Craig, Ian K.; Bendixen, Mike T.

Article

Risk estimation for shares on the Johannesburg Stock Exchange using transfer function modeling

South African Journal of Business Management

Provided in Cooperation with:
University of Stellenbosch Business School (USB), Bellville, South Africa

Suggested Citation: Craig, Ian K.; Bendixen, Mike T. (1998) : Risk estimation for shares on the Johannesburg Stock Exchange using transfer function modeling, South African Journal of Business Management, ISSN 2078-5976, African Online Scientific Information Systems (AOSIS), Cape Town, Vol. 29, Iss. 1, pp. 1-13, https://doi.org/10.4102/sajbm.v29i1.765

This Version is available at:
https://hdl.handle.net/10419/218184

Risk estimation for shares on the Johannesburg Stock Exchange using transfer function modeling

Ian K. Craig[1*]
Department of Electrical and Electronic Engineering, University of Pretoria, Pretoria, Republic of South Africa 0002
ian.craig@eng.up.ac.za

Mike T. Bendixen
Formerly of the Graduate School of Business Administration, University of the Witwatersrand, P.O. Box 98, Wits, 2050 Republic of South Africa

This study investigates whether the estimation of the systematic risk component or the beta of shares on the Johannesburg Stock Exchange (JSE) can be improved using transfer function or MARIMA modeling. Two propositions are tested. Transfer function modeling will result in estimates of systematic risk which are different from those obtained using conventional OLS regression methods. Transfer function models will provide forecasting results which are better than those provided by betas estimated in the conventional way. Proposition 1 cannot be tested using conventional inferential tests as the standard errors of estimate of the betas estimated from MARIMA modeling cannot, in general, be measured. It is found however that 16.9% of the MARIMA beta estimates fall outside the 95% confidence intervals of the respective OLS regression beta estimates. Similar results are obtained when the OLS regression betas are compared to the University of Cape Town (UCT) Financial Risk Service and BFA-NET beta estimates. Proposition 2 can in general not be supported as the MARIMA and OLS regression forecasts are found not to be statistically significantly different. Cross correlations between index and share returns are in many cases found not to be statistically significant. In such cases one is probably better off using OLS regression. Resulting beta estimates should be used with caution.

*Author to whom correspondence should be addressed.

Introduction

The Capital Asset Pricing Model (CAPM), developed by Sharpe (1964), Lintner (1965) and others, and the market model (Markowitz, 1959; Sharpe, 1963), are widely used products of Capital Market Theory. These models essentially relate returns to the systematic risk taken to achieve such returns. Non-systematic risk can be effectively eliminated at virtually no cost through diversification (for example Fama, 1965; Statman, 1987), i.e. investors do not get paid for bearing risk that can be diversified away.

The estimation of the systematic risk component of shares have been the focus of many researchers in the field of financial economics. It is traditionally estimated from the market model (Bodie, Kane & Marcus, 1996: 279) which postulates a simple linear relationship between the return on the share and the return on the market. The beta coefficient of a share, which is the measure used to express systematic risk, is simply the slope coefficient of this linear relationship. The beta of a share is therefore obtained by means of an Ordinary Least Squares (OLS) regression of the total returns of the share against the total market index returns. In the CAPM framework, excess returns (deviations from the risk free rate) would typically be used in the regressions.

Efficient beta estimates are useful for portfolio design, asset pricing decisions, cost of capital calculations, and efficient testing of financial models such as the CAPM (for example Firer, 1993). Due to the importance of the beta coefficient to the investment community, much empirical research has focussed on the problems associated with estimating beta (Bowie & Bradfield, 1993).

This study investigates whether the estimation of the systematic risk component of shares on the JSE can be improved using transfer function modeling as described in Box, Jenkins & Reinsel (1994) and Makridakis, Wheelright & McGee (1983). Box-Jenkins transfer function or Multivariate Autoregressive Integrated Moving Average (MARIMA) models use the same dependent and independent variables that are used in the OLS regression estimation of beta.

Transfer function models have the advantage that they account for leads and lags between the share price returns and the market index returns in a systematic way, unlike conventional OLS regression which only account for share price return and market index return movements which occur simultaneously. It is well known for example that the prices of all classes of shares lead the business cycle whereas the non-residential building sector lags the business cycle (Van der Walt & Pretorius, 1994). Conventional OLS regression estimates of beta do not account for such a phenomenon.

Earlier attempts at incorporating both leads and lags of the relevant return series in the estimation of beta, can be found in Dimson (1979), Cohen, Hawawini, Maier, Schwartz & Whitcomb (1983), and Bradfield & Barr (1989). These papers address risk estimation in thinly traded markets.

In this article, two propositions are tested:

1. Transfer function modeling will result in estimates of systematic risk which are different from those obtained using conventional OLS regression methods.
2. Transfer function models will provide forecasting results which are better than those provided by betas estimated in the conventional way.

Procedure for estimating beta

This section describes how beta is estimated using both an OLS regression and a transfer function modeling approach. For a more comprehensive account see Box *et al.* (1994) and Makridakis *et al.* (1983).

Data preparation

Data preparation is required before transfer function modeling and OLS regression can commence. Firstly, a meaningful share and index sample is required from which returns can be calculated. The resulting time series then needs to be made stationary via differencing before the Box-Jenkins methodology is applied. Finally, the input and output series are prewhitened to simplify the MARIMA modeling procedure, and to make the results easier to interpret.

Sample selection

Weekly closing prices of the industrial index and all the shares listed on the industrial board of the JSE were obtained from BFA-NET, the financial database system of the University of Pretoria's Bureau for Financial Analysis, for the period 1 January 1988 to 31 December 1995. The industrial index was used as a proxy for the market. All shares which did not trade for the full eight-year period were eliminated from the data set. This left 250 shares from which a random sample of 60 shares were taken and on which transfer function modeling and OLS regression were performed.

Eight years worth of four-weekly share and index returns were calculated over individual four-week periods by dividing the price difference of the fourth and first week closing prices by the closing price of the first week. This resulted in 13 equispaced return data points for each share and the index per annum.

Sixty months of data have traditionally been considered a reasonable compromise between stability of the underlying betas and a sufficient number of data points for efficient estimation (Dimson & Marsh, 1983; Bowie & Bradfield, 1993). The eight years worth of data were therefore partitioned such that the first five years of data could be used for model or beta estimation. The remaining three years worth of data was used for model validation, as it is standard practice in the field of system identification to perform model estimation on two thirds of the data set and model validation on the other third (Ljung, 1995).

Share returns were not corrected for dividends as this was deemed not to have a significant impact on the findings of this work (Brummer, 1997).

Differencing

Many empirical time series, such as the share and index returns discussed here, behave as though they do not have a fixed mean. Models for such non-stationary series can however be obtained by supposing some suitable difference of the process to be stationary. For the time series under discussion it was found that first differencing suffices, that is the estimated auto and cross correlations of the differenced series, damp out quickly. First differencing of the index returns (X_t) and share returns (Y_t) respectively, are shown below:

$$X_t' = (1 - B)X_t$$
$$Y_t' = (1 - B)Y_t \quad (1)$$

with X_t' and Y_t' representing the differenced index and share returns respectively. B is the backshift operator defined by:

$$B_n X_t = X_{t-n} \quad (2)$$

B_n, operating on X_t has the effect of shifting the data back n periods.

After the differencing performed in equation (1), the means of X_t' and Y_t' were subtracted from the differenced index and share return series, as the focus of this article is on estimating the slope coefficient or beta in the market model and not the constant alpha. The resulting zero mean series were used in the steps that follow.

Prewhitening

It is always desirable for building models from time series data to be able to have some control over the input of the process. When dealing with physical processes, this is quite a feasible objective. In business and economics the input can usually not be controlled, but the input series can be prewhitened to simplify the identification process. Prewhitening of a suitably differenced input process X_t' can be performed by obtaining an ARMA (Autoregressive Moving Average) model which transforms the correlated input series to an uncorrelated white noise series $e_{x,t}$ as follows:

$$X_t^1 = \frac{\theta_x(B)}{\phi_x(B)} e_{x,t} \quad (3)$$

Usually the same transformation is applied to the output series Y_t' such that the cross correlation function between the prewhitened input and correspondingly transformed output is directly proportional to the impulse response function of the system being modeled. In this article however, the Granger & Newbold methodology (1977: 232, 237) was applied, where in addition to prewhitening the input series, the output series was prewhitened by its own ARMA model:

$$Y_t^1 = \frac{\theta_y(B)}{\phi_y(B)} e_{y,t} \quad (4)$$

Care should be applied when interpreting parameters of MARIMA models as they reflect both the relationships between the variables concerned, as well as the influence of all other variables which are modeled through the MARIMA process (Nicholson, 1991). The Granger-Newbold methodology overcomes this problem as it makes the cross correlogram easier to interpret by eliminating spurious regression and by taking care of possible feedback in the system.

The autocorrelation and partial autocorrelation functions (ACFs and PACFs) are typically calculated and plotted to determine whether the particular time series has been prewhitened to a sufficient degree (usually to 95% confidence). The Box-Jenkins MARIMA models are then developed from the residual series with $e_{x,t}$ as input and $e_{y,t}$ as output.

Box-Jenkins methodology

The general format of Box-Jenkins MARIMA models is given by:

$$Y_t = \frac{\omega^*(B)}{\delta^*(B)} X_{t-b} + N_t \quad (5)$$

The polynomials ω^* and δ^* and are given by:

$$\omega^*(B) = \omega_0^* - \omega_1^* B - \omega_2^* B^2 - \ldots \quad (6)$$

$$\delta^*(B) = 1 - \delta_1^* B - \delta_2^* B^2 - \ldots$$

where

$$\frac{\omega^*(B)}{\delta^*(B)}$$

is known as the transfer function model. N_t is the noise model and b is the number of sampling intervals that elapse before the output starts responding to a change in the input. It is assumed that the pairs of observations (X_t, Y_t) are sampled at equispaced intervals of time and that X_t is the 'input' and Y_t the 'output' of some dynamic system. For the purposes of this article, the dynamic relationship between share returns (outputs) and market index returns (inputs) were described.

The MARIMA model as shown in equation (5) has the advantage that it accounts for leads and lags between the share price returns and the market index returns in a systematic way. It does this by recognizing that index return values can be dynamically distributed over past and future time periods. This information is then captured in a transfer function model. Conventional OLS regression only accounts for share price return and market index return movements which occur simultaneously.

Now that the share and index returns have been calculated, the return series have been made stationary by differencing, and the differenced series have been prewhitened, MARIMA models can be estimated. The MARIMA modeling process consists of an identification stage in which the model order and time delay is estimated, a model fitting stage in which the transfer function parameters are determined, and a validation stage where the ability of the transfer function to forecast returns, is tested. These stages will be described in more detail in what follows.

Model identification

Before fitting a transfer function model to the data, it is useful to have an indication of the size of the delay b, and the order of the numerator and denominator polynomials. In the model identification stage it is also necessary to determine if share returns lead the index returns. In such a case the share returns need to be shifted relative to the index returns before MARIMA modeling can commence. The size of this shift must therefore be determined.

The order of a linear system can be estimated in many different ways (Ljung, 1987: 413). For example, the estimation data set could be used to examine the spectral analysis estimate of the transfer function, to test the rank of the sample covariance matrix, and to examine the information matrix.

In the forecasting literature it is common to use the cross correlogram between the input ($e_{x,t}$) and output ($e_{y,t}$) series to estimate model order, as well as the size of the delay parameter b in equation (5). In addition, the cross correlogram between the residual series $e_{x,t}$ and $e_{y,t}$ is also useful for determining whether a particular share return series actually leads the index return series, and what the size of the shift of the share return series relative to the index return series needs to be.

An estimate of the order of the model can be obtained from the number of significant (at the 95% confidence level) cross correlations, as they are approximately equal to the sum of the orders of the numerator and denominator polynomials, ω^* (B) and δ^* (B). The standard error for cross correlations of lag k were calculated in this work using Bartlett's formulation (Makridakis *et al.*, 1983: 495): std. error of

$$r_{XY}(k) = \sqrt{\frac{1}{n-k}}$$

The MATLAB® System Identification Toolbox (Ljung, 1995) was used as a basis for obtaining the MARIMA models. In MATLAB notation the model order and the size of the delay are specified as:

nn=[nb nc nd nf nk]

with nk being the number of sampling periods that make up the delay (b in equation 5), nb and nf being the model orders of the numerator and denominator polynomials ω(B) and δ(B) respectively, and nc and nd being the model orders of the numerator and denominator noise polynomials, θ(B) and ϕ(B) respectively. This notation corresponds to that used in equation (7).

The impulse response weights, which are calculated from the cross correlation coefficients, are used to obtain a preliminary estimate of the noise series. The orders of the noise polynomials (nc and nd) can then be estimated from the corresponding ACF and PACF derived from this noise series.

The best model structure is a trade-off between flexibility and parsimony. A flexible model can be obtained by using many parameters such that many different possible systems can be described whereas a parsimonious model will use only as many parameters as is necessary. The motto is to try simple things first.

Model fitting

Model parameters were calculated from the estimation range of the data set (1 January 1988 to 31 December 1992), and were developed from the residual series with $e_{x,t}$ as input and $e_{y,t}$ as output:

$$e_{y,t} = \frac{\omega(B)}{\delta(B)} e_{x,t-b} + \frac{\theta_n(B)}{\phi_n(B)} e_t \quad (7)$$

To estimate the model parameters, a robustified quadratic prediction error criterion was minimized using an iterative Gauss-Newton algorithm (Ljung, 1995: 4–11). Standard errors were calculated for each model parameter and only models with statistically significant parameters were used for further evaluation. In addition to the loss function, Akaike's final prediction error (FPE) criterion (Ljung, 1987: 420) was calculated to facilitate comparison between different models.

Before validating a model as described in the next section, preliminary tests were performed using the estimation data set to determine whether the model should be discarded or

not. Firstly, the model output, as generated by the actual input series, was visually compared to the actual output. This gives a good indication of the quality of the fit. Subsequently, the autocorrelation of residuals, and the cross correlation of the residuals and the input series, were computed and tested for statistical significance. If the model is a good one, the residuals should be a white noise series, that is there should be no statistically significant autocorrelations except for lag 0. Also the input and residual series should be independent, that is there should be no statistically significant cross correlation between these two series. If significant cross correlations do exist, it implies that there is information in the output which originates from the input but which is not explained by the model. Finally, the residuals were plotted and inspected for outliers.

Model validation

Models will in general perform well when evaluated on the data set to which it was adjusted. The real test for a model is whether it is capable of also describing different data sets from the same process. Model validation was therefore done using the validation range of the data set (1 January 1993 to 31 December 1995). When determining if the model is good enough, the following questions arise (Ljung, 1987: 424):

- Does the model agree sufficiently well with the observed data?
- Is the model good enough for my purpose?
- Does the model describe the true system?

These questions are addressed by subjecting the model to as much information about the true system as is possible. This could include prior knowledge, experimental data, and experience of using the model.

Before model validation was done, the same data manipulation was performed on the validation data as was applied to the estimation data, that is differencing, prewhitening (using the ARMA models found in the estimation stage), and shifting of shares whose returns lead the index returns.

The first validation step was to visually compare the model output to the actual output. The quality of the fit was typically worse than that obtained using the estimation data. Subsequently, the autocorrelation of residuals and cross correlation of the residuals as well as the input series were computed and tested for statistical significance as described earlier.

In addition to these tests, Theil's U-statistic (Makridakis *et al.*, 1983: 50) was calculated from the following equation to evaluate the forecasting ability of the MARIMA models relative to a naïve approach.

$$U = \frac{\sqrt{\sum_{i=1}^{n-1}\left(\frac{\hat{Y}_{i+1} - Y^{1}_{i+1}}{Y^{1}_{i}}\right)^{2}}}{\sqrt{\sum_{i=1}^{n-1}\left(\frac{Y^{1}_{i+1} - Y^{1}_{i}}{Y^{1}_{i}}\right)^{2}}} \qquad (8)$$

where $\hat{Y}_i$ is the forecast and Y^1_i the observation (differenced share return) values.

Finally Pearson's correlation coefficient (Makridakis *et al.*, 1983: 37), which is a measure of the correlation between the actual validation output (Y^1_i) and the validation output as estimated by the model ($\hat{Y}_i$), was calculated as follows

$$r = \frac{\sum (\hat{Y}_i - \bar{\hat{Y}})\left(Y^1_i - \bar{Y}^1\right)}{\sqrt{\sum (\hat{Y}_i - \bar{\hat{Y}})^2}\sqrt{\sum\left(Y^1_i - \bar{Y}^1\right)^2}} \qquad (9)$$

The measures described above, together with more subjective measures such as experience in model building, were used to determine how useful the models are.

Calculating the MARIMA beta

The resulting MARIMA beta is essentially the steady-state gain of the estimated transfer function model (Leskinen & Tersvirta, 1976). Care must however be taken that the correct transfer function be used when beta is calculated. The transfer function to be used can be obtained by substituting equations (3) and (4) into equation (7) as follows:

$$\frac{\phi_y(B)}{\theta_y(B)}Y^1_t = \frac{\omega(B)\phi_x(B)}{\delta(B)\theta_x(B)}X^1_{t-b} + \frac{\theta_n(B)}{\phi_n(B)}e_t \qquad (10)$$

or

$$Y^1_t = \frac{\theta_y(B)\omega(B)\phi_x(B)}{\phi_y(B)\delta(B)\theta_x(B)}X^1_{t-b} + \frac{\theta_y(B)\theta_n(B)}{\phi_y(B)\phi_n(B)}e_t \qquad (11)$$

To make comparisons with the OLS regression results meaningful, the MARIMA beta estimate was taken as the steady-state gain of the transfer function

$$\frac{\omega^*(B)}{\delta^*(B)} = \frac{\theta_y(B)\omega(B)\phi_x(B)}{\phi_y(B)\delta(B)\theta_x(B)}$$

such that:

$$\beta_{MARIMA} = \frac{\omega^*_0 - \omega^*_1 - \omega^*_2 - \ldots}{1 - \delta^*_1 - \delta^*_2 - \ldots} \qquad (12)$$

OLS regression

The beta coefficient is normally estimated by means of an Ordinary Least Squares (OLS) regression of the share returns against the market index returns. (See Makridakis *et al.*, 1983 or Levin & Rubin, 1991, for a discussion on OLS regression.) The dependent and independent variables used in the OLS regression estimation of beta were the same as used for MARIMA modeling. OLS regression was therefore carried out using the first differenced series X^1_t (without the delay b taken into account) and Y^1_t. In addition, the estimation range of the data set (1 January 1988 to 31 December 1992) was used.

The regression was performed in MATLAB for each of the 60 randomly selected share returns, that is a polynomial of degree one was fitted to the data in a least squares sense. The following parameters were calculated: the slope (beta),

95% confidence intervals for beta (estimated using the appropriate t-test), the residuals, and statistics describing the quality of the fit – the coefficient of determination R^2 and the F-statistic for overall significance of the fit.

A model validation exercise was carried out on the OLS models. The autocorrelations of the residuals were inspected, the raw Theil's U-statistic, and Pearson's correlation coefficient, were calculated. Finally, a significance test was performed to establish whether the correlation coefficients calculated for the MARIMA and OLS models, respectively, were significantly different (Kanji, 1993: 35). When performing the Z-test for two sample correlation coefficients (r_i), the hypothesis is made that the two population correlation coefficients are equal, i.e.:

$$H_0: \rho_1 = \rho_2 \qquad (13)$$
$$H_A: \rho_1 \neq \rho_2$$

This hypothesis is tested by calculating:

$$Z = \frac{Z_1 - Z_2}{\sigma}$$

with

$$Z_i = \frac{1}{2} ln\left(\frac{1 + r_i}{1 - r_i}\right) \qquad (14)$$

and

$$\sigma = \sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}$$

Z is assumed to be normally distributed.

Results

This section describes the results of OLS regression and MARIMA beta estimation, based on the estimation range of the data set (1 January 1988 to 31 December 1992). In addition, validation results for the OLS regression and MARIMA estimates are given, and the two propositions stated in the introductory section are tested. Model validation is done using the validation range of the data set (1 January 1993 to 31 December 1995). Finally, the betas obtained in this work are compared to those that are available commercially.

Estimation of OLS betas

The results from the estimation of OLS betas are given in Table 1. This table shows, in order of appearance, the name of the share, its beta estimate, the lower and upper bounds of the 95% confidence interval for beta, the coefficient of determination R^2, the F-statistic, and the accompanying p value of the regression.

From inspecting the F-statistic it is clear that only 31 of the 60 shares returns yielded fits that are significant at the 0.05 level (F>4.00). It is therefore not surprising that there are relatively wide 95% confidence intervals on most of the OLS betas. As one would expect, the shares of well-traded companies, such as MALBAK (no. 30) and CG Smith (no. 11), have significant F-statistics, and relatively large R^2 values.

The JSE is notorious for thin trading which results in estimation problems when the OLS methodology is used (Bowie & Bradfield, 1993). Bad fits (small R^2 and F-statistics) are therefore to be expected for thinly traded shares such as RENBEL (no. 49).

Estimation using MARIMA models

The results obtained from estimating MARIMA models are given in Table 2. This table shows, in order of appearance, the name of the share, the order of the ARMA model used to prewhiten the output Y_t^I (na=0 and nc=1 corresponds to a ARMA (0,1) model), the number of months that Y_t^I was shifted relative to X_t^I, the order of the numerator and denominator polynomials of the transfer function (nb and nf respectively), the noise model numerator and denominator order (nc and nd respectively), the delay in four-week periods (nk), and finally the MARIMA beta (as calculated from equation 12). The industrial index return series (X_t^I) was prewhitened with a ARMA (0,1) model.

Some general comments can be made regarding the data shown in Table 2. Detailed comments on each fit are given in Table 3.

- The return series of most shares can be prewhitened adequately using an ARMA (0,1) model. An exception is for example LOHNRO where an ARMA (2,0) model is required.
- The return series of most shares do not have significant negative cross correlations with the index returns – they do not lead the index returns and therefore required no shifting. The ones that do, were shifted relative to the index, and the number of four-week periods that they were shifted are shown in the shift column. For example, HORTORSs returns (no. 23) were shifted eight weeks or two data points. The return series of shares such as MACMED, NORBAKE, NATRAWL and RENBEL required large shifts (7, 15, 8 and 13 data points respectively) in order to obtain fits. This could be due to thin trading, as there is no apparent reason for the return series of these four shares to lead the index returns by such large time periods.
- The return series of some shares lagged the index returns, that is, the only significant cross correlations were found at lag k where $k \geq nk$. Such delays were incorporated into the model and are shown in the nk column. For example, AUKLAND's returns (no. 6) lag the index returns by 12 weeks or three data points.
- Share returns with very significant cross correlations at k=0 and nowhere else, are actively traded and generally modeled well with MARIMA models of order nn= [1 0 0 0 0] (MATLAB notation). The result is a single parameter (ω_0) model which is the same (apart from the effects of prewhitening) as determined by OLS regression. For example, such a model was obtained for MALBAK ($\beta_{MARIMA} = 1.291$), and comparing its beta to that obtained with OLS regression ($\beta_{OLS} = 1.725$), it can be seen that prewhitening can have an appreciable effect on the absolute value of beta.
- The fits were in general not very good. Parsimonious models were therefore used in most cases. For example,

Table 1 OLS regression results

	Share name	β_{OLS}	cnfi95%	cnfi95%	R^2	F-stat	p(F)
1 AFX	afrox	0.948	0.400	1.497	0.162	11.947	0.001
2 AKJ	arthur kaplan	1.593	-0.185	3.372	0.049	3.208	0.078
3 ALX	alexwyt	1.398	0.202	2.595	0.081	5.455	0.023
4 AOO	afr&ovr	0.380	-0.150	0.910	0.032	2.058	0.156
5 ATN	altron	1.547	0.919	2.176	0.281	24.235	0.000
6 AUK	aukland	0.310	-3.359	3.979	0.000	0.029	0.866
7 BIV	bivec	0.372	-0.243	0.986	0.023	1.463	0.231
8 BSR	basread	0.372	-1.270	2.014	0.003	0.205	0.652
9 BWR	bolwear	1.418	0.130	2.705	0.072	4.844	0.031
10 BZK	berzack	0.183	-0.479	0.846	0.005	0.306	0.582
11 CGS	cgsmith	1.032	0.725	1.338	0.422	45.300	0.000
12 CHR	charter	0.456	-0.157	1.069	0.034	2.214	0.142
13 CMI	cmi	0.085	-0.700	0.871	0.001	0.047	0.828
14 COT	coates	0.342	-0.682	1.367	0.007	0.447	0.506
15 CRG	cargo	1.490	0.306	2.674	0.093	6.332	0.014
16 CTY	cityhold	-0.898	-1.692	-0.105	0.076	5.120	0.027
17 DLC	delcorp	1.948	0.917	2.979	0.187	14.264	0.000
18 DON	don group	0.145	-0.462	0.752	0.004	0.227	0.635
19 EUR	eureka	1.296	0.347	2.244	0.107	7.449	0.008
20 FOS	foschini	0.105	-0.374	0.585	0.003	0.193	0.662
21 GNR	grincor	0.688	-0.106	1.481	0.046	3.004	0.088
22 HDC	hudaco	0.942	0.577	1.307	0.301	26.637	0.000
23 HTR	hortors	-0.301	-3.011	2.408	0.001	0.049	0.825
24 IBJ	ib joffe	-0.344	-0.838	0.150	0.030	1.938	0.169
25 ITL	inteles	1.796	1.047	2.545	0.270	22.967	0.000
26 LNM	lion	1.189	0.517	1.860	0.168	12.533	0.001
27 LON	lonrho	1.158	0.461	1.856	0.151	11.023	0.002
28 LSU	lonsugr	-0.073	-1.886	1.739	0.000	0.007	0.936
29 LTA	lta	1.582	0.480	2.684	0.117	8.232	0.006
30 MLB	malbak	1.725	1.321	2.128	0.541	73.039	0.000
31 MLT	multisource	0.336	-1.404	2.076	0.002	0.149	0.701
32 MMD	macmed	-2.362	-5.115	0.391	0.045	2.940	0.091
33 MTC	metcash	0.667	-0.030	1.364	0.056	3.660	0.060
34 MUR	m&r hld	1.460	0.921	2.000	0.321	29.305	0.000
35 NEH	neihold	0.380	-0.291	1.050	0.020	1.282	0.262
36 NPK	nampak	1.112	0.680	1.545	0.299	26.454	0.000
37 NRK	norbake	-0.170	-1.209	0.869	0.002	0.107	0.745
38 NTR	natrawl	2.860	0.150	5.570	0.067	4.452	0.039
39 OCF	ocfish	0.060	-0.688	0.808	0.000	0.026	0.873
40 OMA	omega	1.636	0.450	2.822	0.109	7.600	0.008
41 PEG	pepgro	0.921	0.223	1.620	0.101	6.956	0.011
42 PEP	pepkor	0.132	-0.385	0.648	0.004	0.260	0.612
43 POR	port	-0.153	-0.740	0.433	0.004	0.274	0.603
44 PPC	ppc	0.626	0.358	0.895	0.259	21.723	0.000
45 PTC	putco	0.241	-3.061	3.542	0.000	0.021	0.885
46 PWK	pikwik	1.972	1.389	2.555	0.424	45.661	0.000
47 QDT	q data	-0.098	-1.801	1.604	0.000	0.013	0.909
48 REG	reggies	2.619	0.732	4.507	0.110	7.696	0.007
49 RNT	renbel	-0.074	-0.788	0.641	0.001	0.042	0.838
50 SAP	sappi	1.046	0.743	1.348	0.435	47.795	0.000
51 SPU	spur	1.742	1.041	2.443	0.285	24.692	0.000
52 TEG	teglor	1.527	1.154	1.899	0.520	67.232	0.000
53 TPC	tnpaco	-0.232	-1.759	1.296	0.001	0.092	0.763
54 TRH	tradhld	0.795	-0.039	1.628	0.055	3.635	0.061
55 TUN	t&n holdings	0.899	0.019	1.779	0.063	4.174	0.045
56 UTR	unitran	0.179	-0.506	0.865	0.004	0.274	0.603
57 VLX	voltex	1.496	0.459	2.533	0.118	8.318	0.005
58 VNT	ventron	0.861	0.307	1.415	0.135	9.649	0.003
59 WAL	waltons	1.722	1.151	2.294	0.369	36.332	0.000
60 WLH	walhold	1.614	1.026	2.203	0.327	30.082	0.000

Table 2 MARIMA modeling results

Share name		na	nc	shift	nb	nc	nd	nf	nk	β_{MARIMA}
1 AFX	afrox	0	1	0	1	0	0	0	0	0.837
2 AKJ	arthur kaplan	3	1	0	1	0	0	0	0	0.344
3 ALX	alexwyt	0	1	0	1	0	0	0	0	0.670
4 AOO	afr&ovr	0	1	5	1	0	0	0	0	0.452
5 ATN	altron	0	1	0	1	0	0	0	0	1.218
6 AUK	aukland	2	1	0	2	0	0	0	3	-1.034
7 BIV	bivec	0	1	0	1	0	0	3	0	0.515
8 BSR	busread	0	1	0	1	0	0	0	0	0.844
9 BWR	bolwear	0	1	0	1	0	0	0	0	0.829
10 BZK	berzack	0	1	0	1	0	0	0	0	0.227
11 CGS	cgsmith	0	1	0	2	0	0	0	0	0.957
12 CHR	charter	0	1	0	1	0	0	1	4	0.710
13 CMI	cmi	0	1	0	1	0	0	2	1	0.215
14 COT	coates	0	1	0	1	0	0	0	2	0.370
15 CRG	cargo	0	1	0	1	0	0	0	0	0.689
16 CTY	cityhold	0	2	1	1	0	0	2	0	0.128
17 DLC	delcorp	0	1	0	1	0	0	0	0	1.294
18 DON	don group	0	1	0	1	0	0	0	2	-0.257
19 EUR	eureka	1	1	0	1	0	0	0	0	0.976
20 FOS	foschini	0	1	0	1	0	0	0	1	0.335
21 GNR	grincor	2	1	0	1	0	0	0	0	0.812
22 HDC	natrawl	0	1	0	2	0	0	1	0	0.845
23 HTR	hortors	0	1	2	1	0	0	0	0	0.481
24 IBJ	ib joffe	0	1	0	1	0	0	0	1	-0.159
25 ITL	inteles	0	1	0	1	0	0	0	0	1.107
26 LNM	lion	0	1	0	1	0	0	0	0	0.856
27 LON	lonrho	2	0	2	1	0	0	0	0	-0.574
28 LSU	lonsugr	0	1	0	-	-	-	-	-	no fit
29 LTA	lta	0	1	0	1	0	0	0	0	1.183
30 MLB	malbak	0	1	0	1	0	0	0	0	1.291
31 MLT	multisource	0	1	1	1	0	0	1	0	0.635
32 MMD	macmed	0	1	7	1	0	0	0	0	-1.285
33 MTC	metcash	0	1	0	2	0	0	1	0	0.642
34 MUR	m&r hld	1	1	0	1	1	1	1	0	1.068
35 NEH	neihold	0	1	0	1	0	0	0	1	0.515
36 NPK	nampak	0	1	0	1	0	0	1	0	1.034
37 NRK	norbake	0	1	15	1	1	1	0	0	-0.696
38 NTR	natrawl	0	1	8	1	0	0	0	0	-0.993
39 OCF	ocfish	0	1	0	1	1	1	0	1	0.607
40 OMA	omega	0	1	3	1	0	0	0	0	0.991
41 PEG	pepgro	0	1	0	1	0	0	1	0	0.753
42 PEP	pepkor	0	1	0	1	0	0	2	0	0.561
43 POR	port	0	1	0	1	0	0	0	1	0.617
44 PPC	ppc	1	2	0	1	0	0	0	0	0.694
45 PTC	putco	1	1	0	1	0	0	4	0	1.594
46 PWK	pikwik	1	1	0	1	0	0	0	0	1.090
47 QDT	q data	0	1	0	1	1	1	2	0	0.340
48 REG	reggies	0	1	0	1	0	0	0	0	0.903
49 RNT	renbel	0	1	13	1	0	0	1	0	0.688
50 SAP	sappi	0	1	0	1	0	0	0	0	0.945
51 SPU	spur	2	1	0	1	1	1	1	0	0.702
52 TEG	tegkor	0	1	0	1	0	0	0	0	1.260
53 TPC	tmpaco	0	1	0	1	0	0	0	0	0.472
54 TRH	tradhld	0	1	3	2	0	0	2	0	-0.848
55 TUN	t&n holdings	0	1	0	2	0	0	1	0	0.946
56 UTR	unitran	0	1	0	1	0	0	0	1	0.711
57 VLX	voltex	0	1	0	1	0	0	0	0	0.902
58 VNT	ventron	0	1	0	1	0	0	0	0	0.913
59 WAL	waltons	0	1	0	1	1	1	0	0	1.150
60 WLH	walhold	0	1	0	1	1	1	1	0	1.081

there exists no significant cross correlations between the returns of ARTHUR KAPLAN and the index returns. In such a case there is no point in fitting a high order model to the data. In fact, an argument can be made for not fitting any model to such data.

- No MARIMA model was estimated for LOHNRO SUGAR as no significant model parameters could be obtained. This is not surprising as its share price changed very infrequently during the estimation period (1 January 1988 to 31 December 1992).

Validation results

The validation results for the OLS regression and MARIMA models are given in Table 3. Table 3 shows, in order of appearance, the name of the share, the MARIMA beta estimates, the mean squared error (mse), Theil's U-statistic (equation 8), and the correlation estimates (r in equation 9) for the MARIMA models. This is followed by the OLS beta estimates, the mse, Theil's U-statistic, and the correlation estimates for the OLS regression models. The Z values shown are calculated according to equation (14) and are used to test the hypothesis that the correlation estimates for OLS regression and MARIMA modeling are the same. The column with heading H shows a '0' if the null hypothesis cannot be rejected and '1' if it can. Finally comments are given regarding each fit.

These results can be used to test proposition 2, that transfer function models will provide forecasting results which are better than those provided by betas estimated in the conventional (OLS regression) way.

One can start testing this proposition by looking at the means of the mean squared error (mse), Theil's U-statistic, and the correlation coefficient r as shown in Table 3 at the bottom of the respective columns. From these means it would appear that the MARIMA models forecast returns slightly better than OLS regression models. For example the MARIMA mse mean is 0.0371 as opposed to the OLS regression mse mean of 0.0378. Similarly, on average the MARIMA models fare slightly better on the Theil's U-statistic. However, both the MARIMA (U>1 and Theil's U= 1.0369) and the OLS regression (U>1 and Theil's U= 1.0840) models on average fare worse than the naïve model. The mean absolute value of the correlation coefficient r is lower and therefore worse in the MARIMA modeling case than it is for OLS regression (0.2815 compared to 0.2936).

The question arises whether the differences in the means given above are statistically significant. To answer this question a hypothesis test for the difference in means of the two samples can be performed (Jones, 1996: 2–209). Hypothesis tests were performed for the two mse, Theil's U, and r samples. It was found that none of these means are statistically significantly different at the 0.05 significance level.

The hypothesis tests on the correlation coefficients show that the null hypothesis (equation 13) is found not to be true in seven out of the 60 cases. In all of these cases the correlation coefficients have different signs.

There are some important validation results which are not captured in Table 3. The first validation step is to visually compare the model output to the actual output. Subsequently, the autocorrelation of residuals and cross correlation of the residuals and the input series are computed and tested for statistical significance. In general the residuals generated by the OLS regression models are not white, that is, there are significant residual autocorrelations. The MARIMA models were chosen such that the residuals generated from the estimation and validation data had no significant autocorrelations or cross correlations with the input series. In this sense the MARIMA models are statistically better than the OLS regression models.

Beta comparisons

The data shown in Table 4 can be used to indicate whether transfer function modeling will result in estimates of systematic risk which are different from those obtained using conventional OLS regression methods. This table shows, in order of appearance, the name of the share, the OLS regression beta estimate, the lower and upper bounds of the 95% confidence interval for the OLS regression beta, the MARIMA beta estimate, the UCT Financial Risk Service beta estimate for selected shares only, and the BFA-NET beta estimate, all for the period 1 January 1988 to 31 December 1992. The last three columns show whether the MARIMA, UCT, and BFA betas respectively fall outside of the 95% confidence intervals of the OLS regression betas. A 'n' is indicated if they do not and a 'y' if they do. The percentages of y's and n's are given at the bottom of the respective columns.

The standard errors of estimate of the betas estimated from MARIMA modelling cannot, in general, be measured because of the complex way in which they are calculated, namely the ratio of polynomials. Thus, statistically testing the first proposition, which states that transfer function modeling will result in estimates of systematic risk which are different from those obtained using conventional OLS regression methods, cannot be achieved using conventional inferential tests. Albeit weak, some evidence of the differences between the betas can be gleaned by examining whether or not the MARIMA beta estimates fall within the 95% confidence limits of the OLS regression beta estimates, the standard errors of estimate being readily assessed in this case. It can be seen from Table 4 that for ten out of 59 shares the MARIMA beta fall outside of the 95% confidence interval of the OLS beta.

Finally, the betas obtained in this work are compared to those that are available commercially. Significant adjustments are made for thin trading in the UCT beta estimation process which would account for some of the vast differences between the UCT betas and some of the OLS beta estimates in Table 4. Some of these adjustments are summarised in a review paper on systematic risk estimation on the JSE by Bowie & Bradfield (1993). It can be seen from Table 4 that for four out of the 23 shares (17.4%) for which UCT betas are available, the UCT beta fall outside of the 95% confidence interval of the OLS beta. This is similar to the proportion of MARIMA betas (16.9%) which fall outside this interval.

The BFA-NET beta estimates were calculated using OLS regression on undifferenced index (X) and share (Y) returns. Returns were not adjusted for dividends and the industrial index was used as market proxy. These betas differ from the OLS regression betas obtained in this study mainly due to the non-linear way in which these input and output variables enter the equation from which beta is calculated (Makridakis

Table 3 Validation results for OLS regression and MARIMA models

Share name		β_{MARIMA}	mse	Theilu	r	Zval	β_{OLS}	mse	Theilu	r	Zval	H	comments
1 AFX	afrox	0.837	0.0062	2.0113	0.4411	0.4736	0.948	0.0064	2.0779	0.4422	0.4750	0	good fits
2 AKJ	arthur kaplan	0.344	0.0980	1.0266	0.2269	0.2310	1 593	0.0943	1.1902	0.2594	0.2655	0	no significant cross correlations
3 ALX	alexwyt	0.670	0.0274	1.1855	-0.3684	-0.3866	1.398	0.0369	1.4508	-0.3682	-0.3864	0	cross correlations barely significant; inferior fits
4 AOO	afr&ovr	0.452	0.0165	1.3319	0.1028	0.1031	0.380	0.0156	1.1186	-0.0763	-0.0764	0	inadequate fits; no model without shift
5 ATN	altron	1.218	0.0202	1.4958	0 2997	0.3092	1.547	0.0210	1.6724	0.3213	0.3331	0	adequate fits; Theilu misleading
6 AUK	aukland	-1.034	0.3075	0.9713	-0.1165	-0.1171	0 310	0.2846	0.9790	0.1965	0.1991	0	inferior fits
7 BIV	bivec	0.515	0.0225	0.8831	-0.0448	-0.0449	0.372	0.0228	1.0347	-0.2060	-0.2090	0	MARIMA estimation adequate
8 BSR	basread	0.844	0.0749	1.4369	0.0534	0.0535	0 372	0 0739	1.1761	0.0423	0.0424	0	no significant cross correlations
9 BWR	bolwear	0.829	0.0355	1.4201	0.3695	0.3878	1 418	0.0348	1.7490	0.3695	0.3878	0	good fits
10 BZK	berzack	0.227	0.0184	0.7824	0.2070	0.2100	0.183	0 0185	0.7980	0.1942	0.1967	0	no significant cross correlations; insignificant model parameter nb=1
11 CGS	cgsmith	0 957	0.0092	0.8279	0.5014	0.5512	1.032	0.0093	0.8157	0.5013	0.5510	0	adequate fits
12 CHR	charter	0.710	0.0119	0.4025	-0.0755	-0.0756	0.456	0.0089	0.8357	0.0581	0.0582	0	cross correlation at -2 ; no fit with shift; awkward cross correlation at 0; inadequate fit; nk=4 not justifiable from cross correlation
13 CMI	cmi	0.215	0.1753	***			0 085	0.1695					no significant cross correlations: bad validation fits; good MARIMA estimation fit
14 COT	coates	0.370	0 0059	0.9472	0.1065	0.1069	0.342	0.0055	0.9195	0.2211	0.2248	0	nb=1 not significant; inadequate fits
15 CRG	cargo	0.689	0.0562	0 9760	-0.0375	-0.0375	1.490	0.0626	0.9846	-0.0363	-0.0364	0	inadequate fits
16 CTY	cityhold	0.128	0.1162	0.7910	-0.1214	-0.1220	-0.898	0.1360	1.1190	-0.1507	-0.1518	0	inferior fits; Theilu misleading; prewhitening of output not 100%
17 DLC	delcorp	1.294	0.0210	0 5657	0.4014	0.4254	1.948	0.0231	0.3521	0 4044	0.4289	0	adequate fits
18 DON	don group	-0.257	0.0280	0.6856	0.0256	0.0256	0.145	0.0280	1.0574	-0.0625	-0.0626	0	inferior fit
19 EUR	eureka	0.976	0.0627	1.1116	0.5264	0.5852	1 296	0.0608	1.1302	0.5148	0.5693	0	better without shifting
20 FOS	foschini	0.335	0.0172	0.8630	-0.1532	-0.1544	0.105	0.0154	1.0432	0.4886	0.5342	1	cross correlation at +14
21 GNR	grincor	0.812	0.0145	2.9555	0.1514	0.1525	0.688	0.0135	2.1103	0.1077	0.1082	0	adequate fits
22 HDC	hudaco	0.845	0.0179	1.1321	0.1408	0.1417	0.942	0.0183	1.1446	0.1416	0.1426	0	adequate fits
23 HTR	hortors	0.481	0.0307	0.9864	-0.0316	-0.0316	-0 301	0.0317	0.9845	0.0116	0.0116	0	inferior fit: nb=1 not sign;
24 IBJ	ib joffe	-0.159	0.0076	***	-0.2685	-0.2753	-0.344	0.0068	***	0.2190	0.2226	1	inferior fits: /0 in calculating Theilu due to thin trading
25 ITL	inteles	1.107	0.0096	0.9663	0.1807	0.1827	1 796	0.0147	1.3829	0.1929	0.1953	0	inadequate fits
26 LNM	lion	0.856	0.0353	1.1473	0.2383	0.2430	1.189	0.0360	1.2095	0.2313	0.2355	0	good fits
27 LON	lonrho	-0.574	0.0301	0.5052	-0.0666	-0.0667	1.158	0.0320	0.7928	-0.1846	-0.1867	0	similar to OLS if no shift
28 LSU	lonsugr	no fit					-0.073						no fit - insignificant model parameters for no shift and shift=7
29 LTA	lta	1 183	0.0176	0.6773	0.5775	0.6587	1.582	0.0170	0.5984	0.5771	0.6581	0	good fits
30 MLB	malbak	1.291	0 0151	0.5089	0.4587	0.4956	1.725	0.0160	0.6540	0.4686	0.5083	0	adequate fits
31 MLT	multisource	0.635	0.0133	0.9111	0.1180	0.1186	0.336	0.0123	1.0486	0.2383	0.2430	0	nb=1 not significant, no model without shift: no significant cross correlations
32 MMD	macmed	-1.285	0.0651	0.5168	0.3015	0.3111	-2.362	0.0995	1.4197	-0.3269	-0.3393	1	lag 7 marginally significant; adequate MARIMA fit compared to OLS

33 MTC	metcash	0.642	0.0093	0.9616	0.4146	0.4411	0.667	0.0098	0.9347	0.3589	0.3756	0 adequate fits
34 MUR	m&r hld	1.068	0.0085	0.8049	0.3576	0.3742	1.460	0.0102	0.7641	0.3577	0.3743	0 adequate fits
35 NEH	neihold	0.515	0.0180	1.1047	0.1788	0.1807	0.380	0.0198	0.8789	-0.1253	-0.1260	0 adequate fits
36 NPK	nampak	1.034	0.0054	0.6925	0.6377	0.7542	1.112	0.0050	0.7091	0.6645	0.8008	0 good fits: nf=1 required for white residuals in validation
37 NRK	norbake	-0.696	0.0060	0.7036	0.3623	0.3795	-0.170	0.0093	0.9727	-0.0336	-0.0336	0 adequate fits; cross correlation at 15 barely significant
38 NTR	natrawl	-0.993	0.0776	0.9708	0.4607	0.4983	2.860	0.0991	2.0835	-0.0058	-0.0058	1 nb=1 not significant, no model without shift; inferior fits; cross correlation at 8 barely significant
39 OCF	ocfish	0.607	0.0102	0.9487	-0.1402	-0.1411	0.060	0.0080	0.9578	0.3025	0.3123	0 Output does not prewhiten perfectly. MARIMA better than statistics suggest
40 OMA	omega	0.864	0.1487	0.9112	0.4534	0.4889	1.636	0.1558	1.3561	0.0920	0.0922	0 MARIMA moves in right direction in simulation
41 PEG	pepgro	0.753	0.0126	1.1088	0.4959	0.5438	0.921	0.0120	1.0600	0.5155	0.5702	0 inadequate fits- shift=2 results in worse model
42 PEP	pepkor	0.561	0.0136	1.3126	0.4425	0.4754	0.132	0.0152	0.8805	0.6400	0.7582	0 adequate fits; Theilu misleading; MARIMA cross correlation between residuals and input not independent at 0
43 POR	port	0.617	0.0319	0.9201	-0.2163	-0.2198	-0.153	0.0291	1.0104	-0.3320	-0.3451	0 inferior fits
44 PPC	ppc	0.694	0.0241	0.7482	0.5917	0.6803	0.626	0.0236	0.7788	0.6592	0.7914	0 good fits: valid residuals not quite white
45 PTC	putco	1.594	0.0633	3.4252	-0.0906	-0.0908	0.241	0.0261	1.1220	-0.0842	-0.0844	0 MARIMA better than statistics suggests; significant cross correlation at 13
46 PWK	pikwik	1.090	0.0142	0.4691	0.3445	0.3591	1.972	0.0154	1.1805	0.4520	0.4871	0 OLS looks better than statistics suggest
47 QDT	q data	0.340	0.0207	1.0780	0.2105	0.2137	-0.098	0.0222	1.0145	-0.3791	-0.3990	1 inadequate fit: no significant cross correlations
48 REG	reggies	0.903	0.0722	0.9358	0.5408	0.6053	2.619	0.0609	0.8186	0.5407	0.6051	0 OLS much better in validation; [1 0 0 5 0] all significant but inferior in validation
49 RNT	renbel	0.688	0.0572	0.9680	0.0827	0.0829	-0.074	0.0533	1.0010	-0.2703	-0.2772	0 inadequate fits; significant cross correlation at 13
50 SAP	sappi	0.945	0.0141	0.9021	0.4977	0.5463	1.046	0.0140	0.8870	0.5008	0.5504	0 adequate fits
51 SPU	spur	0.702	0.0100	1.3233	0.5726	0.6513	1.742	0.0091	1.5806	0.6376	0.7540	0 OLS much better than MARIMA in validation
52 TEG	tegkor	1.260	0.0053	0.9232	0.5781	0.6596	1.527	0.0062	1.0745	0.5815	0.6647	0 adequate fits
53 TPC	tmpaco	0.472	0.0423	****			-0.232	0.0450				no significant cross correlations and no fit
54 TRH	tradhld	-0.848	0.0243	1.0723	0.2833	0.2913	0.795	0.0291	1.0976	-0.2025	-0.2053	1 prewhitening of output not perfect; inadequate fits
55 TUN	t&n holdings	0.946	0.0186	0.9694	0.3064	0.3166	0.899	0.0183	0.9906	0.2986	0.3080	0 [1 0 0 2 0] also gives adequate fit
56 UTR	unitran	0.711	0.0128	0.9832	-0.2260	-0.2300	0.179	0.0089	0.8922	0.2639	0.2702	1 very good prewhitening of output; MARIMA closer to share volatility; inferior fits
57 VLX	voltex	0.902	0.0242	0.9120	0.2388	0.2435	1.496	0.0263	0.9287	0.2387	0.2434	0 good fits
58 VNT	ventron	0.913	0.0094	1.1360	0.2027	0.2056	0.861	0.0091	1.1150	0.2153	0.2187	0 OLS than MARIMA in validation
59 WAL	waltons	1.150	0.0291	0.7732	0.2246	0.2284	1.722	0.0318	0.7012	0.2280	0.2321	0 adequate fits; OLS closer to share volatility
60 WLH	walhold	1.081	0.0284	0.9889	0.1799	0.1819	1.614	0.0327	1.0633	0.1409	0.1418	0 noise model parameters not quite significant. Required for white MARIMA model residuals in validation
	Mean values		**0.0373**	**1.0369**	**0.2815**			**0.0378**	**1.0840**	**0.2936**		

Table 4 Beta comparisons

Share name		β_{OLS}	cnfi95%	cnfi95%	β_{MARIMA}	β_{UCT}	β_{BFA}	MAR	UCT	BFA
1 AFX	afrox	0.948	0.400	1.497	0.837	0.910	0.883	n	n	n
2 AKJ	arthur kaplan	1.593	-0.185	3.372	0.344		0.276	n		n
3 ALX	alexwyt	1.398	0.202	2.595	0.670		0.501	n		n
4 AOO	afr&ovr	0.380	-0.150	0.910	0.452	0.830	0.193	n	n	n
5 ATN	altron	1.547	0.919	2.176	1.218	1.610	0.804	n	n	y
6 AUK	aukland	0.310	-3.359	3.979	-1.034		1.025	n		n
7 BIV	bivec	0.372	-0.243	0.986	0.515	0.960	0.488	n	n	n
8 BSR	basread	0.372	-1.270	2.014	0.844		0.544	n		n
9 BWR	bolwear	1.418	0.130	2.705	0.829		0.811	n		n
10 BZK	berzack	0.183	-0.479	0.846	0.227	0.850	0.725	n	y	n
11 CGS	cgsmith	1.032	0.725	1.338	0.957	0.910	1.013	n	n	n
12 CHR	charter	0.456	-0.157	1.069	0.710	0.750	0.664	n	n	n
13 CMI	cmi	0.085	-0.700	0.871	0.215		0.476	n		n
14 COT	coates	0.342	-0.682	1.367	0.370	0.890	0.342	n	n	n
15 CRG	cargo	1.490	0.306	2.674	0.689		0.544	n		n
16 CTY	cityhold	-0.898	-1.692	-0.105	0.128		-0.375	y		n
17 DLC	delcorp	1.948	0.917	2.979	1.294		0.669	n		y
18 DON	don group	0.145	-0.462	0.752	-0.257		0.506	n		n
19 EUR	eureka	1.296	0.347	2.244	0.976	1.260	0.417	n	n	n
20 FOS	foschini	0.105	-0.374	0.585	0.335	0.970	0.658	n	y	y
21 GNR	grincor	0.688	-0.106	1.481	0.812		0.826	n		n
22 HDC	hudaco	0.942	0.577	1.307	0.845		0.981	n		n
23 HTR	hortors	-0.301	-3.011	2.408	0.481	0.820		n	n•	
24 IBJ	ib joffe	-0.344	-0.838	0.150	-0.159		0.118	n		n
25 ITL	inteles	1.796	1.047	2.545	1.107		0.987	n		y
26 LNM	lion	1.189	0.517	1.860	0.856	0.680	0.541	n	n	n
27 LON	lonrho	1.158	0.461	1.856	-0.574		0.997	y		n
28 LSU	lonsugr	-0.073	-1.886	1.739	no fit		0.211			n
29 LTA	lta	1.582	0.480	2.684	1.183	0.750	0.610	n	n	n
30 MLB	malbak	1.725	1.321	2.128	1.291	1.340	1.174	y	n	y
31 MLT	multisource	0.336	-1.404	2.076	0.635		0.089	n		n
32 MMD	macmed	-2.362	-5.115	0.391	-1.285		-0.542	n		n
33 MTC	metcash	0.667	-0.030	1.364	0.642		0.373	n		n
34 MUR	m&r hld	1.460	0.921	2.000	1.068		0.967	n		n
35 NEH	neihold	0.380	-0.291	1.050	0.515		0.272	n		n
36 NPK	nampak	1.112	0.680	1.545	1.034	0.950	0.950	n	n	n
37 NRK	norbake	-0.170	-1.209	0.869	-0.696		0.256	n		n
38 NTR	natrawl	2.860	0.150	5.570	-0.993		-0.090	y		y
39 OCF	ocfish	0.060	-0.688	0.808	0.607	0.350	0.335	n	n	n
40 OMA	omega	1.636	0.450	2.822	0.991		0.897	n		n
41 PEG	pepgro	0.921	0.223	1.620	0.753		0.605	n		n
42 PEP	pepkor	0.132	-0.385	0.648	0.561	1.010	0.569	n	y	n
43 POR	port	-0.153	-0.740	0.433	0.617	0.150	0.366	y	n	n
44 PPC	ppc	0.626	0.358	0.895	0.694	0.850	0.768	n	n	n
45 PTC	putco	0.241	-3.061	3.542	1.594		1.213	n		n
46 PWK	pikwik	1.972	1.389	2.555	1.090		0.996	y		y
47 QDT	q data	-0.098	-1.801	1.604	0.340		0.547	n		n
48 REG	reggies	2.619	0.732	4.507	0.903		0.485	n		y
49 RNT	renbel	-0.074	-0.788	0.641	0.688	0.550		y	n	
50 SAP	sappi	1.046	0.743	1.348	0.945	1.260	1.068	n	n	n
51 SPU	spur	1.742	1.041	2.443	0.702		0.953	y		y
52 TEG	tegkor	1.527	1.154	1.899	1.260	1.190	1.488	n	n	n
53 TPC	tmpaco	-0.232	-1.759	1.296	0.472		0.019	n		n
54 TRH	tradhld	0.795	-0.039	1.628	-0.848		0.331	y		n
55 TUN	t&n holdings	0.899	0.019	1.779	0.946		1.068	n		n
56 UTR	unitran	0.179	-0.506	0.865	0.711		0.505	n		n
57 VLX	voltex	1.496	0.459	2.533	0.902		0.740	n		n
58 VNT	ventron	0.861	0.307	1.415	0.913	1.450	0.590	n	y	n
59 WAL	waltons	1.722	1.151	2.294	1.150		1.090	y		y
60 WLH	walhold	1.614	1.026	2.203	1.081		1.017	n		y
							n=	49	19	47
							y=	10	4	11
							n%	83.1	82.6	81.0
							y%	16.9	17.4	19.0

et al., 1983: 231). Altogether nineteen per cent of the BFA-NET betas fall outside of theninety five per cent confidence interval of the OLS beta.

The standard errors of estimate for the University of Cape Town Financial Risk Service and the BFA-NET beta estimates were not available for this study, and consequently, conventional inferential tests were again not possible.

Summary and conclusions

This study investigated whether the estimation of the systematic risk component of shares on the JSE can be

improved using transfer function modeling or Multivariate Autoregressive Integrated Moving Average (MARIMA) models. In particular, two propositions were tested.

1. Transfer function modeling will result in estimates of systematic risk which are different from those obtained using conventional OLS regression methods.
2. Transfer function models will provide forecasting results which are better than those provided by betas estimated in the conventional way.

The first proposition could not be tested using conventional inferential tests as the standard errors of estimate of the betas estimated from MARIMA modelling cannot, in general, be measured because of the complex way in which they are calculated, namely the ratio of polynomials. It was found however that 16.9% of the MARIMA beta estimates fell outside the 95% confidence intervals of the respective OLS regression beta estimates. Similar results were obtained when the OLS regression betas were compared to the UCT Financial Risk Service and BFA-NET betas.

It was found that proposition 2 can in general not be supported as it only holds in certain individual cases. This proposition was tested by studying the means of the mean squared error (mse), Theil's U-statistic, and the correlation coefficients of all MARIMA and OLS regression forecasts. It was found that these means were not statistically significantly different at the 0.05 significance level. A factor in favour of the MARIMA modeling process is that the models it produced generated uncorrelated residuals (white noise) when forecasting, as opposed to the OLS regression models which generated correlated residuals. In this sense the MARIMA models are statistically better than the OLS regression models.

The results of this study show that the cross correlations between the differenced and prewhitened index and share returns are in general not statistically significant or only barely so. Table 1 also shows that there are 29 shares for which the OLS regression models are not statistically significant – $p(F)>0.05$. In such cases it is difficult to justify fitting MARIMA, or any other models to the data. If a model has to be obtained however, one is probably better off using OLS regression as it is easier to use and understand, and as the results are similar to those obtained by MARIMA modeling.

For shares with insignificant index correlations and OLS models, one can argue from a statistical point of view that beta equal to zero should be used in the share return forecasts. Such share return forecasts would then not depend on the index returns. One should rather look at the intended use of beta and the consequences of setting it equal to zero. If beta is to be used in the CAPM framework, setting beta equal to zero would not make much sense as this would imply that the expected share return is equal to the risk free rate. If one has no confidence in the estimated beta, it would perhaps be better to set beta equal to one rather than zero, as the beta of all shares tend to one over time (Bodie *et al.*, 1996: 282). The CAPM estimates would then also make more sense as the expected share return would then be equated to the expected market return and not the risk free rate.

This does not mean however that the MARIMA modeling technology does not have a role to play in beta estimation. The OLS regression process can benefit significantly from the use of some of the MARIMA model identification tools such as the cross correlogram. It can point out the size of leads or lags, as well as where the most significant cross correlations are. The share and index return series can then be shifted relative to each other, using the information provided by the cross correlogram. Normal OLS regression can then be performed on the shifted data series. One can even go as far as prewhitening the input and output series before performing OLS regression, but this would make the procedure more complex. There are also cases where a more complex methodology such as MARIMA modeling is required, for example when there are many significant cross correlations between the share and index returns.

Note

1. In lieu of research done in the Faculty of Management, University of the Witwatersrand, Johannesburg, in partial fulfillment of the requirements for the degree of Master of Business Administration.

References

Bodie, Z., Kane, A. & Marcus, A.J. 1996. *Investments.* 3rd Edition. Chicago: Irwin, 937p.

Bowie, D.C. & Bradfield, D.J. 1993. A review of systematic risk estimation on the JSE, *De Ratione*, 7(1):6–22.

Box, G.E.P., Jenkins, G.M. & Reinsel, G.C. 1994. *Time series analysis: forecasting and control.* 3rd ed. Englewood Cliffs: Prentice Hall, 598p.

Bradfield, D.J. & Barr, G.D.I. 1989. Risk estimation in the thinly traded JSE environment, *South African Journal of Business Management*, 20(4):169–173.

Brummer, L. 1997. Personal communication, Director: Bureau for Financial Analysis, University of Pretoria, Pretoria, 5 February.

Cohen, K.J., Hawawini, G.A., Maier, S.F., Schwartz, R.A. & Whitcomb, D.K. 1983. Friction in the trading process and the estimation of systematic risk, *Journal of Financial Economics*, 12:263–278.

Dimson, E. 1979. Risk measurement when shares are subject to infrequent trading, *Journal of Financial Economics*, 7:197–226.

Dimson, E. & Marsh, P.R. 1983. The stability of UK risk measures and the problem of thin trading, *Journal of Finance*, 38:753–783.

Fama, E.F. 1965. The behaviour of stock market prices, *Journal of Business,* 38:34–105.

Firer, C. 1993. Estimating the return parameters of the Capital Asset Pricing Model, *De Ratione*, 7(1):23–39.

Granger, C.W.J., & Newbold, P. 1977. *Forecasting economic time series.* New York: Academic Press, 333p.

Jones, B. 1996. *Statistics TOOLBOX. For use with MATLAB®.* Natick: The MathWorks, Inc.

Kanji, G. K. 1993. *100 statistical tests.* London: Sage Publications.

Leskinen, E. & Tersvirta, T. 1976. Forecasting the consumption of alcoholic beverages in Finland, *European Economic Review,* 8:349–369.

Levin, R.I., & Rubin, D.S. 1991. *Statistics for management.* 5th ed. Englewood Cliffs: Prentice Hall, 862p.

Lintner, J. 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics*, 47:13–37.

Ljung, L. 1987. *System identification. Theory for the user.* Englewood Cliffs: Prentice Hall, 519p.

Ljung, L. 1995. *System identification TOOLBOX. For use with MATLAB®*. Natick: The MathWorks, Inc.

Makridakis, S., Wheelwright, S.C. & McGee, V.E. 1983. *Forecasting: methods and applications*. 2nd ed. New York: John Wiley & Sons, 926p.

Markowitz, H.M. 1959. *Portfolio selection: efficient diversification of investments.* New York: John Wiley & Sons.

Nicholson, M. 1991. Modeling egg consumption. Unpublished MBA research report, Johannesburg: University of the Witwatersrand.

Sharpe, W.F. 1963. A simplified model for portfolio selection, *Management Science*, 9:277–293.

Sharpe, W.F. 1964. Capital asset prices: a Stheory of market equilibrium under conditions of risk, *Journal of Finance*, 19:425–442.

Statman, M. 1987. How many stocks make a diversified portfolio?, *Journal of Financial and Quantitative Analysis*, 22(3):353–363.

Van der Walt, B.E. & Pretorius, W.S. 1994. Notes on revision of the composite business cycle indicators, *SA Reserve Bank Quarterly Bulletin*, September: 29–32.