t_2=0}}$$

with

$$M(t_1,t_2) = E[\exp(t_1 x + t_2 y)]$$

and

$$F_0(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} \left(2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}\right)^{-1} \exp\left(-\frac{1}{2(1-\rho^2)}\left[\left(\frac{w_1-\mu_x}{\sigma_x}\right)^2 + \left(\frac{w_2-\mu_y}{\sigma_y}\right)^2 - 2\rho\left(\frac{w_1-\mu_x}{\sigma_x}\right)\left(\frac{w_2-\mu_y}{\sigma_y}\right)\right]\right) dw_2\, dw_1 .$$

As in the univariate case, we can standardise the expansion by noting that:

$$F_0(x,y) = \int_{-\infty}^{\frac{x-\mu_x}{\sigma_x}}\int_{-\infty}^{\frac{y-\mu_y}{\sigma_y}} \phi(w,u;\rho)\, du\, dw$$

$$= \Phi\left(\frac{x-\mu_x}{\sigma_x}, \frac{y-\mu_y}{\sigma_y}; \rho\right).$$

Consequently:

$$\frac{\partial^{i+j} F_0(x,y)}{\partial x^i \partial y^j} = \frac{1}{\sigma_x^i \sigma_y^j} \frac{\partial^{i+j}}{\partial x_1^i \partial x_2^j} \Phi(x_1,x_2;\rho) \Bigg|_{x_1 = \frac{x-\mu_x}{\sigma_x},\ x_2 = \frac{y-\mu_y}{\sigma_y}}$$

Further, noting that:

$$\Phi(x_1,x_2;\rho) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} \left(2\pi\sqrt{1-\rho^2}\right)^{-1} \exp\left(-\frac{1}{2(1-\rho^2)} (w_1^2 + w_2^2 - 2\rho w_1 w_2)\right) dw_2\, dw_1 :$$

$$=\int_{-\infty}^{x_1}\frac{1}{\sqrt{2\pi}}e^{-w_1^2/2}\int_{-\infty}^{x_2}\left(2\pi\sqrt{1-\rho^2}\right)^{-1}\exp\left(-\frac{1}{2(1-\rho^2)}(w_2-\rho w_1)^2\right)dw_2\,dw_1$$

$$=\int_{-\infty}^{x_2}\frac{1}{\sqrt{2\pi}}e^{-w_2^2/2}\int_{-\infty}^{x_1}\left(2\pi\sqrt{1-\rho^2}\right)^{-1}\exp\left(-\frac{1}{2(1-\rho^2)}(w_1-\rho w_2)^2\right)dw_1\,dw_2$$

all the derivatives are readily calculated. After some tedious but straightforward calculations (see the Appendix), we find that:

$$\frac{\partial^3 F_0}{\partial x^3}(x,y)=\frac{1}{\sigma_x^3}\left(D_1^{1,1,1}\frac{\partial\Phi}{\partial x_1}+D_2^{1,1,1}\frac{\partial^2\Phi}{\partial x_1\partial x_2}\right)$$

$$\frac{\partial^3 F_0}{\partial y^3}(x,y)=\frac{1}{\sigma_y^3}\left(D_1^{2,2,2}\frac{\partial\Phi}{\partial x_2}+D_2^{2,2,2}\frac{\partial^2\Phi}{\partial x_1\partial x_2}\right)$$

$$\frac{\partial^3 F_0(x,y)}{\partial x_1\partial y^2}=\frac{1}{\sigma_x\sigma_y^2}D_2^{1,2,2}\frac{\partial^2\Phi}{\partial x_1\partial x_2}$$

$$\frac{\partial^3 F_0(x,y)}{\partial x_1^2\partial y}=\frac{1}{\sigma_x^2\sigma_y}D_2^{1,1,2}\frac{\partial^2\Phi}{\partial x_1\partial x_2}$$

$$\frac{\partial^4 F_0}{\partial x^4}(x,y)=\frac{1}{\sigma_x^4}\left[D_1^{1,1,1,1}\frac{\partial\Phi}{\partial x_1}+D_2^{1,1,1,1}\frac{\partial^2\Phi}{\partial x_1\partial x_2}\right]$$

$$\frac{\partial^4 F_0}{\partial y^4}(x,y)=\frac{1}{\sigma_y^4}\left[D_1^{2,2,2,2}\frac{\partial\Phi}{\partial x_2}+D_2^{2,2,2,2}\frac{\partial^2\Phi}{\partial x_1\partial x_2}\right]$$

$$\frac{\partial^4 F_0}{\partial x_1\partial y^3}(x,y)=\frac{1}{\sigma_x\sigma_y^3}\left[D_2^{1,2,2,2}\frac{\partial^2\Phi}{\partial x_1\partial x_2}\right]$$

$$\frac{\partial^4 F_0}{\partial x_1^3\partial y}(x,y)=\frac{1}{\sigma_x^3\sigma_y}\left[D_2^{1,1,1,2}\frac{\partial^2\Phi}{\partial x_1\partial x_2}\right]$$

$$\frac{\partial^4 F_0}{\partial x_1^2\partial_y^2}(x,y)=\frac{1}{\sigma_x^2\sigma_y^2}\left[D_2^{1,1,2,2}\frac{\partial^2\Phi}{\partial x_1\partial x_2}\right].$$

Consequently, using equation (4) and the above:

$$F(x_1,x_2)=\Phi(x_1,x_2,\rho)+$$

$$-\frac{1}{6}\left[\frac{K^{1,1,1}}{\sigma_{x_1}^{3}}D_1^{1,1,1}\frac{\partial\Phi}{\partial x_1}+\frac{K^{2,2,2}}{\sigma_{x_2}^{3}}D_1^{2,2,2}\frac{\partial\Phi}{\partial x_2}\right]$$

$$-\frac{1}{6}\left[\frac{K^{1,1,1}}{\sigma_{x_1}^{3}}D_2^{1,1,1}+\frac{K^{2,2,2}}{\partial_{x_2}^{3}}D_2^{2,2,2}+\frac{3K^{1,1,2}}{\sigma_{x_1}^{2}\sigma_{x_2}}D_2^{1,1,2}+\frac{3K^{1,2,2}}{\sigma_{x_2}^{2}\sigma_{x_1}^{3}}D_2^{1,2,2}\right]\frac{\partial^2\Phi}{\partial x_1\partial x_2}$$

$$+\frac{1}{24}\left[\frac{K^{1,1,1,1}}{\sigma_{x_1}^{4}}D_1^{1,1,1,1}\frac{\partial\Phi}{\partial x_1}+\frac{K^{2,2,2,2}}{\sigma_{x_2}^{4}}D_1^{2,2,2,2}\frac{\partial\Phi}{\partial x_2}\right]$$

$$+\frac{1}{24}\left[\frac{K^{1,1,1,1}}{\sigma_{x_1}^{4}}D_2^{1,1,1,1}+\frac{K^{2,2,2,2}}{\sigma_{x_2}^{4}}D_2^{2,2,2,2}+\frac{4K^{1,2,2,2}}{\sigma_{x_1}\sigma_{x_2}^{3}}D_2^{1,2,2,2}+\frac{4K^{2,1,1,1}}{\sigma_{x_1}^{3}\sigma_{x_2}^{4}}D_2^{2,1,1,1}\right.$$

$$\left.+\frac{6K^{1,1,2,2}}{\sigma_{x_1}^{2}\sigma_{x_2}^{2}}D_2^{1,1,2,2}\right]\frac{\partial^2\Phi}{\partial x_1\partial x_2} \qquad \text{(A.1)}$$

where:

$$D_1^{1,1,1}=H_2(x_1)$$

$$D_1^{2,2,2}=H_2(x_2)$$

$$D_1^{1,1,1,1}=-H_3(x_1)$$

$$D_1^{2,2,2,2}=-H_3(x_2)$$

$$D_2^{1,1,1}=\rho(H_1(x_1)+h_1(x_1,x_2)); \qquad (6)$$

$$D_2^{2,2,2}=\rho(H_1(x_2)+h_2(x_1,x_2));$$

$$D_2^{i,j,j}=-h_j(x_1,x_2)$$

$$D_2^{1,1,1,1}=\frac{\rho(2-\rho^2)}{1-\rho^2}-\rho H_2(x_1)-\rho h_1(x_1,x_2)(H_1(x_1)+h_1(x_1,x_2));$$

$$D_2^{2,2,2,2}=\frac{\rho(2-\rho^2)}{1-\rho^2}-\rho H_2(x_2)-\rho h_2(x_1,x_2)(H_1(x_2)+h_2(x_1,x_2));$$

$$D_2^{i,j,j,j} = -\frac{1}{1-\rho^2} + h_j^2(x_1, x_2)$$

$$D_2^{1,1,2,2} = \frac{\rho}{1-\rho^2} + h_1(x_1, x_2) h_2(x_1, x_2),$$

with $H_j(x_i)$ the standard Hermite polynomial and:

$$h_1(x_1, x_2) = \frac{x_1 - \rho x_2}{1-\rho^2}$$

$$h_2(x_1, x_2) = \frac{x_2 - \rho x_1}{1-\rho^2}.$$

Also:

$$\frac{\partial \Phi}{\partial x_1} = \varphi(x_1)\Phi\left(\frac{x_2 - \rho x_1}{\sqrt{1-\rho^2}}\right)$$

$$\frac{\partial \Phi}{\partial x_2} = \varphi(x_2)\Phi\left(\frac{x_1 - \rho x_2}{\sqrt{1-\rho^2}}\right)$$

$$\frac{\partial^2 \Phi}{\partial x_1 \partial x_2} = \varphi(x_1, x_2; \rho).$$

Further, noting that:

$K_3 = E(z - \mu)^3 = K^{1,1,1}$, for $z = x_1$

$= K^{2,2,2}$, for $z = x_2$

and

$K_4 = K^{1,1,1,1}$, for $z = x_1$

$= K^{2,2,2}$, for $z = x_2$,

we have:

$$F_1(x_1) = \Phi(x_1) - \frac{1}{6}\frac{K^{1,1,1}}{\sigma_{x_1}^3} H_2(x_1)\varphi(x_1) - \frac{1}{24}\frac{K^{1,1,1,1}}{\sigma_{x_1}^4} H_3(x_1)\varphi(x_1) \qquad \text{(A.2)}$$

$$F_2(x_2) = \Phi(x_2) - \frac{1}{6}\frac{K^{2,2,2}}{\sigma^3_{x_2}} H_2(x_2)\varphi(x_2) - \frac{1}{24}\frac{K^{2,2,2,2}}{\sigma^4_{x_2}} H_3(x_2)\varphi(x_2).$$

Combining the univariate and bivariate expansions from equations (A.1) and (A.2), we have:

$$H = 1 - \Phi(x_1) - \Phi(x_2) + 2\Phi(x_1, x_2; \rho) \qquad \text{(A.3)}$$

$$+\frac{1}{24}\left[\left(\frac{4K^{1,1,1}}{\sigma^3_{x_1}} H_2(x_1) + \frac{K^{1,1,1,1}}{\sigma^4_{x_1}} H_3(x_1)\right)\varphi(x_1)\left(1 - 2\Phi\left(\frac{x_2 - \rho x_1}{\sqrt{1-\rho^2}}\right)\right)\right.$$

$$\left.+\left(\frac{4K^{2,2,2}}{\sigma^3_{x_2}} H_2(x_2) + \frac{K^{2,2,2,2}}{\sigma^4_{x_2}} H_3(x_2)\right)\varphi(x_2)\left(1 - 2\Phi\left(\frac{x_1 - \rho x_2}{\sqrt{1-\rho^2}}\right)\right)\right]$$

$$-\frac{1}{3}\left[\left(\frac{K^{1,1,1}}{\sigma^3_{x_1}} D_2^{1,1,1} + \frac{K^{2,2,2}}{\sigma^3_{x_2}} D_2^{2,2,2}\right) + 3\left(\frac{K^{1,1,2}}{\sigma^2_{x_1}\sigma_{x_2}} D_2^{1,1,2} + \frac{K^{1,2,2}}{\sigma_{x_1}\sigma^2_{x_2}} D_2^{1,2,2}\right)\right]\varphi(x_1, x_2; \rho)$$

$$+\frac{1}{12}\left[\frac{K^{1,1,1,1}}{\sigma^4_{x_1}} D_2^{1,1,1,1} + \frac{K^{2,2,2,2}}{\sigma^4_{x_2}} D_2^{2,2,2,2} + 4\left(\frac{K^{1,2,2,2}}{\sigma_{x_1}\sigma^3_{x_2}} D_2^{1,2,2,2} + \frac{K^{2,1,1,1}}{\sigma^3_{x_1}\sigma_{x_2}} D_2^{2,1,1,1}\right)\right.$$

$$\left.+\frac{6K^{1,1,2,2}}{\sigma^2_{x_1}\sigma^2_{x_2}} D_2^{1,1,2,2}\right]\varphi(x_1, x_2; \rho).$$

Substituting for $D_2^{i,j,k}$ and $D_2^{i,j,k,l}$ in equation (A.3), we have:

$$H = 1 - \Phi(x_1) - \Phi(x_2) + 2\Phi(x_1, x_2; \rho) \qquad \text{(A.4)}$$

$$+\frac{1}{24}\left[\left(\frac{4K^{1,1,1}}{\sigma^3_{x_1}} H_2(x_1) + \frac{K^{1,1,1,1}}{\sigma^4_{x_1}} H_3(x_1)\right)\varphi(x_1)\left(1 - \Phi\left(\frac{x_2 - \rho x_1}{\sqrt{1-\rho^2}}\right)\right)\right.$$

$$+\left(\frac{4K^{2,2,2}}{\sigma_{x_2}^3}H_2(x_2)+\frac{K^{2,2,2,2}}{\sigma_{x_2}^4}H_3(x_2)\right)\varphi(x_2)\left(1-\Phi\left(\frac{x_1-\rho x_2}{\sqrt{1-\rho^2}}\right)\right)$$

$$+\frac{1}{12}\left[-4\left(\frac{K^{1,1,1}}{\sigma_{x_1}^3}(\rho(H_1(x_1)+h_1(x_1,x_2)))+\frac{K^{2,2,2}}{\sigma_{x_2}^3}(\rho(H_1(x_2)+h_2(x_1,x_2)))\right)\right.$$

$$+12\left(\frac{K^{1,1,2}}{\sigma_{x_1}^2\sigma_{x_2}}h_1(x_1,x_2)+\frac{K^{1,2,2}}{\sigma_{x_1}\sigma_{x_2}^2}h_2(x_1,x_2)\right)$$

$$+\left(\frac{K^{1,1,1,1}}{\sigma_{x_1}^4}+\frac{K^{2,2,2,2}}{\sigma_{x_2}^4}\right)\frac{\rho(2-\rho^2)}{1-\rho^2}-\frac{\rho}{1-\rho^2}4\left(\frac{K^{1,2,2,2}}{\sigma_{x_1}\sigma_{x_2}^3}+\frac{K^{2,1,1,1}}{\sigma_{x_1}^3\sigma_{x_2}^4}\right)$$

$$+\frac{6\rho}{1-\rho^2}\frac{K^{1,1,2,2}}{\sigma_{x_1}^2\sigma_{x_2}^2}+\frac{K^{1,1,1,1}}{\sigma_{x_1}^4}\left(-\rho H_2(x_1)-\rho h_1(x_1,x_2)(H_1(x_1)+h_1(x_1,x_2))\right)$$

$$+\frac{K^{2,2,2,2}}{\sigma_{x_2}^4}\left(-\rho H_2(x_2)-\rho h_2(x_1,x_2)(H_1(x_2)+h_2(x_1,x_2))\right)$$

$$+4\left(\frac{K^{1,2,2,2}}{\sigma_{x_1}\sigma_{x_2}^3}h_2^2(x_1,x_2)+\frac{K^{2,1,1,1}}{\sigma_{x_1}^3\sigma_{x_2}^3}h_1^2(x_1,x_2)\right)$$

$$\left.+\frac{6K^{1,1,2,2}}{\sigma_{x_1}^2\sigma_{x_2}^2}h_1(x_1,x_2)h_2(x_1,x_2)\right]\varphi(x_1,x_2;\rho)\,.$$

The above expression for H given by equation (9) needs to be evaluated for specific values of x and y. Recalling that $x_1 = \frac{x - \mu_x}{\sigma_x}$ and $x_2 = \frac{y - \mu_y}{\sigma_y}$, when x = y = 0, the expression needs to be evaluated at $x_1 = \frac{-\mu_x}{\sigma_x}$ and $x_2 = \frac{-\mu_y}{\sigma_y}$, the Sharpe ratios of both x and y. ■

Proof of Proposition 4

Since x_t is N(0,1), we have $E(x_t^{2k}) = \frac{(2k)!}{2^k k!}$ and $E(x_t^{2k-1}) = 0$ for k ≥ 1, and using this along with the following, we can readily verify the expressions given in the proposition:

$$E(y_t^4) = E([\beta_1 x_t + \beta_2 (x_t^2 - 1) + \beta_3 x_t^3]^4) + 6\sigma_v^2 E([\beta_1 x_t + \beta_2 (x_t^2 - 1) + \beta_3 x_t^3]^2) + 3\sigma_v^4$$

$$E(x_t y_t^3) = E(x_t[\beta_1 x_t + \beta_2 (x_t^2 - 1) + \beta_3 x_t^3]^3) + 3\sigma_v^2 E(x_t[\beta_1 x_t + \beta_2 (x_t^2 - 1) + \beta_3 x_t^3])$$

$$E(x_t^3 y_t) = E(x_t^3[\beta_1 x_t + \beta_2 (x_t^2 - 1) + \beta_3 x_t^3])$$

$$E(x_t^2 y_t^2) = E(x_t^2[\beta_1 x_t + \beta_2 (x_t^2 - 1) + \beta_3 x_t^3]^2) + \sigma_v^2 .$$

■