TESTING A MEAN REVERSION HYPOTHESIS ON THE TORONTO STOCK EXCHANGE

By

Rowan Peter Williams-Short

B.Sc, University of Cape Town, 1983

B.Sc(Hons), University of Cape Town, 1985

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

(Department of Statistics)

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

September 1989

6

ABSTRACT

Monthly returns data for the 150 companies that have been listed on the Toronto Stock Exchange for at least 95% of the period from January 1963 to December 1987 are transformed into continuously compounded returns, adjusted for inflation, gathered into portfolios, and examined for the property known as mean reversion. The model used to describe this scenario comprises two components, namely a random walk and an autoregressive scheme of order one. The study's foundation is a paper by Fama and French(1988), using data from the New York Stock Excange, but in addition to all the estimation performed in that paper, other parameters are estimated. The parameter originally of prime interest is shown to behave in accordance with model predictions, in the case of nearly all 28 portfolios constructed. However, this interpretation is shown to be seriously jeopardised in view of closer scutiny of model parameters, which indicate that the model is a poor approximation to the data.

CONTENTS

List of Tables

List of Figures

1.0 INTRODUCTION

Countless random walk tests for stock markets have been conducted on stock exchanges around the world, and yet the debate over whether the random walk model adequately describes price behaviour seems unresolved, although more recent evidence, sometimes resulting from tests more statistically sophisticated than has been the norm, has tended to dispute the pure random walk model (e.g. Lo and MacKinlay 1988). Many of these tests have been primarily concerned with the fundamental efficient market hypothesis. Jensen (1978) believes that "there is no other proposition in economics which has more solid empirical evidence supporting it than the efficient market hypothesis". An efficient capital market is one that is efficient in processing information. Most importantly, in this scenario prices fully reflect all available information, so that they are fair indicators of firms' true value. Amongst other things, an efficient market hypothesis says that expected returns are positive and constant and that if the true joint distribution of prices of different securities at time t is multivariate normal, then the joint distribution of security returns at time t is multivariate normal. The implication which is of most concern here though, is the capacity for enhanced investment adeptness amongst sectors of the investment community. Under the efficient market hypothesis, there ought not to be any such sector consistently able to make investment decisions resulting in higher returns than comparable decisions made by the rest of the investment community, since this would be implicative of unequal dissemination of relevant information, or disparate utilization of such information.

Now in the random walk model, it is believed that price changes and rate of return changes are simply a series of random numbers, following no discernible pattern. Also, any security price is an unbiased estimate of the true intrinsic value of the security at the particular time, i.e. the price itself is not a random number. Of course this model does not preclude the existence of trend, but does assert that price movements are unpredictable and that hence, in the long run, chartists will not be capable of beating the market. Essentially, the probability distribution for prices, conditional on previous price changes and volumes, equals the marginal price distribution.

In a 1988 paper, Fama and French dispute the position that stock returns are unpredictable, based on a mean reversion hypothesis, which they tested on the New York Stock Exchange (N.Y.S.E.). They estimated that 25 to 45 percent of the variation of three to five year stock returns is predictable from past returns, and add "There is little in the literature that foreshadows our estimates". The idea is that prices take long but temporary swings away from fundamental or intrinsic values (the rational 'means' or expected values). Summers (1986) first translated this economic theory into the statistical hypothesis that prices have slowly decaying stationary components, which Fama and French modelled with the most natural choice, namely an autoregressive scheme of order one, or an AR(1). Previous tests had only calculated autocorrelation between daily or weekly returns. The huge resulting sample sizes made it easy, if somewhat misleading, to find evidence against the hypothesis that the true autocorrelations were zero, even though the estimates were so close to zero as to be of no viable interest to potential investors, who face such factors as commissions on each transaction. Such brief lags understate or fail to detect the importance of mean reversion, which typically unfolds in units of years.

This study, which uses data from the Toronto Stock Exchange (T.S.E.), shows how a structural model for prices, comprising both a random walk and an AR(1) component, may be evolved, chiefly by manipulation of relevant covariance terms, into an empirical test based on continuously compounded returns, for the existence of mean reversion. It shows how the stationary component induces strong negative autocorrelations in long horizon returns, and then examines the data for actual existence of such values.

The methodology here is based largely on the paper by Fama and French but apart from providing a more detailed theoretical development, also establishes a fuller validation of model premises and estimates more model parameters, partly as a means to diagnostic checking. One drawback of this study is that only 25 years of data were available for the T.S.E., whereas 61 years were used on the N.Y.S.E. This means that return horizons as long as those used in the N.Y.S.E. study (up to ten years) are basically intractable without risking serious estimation errors. Nevertheless, horizons of up to eight years were investigated and the ensuing results for the parameter of prime interest are quite encouraging.

In the spirit of the paper by Fama and French, portfolios will be constructed on the basis of firm size and industry, since these factors are known to bear some association to return behaviour. This in no way suggests that traditional methods of portfolio construction (in which an 'optimal' payoff between expected return and risk is strived after) should not be used in conjunction with mean reversion ideas. The latter will assist in the timing decision that must be made by all portfolio investors, and is suggested as a useful complemetary tool in the investor's arsenal.

2.0 THE DATA AND MISSING VALUES

All the data stems from the Toronto Stock Exchange (TSE), and was retrieved from the Laval Master Tape. This data is monthly, beginning, for those companies which were already trading by then, in January 1963. The most recently available version of the Laval tape encompasses data up to and including, again where applicable, December 1987, for a total of 25 years' worth of data. All common stocks that were traded during part or all of this period are initially considered as possibly suitable candidates for inclusion in the study. An exception, however, is a small number of stocks classified under mines (M) or oils (O), which never attained a price level of $5.00 or more at any stage during which they were active. However, all industrial (I) stocks that were traded for any segment of the 1963 - 1987 period are included in the data base, regardless of their price at any time.

It is due to the well documented "noisy" behaviour of daily stock prices that monthly rather than daily data is considered preferential for use in most empirical studies. Furthermore, this study is aimed at detecting the existence of predictable components in fairly long horizon investment schemes and consequently, the discovery of any substantially non-zeroautocorrelations between daily prices at whatever lag, is not of interest here. Such precisely defined lags would surely be spurious, unlikely to be repeated, and intractable to economic theory. On the other hand, investment horizons measured with months or even years as units, are more stable and statistically reliable, and are of more practical use to the investor. The enhanced statistical reliability stems directly from the reduced variability associated with monthly returns.

Under the above definition for the scope of stocks eligible for the Laval tape, 1433 companies are included. Each company is included once only, irrespective of whether or not it was traded under more than one name during the 25 year period. The large majority of companies however, do not have actual data for the entire 25 year period, either because they only began trading on the TSE after the beginning of 1963, or ceased to be listed before the end of 1987, or were for one reason or another delisted during some intervening period. Furthermore, for many companies there are spurious single months, sometimes occurring more than once, for which no datum is available. Since it is crucial to have data that is as reliable as is feasibly attainable without the adoption of an overly exclusive policy whose residual base is statistically unworkable, and since long horizons will be involved in the testing procedures, a compromising scheme of selecting only those companies for which at least 95% of the 25 year period's data is available, was decided upon. This translates into a total number of permissable missing values corresponding to at most 15 months for each company selected. This figure includes, for the trading history of each company, the very first month, which is necessarily missing due to the form of the data actually utilised, as described below. After examination of each of the companies on the Laval tape for admission under these criteria, 150 companies remained. For these companies, the few remaining missing values were replaced by a technique that admits to the time series structure of the data. This is described later.

Although the statistical model to be presented relates primarily to the stock prices themselves, as will be shown in the theoretical development, the testing procedures and hence the basic data, involve monthly returns. The Laval tape provides data for the monthly **rates** of returns, r_t^*, adjusted for capital changes, namely

$$r_t^* = \frac{P_t + DIV_t + DIS_t - P_{t-1}}{P_t}$$

where P_t = share price taken as month-end price for month t and split adjusted as needed eg. if a stock splits 2-for-1 for the first time in month t, the split factor is 0.5 ;

on the second 2-for-1 split, the split factor becomes 0.5x(0.5) = 0.25 ;

if subsequently a 1% stock dividend was distributed, the split factor becomes 100/101x0.25 = 0.2475, etc.

DIV_t = cash dividend, split adjusted as needed, in (ex-dividend) month t

DIS_t = distribution of capital (eg. value of distributed rights, per share), also split adjusted where necessary.

Each of these quantities is adjusted for splits or reverse splits, by dividing by the appropriate split factor.

The definition of the monthly rate of return explains why the first month's return data for each company is necessarily missing, as the value "P_0" is non-existent. It would clearly be a meaningless exercise to backcast such a value. Whilst care is seen to have been taken with regard to most extraordinary events during any month for each stock, some factors, whose impact is difficult to gauge, such as the actual number of trading days in each month, are omitted, resulting in a not entirely 'crisp' data set.

2.1 DOUBLE EXPONENTIAL SMOOTHING

This methodology, also known as discounted least squares, was used to estimate any spurious missing values still encountered in the data. It was selected because of its suitability in many practical time series frameworks, and because of its outstanding amenability to computing, based on a recursive algorithm. Simply put, exponentially successively less weight is attached to 'older' observations when forecasting ; the rate of exponential decline depends on the local nature of the data. Since the technique is widely known, only a brief overview is presented here. Generally, we consider models of the form

$$r_n^*(p) = \sum_{i=1}^{m} \beta_i f_i(p) = f(p)\beta$$

where

m is the dimension of the vector ß (m=2 for double exponential smoothing),

n is the length of the series up to, but not including, the missing value,

p is the number of periods ahead that we wish to forecast (usually just 1).

In general exponential smoothing, we then estimate the parameters by minimizing

$$\sum_{j=0}^{n-1} \omega^j [r^*_{(n-j)} - f(-j)\beta]^2$$

The "known" constant ω ($|\omega| < 1.0$) is a discount factor that discounts past observations exponentially. In double exponential smoothing, we assume that any trend is only locally linear. It may easily be shown that this leads to computing forecasts from the function

$$\hat{r}_n^*(p) = (2 + \frac{1-\omega}{\omega}p)S_n^{[1]} - (1 - \frac{1-\omega}{\omega}p)S_n^{[2]}$$

where

$$S_n^{[1]} = (1-\omega)\sum_{j=0}^{n-1} \omega^j r^*_{(n-j)} = (1-\omega)r^*_n + \omega S_{n-1}^{[1]}$$

$$S_n^{[2]} = (1-\omega)^2 \sum_{j=0}^{n-1} (j+1)\omega^j r^*_{(n-j)} = (1-\omega)S_n^{[1]} + \omega S_{n-1}^{[2]}$$

Initial estimates are usually obtained from the two equations

$$S_0^{[1]} = \hat{\beta}_0 - \frac{\omega}{1-\omega}\hat{\beta}_1$$

$$S_0^{[2]} = \hat{\beta}_0 - 2\frac{\omega}{1-\omega}\hat{\beta}_1$$

where

$$\hat{\beta}_1 = \frac{12\sum_{t=1}^{n} (t - \frac{n+1}{2})r^*_n}{(n^3 - n)}$$

$$\hat{\beta}_0 = \overline{r^*} - \hat{\beta}_1\frac{n+1}{2}$$

The algorithm is due mainly to G.M Brown (1962), who has suggested that ω be chosen to lie between 0.84 and 0.97. This is primarily because, the closer ω is to 1.0, the more likely it is that the data <u>are</u> described by a locally linear trend model, as assumed by double exponential smoothing. In the limiting case, when $\omega = 1.0$, the best forecast is achieved by fitting a constant linear trend model to all available observations and giving each observation the same weight. Furthermore, if the underlying mean reversion is long-term, this is supportive of a larger ω, in which case relatively more weight is attached to 'older' observations.

In fact, ω is chosen by minimizing the squared one - step - ahead forecast errors, i.e. the function

$$SSE(\omega) = \sum_{t=1}^{n} [r_t^* - \hat{r}_t^*(p)]^2$$

So, small increments in ω are tested and then that value of ω which minimizes this SSE is used in the forecasting algorithm.

In all cases, the optimal ω indeed turned out to be close to 1.0. Whenever more than one missing value occurred in any series (not counting the very first "observation"), ω was found from the data up to the first missing value, which was subsequently forecast, and then ω was re-estimated up to the next occurrence of a missing value.

Technical note. In standard ARIMA notation, double exponential smoothing can be shown to be equivalent to the process

$$(1 - B)^2 r_t = (1 - \omega B)^2 \varepsilon_t$$

2.2 CONTINUOUSLY COMPOUNDED RETURNS

Prices (split and capital adjusted) for months t and (t-1) are $P_{(t)}$ and $P_{(t-1)}$.

If interest were compounded at nominal rate i, m times per month, then

$$P_{(t)} = P_{(t-1)}(1.0 + \frac{i}{m})^m$$

Therefore, as m --> ∞ , i.e. when we have continuous compounding,

$$P_{(t)} = P_{(t-1)} \cdot e^i$$

$$\Rightarrow i = \ln \frac{P_{(t)}}{P_{(t-1)}} = \ln P_{(t)} - \ln P_{(t-1)} \equiv r_{(t)}$$

i.e. $r_{(t)}$ is the continously compounded nominal return in month t

(The effective return described before is

$$r^*_{(t)} = \frac{P_{(t)} - P_{(t-1)}}{P_{(t-1)}} = e^i - 1.0 \text{)}$$

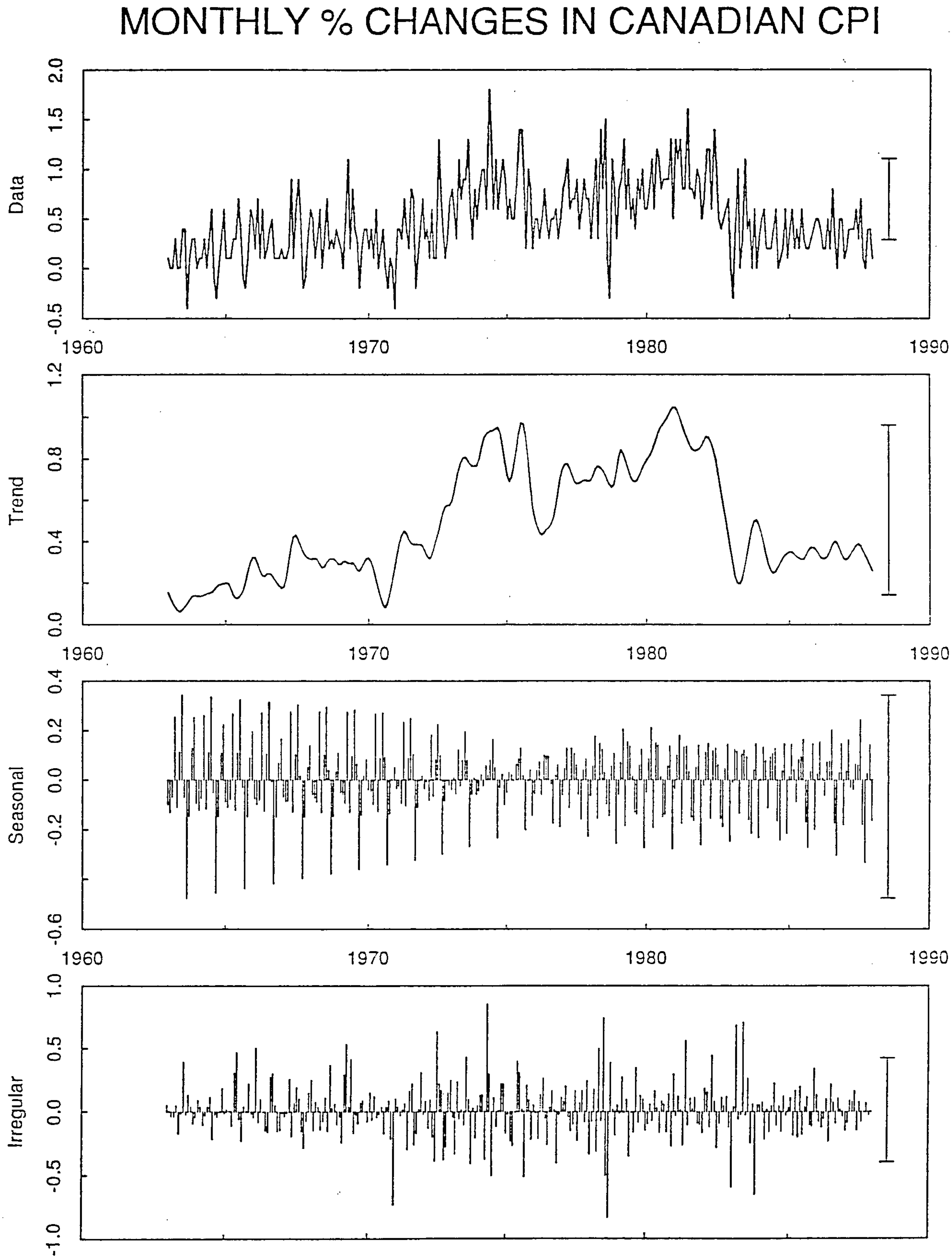
$$\therefore\ r_{(t)} = \ln(1.0 + r^{*}_{(t)})$$

So each return recorded on the Laval tape is transformed according to this rule. (Note that in subsequent sections, it will be convenient to alter the return notation slightly; there r_t will be written as $r_{(t-1,t)}$, as a special case of $r_{(t-T,t)}$ which is just the return from time (t-T) to time t.) The chief advantage of using continuously compounded returns is that it becomes legitimate to simply add consecutive monthly returns to obtain longer horizon returns.

i.e. $r_{(t-T,t)} = r_{(t-T,t-T+1)} + \ldots + r_{(t-1,t)}$. This would not be correct for $r_{(t)}^{*}$.

These returns are then adjusted for inflation by using the monthly increases in the Canadian Consumer Price Index (CPI), (Source : The CPI, Statistics Canada), over the period 1963 to 1987, to obtain real rates of return. Before adjustment was made with the raw inflation figures, the whole inflation series was examined for seasonality, the presence of which would interfere with recovery of true returns at any time. The plots shown overleaf, however, indicate that the seasonality is not serious, and that transformations of the inflation series are not necessary. The three plots below that of the data itself show a decomposition of monthly CPI percentage changes into the following components - trend, seasonal, irregular. To the right of each plot is a bar which portrays the relative vertical scaling of that plot. In the third plot, horizontal lines are drawn at the midmean of the seasonal effect for each month, giving a general picture of the seasonal effect. The vertical lines emanating from these midmeans show the actual seasonal effects computed for the various years of data.

figure 1


MONTHLY % CHANGES IN CANADIAN CPI
Data
-0.5
0.0
0.5
1.0
1.5
2.0
Trend
0.0
0.4
0.8
1.2
Seasonal
-0.6
-0.2
0.0
0.2
0.4
Irregular
-1.0
-0.5
0.0
0.5
1.0
1960
1970
1980
1990

3.0 THEORY AND METHODOLOGY

The model proposed here is one for stock prices and comprises not only the traditional random walk component, but also a stationary component exemplified by a Box-Jenkins type autoregressive scheme of order one. The belief is that instead of the market being entirely efficient, prices in fact tend to stray from fundamental values, but that such straying is only temporary and ends with a reversion to the value consistent with rational expectations. Summers (1986) first equated this with the statistical hypothesis that prices have slowly decaying stationary components. However, this mean reversion might occur slowly and thus only be revealed in horizons of several years. In the technical development of this section, it will be shown that we may test empirically for the existence or absence of a stationary component by regressing T-period returns on preceeding T-period returns and examining whether the estimated slope parameters (which are also the estimated autocorrelations) are significantly different from zero. If these slopes are found to be non-zero then some fraction of the variation of returns is explained by mean reversion; the greater this fraction the better, from a viewpoint of predictability. The value of T will be varied in a logical attempt to identify the length of the reversion period for different portfolios of stocks.

It will also be shown that the presence of a slowly decaying stationary component induces negative autocorrelations in long horizon returns. At the same time, it will be explained why these autocorrelations are weak (approximately zero) for short horizon returns, for instance from day to day or week to week.

To quote from Fama and French (1988) for the economic theory, "There are two competing economic stories for strong predictability of long horizon returns due to slowly decaying price components. Such price behaviour is consistent with common models of an irrational market in which stock prices take long temporary swings away from fundamental values. But the predictability of long horizon returns can also result from time-varying equilibrium expected returns generated by rational pricing in an efficient market. Poterba and Summers (1987) show formally how these opposite views can imply the same price behaviour. Expected returns correspond roughly to the discount rates that relate a current stock price to expected future dividends. Suppose that investor tastes for current versus risky future consumption and the stochastic evolution of the investment opportunities of firms result in time-varying equilibrium expected returns that are highly autocorrelated but mean reverting. Suppose that shocks to expected returns are uncorrelated with shocks to rational forecasts of dividends. Then a shock to expected returns has no effect on expected dividends or expected returns in the distant future. Thus the shock has no long term effect on expected prices. The cumulative effect of a shock on expected returns must be exactly offset by an oppposite adjustment in the current price."

The foregoing theory leads (not uniquely) to the following model.

(1) $P_{(t)} = q_{(t)} + z_{(t)}$ where $q_{(t)}$ is a Random Walk

(and $P_{(t)}$ is the natural logarithm of the price at time t , as explained above)

(2) $q_{(t)} = \mu + q_{(t-1)} + \eta_{(t)}$ where μ is drift and $\eta_{(t)} \sim$ W.N.$(0,\psi^2)$

(3) $z_{(t)} = \Phi z_{(t-1)} + \varepsilon_{(t)}$ where $\varepsilon_{(t)} \sim$ W.N.$(0,\sigma^2)$ and $|\Phi| < 1.0$

The process (3) is autoregressive of order 1, or AR(1), so this model is a mixture of random walk and stationary components. Now,

$r_{(t,t+T)} = P_{(t+T)} - P_{(t)}$ ($P_{(t)}$ is capital adjusted $\forall$ t)

$$= [q_{(t+T)} - q_{(t)}] + [z_{(t+T)} - z_{(t)}]$$

(but from (2), $q_{(t)} - q_{(t-1)} = \mu + \eta_{(t)}$ i.e. 1st difference of a R.W. produces W.N.)

$$= T\mu + (\eta_{(t+T)} + \eta_{(t+T-1)} + \ldots + \eta_{(t+1)}) + \Phi(z_{(t+T-1)} - z_{(t-1)}) + \varepsilon_{(t+T)} - \varepsilon_{(t)}$$

i.e. the random walk component of the price produces white noise in the returns, a desirable feature of the model, from empirical considerations. Next, consider the slope in the regression of $z_{(t+T)} - z_{(t)}$ on $z_{(t)} - z_{(t-T)}$. This is of course just the autocorrelation of consecutive T-period changes in $z_{(t)}$,

$$\rho(T) = \frac{Cov[z_{(t+T)} - z_{(t)} \,;\, z_{(t)} - z_{(t-T)}]}{\sqrt{Var[z_{(t+T)} - z_{(t)}] \cdot Var[z_{(t)} - z_{(t-T)}]}} = \frac{Cov[z_{(t+T)} - z_{(t)} \,;\, z_{(t)} - z_{(t-T)}]}{Var[z_{(t+T)} - z_{(t)}]}$$

since $z_{(t)}$ is a stationary process.

Two interpretations of this numerator term are now given.

Firstly,

$$Cov[z_{(t+T)} - z_{(t)} \,;\, z_{(t)} - z_{(t-T)}]$$

$$= Cov[z_{(t)} \,;\, z_{(t+T)}] - Var[z_{(t)}] - Cov[z_{(t-T)} \,;\, z_{(t+T)}] + Cov[z_{(t-T)} \,;\, z_{(t)}]$$

(4) $$= 2Cov[z_{(t)} \,;\, z_{(t+T)}] - \frac{\sigma^2}{1 - \Phi^2} - Cov[z_{(t)} \,;\, z_{(t+2T)}]$$

due to the stationarity of $z_{(t)}$.

But since $z_{(t)}$ is an AR(1) process,

$$Cov[z_{(t)} \,;\, z_{(t+T)}] = \frac{\sigma^2\Phi^T}{1 - \Phi^2} \to 0 \text{ as } T\uparrow\infty$$

since $|\Phi| < 1.0$ by assumption.

It follows that

$$\mathrm{Cov}[z_{(t+T)} - z_{(t)} \,;\, z_{(t)} - z_{(t-T)}] \to \frac{-\sigma^2}{1 - \Phi^2} \text{ as } T\uparrow\infty$$

And examining the denominator in the autocorrelation term,

$$\mathrm{Var}[z_{(t+T)} - z_{(t)}] = \mathrm{Var}[z_{(t+T)}] + \mathrm{Var}[z_{(t)}] - 2\mathrm{Cov}[z_{(t+T)} \,;\, z_{(t)}]$$

$$= \frac{2\sigma^2}{1 - \Phi^2} - 2\mathrm{Cov}[z_{(t+T)} \,;\, z_{(t)}] \to \frac{2\sigma^2}{1 - \Phi^2} \text{ as } T\uparrow\infty$$

Thus the slope in the regression of $z_{(t+T)} - z_{(t)}$ on $z_{(t)} - z_{(t-T)}$ converges to $\{\frac{-\sigma^2}{1 - \Phi^2}\} \div \{\frac{2\sigma^2}{1 - \Phi^2}\} = -0.5$ for large T.

Secondly, consider the expected value of the T-period change in $z_{(t)}$, as exemplified by the change from time t to time $(t + T)$ as before. If $E_t[\,f\,]$ is the conditional expectation of the function f, given $\{z_{(s)}\,,\, s \leq t\}$ with respect to time t, then

$$E_t[z_{(t+T)} - z_{(t)}] = E_t[\Phi z_{(t+T-1)} + \varepsilon_{(t+T)} - z_{(t)}]$$

(but $E_t[\varepsilon_{(s)}] = 0 \;\forall\; s > t$ since $\varepsilon_{(s)}$ is indep. of $\{z_{(\gamma)}\,,\, \gamma \leq t\}$ for $s > t$)

$$= E_t[\Phi(\Phi z_{(t+T-2)} + \varepsilon_{(t+T-1)}) - z_{(t)}]$$

$$\vdots$$

(5) $$= E_t[\Phi^T z_{(t)} - z_{(t)}] = (\Phi^T - 1)z_{(t)}$$

$$\therefore \mathrm{Var}\{E_t[z_{(t+T)} - z_{(t)}]\} = (\Phi^T - 1)^2 \mathrm{Var}\{z_{(t)}\} = (\Phi^T - 1)^2 \frac{\sigma^2}{1 - \Phi^2}$$

But, from (4) the numerator in the slope of the regression is

$$2\mathrm{Cov}[z_{(t)}\ ;\ z_{(t+T)}] - \frac{\sigma^2}{1 - \Phi^2} - \mathrm{Cov}[z_{(t)}\ ;\ z_{(t+2T)}]$$

$$= \frac{2\sigma^2\Phi^T}{1 - \Phi^2} - \frac{\sigma^2}{1 - \Phi^2} - \frac{\sigma^2\Phi^{2T}}{1 - \Phi^2}$$

$$= -(1 - \Phi^T)^2 \frac{\sigma^2}{1 - \Phi^2} \tag{6}$$

$$= -\mathrm{Var}[E_t\{z_{(t+T)} - z_{(t)}\}]$$

i.e. the covariance term is minus the variance of the expected T-period change in the AR(1) component of the model. Thus, the slope in the regression of $z_{(t+T)} - z_{(t)}$ on $z_{(t)} - z_{(t-T)}$ is

$$\rho(T) = \frac{-\mathrm{Var}[E_t\{z_{(t+T)} - z_{(t)}\}]}{\mathrm{Var}[z_{(t+T)} - z_{(t)}]}$$

$$\rightarrow -0.5 \text{ as } T\uparrow\infty \quad ,$$

as demonstrated before.

(From the last expression, it is seen that the slope parameter may be interpreted as the proportion of variability of T-period changes in $z_{(t)}$ that is due to expected change.)

But from (6), we see that the slope parameter is approximately zero for small T, presuming, from the economic theory, that Φ is close to 1.0. So these two interpretations highlight a fundamental point, namely that slow mean reversion in stock prices can easily go undetected with short return horizons.

The above is merely a theoretical development, and does not yet provide a framework for any empirical testing, since $z_{(t)}$ is of course unobservable. It is the monthly returns, $r_{(t)}$, which have been observed for each stock. Consider again the slope in a T-period autoregression, but now dealing with the regression of $r_{(t,t+T)}$ on $r_{(t-T,t)}$

$$\beta_{(T)} = \frac{\mathrm{Cov}[r_{(t,t+T)} \,;\, r_{(t-T,t)}]}{\sqrt{\mathrm{Var}[r_{(t,t+T)}]\cdot \mathrm{Var}[r_{(t-T,t)}]}}$$

But,

$$\mathrm{Var}[r_{(t,t+T)}] = \mathrm{Var}[P_{(t+T)} - P_{(t)}]$$

because $P_{(t)}$ is the natural logarithm of the price.

$$=\mathrm{Var}[T\mu + \{\eta_{(t+T)} + \ldots + \eta_{(t+1)}\} + \Phi\{z_{(t+T-1)} - z_{(t-1)}\} + \varepsilon_{(t+T)} - \varepsilon_{(t)}]$$

$$= T\psi^2 + 2\Phi^2[\frac{\sigma^2}{1-\Phi^2} - \frac{\sigma^2\Phi^T}{1-\Phi^2}] + 2\sigma^2$$

(recall that ψ^2 is the variance of the White Noise process $\eta_{(t)}$)

$$= T\psi^2 + \frac{2\sigma^2\Phi^2}{1-\Phi^2}[1 - \Phi^{T-2}] + 2\sigma^2$$

$$= \mathrm{Var}[r_{(t-T,t)}]$$

$$\therefore \beta_{(T)} = \frac{\mathrm{Cov}[\{(q_{(t+T)} - q_{(t)}) + (z_{(t+T)} - z_{(t)})\} \,;\, \{(q_{(t)} - q_{(t-T)}) + (z_{(t)} - z_{(t-T)})\}]}{\mathrm{Var}[r_{(t-T,t)}]}$$

$$= \mathrm{Cov}\{(T\mu + \eta_{(t+T)} + \cdots + \eta_{(t+1)}) + (z_{(t+T)} - z_{(t)})\};$$
$$\{(T\mu + \eta_{(t)} + \cdots + \eta_{(t-T+1)}) + (z_{(t)} - z_{(t-T)})\} \div \{\mathrm{Var}[r_{(t-T,t)}]\}$$

but $\eta_{(t)} \sim$ W.N. $\Rightarrow$ $\mathrm{Cov}[\eta_{(t)} \,;\, \eta_{(s)}] = 0 \;\; \forall \;\; t \neq s$, T is fixed ,

and $\eta_{(t)}$ is independent of $\varepsilon_{(t)}$

$$= \frac{\mathrm{Cov}[\{z_{(t+T)} - z_{(t)}\} ; \{z_{(t)} - z_{(t-T)}\}]}{\mathrm{Var}[r_{(t-T,t)}]}$$

$$= \frac{\rho(T) \cdot \mathrm{Var}[z_{(t+T)} - z_{(t)}]}{\mathrm{Var}[r_{(t-T,t)}]}$$

$$(7) \qquad = \frac{-\mathrm{Var}[E_t\{z_{(t+T)} - z_{(t)}\}]}{\mathrm{Var}[r_{(t-T,t)}]}$$

From expression (7), it is clear why it is claimed that $ß(T)$ may be interpreted as the proportion of the variance of T-period returns that is explained by the mean reversion of a slowly decaying stationary component of the price, exemplified here by an AR(1). Furthermore, since we can re-write expression (7) as :

$$\frac{\rho(T) \cdot \mathrm{Var}[z_{(t+T)} - z_{(t)}]}{\mathrm{Var}[z_{(t+T)} - z_{(t)}] + \mathrm{Var}[q_{(t+T)} - q_{(t)}]} ,$$

we may immediately assess two extreme cases :

(1) The price has no stationary component
$\Rightarrow z_{(t)} \equiv 0 \quad \forall t \Rightarrow \rho(T) = 0 \Rightarrow \beta_{(T)} = 0$

(2) The price has no random walk component
$\Rightarrow q_{(t)} \equiv 0 \ , \forall t \Rightarrow \beta_{(T)} = \rho(T) \Rightarrow \beta_{(T)} \rightarrow -0.5$ as $T \uparrow \infty$

However, when the price consists of a mixture of stationary and random walk components, we anticipate that the behaviour of $ß_{(T)}$ will vary as T varies. It has been established that

(i) $\mathrm{Var}[q_{(t+T)} - q_{(t)}] = T\psi^2$
(ii) $\mathrm{Var}[z_{(t+T)} - z_{(t)}] = \frac{2\sigma^2}{1 - \Phi^2}[1 - \Phi^T]$
(iii) $\rho(T) \cong 0$ for small T, and $\rho(T) \rightarrow -0.5$ for large T

So, for small T, we expect $ß_{(T)}$ to be approximately zero and for large T, we expect the quantity (i) to dominate, leading again to $ß_{(T)}$ being close to zero. But, for some intermediate values of T, we expect $\beta_{(T)}$ to reach a (negative) minimum. Overall, a plot of $\beta_{(T)}$ versus T ought to yield a parabola-like curve when the stock price is a mixture of random walk and sationary components. By observing the behaviour of $\beta_{(T)}$, it will thus be possible to gauge the relevance of including a stationary component in the model for stock prices. If $\beta_{(T)}$ is non-zero, not only is the inclusion of a stationary component significant, but also, the magnitude of $\beta_{(T)}$ is a direct measure of the predictable proportion of variance in portfolio returns.

3.1 CONSTRUCTING PORTFOLIOS AND SECTOR INDICES

It is chiefly to combat the problem of high variability associated with single firm returns that empirical analysis will be undertaken for portfolios and sector indices. Furthermore, inferences made from portfolio and sector indices rather than individual stocks are clearly concomitant with enhanced reliability. Of course reasonable portfolios may be constructed in many ways, such as those with the commonly desired property of maximum expected return for minimum expected variance, for a fixed number of firms to be included. However, a simpler principle for portfolio construction, with which easier interpretation is associated, will suffice here. Several studies (eg, King 1966, Banz 1981, Huberman and Kandel 1985) have demonstrated substantial associations between firm size and return behaviour. Also, the homogeneity of firms within a sector or industry tends to lead to a degree of similarity in return behaviour, obviously as a result of industry-wide factors such as new legislation peculiar to one industry. In this light, it is somewhat regrettable that T.S.E. firms are classified into only three sectors. (Within the 150 most long-lived firms on the T.S.E. the majority are classified under Industrials.)

Ten size-defined (or decile) portfolios were constructed; size is determined by the simple definition of price times number of outstanding shares. For each of these ten groups of fifteen companies, as well as for the three sector indices, both equally weighted and value(size) weighted monthly returns were computed. Within each such portfolio, the value weighted returns assign greater representation of the over-all portfolio behaviour to the larger firms. In fact, for some value weighted decile portfolios, the behaviour is heavily dominated by two or three firms. Equally weighted portfolios give a fair representation to all firms in any decile portfolio, but are less realistically representative of actual investors' returns.

As a final 'safety net', both an equally and value weighted all-share index were created. This protects against the tendency, especially of stocks with unusually high or low returns, to move across deciles over the years. Thus a total of 28 return sequences are examined in detail.

4.0 CHECKING ASSUMPTIONS AND THE CONTRIBUTION OF THE STATIONARY COMPONENT, AND ESTIMATING Φ

None of the model predictions are of any worth if the assumptions are invalid. It must also be established that the inclusion of the stationary component appears to be statistically worthwhile. Finally, parameter estimation not only leads to a fuller appreciation of the data, but serves as a further indication of model validity. These three topics are dealt with in this section.

Firstly, from the expressions already derived for the returns, it is seen that the model assumes $r_{(t)}$ to be stationary (not $P_{(t)}$). Second order or weak stationarity may be tested in several ways, and two of these are used here. Since stationary processes should exhibit non-changing means, we may simply fit a linear regression, with t as the independent variable, to the sequences of returns, and ascertain that the slope parameter is zero. Also, if we fit an autoregressive model to the series, the residuals from this regression should display no serial correlation. Both the Durbin-Watson statistic and the first order autocorrelation of the residuals will be calculated for each portfolio. Note that in the absence of any serial correlation, the former of these should equal 2.0 and the latter 0.0. The order of the autoregression is robust in this test, and 10 is used here.
To facilitate interpretation of the slope, the ratio of estimated slope to the standard error of the estimated slope is given for each portfolio. We would have cause for concern if any of these ratios were greater than, say, 2.0. The results for both the Durbin-Watson statistic and the A.C.F.(1) of the residuals are so clear cut as to make discussion unnecessary. The estimated slopes indicate that the data do indeed seem to have constant means. Overall, it would appear to be quite safe to proceed with model estimation, content that we are working with stationary data.

Table (I)

Durbin-Watson statistics and ACF(1)'s of the autoregression residuals, and slope estimates from simple linear regression.

(A) Equal weighting

PORTFOLIO	D-W	A.C.F.(1)	SLOPE	SLOPE/(S.E. SLOPE)
1	1.996	0.0001	0.0	0.0
2	1.994	0.0021	0.000005	0.14
3	1.988	0.0042	0.000016	0.48
4	1.989	0.0049	-0.000005	0.15
5	2.006	-0.0039	0.000018	0.55
6	1.991	0.0039	-0.000016	0.53
7	2.001	-0.0007	-0.000019	0.57
8	1.997	0.0107	-0.000015	0.42
9	1.999	-0.0005	0.000012	0.33
10	2.007	-0.0039	0.000067	1.58
Industrials	1.993	0.0029	0.000009	0.32
Mines	1.999	-0.0005	-0.000009	0.18
Oils	1.998	0.0004	-0.000016	0.27
All-shares	1.995	0.0011	-0.000002	0.57

(B) Value weighting

PORTFOLIO	D-W	A.C.F.(1)	SLOPE	SLOPE/(S.E. SLOPE)
1	2.002	-0.0031	0.000001	0.04
2	1.994	0.0021	0.000005	0.13
3	1.998	0.0042	0.000017	0.48
4	1.989	0.0051	-0.000007	0.19
5	2.005	-0.0033	0.000017	0.50
6	1.991	0.0041	-0.000017	0.54
7	2.000	-0.0040	-0.000015	0.43
8	1.979	0.0096	-0.000018	0.50
9	2.000	-0.0008	0.000013	0.35
10	2.007	-0.0066	0.000021	0.59
Industrials	1.997	0.0004	0.0	0.0
Mines	2.001	-0.0012	-0.000004	0.07
Oils	2.002	-0.0017	-0.000014	0.21
All-shares	1.997	0.0002	0.000004	0.12

In checking the suitability of the model proposed, the element of prime interest is whether the inclusion of a stationary component in addition to the random walk component is justified. More specifically, we may consider the hypotheses

$$H_0 : P_{(t)} = \mu + P_{(t-1)} + \eta_{(t)} \qquad \text{i.e. a pure random walk}$$

$$\begin{aligned} H_a : P_{(t)} &= q_{(t)} + z_{(t)} \\ q_{(t)} &= \mu + q_{(t-1)} + \eta_{(t)} \\ z_{(t)} &\text{ is non-degenerate and stationary} \end{aligned}$$

But a necessary and sufficient condition for the model H_a to be identical to H_0 is that $z_{(t)}=0$, a.s , for all t. In this case ,

$$r(t)=P(t)-P(t-1)=\mu+\eta(t)$$

i.e., $\{r_{(t)}-m\}$ is a $WN(0,\sigma_\eta^2)$.

The graphs(figures 1 - 3) and table 2 presented overleaf show A.C.F.'s estimated up to lag 10, for each of the 28 portfolios or sector indices. These quantities are estimated in the usual way, namely

$$\widehat{ACF}(\alpha) = \frac{\sum_{t=\alpha}^{n} (r_{(t)} - \bar{r})(r_{(t-\alpha)} - \bar{r})}{\sqrt{\sum (r_{(t)} - \bar{r})^2 \cdot \sum (r_{(t-\alpha)} - \bar{r})^2}} \cong \frac{\sum_{t=\alpha}^{n} (r_{(t)} - \bar{r})(r_{(t-\alpha)} - \bar{r})}{\sum (r_{(t)} - \bar{r})^2}$$

and $\bar{r} = \frac{1}{n}\sum_{t=1}^{n} r_{(t)}$

The dashed lines in each plot are drawn at +1,-1,+2 and -2 times the standard errors (which are $(n)^{-1/2}$ under the hypothesis of white noise), to permit objective consideration of the hypothesis in question. As may have been anticipated from the N.Y.S.E. study, it is the smaller firms, namely those in portfolios 6 to 10, which especially violate the pure random walk model, as seen particularly from the A.C.F.'s and P.A.C.F.'s at lag one. For each of these portfolios, these two functions (which are by definition equal at lag one) are well above two standard errors from zero, leading to the conclusion that Φ is not equal to zero. Looking at the A.C.F.'s at lag 2, we see that in the case of portfolios 8 to 10, (which represent the smallest 45 companies) the hypothesis that $\Phi = 0$ is also clearly violated. For the portfolios of larger firms, it is not clear whether a pure random walk model can be

Table (II)

Autocorrelations and partial autocorrelations for each of the 28 portfolios or sector indices.

(A) Equal weighting

PORTFOLIO	A.C.F.(1)=P.A.C.F.(1)	A.C.F.(2)	P.A.C.F.(2)
1	0.0443	-0.0513	-0.0533
2	0.1062 *	-0.1018 *	-0.1143 *
3	0.1412 **	-0.1022 *	-0.1246 **
4	0.0559	-0.0971 *	-0.1005 *
5	0.1127 *	-0.0208	-0.0339
6	0.1581 **	-0.0205	-0.0467
7	0.2046 **	-0.0466	-0.0923 *
8	0.1677 **	-0.0818 *	-0.1131 *
9	0.1726 **	0.0791 *	0.0508
10	0.3345 **	0.0522	-0.0672 *
Industrials	0.2412 **	-0.0210	-0.0840 *
Mines	0.1042 *	-0.0960 *	-0.1080 *
Oils	0.0022	0.0003	0.0003
All-shares	0.1895 **	0.0208	-0.0156

(B) Value weighting

PORTFOLIO	A.C.F.(1)=P.A.C.F.(1)	A.C.F.(2)	P.A.C.F.(2)
1	0.0441	-0.0493	-0.0514
2	0.1045 *	-0.1031 *	-0.1153 *
3	0.1412 **	-0.1022 *	-0.1246 **
4	0.0574	-0.0979 *	-0.1016 *
5	0.1145	-0.0264	-0.0400
6	0.1634 **	-0.0071	-0.0348
7	0.2046 **	-0.0458	-0.0905 *
8	0.1454 **	-0.0979 *	-0.1216 **
9	0.1602 **	0.0671 *	0.0426
10	0.1562 **	-0.0528	-0.0791 *
Mines	0.0566	-0.1073 *	-0.1108 *
Oils	0.0158	-0.0483	-0.0485
All-shares	0.0971 *	-0.0761 *	-0.0864 *

Legend * The value is above or below one standard error from zero
** The value is above or below two standard errors from zero

Autocorrelations Of Equally Weighted Portfolio Returns

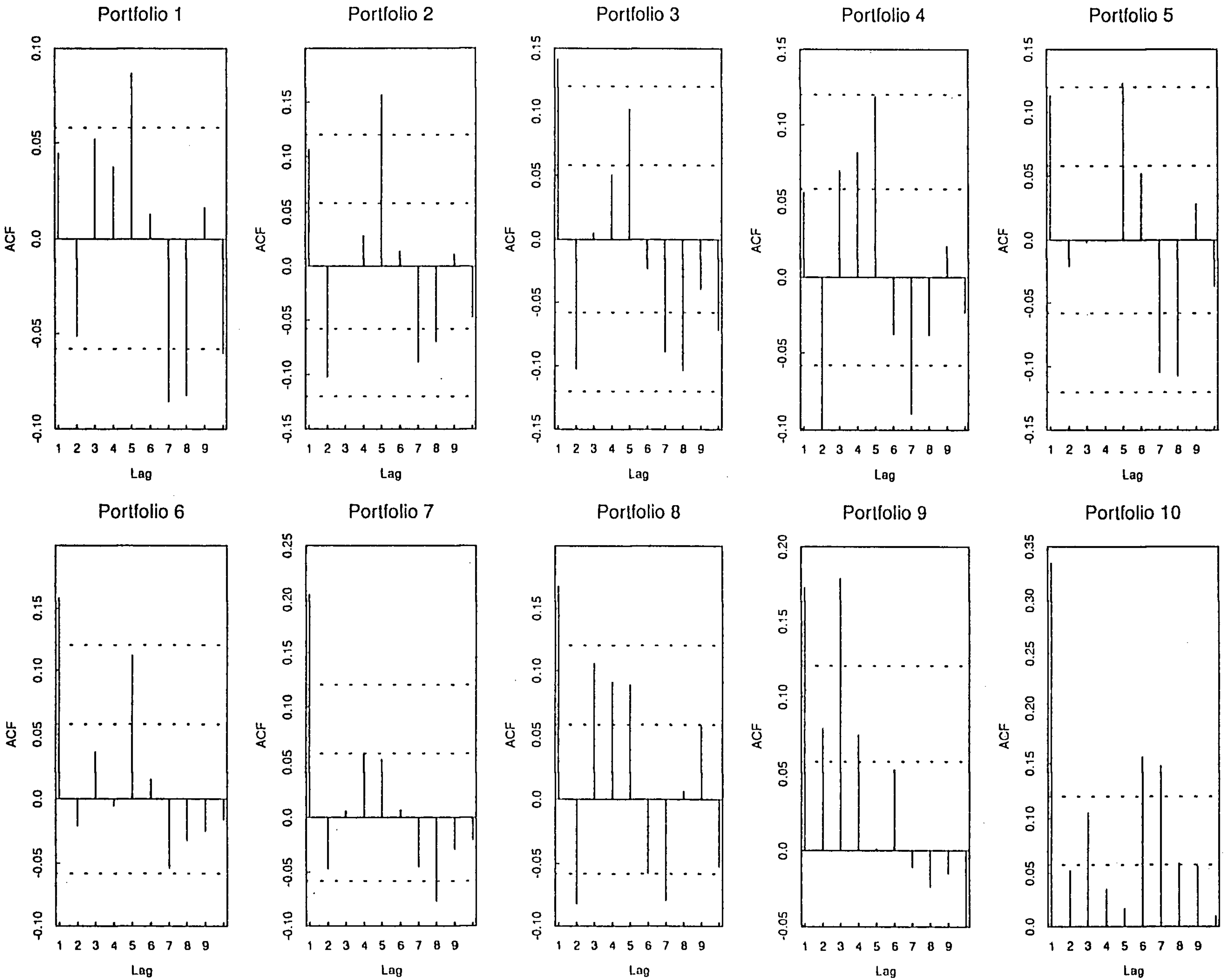


figure 2

Autocorrelations Of Value Weighted Portfolio Returns

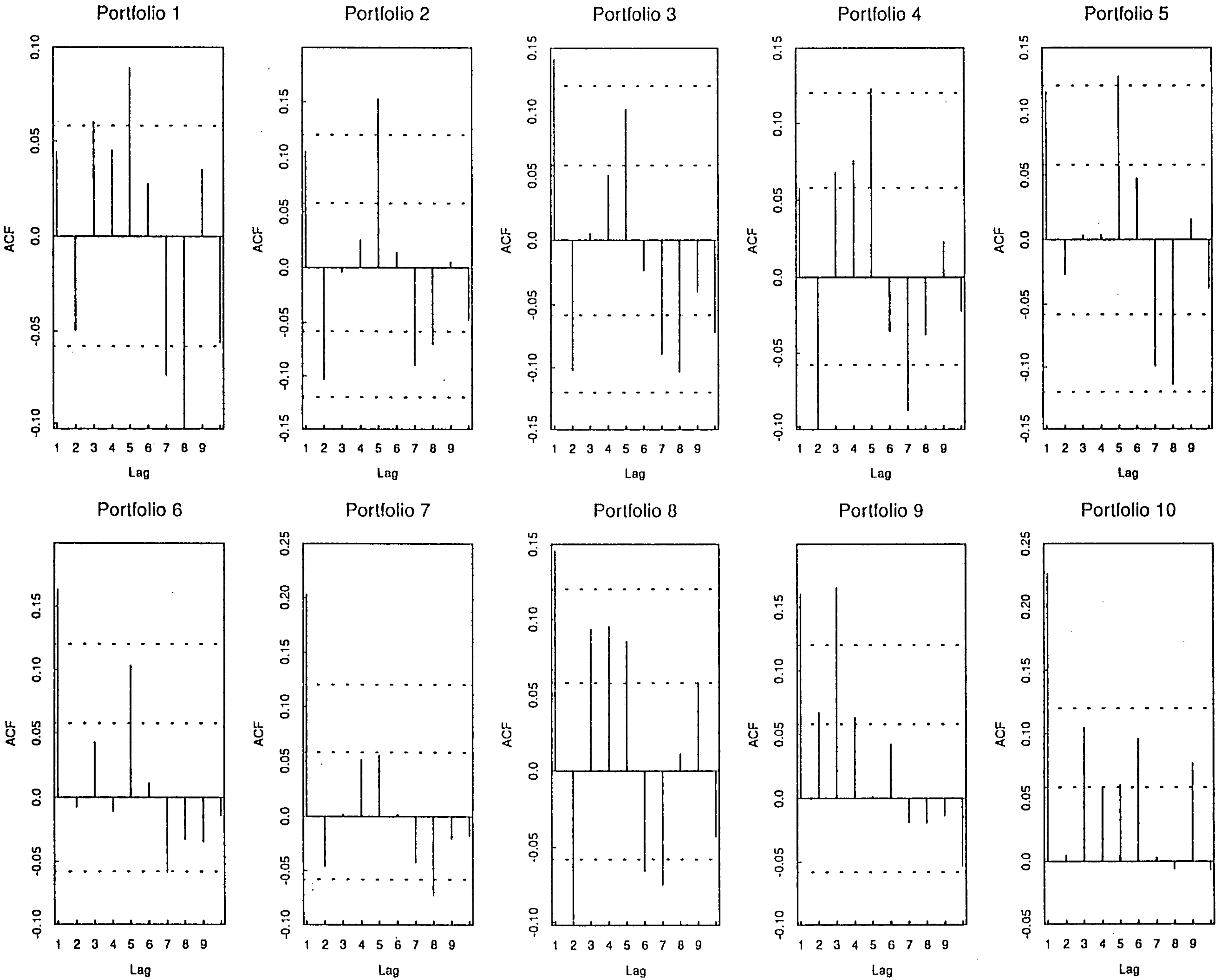


figure 3

Autocorrelations Of Sector Returns

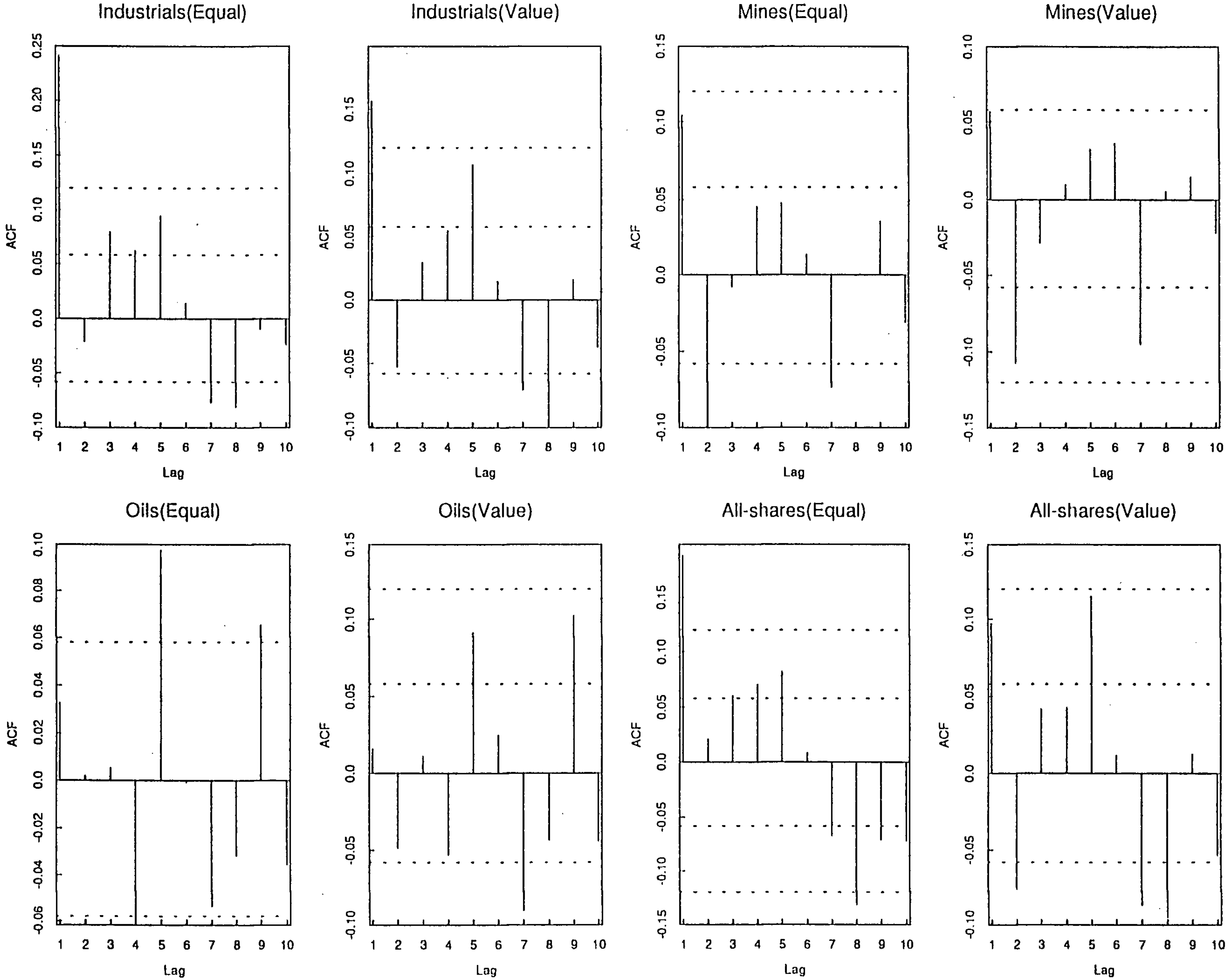


figure 4

rejected in favour of the mixed model. Further confirmation will come from estimation of the β's in the main testing procedure. In the case of the industrial, mining and all-share indices, we are again led to the rejection of a pure random walk model. This does not appear to be true of the oil index, but it should be borne in mind that only three of the 150 companies were classified under Oils (17 were Mines and 130 were Industrials) so the dependability of this index is jeopardised. Furthermore, the value of those three companies differed vastly.

4.1 ESTIMATING MODEL PARAMETERS

Although the foregoing plots provide a good indication of the merit of the inclusion of a stationary component, it is not clear if the AR(1) component is a suitable candidate. To see the goodness of fit of the AR(1) model,we begin by actually estimating Φ for each portfolio. Two special cases of result (4) provide a pair of equations that can be solved simultaneously to yield an estimate of Φ . Recall that

$$\mathrm{Cov}[r_{(t,t+T)} \, ; \, r_{(t-T,t)}] = \mathrm{Cov}[(z_{(t+T)} - z_{(t)}) \, ; \, (z_{(t)} - z_{(t-T)})]$$

$$= \frac{2\sigma^2\Phi^T}{1-\Phi^2} - \frac{\sigma^2}{1-\Phi^2} - \frac{\sigma^2\Phi^{2T}}{1-\Phi^2}$$

To be able to solve the equations explicitly, one value of T should be double the other. It is however not at all clear which values of T should be used in the estimation. Throughout the model development it was implicitly assumed that Φ was constant i.e. that it was not time varying. Thus, if the model is a good approximation to the data, it should not matter much which values of T we choose. This is now shown not to be true. For each portfolio or index, six estimates of Φ are attempted by the following general method.

$$(T = m) \qquad = \frac{2\sigma^2\Phi^m}{1-\Phi^2} - \frac{\sigma^2}{1-\Phi^2} - \frac{\sigma^2\Phi^{2m}}{1-\Phi^2}$$

$$(T = 2m) \qquad = \frac{2\sigma^2\Phi^{2m}}{1-\Phi^2} - \frac{\sigma^2}{1-\Phi^2} - \frac{\sigma^2\Phi^{4m}}{1-\Phi^2}$$

But $Cov[r_{(t,t+T)} ; r_{(t-T,t)}]$ is just the autocovariance function of the continuously compounded series (where T monthly returns are summed) r_t at lag 1. Summing returns over T months is a good reason for not choosing T equal to one or two months, since we know that the mean reversion model is quite inappropriate for such short periods of compounded returns. Now suppose the first and second covariance terms above are estimated as x and y respectively. Then

$$\frac{\hat{\sigma}^2}{1-\hat{\Phi}^2}[2\hat{\Phi}^m - 1 - \hat{\Phi}^{2m}] = x , \quad \text{and} \quad \frac{\hat{\sigma}^2}{1-\hat{\Phi}^2}[2\hat{\Phi}^{2m} - 1 - \hat{\Phi}^{4m}] = y$$

$$\therefore \frac{x}{y} = \frac{-(1-\hat{\Phi}^m)^2}{-(1-\hat{\Phi}^{2m})^2} = [\frac{1-\hat{\Phi}^m}{1-\hat{\Phi}^{2m}}]^2 = [\frac{1}{1+\hat{\Phi}^m}]^2$$

$$\therefore \hat{\Phi} = (\sqrt{\frac{y}{x}} - 1)^{1/m}$$

So we at least need $|y| > |x|$, $|y| < 4|x|$, $y < 0$, $x < 0$ for the model to be feasible.(The negativity of x and y is clear from the third last equation since $0.0 < |\Phi| < 1.0$).

In table 3, several estimates for Φ are absent. In each case this is because one of these conditions has not been met. Apart from these obvious model violations, the fact that the estimates of Φ vary according to T points to a more subtle model mis-specification, namely that Φ is time-varying. The model assumed a constant value of Φ in each portfolio.

It immediately follows that

$$\hat{\sigma}^2 = \frac{x(1 - \hat{\Phi}^2)}{(2\hat{\Phi} - 1 - \hat{\Phi}^2)}$$

Also,

$$\text{Var}[r_{(t-1,t)}] = \psi^2 + \sigma^2(\frac{1-\Phi}{1+\Phi} + 1) = \text{v say}$$

$$\Rightarrow \hat{\psi}^2 = v - \hat{\sigma}^2(\frac{1-\hat{\Phi}}{1+\hat{\Phi}} + 1)$$

And lastly, the obvious method of moments estimate for μ is

$$\hat{\mu} = \overline{r_{(t)}} = \frac{1}{n}\sum_{t=1}^{n} r_{(t)}$$

(All of these estimates are just method of moments estimates, and may thus be biased.)

The estimates of these quantities are presented in table 3(ii).

From this table, we see that the model cannot stand up to close scrutiny. The estimates are based entirely on the model structure which can now be seen to be a poor approximation to the data. Two general points can be made. Firstly, the estimates of the random walk error variances are, without exception, negative. Secondly, the estimates of the AR(1) error variances are huge compared to the estimates of the return variances, although, according to the model, they are supposed to be only one component of these variances. This can only point to model mis-specification, a factor that will have to be borne in mind when proceeding with estimating the β's. Arithmetically what is happening is that the AR(1) error variances are so large compared to $\text{Var}(r_{(t)})$ that the random walk error variance is forced to be negative to comply with the above equations.

Table (III)(i)

Estimated values of Φ for each of the portfolios and indices.

(A) Equal weighting

m / PORTFOLIO	6	12	18	24	30	36
1	-	0.983	0.916	-	0.938	0.926
2	-	0.988	0.964	0.953	0.951	0.972
3	1.11	0.949	0.917	-	-	-
4	-	0.989	0.962	0.954	0.861	-
5	-	0.973	0.891	-	-	0.950
6	-	0.995	0.936	0.843	-	0.958
7	-	0.989	0.795	-	-	0.954
8	-	0.990	0.965	0.944	-	-
9	-	0.978	0.914	0.852	-	-
10	-	-	1.019	-	-	-
Industrials	-	0.985	0.952	0.865	-	-
Mines	-	0.988	0.870	-	0.914	-
Oils	-	1.090	0.928	0.873	0.954	0.991
All-shares	-	0.995	0.881	-	-	-

(B)Value weighting

m / PORTFOLIO	6	12	18	24	30	36
1	-	0.985	0.943	0.849	0.932	0.958
2	-	0.994	0.959	0.950	0.951	0.927
3	-	0.950	0.914	-	-	-
4	-	0.992	0.963	0.954	0.863	-
5	-	0.947	0.796	-	0.912	0.964
6	-	0.993	0.932	0.826	-	0.963
7	-	0.992	0.850	-	-	0.947
8	-	0.993	0.974	0.953	-	-
9	-	0.979	0.908	-	-	-
10	-	-	1.005	0.871	-	-
Industrials	-	0.985	0.951	0.857	-	0.962
Mines	-	0.970	-	-	0.941	0.959
Oils	-	1.110	0.956	0.844	0.890	0.966
All-shares	-	0.995	0.942	0.857	0.837	0.923

Table (III)(ii)

Estimated model variances for each of the portfolios and indices.

(A) Equal weighting

PORTFOLIO	$Var(r_{(t)})$	$\hat{\sigma}^2$	$\hat{\psi}^2$	$\hat{\mu}$
1	0.0023	0.700	-0.703	0.0044
2	0.0028	0.762	-0.764	0.0045
3	0.0025	0.310	-0.315	0.0026
4	0.0026	1.140	-1.143	0.0052
5	0.0023	0.227	-0.227	0.0030
6	0.0018	1.638	-1.639	0.0046
7	0.0023	1.193	-1.197	0.0036
8	0.0024	0.935	-0.937	0.0014
9	0.0025	1.132	-1.143	0.0033
10	0.0035	-	-	0.0053
Industrials	0.0017	0.807	-0.811	0.0046
Mines	0.0055	2.220	-2.230	0.0013
Oils	0.0065	-0.098	0.099	0.0033
All-shares	0.000022	0.036	-0.036	0.0003

PORTFOLIO	$Var(r_{(t)})$	$\hat{\sigma}^2$	$\hat{\psi}^2$	$\hat{\mu}$
1	0.0022	0.715	-0.717	0.0045
2	0.0027	1.495	-1.517	0.0044
3	0.0025	0.315	-0.322	0.0026
4	0.0026	1.444	-1.447	0.0052
5	0.0023	0.143	-0.145	0.0027
6	0.0019	1.395	-1.399	0.0048
7	0.0022	1.518	-1.522	0.0037
8	0.0026	1.395	-1.400	0.0010
9	0.0027	1.197	-1.207	0.0035
10	0.0028	-	-	0.0042
Industrials	0.0020	0.794	-0.798	0.0049
Mines	0.0062	0.985	-0.994	0.0045
Oils	0.0093	-0.091	0.091	0.0021
All-shares	0.0022	0.192	-0.190	0.0043

5.0 RESULTS FOR THE β's

Table (IV)

Estimated β'S

(A) Equal weighting

	PORTFOLIO									
T in months	1	2	3	4	5	6	7	8	9	10
6	0.01	0.04	-0.05	0.03	0.03	0.05	0.01	0.07	0.08	0.23
12	-0.19	-0.09	-0.24	-0.16	-0.08	-0.12	-0.14	-0.09	-0.22	0.19
18	-0.32	-0.27	-0.34	-0.32	-0.16	-0.29	-0.43	-0.26	-0.39	-0.07
24	-0.33	-0.34	-0.38	-0.38	-0.15	-0.37	-0.51	-0.28	-0.42	-0.24
30	-0.33	-0.35	-0.38	-0.40	-0.13	-0.37	-0.50	-0.34	-0.43	-0.25
36	-0.36	-0.44	-0.39	-0.55	-0.13	-0.36	-0.50	-0.43	-0.46	-0.21
42	-0.44	-0.52	-0.40	-0.65	-0.14	-0.42	-0.41	-0.47	-0.44	-0.16
48	-0.34	-0.46	-0.25	-0.54	-0.01	-0.31	-0.37	-0.34	-0.39	-0.10
54	-0.33	-0.44	-0.08	-0.38	0.01	-0.24	-0.22	-0.29	-0.28	-0.01
60	-0.31	-0.37	-0.09	-0.36	-0.09	-0.22	-0.34	-0.29	-0.28	0.00
66	-0.33	-0.36	-0.09	-0.36	-0.06	-0.24	-0.32	-0.27	-0.32	0.00
72	-0.34	-0.36	-0.11	-0.35	-0.12	-0.26	-0.30	-0.21	-0.28	-0.06
78	-0.34	-0.32	-0.12	-0.34	-0.12	-0.26	-0.31	-0.04	-0.18	-0.09
84	-0.33	-0.30	-0.12	-0.31	-0.24	-0.24	-0.24	-0.30	-0.01	-0.13
90	-0.32	-0.28	-0.13	-0.30	-0.28	-0.22	-0.28	-0.02	-0.12	-0.23
96	-0.32	-0.26	-0.14	-0.19	-0.32	-0.19	-0.23	-0.07	-0.12	-0.27

(B) Value weighting

	PORTFOLIO									
T in months	1	2	3	4	5	6	7	8	9	10
6	-0.04	0.03	-0.05	0.04	0.02	0.04	0.01	0.06	0.08	0.12
12	-0.14	-0.10	-0.23	-0.15	-0.11	-0.12	-0.13	-0.08	-0.12	0.08
18	-0.28	-0.29	-0.34	-0.31	-0.18	-0.29	-0.38	-0.25	-0.21	-0.12
24	-0.38	-0.36	-0.38	-0.37	-0.15	-0.35	-0.45	-0.27	-0.39	-0.27
30	-0.32	-0.36	-0.39	-0.40	-0.13	-0.36	-0.51	-0.34	-0.42	-0.29
36	-0.36	-0.44	-0.39	-0.55	-0.13	-0.34	-0.43	-0.44	-0.43	-0.31
42	-0.41	-0.53	-0.40	-0.62	-0.15	-0.41	-0.43	-0.48	-0.44	-0.29
48	-0.32	-0.48	-0.29	-0.54	-0.02	-0.29	-0.38	-0.36	-0.42	-0.22
54	-0.33	-0.47	-0.11	-0.38	0.05	-0.22	-0.33	-0.23	-0.36	-0.11
60	-0.35	-0.47	-0.09	-0.35	0.06	-0.20	-0.33	-0.21	-0.26	-0.08
66	-0.41	-0.46	-0.10	-0.36	-0.05	-0.21	-0.34	-0.22	-0.26	-0.16
72	-0.44	-0.46	-0.12	-0.36	-0.12	-0.21	-0.35	-0.22	-0.26	-0.21
78	-0.49	-0.39	-0.12	-0.35	-0.13	-0.26	-0.31	-0.03	-0.18	-0.21
84	-0.41	-0.39	-0.13	-0.32	-0.22	-0.27	-0.30	0.03	-0.17	-0.27
90	-0.40	-0.36	-0.14	-0.30	-0.28	-0.24	-0.28	-0.01	-0.17	-0.33
96	-0.39	-0.33	-0.14	-0.18	-0.32	-0.21	-0.22	0.01	-0.16	-0.31

(C) Equal weighting

	SECTOR			
T in months	Industrials	Mines	Oils	All-shares
6	0.06	0.03	-0.05	0.00
12	-0.09	-0.17	-0.23	-0.18
18	-0.25	-0.45	-0.34	-0.35
24	-0.32	-0.48	-0.38	-0.46
30	-0.32	-0.41	-0.39	-0.55
36	-0.34	-0.50	-0.39	-0.43
42	-0.36	-0.56	-0.40	-0.33
48	-0.26	-0.45	-0.29	-0.22
54	-0.11	-0.44	-0.11	-0.17
60	-0.12	-0.41	-0.09	-0.18
66	-0.12	-0.40	-0.10	-0.19
72	-0.19	-0.42	-0.12	-0.17
78	-0.17	-0.40	-0.12	-0.02
84	-0.14	-0.40	-0.13	-0.01
90	-0.13	-0.37	-0.14	-0.06
96	-0.10	-0.19	-0.14	-0.02

(D) Value weighting

	SECTOR			
T in months	Industrials	Mines	Oils	All-shares
6	0.04	-0.02	0.00	0.02
12	-0.10	-0.21	-0.04	-0.14
18	-0.27	-0.43	-0.28	-0.31
24	-0.35	-0.46	-0.41	-0.39
30	-0.34	-0.43	-0.41	-0.38
36	-0.37	-0.36	-0.39	-0.40
42	-0.41	-0.49	-0.37	-0.46
48	-0.37	-0.39	-0.34	-0.42
54	-0.21	-0.29	-0.34	-0.43
60	-0.22	-0.34	-0.35	-0.42
66	-0.22	-0.35	-0.45	-0.41
72	-0.23	-0.34	-0.52	-0.41
78	-0.22	-0.32	-0.46	-0.33
84	-0.21	-0.33	-0.44	-0.28
90	-0.20	-0.31	-0.43	-0.24
96	-0.19	-0.29	-0.41	-0.21

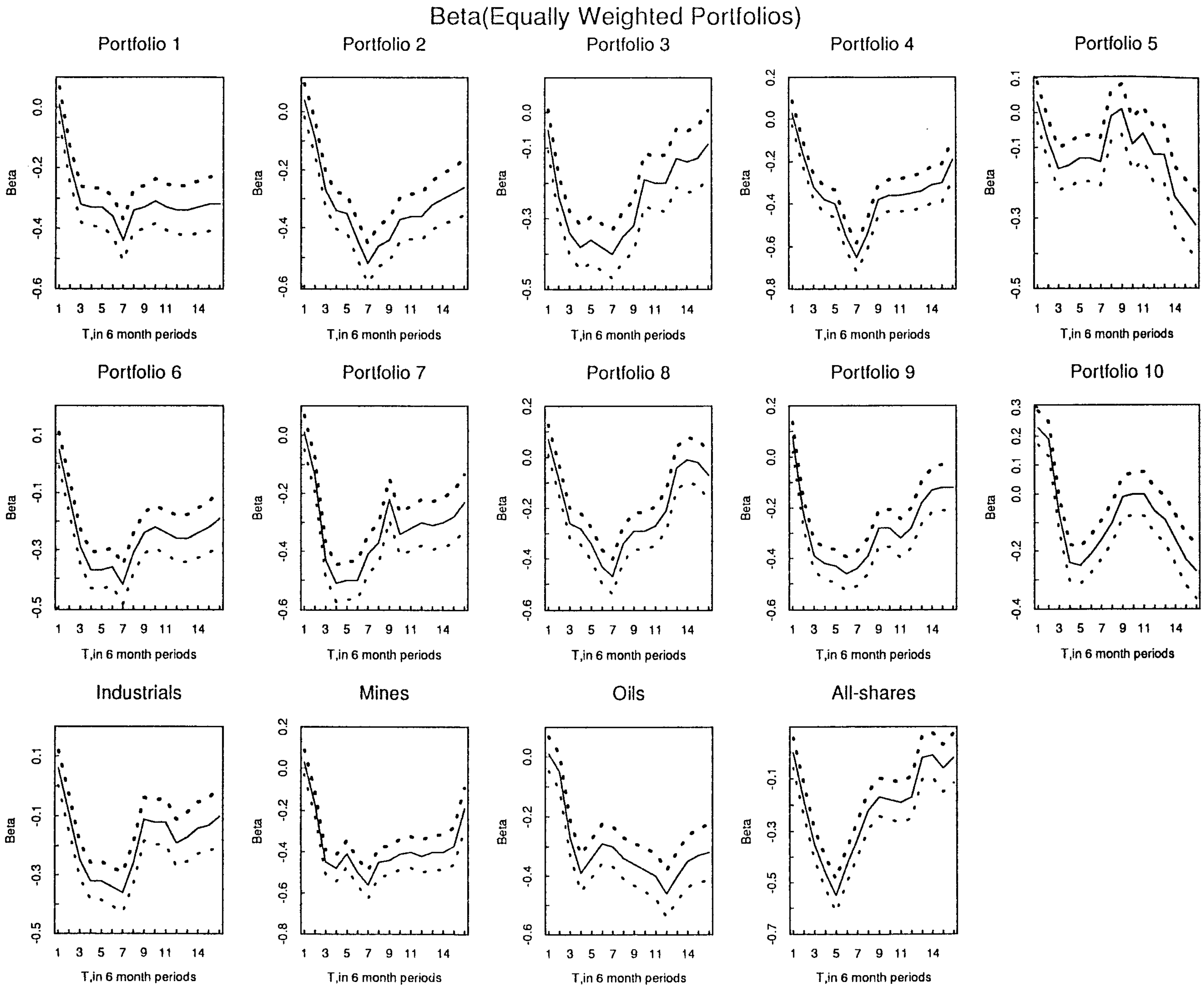
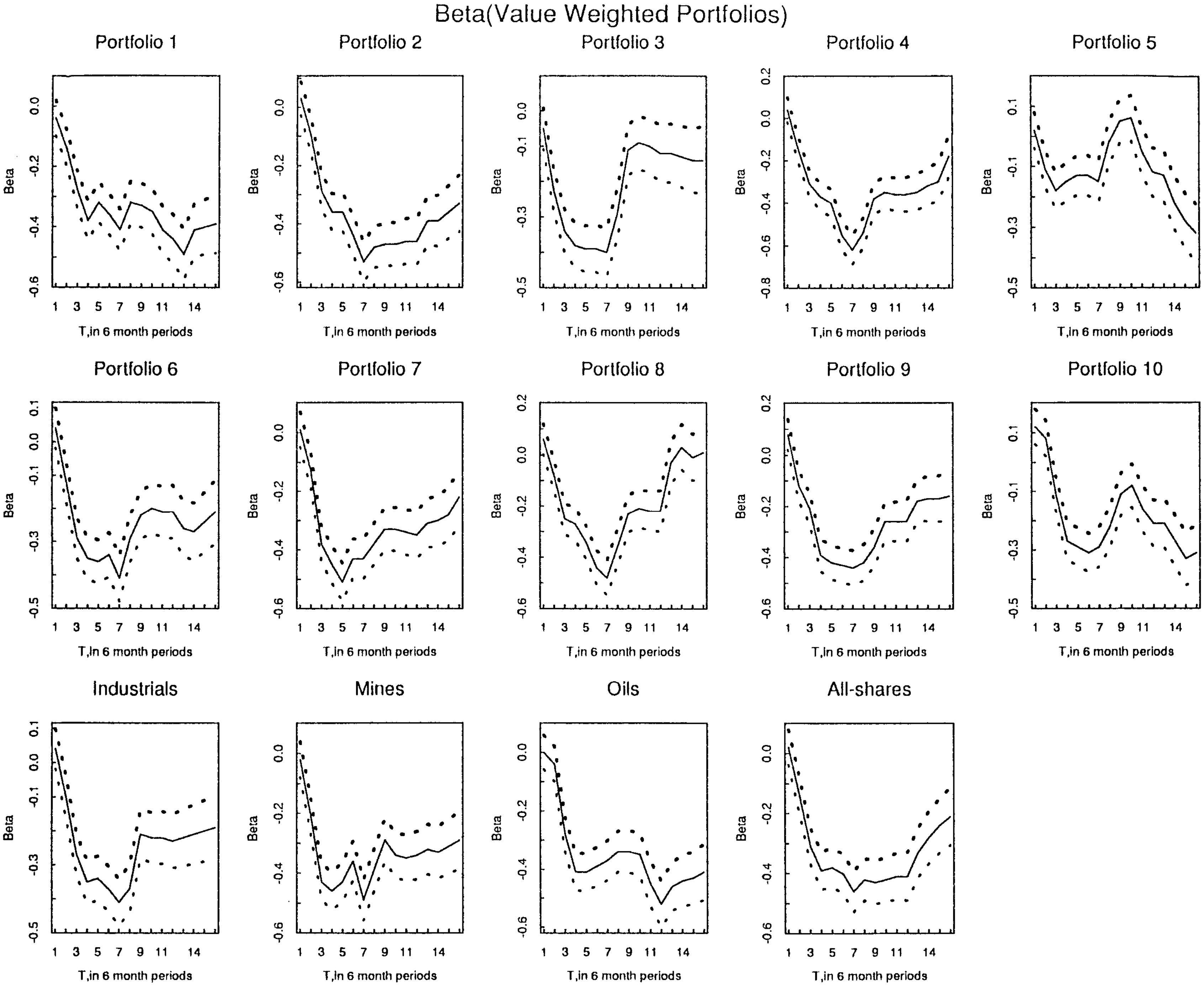

Beta(Equally Weighted Portfolios)
Portfolio 1
Portfolio 2
Portfolio 3
Portfolio 4
Portfolio 5
Portfolio 6
Portfolio 7
Portfolio 8
Portfolio 9
Portfolio 10
Industrials
Mines
Oils
All-shares
Beta
T,in 6 month periods

figure 5

figure 6


Beta(Value Weighted Portfolios)
Portfolio 1
Portfolio 2
Portfolio 3
Portfolio 4
Portfolio 5
Portfolio 6
Portfolio 7
Portfolio 8
Portfolio 9
Portfolio 10
Industrials
Mines
Oils
All-shares
Beta
T,in 6 month periods

The table and plots of the β's show that, as predicted from the model, negative values are very predominant. The only general exception to this occurs for the case T=6 months, but this too may have been expected from the theory developed earlier, when it was pointed out why this study is concerned with longer-horizon returns. The model seems to be further supported by the empirical verification that almost all estimated values of β lie in the range -0.5 to 0.0. Only portfolio 10, consisting of the 15 smallest companies, is in serious violation of model predictions. This portfolio is exceptional in several ways, one of which was pointed out during the discussion of the Φ's. The biggest company there has a value 19 times larger than the smallest, by 1987 - a very much wider range than is encountered in any other portfolio (the average ratio for the other nine portfolios is only 1.93). Some of these smallest of the long-lived companies have also at times been very thinly traded - a factor notoriously associated with unusual price behaviour. In fact, the very interpretation of the β's estimated from portfoilo ten in the same way as for the other portfolios is highly questionable, since the (attempted) estimation of Φ pointed to a mis-specified model for this portfoilo. The same may be said of the oil indeces, where Φ was estimated as being greater than one - a direct violation of model assumptions.

Comparing results across value vs equally weighted portfolios, we naturally observe different magnitudes, but in all cases the general pattern as T increases is the same. Since value weighting favours larger firms, it is not surprising to detect a slight improvement in the performance of the β's here, over equal weighting, for portfolio 10. This is also true of the two next smallest portfolios, 8 and 9.

If there is a disapointment in these results, it is that the pattern of the β's is not as nicely U-shaped as predicted. This <u>could</u> point to insufficient parameterization but there is also a simple statistical problem here. The main departure from the U-shaped prediction is the general failure of the β's to regain a level around zero, after having attained the predicted minimum. However, nothing in the theoretical development suggested that the 'U' would or should be symmetric or even monotonic so it is not surprising to find that the increasing part of the 'U' has a considerably flatter slope than the decreasing part. All indications from the plots of the β's are that a value of 0.0 <u>will</u> again be reached, but that this does not generally occur within 96 months (8 years). Therein lies the simple statistical problem - since data from the T.S.E.reaches back only to 1963, examining summed returns and then autocorrelations for values of T over 96 months becomes futile. If longer sequences of returns were available, it would be a simple matter to confirm whether or not the β's really do return to zero. Several more years of data will have to be patiently accumulated before this can be done reliably. Already, looking at the plots of the β's for all the portfolios or indices, a visual impression is gained of how the reliability of the estimates is dwindling for the larger values of T (one standard error for each estimate of β has been added and subtracted to the estimate, and been plotted as a dotted line in each case). The behaviour of the β's in portfolios 5 and 10 however, can evidently not be explained by a shortage of data. Perhaps the model is simply inappropriate for these portfolios or perhaps mean reversion is just not a relevant factor in all firms. The plots for these two portfolios indicate that another possibility is that β has not <u>yet</u> attained a minimum by 96 months. Indeed, for both cases we see that the values at 96 months are considerably lower than the previous minima attained. At this stage there is no way of telling which, if any, of these scenarios is correct. Of course the results for these two portfolios do not seriously detract from the support afforded the model by the results from the other 12 portfolios or indices.

5.1 BIAS ADJUSTMENT

When the β's were calculated, overlapping observations were used, as a necessary way of providing "sufficient" observations. However this procedure obviously introduces bias into the estimates which then turn out too negative, or optimistic. Now we correct for this bias. Fama and French have shown that the bias is biggest when the process is a pure random walk, so the natural conservative approach is to presume this to be the case when adjusting the estimates.

Suppose prices follow a pure random walk, i.e.

$$P_{(t)} = \mu + P_{(t-1)} + \delta_{(t)} \qquad \text{where } \delta_{(t)} \sim W.N.(0,\xi^2)$$

Then

$$\beta_{(T)} = \mathrm{Corr}[r_{(t-T,t)}\ ;\ r_{(t,t+T)}]$$

$$= \mathrm{Corr}[\{T\mu + \delta_{(t-T+1)} + \cdots + \delta_{(t)}\}\ ;\ \{T\mu + \delta_{(t+1)} + \cdots + \delta_{(t+T)}\}]$$

$$= 0$$

So, if we have a pure random walk $\beta_{(T)}$ <u>should</u> equal zero for all T. Note that this is perfectly general - the values of the parameters are immaterial. Thus we may simulate random walks with say N(0,1) errors, construct overlapping returns and estimate $\beta_{(T)}$ for all values of T of interest. The average values of $\beta_{(T)}$ from very many simulations will give us a good idea of the bias introduced by overlapping, and we thus adjust all previous estimates of $\beta_{(T)}$ appropriately. Fama and French did 10 000 random walk simulations with N(0,1) errors and showed that of course the bias increases as the return horizon (T) increases, since this constitutes larger overlaps. Because these simulations are unrelated to any particular stock exchange, their results are used here for adjusting the estimates of β. Their horizons were only in units of whole years, so the half-yearly biases are simply interpolated.

Since the bias is largest for the largest values of T, looking at figures 7 and 8, we immediately notice an improvement in the β's in respect of regaining a level around zero. In some cases, the effect of the bias results in the upper standard error bound being considerably above zero when T is large. Basically, the interpretation of these β's is unchanged, but the values are more dependable.

Table (V)

Bias adjusted β's

(A) Equal weighting

T in months	PORTFOLIO 1	2	3	4	5	6	7	8	9	10
6	0.03	0.06	-0.03	0.05	0.05	0.07	0.03	0.09	0.10	0.21
12	-0.16	-0.06	-0.21	-0.13	-0.05	-0.09	-0.11	-0.06	-0.19	0.22
18	-0.29	-0.24	-0.31	-0.29	-0.13	-0.26	-0.40	-0.23	-0.36	-0.04
24	-0.29	-0.30	-0.34	-0.34	-0.11	-0.33	-0.47	-0.24	-0.38	-0.20
30	-0.28	-0.30	-0.33	-0.35	-0.08	-0.32	-0.45	-0.29	-0.38	-0.20
36	-0.29	-0.37	-0.32	-0.48	-0.07	-0.29	-0.43	-0.37	-0.39	-0.14
42	-0.36	-0.44	-0.32	-0.57	-0.06	-0.34	-0.33	-0.39	-0.36	-0.08
48	-0.24	-0.36	-0.15	-0.44	0.09	-0.21	-0.27	-0.24	-0.29	0.00
54	-0.22	-0.33	0.03	-0.27	0.12	-0.13	-0.11	-0.18	-0.17	0.10
60	-0.18	-0.24	0.04	-0.23	0.04	-0.09	-0.21	-0.16	-0.15	0.13
66	-0.19	-0.22	0.03	-0.22	0.08	-0.10	-0.18	-0.13	-0.18	0.14
72	-0.18	-0.20	0.05	-0.19	0.04	-0.10	-0.14	-0.05	-0.12	0.08
78	-0.16	-0.14	0.06	-0.16	0.06	-0.08	-0.13	0.14	0.00	0.09
84	-0.13	-0.10	0.08	-0.11	-0.04	-0.04	-0.04	-0.10	0.19	0.07
90	-0.10	-0.06	0.09	-0.08	-0.06	0.00	-0.06	0.20	0.10	-0.01
96	-0.08	-0.02	0.10	0.05	-0.08	0.05	0.01	0.16	0.12	-0.08

(B) Value weighting

T in months	PORTFOLIO 1	2	3	4	5	6	7	8	9	10
6	-0.02	0.05	-0.03	0.06	0.04	0.06	0.03	0.08	0.10	0.14
12	-0.11	-0.07	-0.20	-0.12	-0.08	-0.09	-0.10	-0.05	-0.09	0.11
18	-0.25	-0.26	-0.31	-0.28	-0.15	-0.26	-0.35	-0.22	-0.18	-0.09
24	-0.34	-0.32	-0.34	-0.33	-0.11	-0.31	-0.41	-0.23	-0.35	-0.23
30	-0.27	-0.31	-0.34	-0.35	-0.08	-0.31	-0.46	-0.29	-0.37	-0.24
36	-0.29	-0.37	-0.32	-0.48	-0.06	-0.27	-0.36	-0.37	-0.36	-0.24
42	-0.33	-0.45	-0.32	-0.54	-0.07	-0.33	-0.35	-0.40	-0.36	-0.21
48	-0.22	-0.38	-0.19	-0.44	0.08	-0.19	-0.28	-0.26	-0.32	-0.12
54	-0.22	-0.36	0.00	-0.27	0.16	-0.11	-0.22	-0.12	-0.25	0.00
60	-0.22	-0.34	0.04	-0.22	0.19	-0.07	-0.20	-0.08	-0.13	0.05
66	-0.27	-0.32	0.04	-0.22	0.09	-0.07	-0.20	-0.08	-0.12	-0.02
72	-0.28	-0.30	0.04	-0.20	0.04	-0.05	-0.19	-0.06	-0.10	-0.05
78	-0.31	-0.21	0.06	-0.17	0.05	-0.08	-0.13	0.15	0.00	-0.03
84	-0.21	-0.19	0.07	-0.12	-0.02	-0.07	-0.10	0.23	0.03	-0.07
90	-0.18	-0.14	0.08	-0.08	-0.06	-0.02	-0.06	0.21	0.05	-0.11
96	-0.15	-0.09	0.10	0.06	-0.08	0.03	0.02	0.25	0.08	-0.07

(C) Equal weighting

	SECTOR			
T in months	Industrials	Mines	Oils	All-shares
6	0.08	0.05	-0.03	0.02
12	-0.06	-0.14	-0.20	-0.15
18	-0.22	-0.42	-0.31	-0.32
24	-0.28	-0.44	-0.34	-0.42
30	-0.27	-0.36	-0.34	-0.50
36	-0.26	-0.43	-0.32	-0.36
42	-0.28	-0.48	-0.32	-0.25
48	-0.16	-0.35	-0.19	-0.12
54	0.00	-0.33	0.00	-0.06
60	0.01	-0.28	0.04	-0.05
66	0.02	-0.26	0.04	-0.05
72	-0.03	-0.26	0.04	-0.01
78	0.01	-0.22	0.06	0.16
84	0.06	-0.20	0.07	0.19
90	0.09	-0.15	0.08	0.16
96	0.14	0.05	0.10	0.22

(D) Value weighting

	SECTOR			
T in months	Industrials	Mines	Oils	All-shares
6	0.06	0.00	0.02	0.04
12	-0.07	-0.18	-0.01	-0.11
18	-0.24	-0.40	-0.25	-0.28
24	-0.31	-0.42	-0.37	-0.35
30	-0.29	-0.38	-0.36	-0.33
36	-0.30	-0.29	-0.32	-0.33
42	-0.33	-0.41	-0.29	-0.38
48	-0.27	-0.29	-0.24	-0.32
54	-0.10	-0.18	-0.23	-0.32
60	-0.09	-0.21	-0.22	-0.29
66	-0.08	-0.21	-0.31	-0.27
72	-0.07	-0.18	-0.36	-0.25
78	-0.04	-0.14	-0.24	-0.15
84	-0.01	-0.13	-0.24	-0.08
90	0.02	-0.09	-0.21	-0.02
96	0.05	-0.05	-0.17	0.03

Bias Adjusted Beta(Equally Weighted Portfolios)

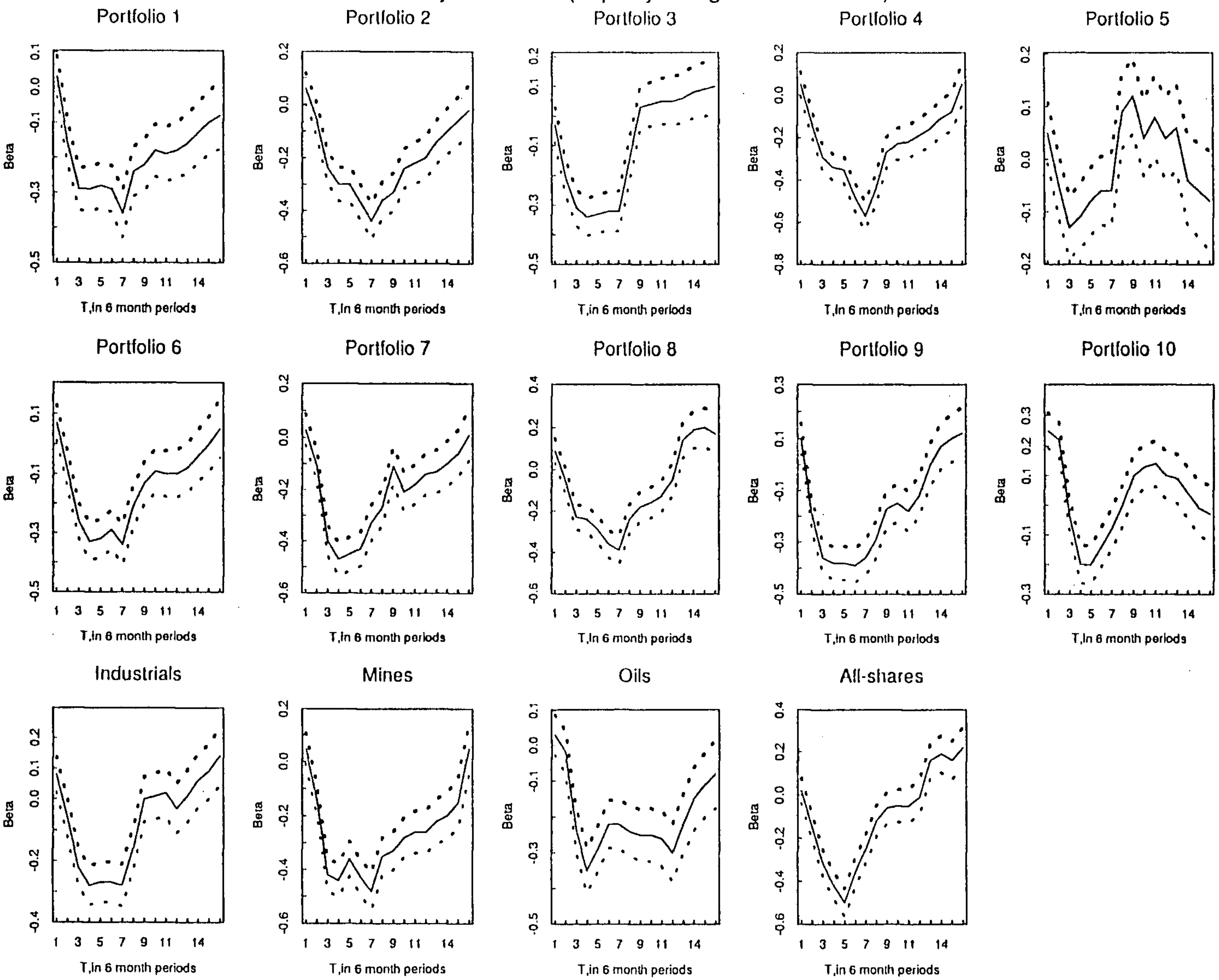


Figure 7

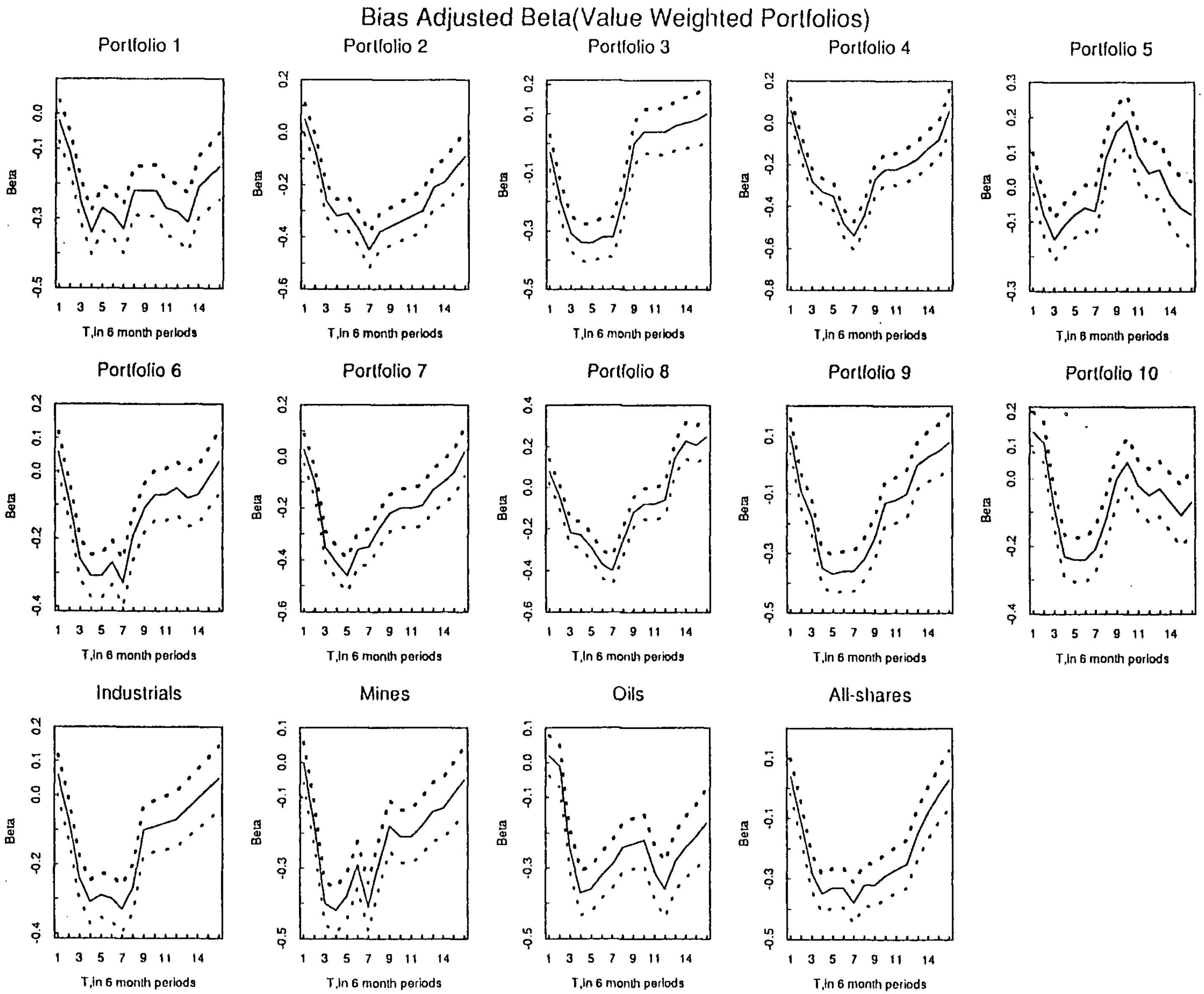

Bias Adjusted Beta(Value Weighted Portfolios)
Portfolio 1
Portfolio 2
Portfolio 3
Portfolio 4
Portfolio 5
Portfolio 6
Portfolio 7
Portfolio 8
Portfolio 9
Portfolio 10
Industrials
Mines
Oils
All-shares
Beta
T,In 6 month periods

Figure 8

5.2 CONCLUSIONS

In simply attempting to repeat the study by Fama and French, we might be led to an overly optimistic interpretation of the β'S. However, closer examination of model assumptions and, in particular additional parameter estimation, show very clearly that a great deal of caution must be exercised. The model appears to represent the data only very weakly, but even more importantly, several model assumptions are seriously violated. In this light, the interpretation of the β'S is no longer as clear as was suggested in the theoretical discussion, where all results followed strictly from the exact nature of the model equations,in particular,the stationary component is AR(1). Whether the fact that the β'S turned out more or less as anticipated can still be attributed to mean reversion is now highly questionable. In fact it seems that their patterns occurred not because of a suitable model, but in spite of a mis-specified model. Of course this does not discount the theory of mean reversion - it simply says that the existence of mean reversion on the T.S.E. could not be convincingly proved from the framework of the model proposed by Fama and French.

References

Abraham B. and Ledolter J. (1983), "Statistical Methods For Forecasting", John Wiley and Sons

Stoodley K.D.C., Lewis T.,Stainton C.L.S. (1980), "Applied Statistical Techniques" , Elis Horwood Series

Kendall M.G. (1973),"Time Series"

Brown R.G. (1961),"The Fundamental Theorem Of Exponential Smoothing" , Cper.Res 9

Fama E.F. and French K.R. (1988),"Permanent and Temporary Components of Stock Prices", Journal of Political Economy, vol 96, no2

Fama E.F. (1976),"Foundations of Finance" Basic Books

Francis J.C. and Archer S.H. (1971),"Portfolio Theory" , Prentice-Hall

Lo A.W. and MacKinlay A.C. (1988),"Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test" , The Review of Financial Studies, vol 1, no1

Brockwell P.J. and Davis R.A. (1987),"Time Series : Theory and Methods",Springer Verlag