CONTRIBUTION N° 13

ESTIMATING RETURNS ON FINANCIAL INSTRUMENTS - STOCHASTIC ANALYSIS

PAR / BY

Mark TIPPETT

Australie / Australia

RENATBILITÉ ESTIMÉE DES INSTRUMENTS FINANCIERS - ANALYSE STOCHASTIQUE
L'article traite de l'estimation de la rentabilité des instruments financiers dans le cas où l'on dispose que d'informations incomplètes sur les cash-flows de ces instruments. Toutefois, à la différence des travaux antérieurs dans ce domaine, on suppose que les cash-flows des instruments financiers peuvent être modélisés comme une fonction stochastique du temps, ce qui implique que l'erreur d'estimation est une variable aléatoire, dont on peut donc faire des évaluations probabilistes de l'amplitude. On montre, en utilisant ces procédures, que les méthodes d'estimation traditionnelles surestiment considérablement les rendements réels. On montre en outre que la distribution de ces rendements est très éloignée de la loi normale. Considérant ces résultats, il est douteux que l'on puisse attacher une crédibilité aux deux méthodes standard d'évaluation des performances d'investissement.
ABSTRACT:

The present paper addresses the problem of estimating returns on financial instruments when there is incomplete information about the instrument's cash flows. However, unlike previous work in this area, it is assumed that the financial instrument's cash flows may be modelled as a stochastic function of time. This implies that the estimation error is a random variable and as a consequence, probability assessments can be made as to its likely magnitude. Using these procedures it is shown that traditional estimating methods considerably overstate actual returns. Further, it is also shown that returns are far from normally distributed. Given these results, it is doubtful if any credibility can be attached to the standard two parameter method of evaluating investment performance.

ACKNOWLEDGEMENT:

The author is indebted to John Bowers for constructive criticism of an earlier draft of the paper and to Des Nicholls and John Morgan of the Department of Statistics at the Australian National University for confirming the correctness of the paper's mathematics. The programming assistance of Kathy MacLaren is also gratefully acknowledged.

1 - INTRODUCTION

In a previous paper, a co-author and I outlined numerical methods for estimating returns on financial instruments when there is incomplete information about the instrument's cash flows (Kelly and Tippett [1989]). It will be recalled that the pioneering work in this area was conducted by the Bank Administration Institute (1968), Dietz (1972) and Eadie (1973) and is based on the assumption that the financial instrument's cash flows may be represented by a smooth deterministic function of time. The implication of this is that returns should be estimated using some form of numerical technique. Unfortunately this approach makes it difficult to obtain concrete estimates of the likely error associated with the estimated return. In the present paper we overcome this difficulty by making the alternative assumption that the financial instrument's cash flows are a stochastic function of time. This approach implies that the return itself is a random variable and, as a consequence, it is possible to make probability assessments as to the magnitude of the estimating technique's error.

The paper begins in section two, by summarizing the error analysis of Kelly and Tippett (1989) since this provides the cornerstone of the stochastic analysis undertaken in later sections of the paper. In section three we then turn to the problems associated with determining an appropriate stochastic process through which to model a financial instrument's cash flows. It turns out that a "bridge" process is the most logical candidate.
After outlining its more important properties, the "bridge" process is then applied to the problem of obtaining probability bounds for the financial instrument’s return. The techniques are then illustrated in an empirical context by re-examining the problem of estimating the returns earned by Australian superannuation funds. Like McCrae and Tippett (1987) and Kelly and Tippett (1989), we find that the funds’ estimated returns are highly sensitive to the technique used to estimate the returns. Finally, section four concludes with some summary comments.

2. ERROR ANALYSIS

Suppose we define \( i^* \) to be the true (but unknown) money weighted return earned by a financial instrument over the time interval \((0,T)\). If we take \( i_a \) to be an approximation to this return, then Kelly and Tippett (1989, p.6) have shown that the error associated with the approximation is given by:

\[
i_a - i^* = \frac{F(i_a)}{F'(i_a)}
\]

where:

\[
F(i) = \int_0^T e^{-it}dC(t) + M(0) - M(T)e^{-iT}
\]

and:

\[
F'(i) = TM(T)e^{-iT} - \int_0^T e^{-it}dC(t)
\]

for \( dC(t) \) the instantaneous new cash flow (Holbrook [1977, p.24]) and \( M(t) \) the financial instrument's market value at time \( t \).

In the analysis of the Bank Administration Institute [1968], Dietz [1972], Eadie [1973] and Kelly and Tippett [1989], the cash flow function \([C(t)]\) is approximated by a smooth deterministic (a non-stochastic) function of time. This in turn implies that \( F(i) \) and \( F'(i) \) must also be approximated, since the in general unknown \( C(t) \) appears in the error expression (1). Kelly and Tippett [1989, p.67] suggest ways in which this may be done, but conclude with the rather uncertain prescription that their procedures "... should provide a reasonable estimate of the error ..." associated with their approximating methods. In short, Kelly and Tippett [1989] were unable to state with any assurance whether one return estimating technique is better than another.
3. STOCHASTIC ANALYSIS

3.1 Cash Flow Specification

As previously noted, the seminal work of the Bank Administration Institute [1968], Dietz [1972] and Eadie [1973] is based on the assumption that the instantaneous cash flow function \( C(t) \) is a deterministic (or non-stochastic) function of time. An alternative and probably far more realistic approach however, is to assume that the cash flow function is stochastic. The error expression (1) can then be viewed as a random variable, and assessments made as to its likely magnitude.

A problem with this approach however is that we seldom have information on an instrument's cash flows. In the assessment of superannuation fund performance for example, fund cash flows are collected on a monthly basis (McCrae and Tippett [1987, p 213]). Hence, whilst there is uncertainty about a fund's cash flows during a month, at the end of the month this uncertainty is resolved. However, since the fund's return hinges critically on the exact timing of its inter-month cash flows, it is still necessary to view the return as a random variable. But the fact that cash flows are known at the end of each month places a restriction on the type of stochastic process which can be used for modelling purposes.

Fortunately, there are standard ways of modelling such processes. To illustrate, suppose a fund's instantaneous cash flow is described by the following stochastic process:

\[
\frac{dC(t)}{b - t} = \left[ \frac{C(b) - C(t)}{b - t} \right] dt + dW(t) \quad (4)
\]

for \( a \leq t < b \) and \( dW(t) \) a white noise process with variance parameter \( \sigma^2 \). Occasionally \( dW(t) \) is given the alternative representation \( dW(t) = \sigma z \sqrt{dt} \) where \( z \) is distributed as a standard normal variate. Using this specification, which is due to Bernstein (Kannan [1979, p. 271]), it follows that:

\[
\frac{dC(t)}{b - t} = \left[ \frac{C(b) - C(t)}{b - t} \right] dt + \sigma dz \quad (5)
\]

where \( dz = z \sqrt{dt} \) is a normal variate with mean zero and variance \( dt \). Irrespective of which expression is used, it follows that the instantaneous mean cash flow is given by:

\[
E[dC(t)] = \left[ \frac{C(b) - C(t)}{b - t} \right] dt \quad (6)
\]

where \( E \) is the expectations operator taken at time \( t \). Note that as \( t \rightarrow b \), the expected instantaneous cash flow becomes infinitely positive or negative according to the sign of \( [C(b) - C(t)] \). This implies that \( C(t) \) is forced to take on a pre-specified value at time \( b \) and is therefore known with certainty at this point. Since \( C(t) \) is also known with certainty at time zero, this means the process is "tied down" at its two end points and for that reason is aptly described as a "bridge process" (Karlin and Taylor [1981, pp. 267-271]).
Similar procedures show that the variance of the instantaneous cash flow is given by:

\[ \text{Var}[dC(t)] = \frac{d^2}{dt^2} \tag{7} \]

where \( \text{Var}(.) \) is the variance of the relevant random variable.

Following Hoel, Port and Short (1972, pp. 152-159), the solution to the differential equation (4) is given by:

\[ C(t) = C(a) + \frac{[C(b) - C(a)](t - a)}{b - a} + (b - t) \int_a^t \frac{dW(s)}{b - s} \tag{8} \]

where \( C(t) \) is the accumulated cash flow as at time \( t \). Taking expectations across the above expression implies that:

\[ \text{E}[C(t)] = C(a) + \frac{[C(b) - C(a)](t - a)}{b - a} \tag{9} \]

is the expected accumulated cash flow at \( t \). Finally, using Hoel, Port and Stone (1972, pp. 144 - 145), the variance of \( C(t) \) is given by:

\[ \text{Var}[C(t)] = \sigma^2 t (b - t)^2 \int_a^t \frac{ds}{(b - s)^2} \tag{10} \]

\[ \text{Var}[C(t)] = \frac{\sigma^2 (b - t)(t - a)}{b - a} \]

Note that at times \( t = a \) and \( t = b \), the variance of \( C(t) \) is zero in accordance with the fact that \( C(t) \) is known at both these points. Hence, as noted previously, the stochastic process formalized by equation (4) is "tied down" at its end points and is aptly described as a bridge process.

3.3 Estimation of Returns

Since the error associated with the estimated return is given by equation (1), it follows that we need information on the distributional properties of both \( F(i) \) and \( F'(i) \). We begin by considering the integral expression in \( F(i) \). Integration by parts implies (Hoel, Port and Stone [1972, p. 142]):

\[ \int_a^b e^{-it}dC(t) = e^{-ib}C(b) - e^{-ia}C(a) + i \int_a^b e^{-it}C(t)\,dt \tag{11} \]
Substituting equation (8), it then follows:

\[
\int_{a}^{b} e^{-it} dC(t) = e^{-ib} C(b) - e^{-ia} C(a) + \int_{a}^{b} e^{-it} \left\{ \frac{C(b) - C(a)}{b - a} \right\} dt + i \int_{a}^{b} \int_{a}^{t} (b - t)e^{-it} dW(s) e^{-it} dt
\]

(12)

Using the procedures laid down in Doob [1942, pp. 364 - 365] and Apostol [1969, p. 373], it may be shown that the above expression simplifies to:

\[
\int_{a}^{b} e^{-it} dC(t) = \frac{C(b) - C(a)}{i(b - a)} \left[ e^{-ia} - e^{-ib} \right] + \int_{a}^{b} \left[ \int_{a}^{t} \left( b - t \right) e^{-it} + \frac{1}{i} \left( e^{-ib} - e^{-ia} \right) \right] dW(t)
\]

(13)

Taking expectations it follows that \( \int e^{-it} dC(t) \) is normally distributed with mean (Hoel, Port and Stone [1972, p. 133]):

\[
E\left[ \int_{a}^{b} e^{-it} dC(t) \right] = \frac{C(b) - C(a)}{i(b - a)} \left[ e^{-ia} - e^{-ib} \right]
\]

(14)

and variance (Hoel, Port and Stone [1972, pp. 144 - 145]):

\[
\text{Var}\left[ \int_{a}^{b} e^{-it} dC(t) \right] = \sigma^2 \int_{a}^{b} \left[ e^{-it} + \frac{\left( e^{-ib} - e^{-ia} \right)}{i(b - t)} \right]^2 dt
\]

(15)

Unfortunately, the above expression cannot be evaluated in terms of elementary functions and so the variance must be estimated using some form of numerical technique.

Similar analysis shows that:

\[
\int_{a}^{b} e^{-it} dC(t) = \frac{[C(b) - C(a)]}{i^2(b - a)} \left[ e^{-ia} - e^{-ib} \right] + \frac{[C(b) - C(a)]}{i(b - a)} \left[ ae^{-ia} - be^{-ib} \right]
\]

\[
+ \int_{a}^{b} \left[ \frac{te^{-it} - be^{-ib}}{i} \right] \left( \frac{e^{-it} - e^{-ib}}{i^2} \right) dW(t)
\]

(16)
from which it follows that $\int te^{-it}dC(t)$ is normally distributed with mean:

$$E[\int_t e^{-it}dC(t)] = \left[\frac{C(b) - C(a)}{i^2(b - a)}\right][e^{-ia} - e^{-ib}] + \frac{[C(b) - C(a)]}{i(b - a)}[ae^{-ia} - be^{-ib}]$$  \hspace{1cm} (17)

and variance:

$$Var[\int_t e^{-it}dC(t)] = \sigma^2 \int_a^b \left\{e^{-it} - \frac{be^{-ib}}{i(b - t)}\right\}^2 dt$$  \hspace{1cm} (18)

Again, the expression for the variance cannot be evaluated in terms of elementary functions and so must be estimated numerically. Finally, for the covariance we have:

$$\text{Cov}\left[\int_t e^{-it}dC(t), \int_t e^{-it}dC(t)\right] =$$  \hspace{1cm} (19)

$$\sigma^2 \int_a^b \left\{e^{-it} + \frac{(e^{-ib} - e^{-it})}{i(b - t)}\right\} \left\{e^{-it} - \frac{(te^{-it} - be^{-ib})}{i(b - t)}\right\}^2 dt$$

where Cov(.,.) is the covariance between the relevant random variables. This expression must also be evaluated numerically.

To illustrate the application of the above results, consider the problem of estimating returns on Australian superannuation funds. McCrae and Tippett [1987, p. 213] note that funds provide details of their cash flows on a monthly basis and on their asset values on a quarterly (calendar) basis. Hence, if we denote a calendar quarter by the interval $[0, 1]$, it follows that the fund's accumulated cash flow will be known at the time points $t_j = j/3$, for $j = 1, 2, 3$. Modelling the fund's cash flows on each of the intervals $((j - 1)/3, j/3), j = 1, 2, 3$, as a bridge process and using equation (2), it then follows that the quarterly return is defined implicitly by the for which:

$$F(i) = \sum_{j=1}^3 \int_{(j-1)/3}^{j/3} e^{-it}dC(t) + M(0) - M(1) e^{-i} = 0$$  \hspace{1cm} (20)

Since the market value of the fund's assets is known at the beginning and end of each quarter, it follows from equation (14) that $F(i)$ is normally distributed with the following mean:

As previously noted, McCrae and Tippett [1987, p. 215] estimated quarterly returns on Australian superannuation funds over the period 1974 - 1981 by determining the root of the above equation, a method they dubbed the linear technique. They concluded (p. 219) that estimated returns "... are sensitive to the quadrature rule utilized ...", but were unable to place a bound on the magnitude of the error from using a given estimating technique. However, taken in conjunction with equation (2), the stochastic specification given in this section can be used to derive the approximate probability density of the error in the estimated return. We begin by deriving the variance of \( F(i) \).

Using Hoel, Port and Stone (1972, pp. 144-145) the variance of \( F(i) \) is given by:

\[
\text{Var}[F(i)] = \text{Var}\left[ \sum_{j=1}^{3} \int_{(j-1)/3}^{j/3} e^{-it} dC(t) \right]
\]

which from equation (15) is seen to reduce to:

\[
\text{Var}[F(i)] = \sigma^2 \sum_{j=1}^{3} \int_{(j-1)/3}^{j/3} e^{-it} \left( \frac{\exp(-ij/3) - \exp(-it)}{i(j/3 - t)} \right)^2 dt
\]

Similarly from equations (3), (17) and (18), it follows that \( F'(i) \) will be normally distributed with mean:

\[
E[F'(i)] = TM(T)e^{-i} - E[\sum_{j=1}^{3} \int_{(j-1)/3}^{j/3} e^{-it} dC(t)]
\]

\[
E[F'(i)] = TM(T)e^{-i} - \frac{3}{i^2} \sum_{j=1}^{3} \{C(i/3) - C(i/3 - 1)} \{\exp[i(1 - j)] - \exp[-i/3]
\]

\[
- \frac{3}{i} \sum_{j=1}^{3} \{C(i/3) - C(i/3 - 1)} \{i - j} \exp[i(1 - j)] - \frac{i}{3} \exp[-ij/3]}
\]
and variance:

\[
\text{Var}[F'(i)] = \text{Var}\left[ \frac{1}{3} \sum_{j=1}^{3} \int t e^{-it} dC(t) \right]
\]

\[
\text{Var}[F'(i)] = \sigma^2 \sum_{j=1}^{3} \int \left\{ t e^{-it} \left[ \frac{\text{exp(-it)} - \frac{1}{3} \text{exp}(-\frac{ij}{3})}{i (\frac{i}{3} - t)} \right] \right\}^2 dt
\]

Finally from equation (19), the covariance between \( F(i) \) and \( F'(i) \) is given by:

\[
\text{Cov}[F(i), F'(i)] = - \text{Cov}\left[ \frac{1}{3} \sum_{j=1}^{3} \int t e^{-it} dC(t), \frac{1}{3} \sum_{j=1}^{3} \int t e^{-it} dC(t) \right]
\]

Suppose we estimate the return by determining root of equation (21). Then since \( F(i) \)
and \( F'(i) \) are limiting sums of normal variates it follows from Lindgren (1968, p. 473)
that the joint distribution of \( F(i) \) and \( F'(i) \) is also normal. To simplify matters, make
the change of variables (Freeman [1963, p. 155]):

\[
F(i) = \frac{F(i)}{\sqrt{\text{Var}[F(i)]}}, \quad R[F(i), F'(i)] F'(i) = \frac{R[F(i), F'(i)]}{\sqrt{\text{Var}[F'(i)]}}
\]

\[ u = \frac{F(i)}{\sqrt{\text{Var}[F(i)]}} - \frac{R[F(i), F'(i)]}{\sqrt{\text{Var}[F'(i)]}} \]

\[ v = \frac{F'(i)}{\sqrt{\text{Var}[F'(i)]}} \]

(27)
where:

\[ R[F(i),F'(i)] = \frac{\text{Cov}[F(i),F'(i)]}{\sqrt{\text{Var}[F(i)]} \sqrt{\text{Var}[F'(i)]}} \]

is the Pearson product moment correlation coefficient between \( F'(i) \) and \( F(i) \). It then follows that \( \text{Cov}(u,v) = 0 \) if and only if \( u \) and \( v \) are independent (Freeman [1963, p.151]). Using the variables \( u \) and \( v \) and the argument of Geary (1930) and Kendall and Stuart [1963, pp. 270-271], we have that the variate:

\[ z = \frac{-E[F(l_a)](l_a - i^*)}{\sqrt{\text{Var}[F(l_a)]} \left[ \{1 - R^2[F(l_a),F'(l_a)]\} + \left\{ \frac{\text{Var}[F'(l_a)]}{\text{Var}[F(l_a)]}(l_a - i^*) - R[F(l_a),F'(l_a)] \right\}^2 \right]} \]  

is normally distributed with mean zero and unit variance, provided:

\[ E[F(l_a)] \geq 3\sqrt{\text{Var}[F'(l_a)]} \]

effectively precluding the possibility of \( F'(l_a) \) being negative. This latter requirement can easily be checked from the foregoing equations. Suffice it to say it is a condition which is likely to be satisfied unless the fund is small or newly established. An interesting consequence of the above analysis is that if the variance in the fund's cash flows is large, this could induce a large confidence interval for the true return.

### 3.3 Estimating the Cash Flow's Instantaneous Variance

The importance of the analysis in the previous section is that it provides a means for determining a confidence interval for a fund's return over a given period and, as a consequence, a direct test of whether one return estimating technique is more reliable than another. However if we are to implement these results, it is first necessary that we have a reliable estimate of the instantaneous variance in the fund's cash flows. We now demonstrate how such an estimate may be obtained.

We begin by noting that from equation (8) it follows:

\[ \Delta^2 C(t) = (b - t) \int_{t + \Delta}^{t + 2\Delta} \frac{dW(s)}{b - s} - (b - t) \int_{t}^{t + \Delta} \frac{dW(s)}{b - s} - 2\Delta t \int_{t}^{t + \Delta} \frac{dW(s)}{b - s} \]  

where:

\[ \Delta^2 C(t) = C(t + 2\Delta) - 2C(t + \Delta) + C(t) \]

represents the second difference in the fund's total cash flows over the interval \([t, t + 2\Delta]\). Letting \( b = t + 3\Delta \) and using equation (30), it follows that:

\[ E[\Delta^2 C(t)]^2 = \sigma^2 \]

provides an unbiased estimate of the cash flow's instantaneous variance.
As previously noted, in the McCrae and Tippett [1987] study monthly cash flows were used in conjunction with funds' quarterly market values to estimate quarterly rates of return. Given this specification it follows that $T = 1$ and $A_t = 1/3$. Hence, supposing we have $n = 2m$ monthly cash flows for a particular fund, it follows that the statistic:

\[ \sigma^2 = \frac{3}{2m} \sum_{j=0,2,4,6,\ldots}^{2(m-1)} [\Delta^2 C(t_j)]^2 \]  

provides an unbiased estimate of $\sigma^2$.

The above procedure was used to estimate the cash flow variance for the 114 Australian superannuation funds employed in the McCrae and Tippett [1987] study. The difference between the estimated return using the linear technique and the technique traditionally used by the investment community was taken to be the error, $i_a - i_t$. As previously noted, the linear technique estimates the return as the root of equation (21) and is the method which arises naturally from the analysis of section 3.2. The return formula traditionally used by the investment community is the following (McCrae and Tippett [1987, p. 214]):

\[ i = \frac{M(1) - [M(0) + C_1 + C_2 + C_3]}{M(0) + [\frac{5}{6}C_1 + \frac{1}{2}C_2 + \frac{1}{6}C_3]} \]  

where, as previously, $M(t)$ is the market value of the fund's assets at time $t$ and $C_j = C (j/3) - C ((j - 1)/3), j = 1, 2, 3$ is the fund's cash flows during the jth of the quarter. Kelly and Tippett [1989] contains a comprehensive discussion of these and several other techniques.

The estimated variance was then used in conjunction with equation (28) to compute the $z$ statistics applicable to each fund. As noted by McCrae and Tippett [1987, p. 218], there are a maximum of 34 quarterly returns for each fund and so an average $z$ statistic was computed across these returns. The results are summarized in Table 1. Note that all but

| TABLE ONE ABOUT HERE |

25 of the 114 average $z$ statistics are significantly different from zero at the one percentage point level. Further, all but two of the average $z$ statistics were negative, thus confirming the suggestion made by McCrae and Tippett [1987, p. 216] that the traditional formula consistently overestimates the returns earned. Hence, for this and other reasons, it is doubtful if any credibility can be attached to superannuation fund evaluation methods based on the return metric (34).
4. SUMMARY AND CONCLUSIONS

The present note derives a general technique for estimating returns on financial instruments, when there is uncertainty about the instrument's cash flows. The method assumes that the instrument's cash flows are generated by a Brownian bridge or "tied down" process. This allows the general error expression of Kelly and Tippett [1989, p. 61] to be interpreted as a random variable, so that its distributional properties can be derived and probabilistic assessments made as to its likely magnitude. Using these procedures and the data set of McCrae and Tippett [1987] it is shown that the traditional estimating methods considerably overstate actual returns.

Consistent with the empirical results of McCrae and Tippett [1987, pp. 218 - 219], our analysis also shows that returns are far from normally distributed. As a consequence, our analysis raises serious doubts about the appropriateness of evaluating the performance of Australian superannuation funds using the standard two parameter model of Sharpe [1964] and Lintner [1965].

TABLE 1

<table>
<thead>
<tr>
<th>z Score</th>
<th>Significance</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(</td>
<td>z</td>
<td>&lt; 1.64</td>
</tr>
<tr>
<td>1.64 \leq</td>
<td>z</td>
<td>1.96</td>
</tr>
<tr>
<td>1.96 \leq</td>
<td>z</td>
<td>&lt; 2.57</td>
</tr>
<tr>
<td>2.57 \leq</td>
<td>z</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FOOTNOTES

1. See Hinkley [1969] for the procedures to use when this assumption is not satisfied.

2. If we expand equation (2) as a Taylor series around the risk free rate of interest and follow procedures similar to those used in deriving equation (28), it may be shown that the distribution of the risk premium, and by implication the return itself, is not normally distributed. Indeed, the mean and variance of the risk premium do not exist. This is consistent with the empirical work of McCrae and Tippett [1987, pp. 218 - 219]. Given these results, it is doubtful if any credibility can be attached to the standard two parameter method of evaluating superannuation fund performance.

3. Equations (23), (25) and (26) were evaluated using ten point Gauss-Legendre Quadrature. Details of this technique are to be found in Carnahan, Luther, and Wilkes [1969, pp. 101-105].

4. See footnote 2.
REFERENCES


Bank Administration Institute., (1968), Measuring the Performance of Pension Funds, Bank Administration Institute, Park Ridge, Illinois.


Freeman, H., (1963), Introduction to Statistical Inference, Addison-Wesley, Reading, Massachusetts.


