A STUDY OF THE MULTIVARIATE DISTRIBUTION OF COMMODITY FUTURES PRICES WITH A VIEW TO THE DEVELOPMENT OF PORTFOLIOS AND TRADING SYSTEMS

Kevin Brendan Connolly B.Sc., B.A., M.Sc.,

Thesis submitted in partial fulfilment of the requirements for a Doctor of Philosophy Degree

Sponsoring establishment: The City of London Polytechnic

Date of Submission: September 1985

ACKNOWLEDGMENTS

I wish to acknowledge the department of Computing Mathematics Statistics and Management Science of The City of London Polytechnic who provided the opportunity for me to undertake this research. I must also acknowledge the assistance of The International Commodity Clearing House in providing me with a large data base without which this project would not have been possible.

I am indebted to my supervisors, Dr Gordon Gemmill and Bernard Ethell who gave me endless encouragement and advice throughout the course of this research. I would also like to thank Professor Conolly for his invaluable assistance and direction in the early stages of this programme.

A special debt of gratitude must be paid to the entire computing staff of the City of London Polytechnic. In particular I would like to thank Paul Andrews, Robin McLauchlin and Derek Salmon for the many hours spent assisting with computing problems.

ABSTRACT

A STUDY OF THE MULTIVARIATE DISTRIBUTION OF COMMODITY FUTURES PRICES WITH A VIEW TO THE DEVELOPMENT OF PORTFOLIOS AND TRADING SYSTEMS

K.B.CONNOLLY

The univariate and multivariate distribution of daily returns on contracts in the London cocoa, coffee, sugar and rubber futures markets over the period 1975-79 are studied. In the analysis, two relatively recent multivariate procedures (the multivariate serial correlation coefficient and the multivariate extension of the W - test for normality) are investigated. The four dimensional vector of returns with one component from each futures market can be viewed as being generated from a serially independent multivariate normal process with non - constant variance/covariance structure and occasional contaminating extreme realisations.

Examining the multivariate distribution in which all the components are returns on contracts in the <u>same</u> futures market, however, produced different and very unexpected results. Highly significant multivariate serial correlation coefficients of lag one day and significant departures from multivariate normality were discovered.

The multivariate temporal dependence was shown to be due to correlation between certain linear combinations of returns on contracts of differing maturities. Studying the distribution of the linear combination estimates led to the discovery that much of the observed phenomenon can be explained by negatively correlated multivariate spread portfolios.

Multivariate trading rules were devised to exploit the observed temporal behaviour and when applied to all four series produced large, positive and highly statistically significant returns. The introduction of non zero transaction costs reduced returns but still produced positive profits in the cocoa and coffee series.

Models of processes that could explain the observed multivariate temporal behaviour and the multivariate non - normality are presented.

A STUDY OF THE MULTIVARIATE DISTRIBUTION OF COMMODITY FUTURES PRICES WITH A VIEW TO THE DEVELOPMENT OF PORTFOLIOS AND TRADING SYSTEMS

CHAPTER 1

INTRODUCTION AND SUMMARY

The aim of this research project was to investigate the possibility of devising a risk minimizing procedure similar to that devised by Markowitz (1952) that would be applicable to the commodity futures markets. In the inital empirical study, however, the discovery of persistent and significant, multivariate serial correlation suggested the presence of a special kind of multivariate market inefficiency and led to the development of a number of multivariate trading rules. These trading rules when applied to historical data produced consistent, significant positive profits. A model explaining the observed multivariate market inefficiency has been proposed. The outline of the work is as follows.

In Chapter 2 we review the literature on the univariate distributions of commodity returns. Many non - standard univariate and multivariate statistical techniques are employed in this thesis. Rather than include them in Chapter 2, they are reviewed briefly in the appropriate chapters.

The International Commodity Clearing House (ICCH) made available a set of daily prices for all the futures contracts traded in the cocoa, coffee, sugar and rubber markets in London from 1974 to 1979. Many anomalies in the data were discovered and the task of editing and subsequent rearrangement into an easily accessible form is described in Appendix A.

In Chapter 3 we outline the results from carrying out all the standard (and some non standard) univariate statistical tests on the data. As is noted in many other works, returns are highly nonstationary in variance and only one series (rubber) could be described as approximately normally distributed.

No other work to date has investigated, in detail, the joint distribution of returns on a set of commodity futures contracts. A necessary condition for the application of the classical (Markowitz type) portfolio theory and many multivariate analyses is that the joint distribution of returns be multivariate normal. Chapter 4 examines whether the commodity returns could be viewed as multivariate normal. In the same chapter we also investigate the possibility of multivariate temporal dependence in futures contracts from different futures markets. This is carried out using a relatively new technique devised by O'Brien (1980). Our conclusions are that joint sets of returns can be described as being generated from a multivariate normal process with occasional contaminating extreme observations. These extreme returns are rare and not predictable. The observed unpredictable variation in the variance/covariance structure and the absence of any significant (non zero) multivariated mean vectors, however, suggested that the development of a classical Markowitz type model would not be successful.

In Chapter 5 the returns to contracts of different maturities (eg. March, September,..) from a given commodity futures market are examined. It was here that we discovered the presence of persistent and consistent multivariate serial correlation. There is also evidence of departure from multivariate normality of a much more extreme nature than reported in Chapter 4.

In Chapter 6 the degree and exact nature of the multivariate serial correlation is outlined in more detail. We show how the multivariate serial correlation is essentially the joint correlation between certain linear combinations of returns on consecutive days. The extreme collinearity of the data caused estimation problems. The sampling distribution of these correlations and linear combination estimates is little understood and we report the results of some simulation experiments aimed at improving the stability of estimates. Attempts at improving estimates using Ridge regression techniques were investigated, unfortunately with little success.

Although the sample results were very varied, a general pattern in the multivariate correlation parameter estimates was perceived and the pooling of correlation matrices enabled the estimation of 'grand average' linear combinations. These 'grand average estimates proved to be very similar across all four commodity series.

A general picture of a very special type of multivariate market inefficiency common to all four commodity futures markets became apparent. This discovery prompted an investigation of multivariate trading rules.

Examination of the linear combinations of returns suggested that by constructing certain complex multivariate spread positions one might be able to derive portfolios of returns that would exhibit significant univariate negative serial correlation. The analytic derivation of these temporally dependent portfolios is outlined in Chapter 7. It was found that 95 of the 96 resultant portfolios exhibited ex ante negative serial correlation.

In Chapter 8 we outline three trading rules of various degrees of sophistication that were specifically designed to exploit the persistent negative serial correlation in the spread portfolios. In all cases the profits obtained from applying these rules were not only positive and highly statistically significant, but in the cocoa and coffee series quite spectacular. The inclusion of transaction costs reduced profits and in the sugar and rubber series resulted in losses. However the most sophisticated rule explicitly incorporates the costs of transaction as part of the strategy and in the cocoa and coffee series produced very encouraging returns.

Careful examination of the linear combination estimates and the results of the multivariate analysis led to the development of a number of models that could explain the observed multivariate inefficiency. In Chapter 9, two multivariate models of commodity futures prices are presented. Both models involve the generation of small perturbations in the multivariate vector of prices. In the first model we show analytically that, under certain simplifying assumptions, the multivariate spread portfolios have serial correlation coefficients with a theoretical lower bound of -0.50. This is consistent with the observed grand average serial correlation coefficients which are approximately -0.45.

In an attempt to incorporate the observed complex 'grand average' correlation matrices, a second and more sophisticated model was developed. Attempts at estimating the parameters of this second model have so far proved unsuccessful. However, the use of trial parameter values have produced a final model that could explain the observed multivariate serial correlation and resultant successful trading rules very well indeed.

CHAPTER 2

LITERATURE REVIEW

This research is concerned with the detailed examination of the empirical distribution of returns on the four major soft commodities traded in London from 1975 to 1979.

In this chapter we review the papers that, to date, have contributed significantly to the body of knowledge relating to the empirical distribution of spot and futures prices. Most works examine returns, R_t,
defined as either:

(i) $R_t = P_t - P_{t-1}$,

(ii) $R_t = P_t / P_{t-1}$,

or its logarithm, i.e.

(iii) $\text{Log}(R_t) = \text{Log}(P_t) - \text{Log}(P_{t-1})$,

in which P_t = price at end of period t.
The period in question can be a day, a week, a month or a year.

Most of the early work on empirical distributions examined stockmarket returns and it will be helpful, briefly, to review this work prior to an analysis of commodity prices.

Much of the literature examining stock and commodity returns is concerned with producing evidence in favour of, or against, some form of efficient market hypothesis. From a purely statistical point of view the works concentrate on two main aspects of the return distribution:

(i) do price series conform to a random walk, (is there temporal dependence in the sequence of returns)? and

(ii) what is the nature of the distribution of returns?

2.1 The Nature of the univariate distribution of stock market returns

Bachelier (1900), first proposed the idea that, for a given stock, if there were a sufficiently large number of transactions per day which were spread uniformly across time, then price changes should be independently and identically distributed (iid) realisations of a Gaussian process. Many researchers, including Kendall (1953) and Osborne (1959), produced evidence in favour of the normal distribution but all noted that the distributions had fatter tails (leptokurtic) than would be expected. These fat tails were due to the occasional very large price changes and gave rise to large estimates of variance.

Mandelbrot (1963) and others suggested that price changes were generated by an infinite variance process. His model (a more general form of Bachelier's model with price changes sampled from the stable Paretian distribution) attempts to explain the observed departures from normality by lifting the finite variance restriction. One of the main reasons for suggesting a stable model for daily or weekly price changes was the fact that, if the sum of iid random variables has a limiting distribution then the random variable must come from a stable distribution (a sort of extension of the classical central limit theorems to cases where the second moment is infinite). No distributional form (except in special cases) exists for the stable family. It is defined by a characteristic function and a given member of the family is specified by the value of three parameters : (i) a location parameter, (ii) a scale parameter and

(iii) a characteristic or shape exponent $\alpha \varepsilon [0,2]$. If $\alpha = 2$ we have the normal distribution and if $\alpha = 1$ we have the Cauchy distribution. The variance (except if $\alpha = 2$) is undefined.

Fama and Roll (1968) studied the stable family and developed techniques for estimating the parameter values.

Praetz (1972) reconsidered Osborne's (1959) Brownian motion model (in which price changes are assumed normal). Praetz introduced the Bayesian concept of placing a prior distribution on the variance. By choosing a suitably vague prior for the variance (inverse gamma) the posterior price change distribution is found to be student - t. Using data on 17 price indices Praetz fitted the student - t and three other distributions : the normal, the compound events model, and the stable distribution. In all cases the student - t gave a superior fit. Blattberg and Gonedes (1974) examined exhaustively the sampling properties of the student - t and the stable distributions and using this experience examined a number of security price series. Strong evidence was produced in favour of the student - t distribution. Praetz (1978), using monthly returns on the Melbourne Stock Exchange, compared the two distributions and found that the student - t was clearly superior with normality a reasonable approximation in some instances. Fama (1970) and others have compared the distribution of returns over various differencing intervals. "All are very roughly normal; the approximation is better, the longer the differencing interval."

2.2 The nature of the distribution of returns on commodity spot and futures series

With the exception of Holthausen and Hughes (1978), all studies examining commodity returns report fat tailed (fatter than normal or leptokurtic) distributions.

Mandelbrot (1963), first suggested that wheat and corn spot returns were stable Paretian distributed. Dusak (1973) fitted a stable distribution to wheat, corn and soybean futures and found estimates of the shape parameter to vary from 1.4 to 1.8. In his paper on the influence of margin levels, Bear (1972) postulated that if margins are set too high, a deficiency in speculative interest would impede the rapid adjustment of prices to new information causing relatively less leptokurtic distributions. Using non parametric techniques Bear finds evidence to support his theory. All distributions were leptokurtic but less so over periods of high margin levels. Loebs (1979) exhaustive empirical study on returns over 20 years and 16 commodities reports that all spot and future returns are leptokurtic, many series being significantly skewed. Loeb noted that logging the returns reduced the degree of lepokurtosis and skewness. Holthausen and Hughes (1978), produce contradictory evidence, reporting thin tailed distributions on 19 spot series.

Taylor and Kingsman (1979), studying spot copper and sugar futures series, note that much of the observed leptokurtosis could be due to non stationarity of variance in the generating process and report that although the distributions are non normal, they are closer to normality than previously suggested and produce evidence to support the hypothesis that a student - t distribution could explain their data reasonably well.

All the empirical studies on distributions of commodity returns mention the difficulties in making any firm conclusions owing to the non stationarity of the series.

2.3 Non Stationarity of commodity returns

If one reads the final paragraph of any article prior to 1978 that examines commodity returns series, invariably it will contain the caveat "owing to the extreme non - stationarity in returns any firm conclusions are very difficult to make".

Loeb (1979), investigated the stationarity of the mean returns and the variance of returns of many spot and futures series over a 20 year period. By dividing the period into four 5 year subperiods he discovered that the mean returns did not vary signficantly but that the variances increased towards to end of the time period considered. Loeb found some evidence that the variance of returns of some agricultural futures was related to seasonal factors. Cox (1976) cites evidence that the suspension of futures trading in some commodities (notably onions) tended to increase the variance in the underlying spot returns series. Taylor and Kingsman (1978) report non stationarity in their long sugar series.

Taylor and Kingsman (1979) appear to be the first authors that attempt to explicitly model and estimate parameters of the fluctuating variance process. Two models that seem to produce results consistent with observations are (i) a simple autoregressive process and (ii) a Markov chain with 3 states (low, medium and high variance).

2.4 Temporal dependence in commodity return series

Much of the published work on the empirical examination of commodity spot and future returns since 1960 have examined the question of temporal dependence. Many formal statistical tests have been applied, some examples being : serial correlation coefficients, runs tests, autoregression analysis and spectral analysis. Non statistical procedures such as filter rule tests and ad hoc chartist type indexes have also been employed. The results of all these studies on different series over various time periods have, to say the least, been mixed.

2.4.1 Serial correlation coefficients

The serial correlation coefficient at lag k defined as r_k in section 4.2.1 is a measure of the correlation between returns distant k periods appart. If there is no temporal dependence in the returns the sampling distribution of r_k , is N (0, SE (r_k)), with $SE(r_k) \sim 1/\sqrt{n}$.

This is true even if the returns are non Gaussian, provided the sample size, n, is large enough and the variance of the returns process is constant. Larson (1960), Smidt (1965), Stevenson and Bear (1970), Cargill and Rausser (1975), Loeb (1979) and Tschoegl (1978) found evidence of significant r_k's but with no apparent consistent pattern. Some studies produced significant negative coefficients, while others produced significant positive coefficients. Dusak (1973) and Praetz (1975) found no evidence of significant coefficients.

Bear (1972) noted that in periods when margin levels were higher than normal, r_k's, tended to be signficiantly positive, supporting his theory that high margins attracted less speculative interest and hence induced a price stickyness in one direction. Conversely, when margins were lower than normal, Bear found most r_k's to be signficantly negative, suggesting that excessive speculative interest induced an excessive number of reversals. When margins were considered normal, r_k's turned out to be near zero.

Loeb (1979) appears to be the first to have studied simple spread returns. He found very significant negative first order daily serial correlation coefficients in all spread series suggesting excessive daily reversalling.

Taylor (1980) used simulations to study the sampling distribution of r_k's with a fluctuating variance processes. It was discovered that a better description of the sampling distribution of r_k is : $N(0, SE(r_k))$ with $SE(r_k) = a/\sqrt{n}$ in which $a \simeq 1.40$.

By developing a returns standardization procedure Taylor shows how to overcome the fluctuating variance problem. Using his recommended technique Taylor finds evidence of small but significantly positive serial correlation coefficients in 8 of 11 of the series studied. Taylor postulates that these small positive coefficients arise out of a model with relatively short lived stochastic trends.

2.4.2 Runs tests

The runs test has always been an attractive alternative to serial correlation analysis (outlined in section 3.2.1.) because no assumptions regarding the distribution of returns or the stationarity of the variance are made. A drawback of the test is that it is not very informative. Unlike serial correlation analysis one obtains a result that there are either too many runs, not enough or that the number of runs is acceptable. Most works that report serial correlations coefficients have also tended to report the results of runs tests. In many of the studies runs tests results back up serial correlation analysis but some eg Bear (1970) and Loeb (1979) produce contradictory results.

2.4.3 Trading rules applied to commodity series

Filter (trading) rules have been applied to stock and commodity price series since as early as 1960 in an (alternative) attempt to make some statement on the efficiency or otherwise of the relevant markets. The rationale of trading rules is that the sequence of prices is non random; particular prices tend to move in trends. Once a trend is established, prices are more likely to move with the trend than against it. A simple filter trading rule involves the choice of an appropriate filter level, say, x (or x%) and the monitoring of the price series. If the price moves up x% or more on a given day then buy and hold until the price falls x%. At this point one may either close out by a sale or close out and go short by selling twice and reverse when the price rises by x%. If prices were to move in well defined trends, trading rules of this sort would outperform simple buy and hold benchmark strategies.

Houthakker (1959), was one of the first to publish a work in which a simple trading rule was used to examine the efficiency of the futures markets. Large gross profits were reported. Stevenson and Bear (1970) included commissions in their study and reported that with small filters (eg one and a half percent) many costly transactions were induced resulting in net losses. Leuthold's (1972) study of live cattle futures reports large gross profits and concludes that "these profits are larger than might be expected under a random walk hypothesis". Martell and Philippatos (1974), introduce the idea of having an adaptive trading rule, one in which the filter size is set to some varying but optimal level as time passes. Net profits from these adaptive filters appears to outperform buy and hold policies but as Pinches (1974) points out, the pooling of profits by Martell and Philippatos (1974) across commodities is misleading.

Applying filter rules of various sizes to 16 futures series, Loeb (1979) gets mixed results but notes that a 5% filter appears to give substantially superior net profits. Following the discovery of consistent negative one day serial correlation coefficients in all spread series, Loeb (1979) constructed a reverse filter rule to take advantage of the observed price reversaling. Consistent net positive profits resulted. However Loeb (1979) notes that the trading rule (set in absolute price differences not % changes) appears to be quite impracticable because it indicates very infrequent trades. For example using a 2·cent filter rule he generated one trade in 4 years.

Praetz (1976a) criticized all work employing returns from filter rules, noting that no well defined distribution for these returns had been proposed. In the absence of any sampling theory with associated standard errors, it is not sufficient to report that filter rule yield large positive returns. In a later paper, Praetz (1976b), develops exact expressions for the mean and standard deviation of returns to filter rules under a null hypothesis of a random walk and shows that comparison with buy and hold policies is unfair. "The returns to the filter man are always biased downwards"; "the situation is so loaded against the filter man that it is like making him play Russian roulette with five live bullets in a six-shooter". Simple reporting of the grand profit (net or otherwise) resulting from the application of a certain filter rule is only as useful as reporting the sample average to a statistician without a measure of accuracy (i.e. standard error).

2.4.4 Forecasting commodity prices

Can commodity prices be forecasted? This of course begs the question of temporal dependence again. Much of the work on commodity price forecasting obviously assumes that there is some (albeit complex) sort of temporal dependence in the sequence of returns.

In an attempt to discover if futures trading in a commodity increases the informational content to dealers in the corresponding actuals market, Cox (1976), fitted autoregressive models to the spot prices of onions (among other commodities) in periods when : (1) futures trading existed and, (2) futures trading was suspended. In many cases he found significant autoregressive coefficients. He also found that the number of significant terms in the regression equations were greater in periods in which futures trading was suspended than when futures trading existed. Cox's work thus suggests that spot market prices can be forecasted and that the forecasting accuracy is increased in periods when no futures market exists.

Labys (1976), considered seven different forecasting schemes ranging from the naive (i.e. the best estimate of tomorrow's price is today's) through exponential smoothing to moving average and autoregressive methods. All seven procedures were executed on 1968 daily data on 8 commodity series. The naive expectation scheme proved to be the most accurate.

Chu (1978) used a Box-Jenkins package in an attempt to forecast, in the short run, monthly prices for 10 spot series. Starting from econometric type supply and demand equations, Chu developed a set of final autoregressive expressions with some of the agricultural series having a seasonal component. Chu notes "These models have, however, only limited capability to predict unusual movements in prices". This is not surprising, since these movements are, as Chu says, "unusual".

Taylor (1980) used his conjectured price model to forecast one - day - ahead returns and one - day - ahead trends. Disappointing results are achieved in returns forecasts but quite encouraging results are obtained for forecasted trends. "It is often quite possible to predict correctly whether tomorrows trend will be positive or negative." Taylor also produced expressions giving the theoretical accuracy of forecasts and noted that there is a consistent and close agreement between his theoretical and the actual forecast errors.

2.5 Multivariate analysis of commodity return series

Very little work has been carried out on the multivariate distribution of commodity returns. We review what little has been done below.

Labys and Perrin (1976), examined the intercorrelations among monthly returns of 31 commodities using 20 years of spot data. Some significant positive correlations were found, in particular amongst the oleaginous products. Other positive correlations were found between copper and aluminium, lead and zinc, copper and tin, and wheat and maize. Principal component analysis gave the oleaginous group as the first principal component, accounting for 12% of the joint variation in returns. The composition of the remaining components could not be associated with any of the usually recognisable commodity groups. Apart from the oils, and a few metals, Labys could find no evidence of any significant covariability amongst commodity returns. Dusak (1973) examined the returns on different contracts of the same agricultural commodities. Correlation coefficients of between 0.85 and 0.95 were reported.

Loeb (1979), examined the joint correlation structure of 16 commodities and identified positive correlations amongst the metals. Cluster analysis confirmed these reports.

No author to date has examined (i) the nature of the joint distribution of commodity returns or (ii) the possibility of multivariate temporal dependence.

2.6 Summary of Chapter 2

In this chapter we reviewed the literature on empirical studies of commodity returns distributions. Most authors find that the returns distributions are longer tailed than the Gaussian distribution and that there is little or no evidence of a consistent temporal pattern. The application of naive trading rules produce mixed results and it is only recently that the statistical validity of trading rules has been investigated. To date no serious empirical multivariate study of commodity returns has been published.

CHAPTER 3

A STUDY OF THE UNIVARIATE DISTRIBUTION OF PRICES

In Chapter two we reviewed the literature dealing with the empirical study of the distribution of commodity spot and futures prices and returns. There are conflicting reports on the presence of serial dependence and most researchers find that returns are non normal and skewed, the degree of skewness depending on the differencing interval. In this study we will limit our interest to daily returns defined as:

$$\text{(i)} \quad x_t = p_t - p_{t-1} ,$$

and the logged returns as:

$$\text{(ii)} \quad x_t = \log(p_t / p_{t-1})$$

3.1 Design of univariate study

Most empirical studies report that the variances of returns vary over time. The tests for serial dependence and of distributional form and stability of population parameters are greatly affected by changing variability. We decided therefore initially to examine the data in a sequence of short time periods in the hope that within these periods the variances would remain fairly constant. The data set covers the period January 1974 to December 1979. However we only have information on the rubber series from March 1975 onwards. We decided therefore to examine all four series for the period of 58 months (1218 days) from March 1975 to December 1979. Each of the years 1976 to 1979 was divided into three periods of four months (approximately 84 days) and 1975 was cut into two periods of five months (approximately 103 days). Thus we have 2 + 3 x 4 = 14 periods of approximately equal duration to examine.

In each period we needed to select one cocoa contract, one coffee contract, one sugar contract and one rubber contract and examine the daily returns. For each commodity we have between 6, 7 or 8 different futures contracts to choose from. As is reported in Chapter 5, all the daily returns on contracts of differing delivery dates for a given commodity appear (not surprisingly) to be highly positively correlated and so it may not matter which one we choose for the univariate examination. However, the behaviour of the price of a futures contract that has entered the delivery month ,the "near" contract, may not be representative of a typical futures price[1]. Also, the volume of trading in those futures contracts that have just started trading ,the "far" contracts, may be so small as to make the prices also non representative. It was decided, therefore, to examine prices that are neither at the beginning nor at the end of the duration of a contract. Thus each of the 14 subperiods chosen contains a four month section of a contract that is roughly in the middle of its life span.

For a list of the contracts chosen with dates for this univariate study see Appendix B. Fig. 3.1 shows pictorially what section of each contract was used. The layout of this chapter is as follows. In section 3.2 the statistical methods are reviewed and the results of all of the procedures are discussed in section 3.3. Section 3.4 gives a summary. In section 3.5 and 3.6 a separate study dealing with the search for long term temporal dependence is carried out.

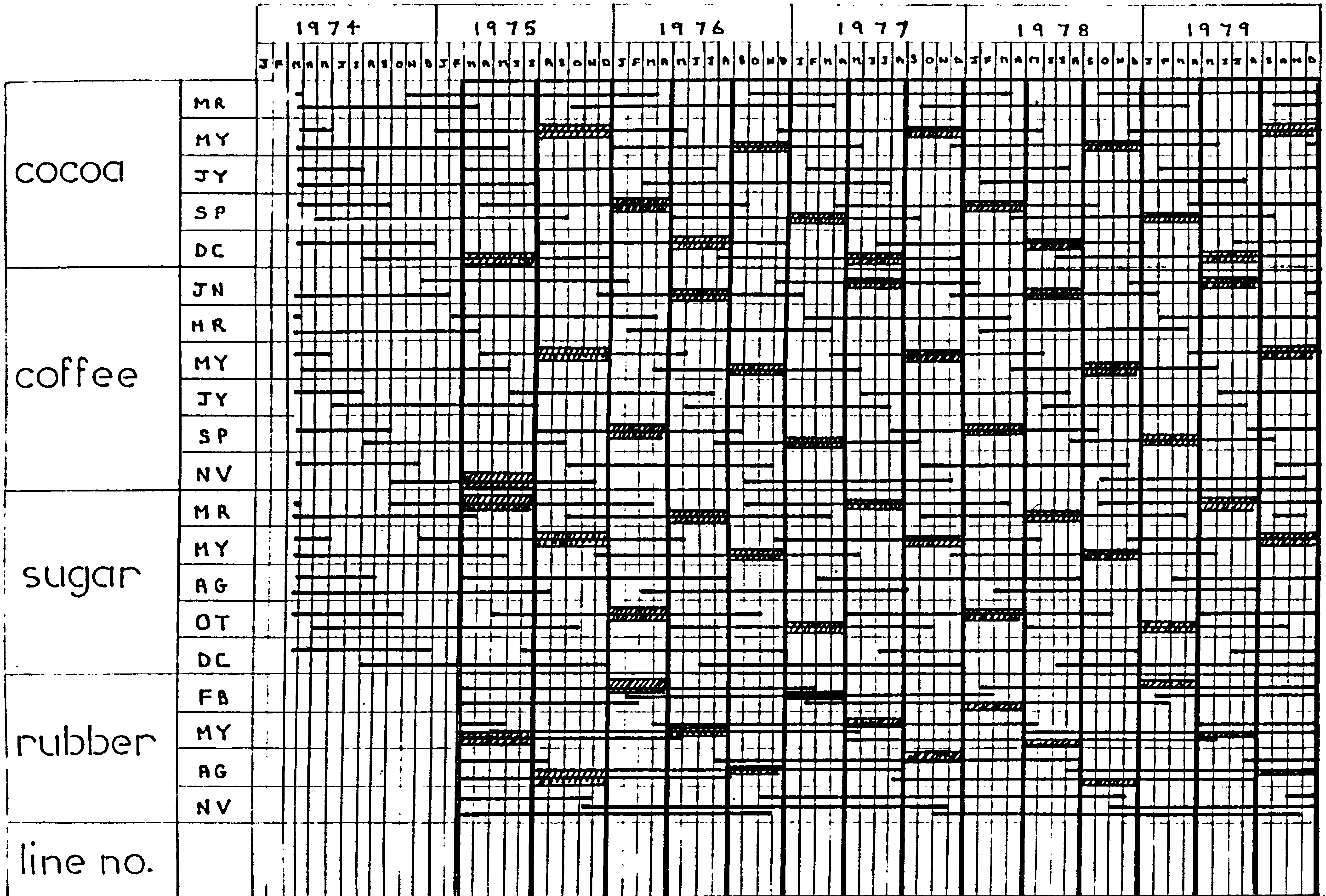

Fig 3.1 Pictorial representation of sections of contracts used in uniuariate study

3.2 Statistical procedures

As in most previous empirical research on commodity futures prices we study the returns and logged returns and examine them in each period for (i) temporal dependence (ii) stability of population parameters and (iii) distributional form. Outlined below is a brief summary of the tests and statistical procedures used. The results of all the tests appear in Tables 3.1 to 3.6. Plots of the unlogged returns appear in Figs. 3.2 to 3.5.

3.2.1 Univariate temporal dependence

The two most commonly used techniques to investigate the presence of temporal dependence are the runs test and the examination of a correlogram i.e. a set of serial correlation coefficients.

The Runs Test

The classical Runs Test examines the sequence of returns. Each return is classified into one of two categories, eg positive or negative. In this study the two categories chosen are: above the median and below or equal to the median. A run is defined as an unbroken succession of outcomes of the same kind. It can be shown[2] that if n_1 = number of outcomes in the first category and n_2 = number of outcomes in the second category and $n = n_1 + n_2$ is large (greater than 20), that the number of runs, r , is approximately normally distributed with mean μ_r and standard

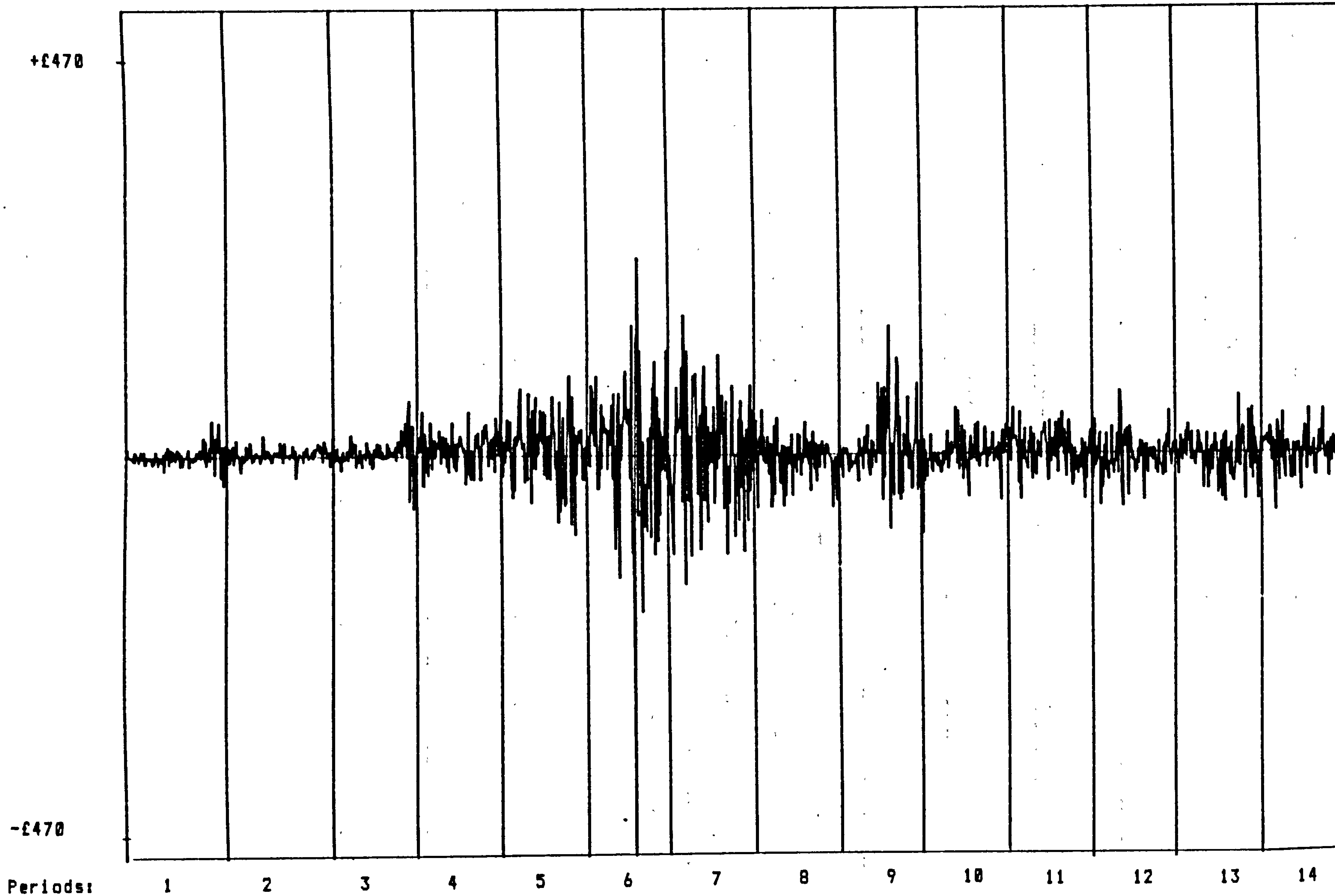

Fig 3.2 Cocoa series returns examined in univariate study

Fig 3.3 Coffee series returns examined in univariate study

Fig 3.4 Sugar series returns examined in uivariate study

Fig 3.5 Rubber series returns examined in univariate study

deviation σ_r given by:

$$\mu_r = \frac{2 \; n_1 n_2}{n_1 + n_2} + 1$$

$$\sigma_r = \sqrt{\frac{2n_1n_2(2n_1n_2-n_1-n_2}{(n_1+n_2)^2(n_1+n_2-1)}}$$

A test of temporal independence is then to compute z_r, where

$$z_r = (r - \mu_r)/\sigma_r$$

in which under the null hypothesis of randomness $z_r \sim N(0,1)$.
The $P(z_r)$ values for each commodity and for all 14 subperiods appear in Tables 3.1 to 3.4. A virtue of the runs test is that no assumptions are made about the distribution of the returns and it is insensitive to the presence of outlying observations and errors in the data.

<u>Serial correlation coefficients</u>

It is standard practice in the study of time series to plot and examine the correlogram or autocorrelogram. The correlogram is a plot of the sample serial correlation coefficients, r_k, at various lags, k, against k. Each r_k is computed using the expression:

$$r_k = \frac{\sum_{i=1}^{n-k} (x_i - \overline{x})(x_{i+k} - \overline{x}) \; / \; (n-k)}{\sum_{i=1}^{n} (x_i - \overline{x})^2 \; / \; n}$$

In the analysis below we computed r_k, for k up to and including r_{20} for each period[3]. If the returns constitute a sequence of serially independent identically normally distributed random variables (the null hypothesis) the r_k values are each normally distributed[4] with a mean of zero and a standard deviation of approximately $1/\sqrt{n}$. Furthermore, under the null hypothesis, the r_k's are mutually independent. A test of serial independence thus involves the computation of

$$z_k = \frac{r_k}{\sqrt{1/n}} \qquad \text{for} \quad k = 1,2,\ldots,20$$

Values of z_k outside the bounds delineated by the normal tables (eg 1.96 for 5% test) are regarded as significant.

In the computation of 20, r_k, values one would of course expect, on average, one (5% of 20) to result in a significantly large z_k value (at the 5% level). In the analysis of time series, interest is usually centred on the r_k values of small lags and in particular correlograms are examined for certain patterns consistent with ARIMA type models[5]. Taylor (1980) also reports that in time series in which the variance is non stable the standard error of $1/\sqrt{n}$ for r_k is no longer valid. We return to this subject again in section 3.5.

3.2.2 Stability of population parameters

An area of central interest in studying the distributions of returns over the 14 subperiods is the stability or otherwise of the population parameters. As mentioned in section 3.2.1 tests for temporal dependence are dependent on the stability of variances.

Recall also that at the outset of this research programme we were interested mainly in the possible application of a Markowitz type Portfolio Analysis to the set of returns. Obviously the stability of the underlying population is crucial for a successful delineation of efficient sets from period to period. It was decided, therefore, to test for the stability of (i) variances and (ii) means. In the foregoing, the test procedures are based on the assumption that the populations are normal but fortunately this assumption can be relaxed with samples as large as the ones being examined in this study.

Comparing variances

If s^2_i and s^2_{i+1} are the usual unbiased estimates of the population variances σ^2_i and σ^2_{i+1} respectively, each estimated on n_i and n_{i+1} observations in periods i and (i+1) then to test the null hypothesis:

$$H_0: \sigma^2_i = \sigma^2_{i+1}$$

against the alternative hypothesis:

$$H_1: \sigma^2_i \neq \sigma^2_{i+1}$$

the sample statistic:

$$F_i = \frac{s^2_i}{s^2_{i+1}}$$

is computed. Under the null, F_i is distributed according to the F distribution on $(n_i - 1)$ and $(n_{i+1} - 1)$ degrees of freedom. Of course, F should be near unity. Very large or very small values of F_i indicate the null hypothesis is probably not true. We use the usual F tables to measure significance.

Comparing Means

The comparison of the means for periods i and i+1 of two (normal) populations is complicated by the dependence of the distributions of the sample means on the population variances.

(a) If the population variances can be considered equal the procedure to test : $H_0\ \mu_i = \mu_{i+1}$,
against $H_1\ \mu_i \neq \mu_{i+1}$
is to calculate T_i where,

$$T_i = \frac{\bar{X}_i - \bar{X}_{i+1}}{\sqrt{s_p^2\left[\dfrac{1}{n_i} + \dfrac{1}{n_{i+1}}\right]}}$$

and s_p = pooled estimate of common variance.

Under the null hypothesis, T_i, is t distributed on $n_i + n_{i+1} - 2$ degrees of freedom. This test was used when the test of equal variances produced a non significant result.

(b) If the population variances cannot be assumed equal, the problem of comparing means is known as the Behrens-Fisher problem. The problem arises since T_1 defined above, does not follow the t distribution[6]. One suggested test[7] of the above null hypothesis uses a critical region of the form.

$$T_2 = \frac{\left|\overline{x}_i - \overline{x}_{i+1}\right|}{\sqrt{\frac{s^2_i}{n_i} + \frac{s^2_{i+1}}{n_{i+1}}}} > C$$

in which the test statistic, T_2, quite reasonably, is a measure of the differences scaled according to an estimate of the standard deviation of the difference in sample means. It can be shown that if C is chosen as the $(100-\alpha/2)$ percentile of the t distribution on $(n_0 - 1)$ degrees of freedom, where $n_0 = \min(n_i, n_{i+1})$ the probability of the critical region under the null hypothesis is at most α. Accordingly, whenever the test of the equality of variances failed we used T_2 with the appropriate degrees of freedom[8].

Are the mean returns significantly different from zero?

In each period $i = 1,2,\ldots,14$, it was decided also to test the hypothesis: $H_0: \mu_i = 0$ against the alternative: $H_1 \mu_i \neq 0$, using the statistic :

$$T_3 = \overline{x}_i \Big/ \sqrt{\sigma^2_i / n_i}$$

Under H_0, T_3 is t distributed on $(n_i - 1)$ degrees of freedom. In essence this is a test for a persistent long term trend in the prices in each subperiod. If T_3 is significantly positive (negative) then the daily returns are, on average, positive (negative) suggesting a significant rise (fall) in the prices over the time period considered.

Tables 3.5 give the results of the tests for stability in variances and means. The values of T_3 appear in Tables 3.1 to 3.4.

3.2.3 Examinations of univariate distributional form

There is an extensive literature on the distributional form of daily, weekly and monthly stock and commodity futures returns. For details see Chapter two. Recall that the initial motive for this research was the application of Markowitz type Portfolio Theory to the commodity futures market. A sufficient condition for this application is that returns be normally distributed. Therefore, rather than spend time on estimating the parameters of various Stable Pareto or t distributions that could explain the returns, we concerned ourselves with one question; are the daily returns normal?

In this study three methods were used to investigate the question of univariate normality :(i) the coefficient of skewness, b_1, (ii) the coefficient of kurtosis, b_2 and (iii) the normal order plot together with the associated Shapiro and Wilk's W-test (1965) for normality.

The coefficients of skewness and kurtosis

If m_k = the kth moment about the mean then

$$\sqrt{b_1} = \frac{m_3}{(m_2)^{3/2}}, \text{ and } \quad b_2 = \frac{m_4}{(m_2)^2}$$

For normal samples:

$$E(\sqrt{b_1}) = 0 , \quad SE(\sqrt{b_1}) \sim \sqrt{6/n}$$
$$E(b_2) = 3 , \quad SE(b_2) \sim \sqrt{24/n}$$

If n is large, $\sqrt{b_1}$ is normally distributed. For small n, Biometrika tables give percentiles for normal samples. Biometrika tables also give percentiles of b_2 for normal samples. For samples of size n = 80 significant values of $\sqrt{b_1}$ could be outside the interval + 0.43 and significant values of b_2 would be outside the interval 2.27 to 3.87.

Distributions resulting in b_2 values less than 3 are termed platykurtic and are characterised by frequency curves more flat-topped and shorter in the tails than the normal distribution. Distributions with b_2 values exceeding 3 are termed leptokurtic and are more sharply peaked and longer in the tails than the normal distribution. Most empirical studies to date find returns leptokurtic.

The Shapiro and Wilk's W - test for normality

A useful first step in studying the question of normality of a sample is to examine the normal order plot: a plot of the ordered sample values $x_{(1)} < x_{(2)} < \ldots < x_{(n)}$ against the expected normal order values $z_{(1)} < z_{(2)} < \ldots < z_{(n)}$. Distributions that are normal result in linear plots with slopes equal to the standard deviation of the sample and vertical intercepts equal to the mean. Non normal distributions result in non linear plots. These plots are also very useful in highlighting errors and/or outlying observations from an otherwise normal sample.

There are a number of formal tests of normality associated with the linearity of this plot, one of which is the Shapiro and Wilk's W - test.

For more details see Royston (1982a). Pearson (1982) et al carried out an extensive study into the power of the various tests for departure from normality and concluded that for symmetrical platykurtic distributions and for most skew distributions, the Shapiro and Wilk's W - test is optimal (most powerful). Here we briefly outline the Shapiro and Wilk's W - test.

Let $\underline{z}^T = (z_{(1)}, z_{(2)}, \ldots, z_{(n)})$ denote the vector of expected values of standard normal order statistics, and let $U = (u_{ij})$ be the corresponding (n x n) covariance matrix. If $\underline{x}^T = (x_1, x_2, \ldots, x_n)$ is a random sample on which the W - test of normality is to be carried out, ordered so that $x_{(1)} < x_{(2)} < \ldots < x_{(n)}$. Then we compute

$$W = \frac{\left(\sum_{i=1}^{n} a_i . x_{(i)} \right)^2}{\sum_{i=1}^{n} (x_i - \overline{x})^2}$$

where $\underline{a}^T = (a_1, a_2, \ldots, a_n)$

$$= \underline{z}^T U^{-1} [(\underline{z}^T U^{-1})(U^{-1} \underline{z})]^{-1/2}$$

Shapiro and Wilks (1965) point out that "the coefficients $\{a_i\}$ are just the normalised 'best linear unbiased' coefficients tabulated in Sarhon and Greenberg (1956). The numerator of W is the best linear unbiased estimate of the squared slope of a linear regression of the ordered observations on the expected values of the standard normal order statistics. Note that if one is indeed sampling from a normal population then the numerator and the denominator of W are both, up to a constant, estimating the same quantity, namely, the variance". For non-normal population, these quantities would not in general be estimating the same thing and the numerator will be less than the denominator. Departure from normality is indicated by small values of W. Shapiro and Wilks give tables for the values of a for samples of size n, up to n = 50 and various percentiles of the distribution of W under the null hypothesis of normality.

Royston (1982a), extends the W - test to samples of size n= 2000 by considering the normalising transformation of W:

$$y = (1 - W)^{\lambda}$$
$$z_w = (y - \mu_y)/\sigma_y ,$$

in which the quantities, λ, μ and σ_y are all funcions of log(n). Under the null, z_w, is standard normal and departures from normality are indicated by large values of z_w. The test is one sided.

Royston (1982b) provides an algorithm to compute the expected normal order statistics and Royston (1982c) provides an algorithm to compute λ, μ_y, W, z_w and $Pr(Z_w > z_w)$ = P(W) for samples of size n up to 2000. The algorithm also computes the appropriate best linear unbiased coefficients $\{a_i\}$.

A routine was constructed that computes and records $\sqrt{b_1}$, b_2,and Pr(W) for each of the 14 subperiods and for each commodity.

3.3 Discussion of results of univariate tests

In this section we discuss the results of the tests presented in Tables 3.1 to 3.5 in the order considered in section 3.2. In these and many of the other tables in this work we use the following notation to indicate the degree of significance of a particular test statistic:

Table 3.1

Univariate tests on cocoa series

Period	Significant r_k'S	$P(Z_r)$	$P(W)$	$\sqrt{b_1}$	b_2	T_3
1	+6 -12	0.37	0.00c	0.83b	6.40c	-0.17
	+6 -12	0.37	0.00c	0.93c	5.90c	-0.15
2	+17	0.43	0.75	-0.04	4.06a	1.97
	+13 +17	0.43	0.86	-0.05	4.06a	1.86
3	+4	0.83	0.00c	-0.22	7.28c	1.57
	+4	0.83	0.08c	-0.11	5.30b	1.70
4	+18 -20	0.59	0.66	0.18	3.27	2.90b
	-2 +9 -20	0.59	0.76	-0.04	3.03	2.81b
5		0.19	0.39	-0.28	2.71	1.42
		0.19	0.22	-0.20	2.53	1.55
6	-3 +13	0.66	0.00c	-1.77	11.31c	0.65
	-3 +13	0.66	0.00c	-1.83	11.49c	0.73
7	+3	0.73	0.75	-0.14	2.61	0.41
	+3	0.73	0.94	-0.12	2.86	0.42
8	-19	0.16	0.01a	-0.44	3.34	-2.48a
	-19	0.16	0.02a	-0.51	3.31	-2.56a
9	+9 +11 +19	0.37	0.00c	1.03c	4.98b	0.86
	+11 +19	0.37	0.00c	1.01c	4.59b	0.88
10		0.28	0.00c	0.62b	8.20c	0.37
		0.28	0.01a	0.59a	7.29c	0.36
11	+3 -17	0.83	0.48	-0.13	2.50	0.68
	+3 -17	0.83	0.60	-0.12	2.53	0.72
12	-2 +5 -11	0.74	0.67	0.19	3.83	-1.74
	-2 +5 -11	0.74	0.70	0.24	3.70	-1.73
13	+13 -20	0.13	0.54	-0.07	3.21	-1.33
	+5 +13	0.13	0.53	0.07	3.56	-1.33
14	-2 +15	0.82	0.22	-0.24	3.88a	0.06
	-2 +15	0.82	0.25	-0.19	3.76	0.06

(Upper figures are the unlogged returns results; the lower figures are the logged returns results)

Table 3.2

Univariate tests on coffee series

Period	Significant r_k'S	$P(Z_r)$	$P(W)$	$\sqrt{b_1}$	b_2	T_3
1	+1	0.92	0.00c	3.87c	29.46c	1.42
	+1	0.92	0.00c	4.23c	31.37c	1.55
2		0.32	0.84	-0.08	3.38	0.59
		0.32	0.87	-0.13	3.37	0.56
3	+9	0.99	0.00c	0.94c	7.47c	2.45b
	+9	0.99	0.00c	0.81b	5.66c	2.70b
4		0.33	0.00c	-1.40c	8.70c	0.50
		0.33	0.00c	-1.33c	8.85c	0.48
5	-4	0.99	0.02a	0.67b	6.32c	3.78c
		0.99	0.94	0.29	4.37a	4.10c
6	-3	0.12	0.68	-0.44a	3.66	0.62
		0.12	0.44	-0.46a	3.54	0.65
7	-6 +9 +18	0.31	0.55	0.23	3.23	-1.33
	-5 -6 +9 -13 +18	0.31	0.51	0.14	3.07	-1.20
8	+15	0.10	0.98	0.20	3.94a	-0.61
	+15	0.10	0.93	0.28	3.85a	-0.57
9		0.18	0.13	0.47a	2.92	-0.80
		0.18	0.33	0.46a	3.02	-0.78
10		0.83	0.04a	-0.45a	5.30b	0.39
		0.83	0.21	-0.25	4.43b	0.40
11	+5	0.66	0.78	-0.10	3.43	-0.31
	+5	0.66	0.77	-0.17	3.61	-0.31
12	-6	0.58	0.42	-0.19	4.49b	2.33a
		0.58	0.93	0.02	4.14a	2.32a
13	+4 +9	0.83	0.00c	2.18c	13.75c	0.73
	+9	0.83	0.00c	2.52c	16.41c	0.76
14	+11 -20	0.18	0.61	-0.46a	3.98a	-1.35
	+11 -20	0.18	0.43	-0.50a	4.19a	-1.39

Table 3.3

Univariate tests on sugar series

Period	Significant r_k's	$P(Z_r)$	$P(W)$	$\sqrt{b_1}$	b_2	T_3
1		0.47	0.59	0.21	3.86	-0.71
		0.47	0.32	0.50a	4.71b	-0.64
2		0.32	0.00c	0.07	7.13c	0.10
		0.32	0.00c	-0.02	6.28c	0.11
3	+9	0.83	0.00c	1.20c	7.48c	0.78
	+9	0.83	0.00c	1.23c	7.52c	0.78
4		0.45	0.56	-0.23	3.28	-0.66
		0.45	0.67	-0.25	3.11	-0.72
5	+10	0.19	0.98	-0.13	3.73	-1.32
	+10	0.19	0.99	-0.09	3.62	-1.28
6	-6	0.66	0.01a	0.84b	5.90c	0.10
	-6	0.66	0.07	0.72b	5.21b	0.10
7	-2 -4 -14 +16	0.45	0.20	-0.17	2.68	-1.46
	-2 -4 -14 +16	0.45	0.22	-0.18	2.61	-1.39
8	-1 +3 -4 +13 -15 +16	0.10	0.16	0.55a	3.73	-0.19
	-1 +3 -4 +13 -15 +16	0.10	0.14	0.54a	3.74	-0.18
9		0.66	0.47	-0.24	4.86b	-1.51
		0.66	0.38	-0.32	5.31b	-1.42
10	+5 -13	0.66	0.64	0.00	2.73	-1.34
	+5 -13	0.66	0.79	0.02	3.06	-1.20
11		0.51	0.14	-0.62b	5.58b	-0.08
		0.51	0.31	-0.52a	5.16b	-0.09
12		0.74	0.07	0.48a	2.82	-0.48
		0.74	0.12	0.51a	2.98	-0.49
13	-12 -18	0.53	0.90	-0.36	4.01a	0.60
	-12 -18	0.53	0.81	-0.41	4.15a	0.60
14	-3	0.18	0.96	0.03	3.95a	1.90
	-3	0.18	0.93	0.12	4.05a	2.02a

Table 3.4

Univariate tests on rubber series

Period	Significant r_k's	$P(Z_r)$	$P(W)$	$\sqrt{b_1}$	b_2	T_3
1	+4-7+10-16+18-20	0.36	0.29	0.00	3.84	1.47
	+4-7+10-16+18	0.36	0.45	-0.01	3.32	1.51
2		0.17	0.77	0.06	3.04	0.47
	-4	0.17	0.86	0.04	3.13	0.45
3	+4 -14 +16	0.83	0.29	0.61a	3.84	1.83
	-14 +16	0.83	0.26	0.49a	3.06	1.91
4	-4 +13	0.75	0.99	-0.05	3.47	0.10
	-4 +13	0.75	0.99	0.06	3.66	0.10
5	-5 +7 +10 +17	0.39	0.55	0.18	3.16	0.28
	-5 +7 +17	0.39	0.65	0.25	2.94	0.30
6		0.65	0.76	-0.21	4.20a	-1.56
		0.65	0.68	-0.20	4.06a	-1.63
7		0.56	0.81	0.08	2.60	-0.35
		0.56	0.87	0.05	2.66	-0.33
8		0.56	0.68	-0.10	3.35	-1.66
		0.56	0.59	-0.20	3.20	-1.73
9		0.37	0.54	0.31	3.61	0.38
		0.37	0.58	0.25	3.48	0.37
10	+10 +14	0.28	0.82	0.19	2.72	0.77
	+10 +14	0.28	0.72	0.19	2.65	0.79
11	-15	0.50	0.90	-0.16	3.10	-0.51
	-15	0.50	0.87	-0.19	3.06	-0.55
12	-9	0.91	0.02a	-0.17	2.07a	0.06
	-9	0.91	0.04a	-0.16	2.11a	0.06
13	+11	0.66	0.89	0.32	3.81	-0.58
	+11	0.66	0.95	0.20	3.59	-0.61
14		0.18	0.65	0.14	2.65	1.09
		0.18	0.60	0.16	2.65	1.12

Table 3.5

Stability of parameter tests

Period Comparison	Cocoa F_1	Cocoa T_2	Coffee F_1	Coffee T_2	Sugar F_1	Sugar T_2	Rubber F_1	Rubber T_2
1 - 2	1.89b	1.28	5.83c	-1.11	2.13c	0.65	1.37	-0.83
	2.36c	1.14	8.33c	-1.29	2.38c	0.60	1.56a	-0.91
2 - 3	4.45c	0.68	5.83c	2.10a	3.85c	0.30	3.19c	1.43
	2.08c	0.49	2.40c	2.07a	3.70c	0.30	1.75a	1.34
3 - 4	1.45	1.21	3.18c	-0.77	1.63a	-1.00	2.45c	-0.91
	1.04	0.74	2.34c	-1.08	1.39	-1.05	1.86b	-1.05
4 - 5	3.52c	-0.12	1.05	2.35a	1.02	-0.45	1.01	0.13
	1.49	-0.62	2.86c	1.71	1.73a	-0.57	1.30	0.12
5 - 6	5.25c	0.04	8.12c	-0.66	2.18c	1.15	2.44c	-1.09
	2.89c	0.14	3.77c	-1.30	2.13c	1.11	2.38c	-1.16
6 - 7	1.98b	-0.30	1.06	-1.38	1.96b	-0.92	1.56	1.01
	2.18c	-0.38	1.98b	-1.35	1.64a	-0.93	1.11	0.97
7 - 8	6.95c	-1.27	5.82c	1.00	1.07	0.87	1.20	-0.99
	4.76c	-1.44	2.50c	0.71	1.17	0.81	1.14	-1.04
8 - 9	2.64c	2.03a	2.73c	0.10	1.14	-0.91	1.27	1.48
	3.28c	2.01a	1.93b	-0.10	1.04	-0.89	1.10	1.49
9 - 10	2.22c	-0.52	2.80c	0.75	1.04	0.17	1.76a	0.38
	2.33c	-0.55	2.86c	0.75	1.22	0.08	1.35	0.35
10 - 11	1.26	0.18	3.52c	-0.49	1.26	0.82	1.04	-0.92
	1.64a	0.16	3.03c	-0.51	1.03	0.79	1.14	-0.95
11 - 12	1.06	-1.71	2.07c	1.60	2.13c	-0.21	1.13	0.39
	1.34	-1.77	2.33c	1.55	2.00c	-0.22	1.14	0.42
12 - 13	1.03	0.29	6.21c	-0.20	1.03	0.77	1.32	-0.47
	1.43	0.11	3.58c	-0.43	1.07	0.77	1.23	-0.49
13 - 14	1.31	1.03	5.40c	-1.22	8.51c	1.61	1.78a	1.13
	1.25	1.02	5.26c	-1.27	4.41c	1.58	1.85b	1.17

F_1 = equality of variances F ratio
T_2 = test of two means statistic outlined in section 3.2.2

(Upper figures are the unlogged returns results, lower figures are the logged returns results)

If the statistic is significant at the 5% level, entry = a

If the statistic is significant at the 1% level, entry = b

If the statistic is significant at the 0.1% level, entry = c

3.3.1 Results of temporal dependence tests

Runs Test

No significant results obtained. The smallest $Pr(z_r)$ value obtained was 0.10.

Serial correlation coefficients

Some series contained one, two or three significant (at 5% level) r_k values, many others contained none. There certainly does not appear to be any consistent positive or negative serial correlation. Table 3.6 below gives a count of the number of significant r_k values over the entire period for each set of returns.

We see that the number of significant r_k values almost exactly equals what one would expect (ie 5%) under the null hypothesis of no temporal dependence. It is interesting to note also on referring to Tables 3.1 to 3.4 that the use of logged returns produces virtually identical results.

In conclusion, therefore, we note that there seems to be no evidence at all of any temporal dependence in any of the series.

Table 3.6

No. of significant serial correlation coefficients at 5% level

	no. of sig. r_k values	no. of possible sig. values	% of sig. r_k values
Cocoa series	22	280	7.9
Coffee series	13	280	4.6
Sugar series	20	280	7.1
Rubber series	20	280	7.1
Total	75	1120	6.7

3.3.2 Results of stability tests

Variance Stability

It is clear, referring to Figs. 3.2 to 3.5, that the variances are non constant. The F ratios in Table 3.5 reinforce the observation. For the cocoa returns, of the 13 subperiods' pairwise comparisons, 8 (7 with logged returns) of the F ratios were significant at the 5% level. This re-affirms the previous findings of excessive fluctuations in variance, and although in many instances the use of logged returns results in a smaller F ratio, the statistics are still what one would consider significant. With the coffee returns, 13 (11 with logged returns) of the 13 comparisons result in significant F ratios. The sugar returns yield 7 (6) significant changes whereas the rubber returns show 5 (5) significant changes.

These results then confirm the previous empirical findings of non constant variance and justifies the splitting up of the data into subperiods. Testing for normality and serial dependence over the whole five year period would certainly result in spurious conclusions.

Stability of Means and Evidence of Trends

These two tests are dealt with together since they are obviously interdependent.

We consider the cocoa series first. In period 4 (middle of 1976) a T_3 value of 2.90 is highly significant. In this period the mean change in the price of a cocoa contract was £6.46 per tonne per day, resulting in a change of £560 per tonne over the 4 month period. In fact most T_3 values for cocoa are positive until period 7, reflecting the rise in cocoa prices from 1975 to mid 1977. In only one pairwise comparison do we get significant T_1 or T_2 values. Similar results are obtained for the coffee returns with significant T_1 values in early and late 1976. Referring to the plot of coffee returns in Fig. 3.3 one can see when and why we get these significant results. Two (one with logged returns) pairwise comparisons yield significant T_2 values indicating the development of, or the disappearance of, a trend from one period to another and these occur in the early and late periods of 1976. Of the sugar results, only 1 of the T values, (the T_1 value) yields a significant result. The rubber series produced no interesting values.

In nearly all situations the absolute values of the T statistics of the logged data were slightly larger than the corresponding T statistics of the unlogged data. This is due to the slight reduction in the variance and associated standard errors afforded by logging returns. However, in most cases the changes caused by the taking of logs had very little effect on the resultant tests of significance.

3.3.3 The question of normality

Tables 3.7 and 3.8 give the number and nature of significant (at 5% level) $\sqrt{b_1}$ and b_2 statistics. Table 3.9 records the number of P(W) values smaller than 0.05. We make the follwing observations:

(i) The cocoa and coffee returns exhibit non-normality most frequently (using any of the measures) and the rubber returns least frequently (only once using the W Test).

(ii) Logging the returns generally tends to reduce the $\sqrt{b_1}$ and b_2 values and increases the P(W) values. Certainly, the use of logs tends to make returns more normal.

(iii) Except in the case of sugar, the significant $\sqrt{b_1}$ values do not show any sign of persistent positive skew.

(iv) Nearly all significant b_2 values are greater than 3 reinforcing previous empirical findings of leptokurtic distributions.

(v) From Tables 3.1 to 3.4, in nearly all cases in which we obtain a significant P(W) statistic the corresponding $\sqrt{b_1}$ or b_2 statistic is significant. There are some situations, however, in which $\sqrt{b_1}$ or b_2 is significant and P(W) is not.

The W - test is therefore the most useful single statistic (compared to $\sqrt{b_1}$ and b_2) for investigating the normality or otherwise of a sample. Under extreme departures from normality, however, as in the early periods of cocoa and coffee returns, the W - test mimics the $\sqrt{b_1}$ and/or b_2 tests but is less informative.

Referring to Figs. 3.2 to 3.5 it can be seen that possibly one reason for the many instances of significant $\sqrt{b_1}$, b_2 and P(W) statistics could be the sudden changes in variance within some periods (eg period one of cocoa and period two of sugar). A second possible explanation is the

Table 3.7

No. and nature of sig. (5% level) skewness statistics (counts in parenthesis relate to logged returns)

No. of sig. $\sqrt{b_1}$'s	Cocoa	Coffee	Sugar	Rubber
$\sqrt{b_1} > 0$	3(2)	4(4)	4(4)	1(1)
$\sqrt{b_1} < 0$	2(2)	5(3)	1(1)	0(0)
Total	5(4)	9(7)	5(5)	1(1)

Table 3.8

No. and nature of sig. (5% level) kurtosis statistics

No. of sig. b_2's	Cocoa	Coffee	Sugar	Rubber
$b_2 > 3$	7(6)	9(8)	8(8)	2(1)
$b_2 < 3$	0(0)	0(0)	0(0)	1(1)
Total	7(6)	9(8)	8(8)	3(2)

Table 3.9

No. of P(W) values less than 0.05

Cocoa	Coffee	Sugar	Rubber
6(5)	6(4)	3(2)	1(1)

presence of a single outlying observation. A detailed examination of each period for anomalous values using normal order plots and a specially constructed outlier detection routine was carried out but we leave the discussion of these techniques together with the results to Chapter 4.

3.4 Summary of univariate tests

For the moment we separate the rubber returns from consideration. The cocoa, coffee and sugar returns exhibit plenty of evidence of non normality - even after logging the data. Variances fluctuate considerably throughout the five year period and for the most part the mean returns in each period are not statistically significantly different from zero. These results lead one to seriously doubt the appropriateness of applying a Markowitz Portfolio type analysis to the data.

However the returns in the rubber series are what could be described as 'well behaved'. Only one period showed evidence of non-normality and although the variances changed in the earlier half of the five year period, reference to Fig. 3.5 shows that this was much more gradual than for the other series considered.

In none of the approximately 84 - day periods examined did we find any evidence of temporal dependence. Taylor (1980), however, examined futures returns over much longer periods and produced quite strong evidence in favour of certain conjectured models for trends that resulted in significant serial correlation. We investigate Taylor's technique in section 3.5.

3.5 A review of Taylors (1980) study of long series of financial prices.

In this section we briefly review Taylor's model and his recommended method for dealing with long series of financial prices. In section 3.6.1 we apply Taylor's technique to series woven together from 5 contracts each approximately one year in length. In section 3.6.2 we apply Taylor's technique to long series that are woven together to approximate to contracts with constant maturity dates.

The Model

Taylor (1980) proposed a number of models of financial prices. Here we consider only the simplest: the basic trend model. If x_t, $t=1,2,\ldots,n$ is the sequence of logged returns then:

$$x_t = \mu_t + e_t$$

with $E(e_t) = 0$ and $E(e_t, e_{t+k}) = 0$ for $k \neq 0$

One usually would have set $\mu_t = 0$ for all t?. Taylors innovation was to cosider μ_t as stochastic, with

$$\mu_t = \begin{cases} \mu_{t-1} & \text{with probability } p \\ \bar{\mu} + \eta_t & \text{with probability } 1-p \end{cases}$$

In which $\mathrm{Cov}(\mu_s, e_t) = 0$ for all s and t, and $E(\mu_t) = \bar{\mu}$.

The η_t are a series of identically distributed uncorrelated random variables with mean zero and each independent of the previous one.

Taylor suggests that we would expect a priori that $1 - p$ and the ratio $Var(\mu_t)/Var(e_t)$ to be small. This would be consistent with small and infrequent changes in the underlying trend, μ_t. Prices, therefore, would tend to move in one direction (the trend) for a period of time and that these trends themselves change in a random and unpredictable fashion. The mean, ζ duration of such trends is shown to be:

$$\zeta = \sum_{k=1}^{\infty} \frac{k(1-p)}{p^{k-1}} = \frac{1}{1-p}$$

The rationale for such a model is that the trends are responses to the anticipated supply and demand for the commodity. New information relating to supply and demand arrives randomly and relatively infrequently and so trends alter in an unpredictable manner.

If indeed returns can be explained by the above model then Taylor shows that the theoretical autocorrelations, π_k, are non zero and are given by:

$$\pi_k = b\,p^k \quad , \quad k = 1, 2, \ldots \quad ; \text{ and } b > 0$$

The constant b is a function of the parameters defining the stochastic processes e_t and η_t and is expected apriori to be small.

The Problem of Fluctuating Variance

All previous research (including this study) report series of returns with fluctuating variances. Until recently the effect of this changing variance on the sampling distribution of the serial correlation coefficients has been unknown and ignored. Taylor and Kingsman (1979) examined the problem and proposed two alternative models that seemed to describe the change in variance quite well. One model specified that the logarithm of the standard deviation follow an autoregressive process of order 1. The second model specified that the standard deviation follow a Markov Chain with three states (low, medium and high values). Extensive simulations showed that either of the two models was a good candidate for explaining the observed variance fluctuations.

Recall that, with constant - variance series in which there is no temporal dependence, the serial correlation coefficients have variances of $1/n$. Simulations by Taylor and Kingsman using samples of size $n = 1000$ with the two postulated variance processes showed that the serial correlation coefficients had variances of $1.34/n$ and $1.47/n$ for the Markov and autoregressive processes respectively. The usual two sided tests using standard errors of $1/\sqrt{n}$ are consequently invalid.

One of the methods Taylor recommends to overcome this problem is to calculate the serial correlation coefficients on the rescaled returns y_t:

$$y_t = x_t / a_t$$

in which a_t is an exponentially smoothed estimate of the average of the absolute changes in returns, computed using

$$a_t = \alpha |x_{t-1}| + (1 - \alpha) a_{t-1}$$

with α set at some suitable value (Taylor suggests 0.1).

In this way the series, y_t, should be of approximately constant variance. Simulations with both of the fluctuating variance processes showed that the, r_k's, calculated by the recommended method had variances very near the expected value of $1/n$.

The test statistics Q , T and U

In order to test the null hypothesis of a random walk (in which all the π_k's are zero) against the alternative hypothesis of a trend model (in which all the π_k, values can be expressed as $\pi_k = bp^k$) Taylor (1980) considered the three test statistics Q , T and U :

$$Q = n \sum_{k=1}^{m} r_k^2$$

$$T = \sum_{k=1}^{m} \phi^k r_k \bigg/ \sqrt{\sum_{k=1}^{m} \phi^{2k}/n}$$

$$U = \sum_{k=2}^{m} \phi^k r_k \bigg/ \sqrt{\sum_{k=2}^{m} \phi^{2k}/n}$$

with $0<\phi<1$

If the null hypothesis is true, each r_k is independently normally distributed with mean zero and variance 1/n and so Q would be asymptotically chi-squared distriubted on m degrees of freedom. The T and U statistics would be asymptotically normally distributed with mean zero and variance unity.

Taylor points out that previous researchers have used Q in testing for temporal dependence but notes that the technique has low power. Under Taylor's alternative hypothesis the r_k's, are expected to be a sequence of monotonically decreasing positive values and has proposed test statistics T , and U , designed to be sensitive to the possibility of such an alternative hypothesis. If errors are present in a time series they will have most influence on r_1 and thus Taylor decides to test his series with U. Experience suggests that suitable values of m and ϕ are 30 and 0.92 respectively. In Taylor's study of 11 series, 8 showed strong evidence of non random behaviour. In nearly all series there was a preponderance of positive r_k values. As an example, his cocoa series (from 1971 to 1976) gave a U value of 3.47 with 21 of the first 30 r_k values positive.

3.6. Taylor's techniques applied to ICCH data

Here we describe the results of applying Taylor's technique to long series (5 years) obtained by weaving together contracts in two completely different ways.

3.6.1 The weaving of annual segments of contracts to form one long series

In line with Taylor's (1980) study, consecutive annual contracts of each commodity were woven together, care being taken not to include periods very near the beginning or the end of a contract for reasons noted in section 3.1. For details of the contracts used with dates see Appendix C. Initial values of a_0 are computed using the first 20 observations from each contract and when evaluating the crossproduct term in each r_k the summation is, of course, limited to those days for which y_t and y_{t-k} are the rescaled returns from the same contract. With the cocoa returns, for example, we had to weave five contracts together resulting in 1213 returns (=1218-5). Subtracting 20 returns from each contract for the estimation of a_0 means that each r_k is computed using 1213 - (5x20) = 1113 returns. The first 30 r_k values, the values of Q , U and a count of the number of positive r_k values are given in Table 3.10. In the computation of y_t and r_k, α and ϕ were set to 0.1 and 0.92 respectively.

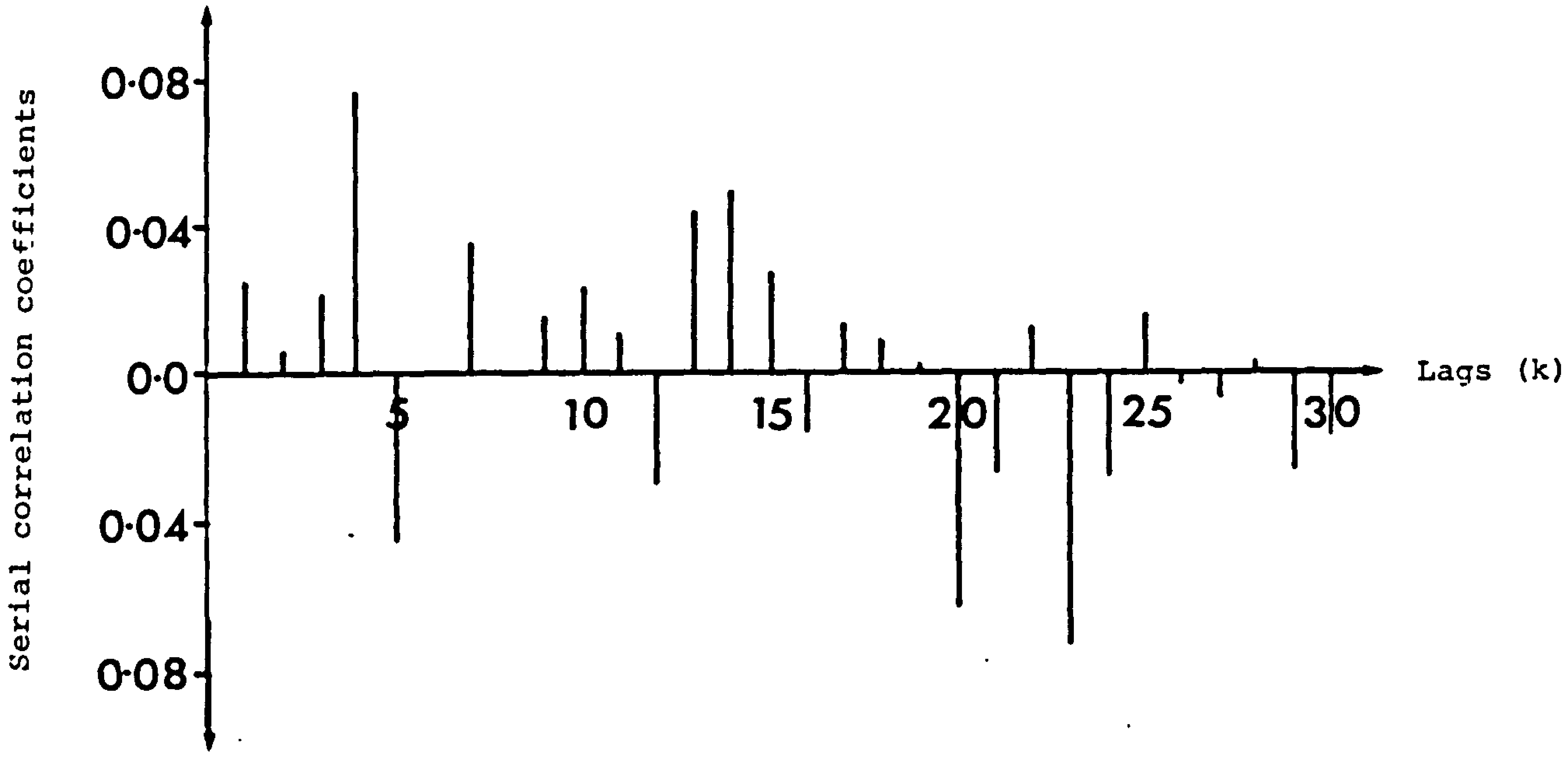

Fig 3.6 Correlogram of long run cocoa series

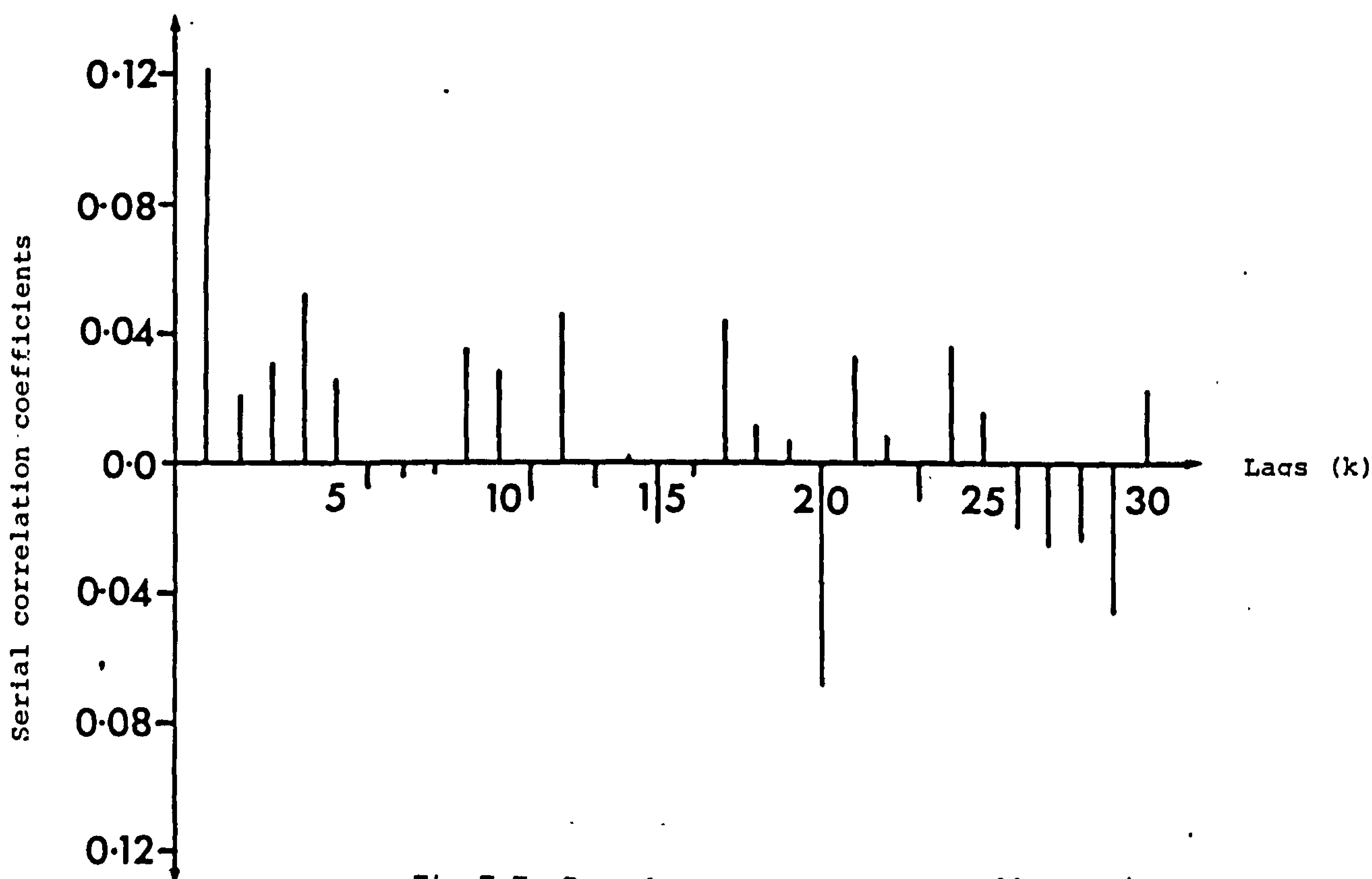

Fig 3.7 Correlogram of long run coffee series

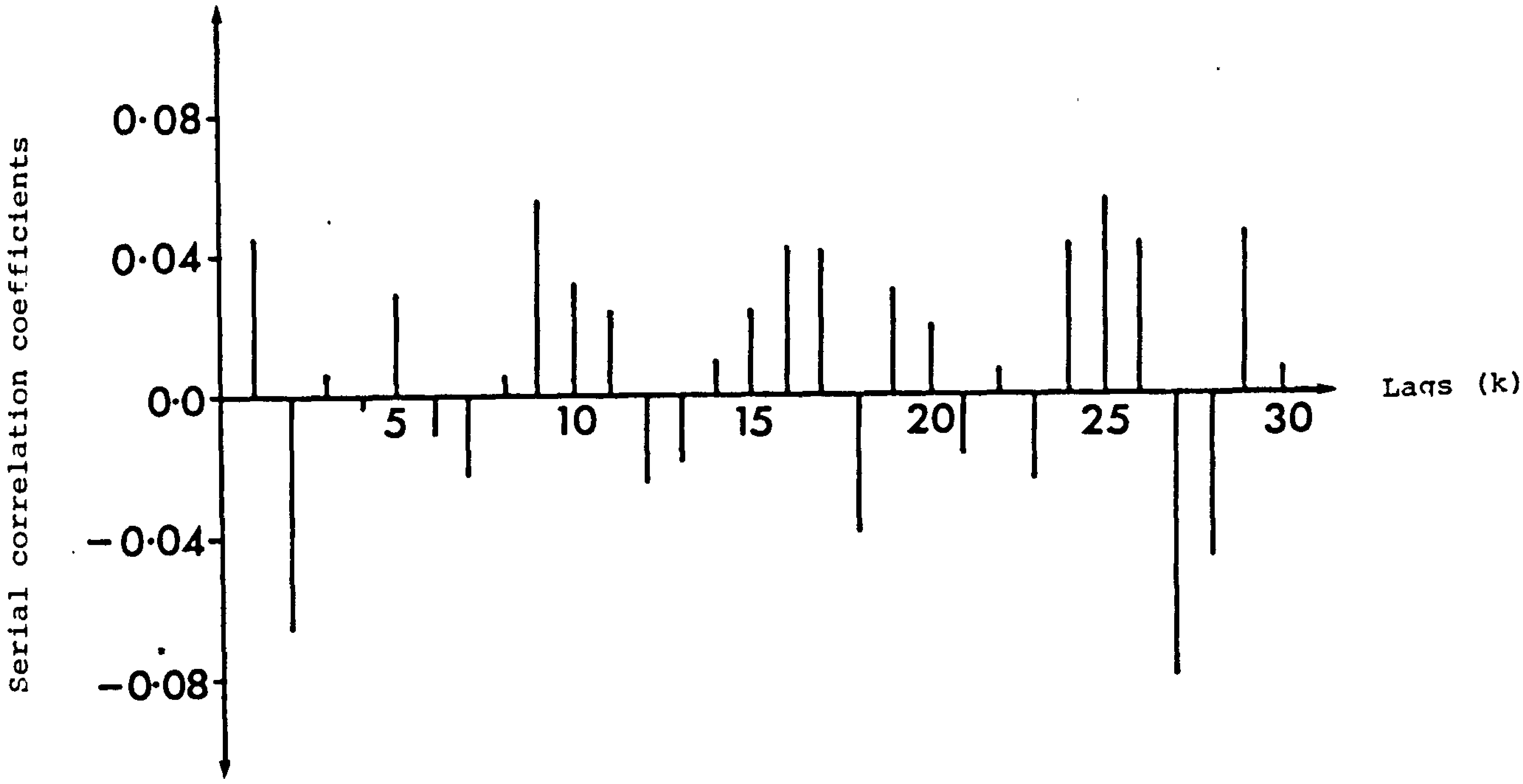

Fig 3.8 Correlogram of long run sugar series

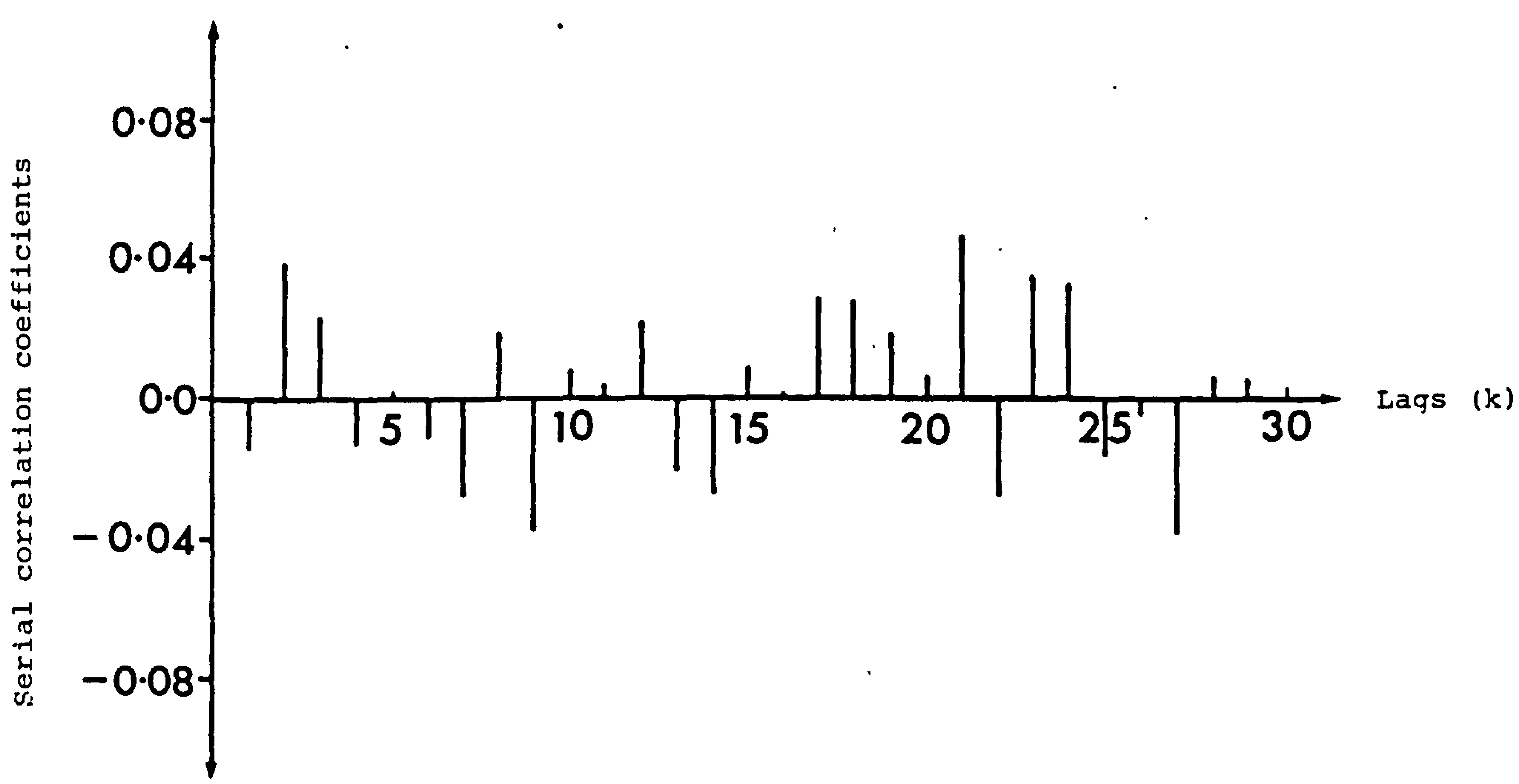

Fig 3.9 Correlogram of long run rubber series

Table 3.10

Results of Taylor's study on annual sections of contracts woven together

Series	no. of returns	no. of positive r_k's at lags 1 to 30: 1-10	11-20	21-30	total	no. of sig. r_k values	Q(30)	U(30)
Cocoa	1113	9	7	3	19	2	33.17	1.37
Coffee	1092	7	5	5	17	2	41.78	1.91
Sugar	1113	6	7	6	19	0	44.46	0.29
Rubber	1117	5	8	6	19	0	17.58	0.59

Referring to Table 3.10 and Figs. 3.6 to 3.9 we note that the number of positive r_k values is greater than 15, half what we would expect under the null, although the number of significant values is small. The correlograms and the U statistics of the cocoa and coffee returns both suggest that an alternative hypothesis of Taylor's type could be true. Note however, that the values of U (1.37 for cocoa, significant at 8.5% level and 1.91 for coffee, significant at the 2.8% level) are much smaller than Taylor's results on cocoa and coffee series from an earlier period (1971 - 1976). In a later work, Taylor (1983), examined cocoa and coffee series over the period (1976 - 1980) and sugar series over the period (1974 - 1980) and produced U values of 3.07, 1.50 and 3.64 respectively. Taylor (1983) also examined the series over the longer periods of (1971 - 1980) for the cocoa and coffee and (1961 - 1980) for the sugar series producing U values of 5.49, 4.83 and 6.58 respectively.

In this study, therefore, two of the four series examined showed evidence of price trends consistent with the models proposed by Taylor.

3.6.2 Taylor's technique applied to all the contracts of a given commodity

In examining the time series woven together in the manner described in section 3.6.1 we are studying the returns of a typical contract over most of its lifespan. If returns are generated by the trend type models suggested by Taylor, then examining a contract over its entire duration may yield misleading results. Consider, for example, the March 1977 contract of cocoa spanning the period January 1976 to December 1976 (255 days). In the early part of the series the prices supposedly represent expectations of the price of cocoa (together with storage costs etc) 15 months into the future. At the end of the series, the prices represent expectations of the cocoa price three months into the future. The responsiveness of prices to anticipated changes in demand and supply of cocoa in the distant months may be different from that when a contract is near to maturity. In other words, it is possible that the stochastic process generating the trend changes may be different at different stages in the life of a contract.

It would be very interesting to examine a series of prices of contracts with delivery dates always a fixed point in the future. This is possible with series of metal futures prices. Each day, for example, a new three month copper futures price is available. One can, however, get an approximation to this situation with the soft futures prices by weaving together contracts in the manner described below.

For simplicity we consider as an example the coffee prices since every two months (about 43 days) a contract expires. Consider Fig.3.10 in which all the futures prices or returns are represented by columns. Each row represents a day. We always have six columns (prices). The prices in

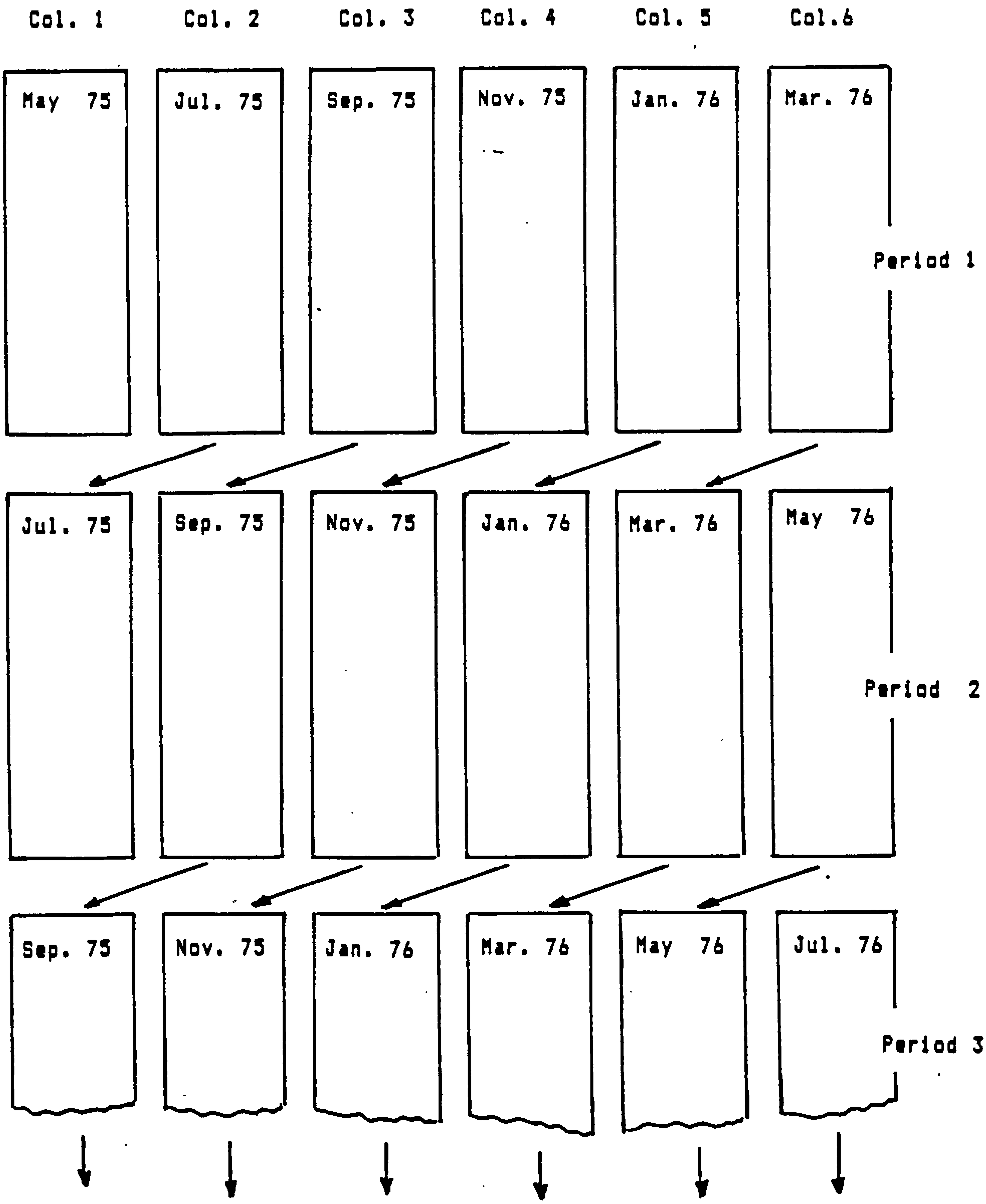

Fig 3.10 Diagramatic representation of the 6 coffee contracts nearest maturity studied in section 3.6.2 and in Chapters 6, 7 and 8

column 1 are the prices of the longest running contract, the one which has the nearest delivery date. Every two months a contract expires and so column 2 becomes column 1 etc etc. At the expiration of a contract all the columns shift one to the left. The prices of the most distant contract are then put into column 6. The data in one of the specially rearranged files is actually set out in this way.

If we weave together all the returns in the first column we will have a sequence of returns (or prices) that will always be between two months to the delivery date and the final delivery date; an average of one month to delivery. Similarly if we weave together all the returns in column 2 we will have a sequence of returns that reflect the situation (trend or whatever) of contracts always between four months and two months of delivery, an average of three months to delivery. Thus we will have six parallel quasi - contracts, each one representing expectations about different points in the future.

We examine all six series of coffee returns looking for positive serial correlation using Taylor's proposed scheme of rescaling returns. Initial estimates of a_0 for each series are obtained using the first 20 observations in period one. At the beginning of period two, three and so on we do not need to put aside the first 20 observations again to re-estimate a_0 as we can use the last estimate from the previous appropriate column corresponding to the same contract. In this way, in the case of coffee, we have 1218 - 29 (periods) = 1189 days with 1189 - 20 = 1169 returns.

An additional problem one encounters when weaving together so many different series of prices (29 series in the case of coffee) is the reduction in the actual number of terms in the cross product expression for r_k. With a lag of 30 and only 43 (say) returns in any given period, there will only be 42 - 30 = 12 terms in the numerator of the

r_{30} expression. Thus in the case of the coffee series there will be a total of 314 cross product terms making up the numerator of r_{30} as compared to 1174 squared terms in the denominator. Obviously the estimates of r_{30} will be affected considerably by this imbalance. It was decided, therefore, that instead of using the usual approximate expression of r_k, we would use the full definition given in Kendall and Stuart (p 375):

$$r_k = \frac{\sum_{t=1}^{n-k} v_t . v_{t+k}}{\sqrt{\sum_{t=1}^{n-k} v^2_t . \sum_{t=1}^{n-k} v^2_{t+k}}}$$

$$\text{where} \quad v_t = y_t - \left[\sum_{t=1}^{n-k} y_t \Big/ (n-k) \right]$$

In this expression, separate means for the series x_t, $t = 1, 2, \ldots, n-k$, and x_{t+k}, $t=k+1,k+2,\ldots,,n$ are required. Note also that the number of terms in the denominator and numerator are equal. Since n_k, the number of terms used in computing r_k, varies considerably with k, the Q and U statistics are now computed using the expressions:

$$Q = \sum_{k=1}^{30} n_k \, r^2_k$$

$$U = \sum_{k=2}^{30} \phi^k \, r_k \Big/ \sqrt{\sum_{k=2}^{30} \phi^{2k}/n_k}$$

Under the null Q and U should be $\chi^2(30)$ and $N(0,1)$ distributed respectively. The number of positive r_k values, the Q and U

Table 3.11

Results of Taylor's study on quasi contracts

Quasi contract		No. of positive r_k values at lags 1 to 30						
		1-10	11-20	21-30	Total	Sig. r_k values	Q(30)	U(30)
Cocoa series	1	4	4	5	13	+4	36.94	-0.80
	2	5	4	4	13	-12	32.19	-0.88
	3	3	4	4	11	-24	37.56	-0.89
	4	4	4	4	12	+4	36.85	-0.68
	5	3	5	5	13	+4	38.14	-0.74
	6	4	6	5	15	-12	34.23	-1.06
Coffee series	1	4	3	3	10	+1 -28	29.29	-2.24
	2	4	2	5	11	+1 -11 -28	34.11	-2.23
	3	3	3	4	10	+1 -11 -28	32.17	-2.20
	4	2	2	4	8	+1 -11 -28	39.37	-2.56
	5	3	1	5	9	+1 -28	41.93	-2.75
	6	4	2	5	11	+1 -6 -11 -28	44.24	-2.63
Sugar series	1	3	3	4	10	-2 -18 -27	29.13	-2.12
	2	4	4	3	11	-2 -18 -29	44.83	-2.11
	3	4	4	4	12	-2 -18 -27 -29	43.09	-1.99
	4	4	4	5	13	-2 -18 -29	46.59	-2.26
	5	5	4	7	16	-2 -18 -29	41.32	-2.03
	6	3	5	7	15	-2 -18	36.61	-1.98
Rubber series	1	4	5	4	13		19.00	-0.79
	2	4	3	7	14		17.53	-1.02
	3	4	4	7	15	-9	14.75	-0.70
	4	4	5	7	16	-9	15.79	-0.62
	5	3	4	7	14	-9	19.86	-0.64
	6	4	4	7	15		21.47	-0.39

statistics are given for each quasi contract in Table 3.11.

3.6.3 Discussion of Table 3.11

The most striking observation from Table 3.11 is that all the U statistics of the quasi contracts of all the commodity series are negative. Furthermore all the statistics are similar for all the quasi contracts of a given commodity. There is an excess of negative serial correlation coefficients. As an example, the fourth coffee series produced 22 negative r_k values. In almost all cases the number of positive r values was less than what we would have expected from a random walk process. Note however that the r_1 value for coffee was significant and positive (= +0.093 for the first quasi contract). This can be compared with an r_1 value of +0.126 from the results of section 3.6.1. None of the U statistics would therefore lead us to reject the null hypothesis in favour of the hypothesis proposed by Taylor. These results are in complete contrast to those produced in section 3.6.1.

Why are there so many negative r_k values, and why are the U statistics all negative? In an attempt to answer this we consider again the approximate formulation of r_k. Recall that the expected value of the approximate r_k is $-1/(n_k-1)$ and the variance is approximately $1/n_k$. So we expect each r_k to be slightly less than zero. If we assume that the above relations are valid for the expected value and variance of r_k computed using the complete expression, we can compute the expected value of U and a 95% probability interval for U.

$$E(U) = \sum_{k=2}^{30} \phi^{k}(-1/(n_k-1)) \Big/ \sqrt{\sum_{k=2}^{30} \phi^{2k}/n_k}$$

$= -0.160$ for $\phi = 0.92$ and using n_k's from the cocoa series

and $\mathrm{Var}(U) = 1.000$ as before

The corresponding 95% probability interval for U, for the coffee series is thus -0.160 ± 1.96, ie -2.12 to $+1.80$. Similar intervals can be calculated for each series.

All the observed U values for the cocoa and rubber series fall inside the 95% probability intervals. All the observed U values for the coffee series and the U value for the fourth sugar series are lower than the lower bounds of the 95% probability intervals. So although we expect each r_k value and the resulting U statistic to be negative, the values observed for the coffee series can be considered as unlikely to result from a random walk process.

Can we explain the conflicting results obtained by examining the coffee returns as described in section 3.6.1 and those obtained using the method described in section 3.6.2? Great care has been taken in the weaving together of all of the periods of each sub-series. There has been no overlapping crossproduct terms used in the computation of the numerator of each r_k. Consequently with lag k = 30 there were only 313 crossproduct and sums of squares terms used in computing r_{30}. This is to be compared with 1125 such terms in the comutation of r_1. Obviously the variance of r_{30} and r_1 will be different and this has been taken into account in the computation of U and Q; but the paucity of terms making up r_{30} for larger lags does not explain why so many values should be negative.

Recall that in Taylor's original model the stochastic process η_t is supposed to represent the random changes in the trend, υ_t. Taylor estimates the mean duration of the trend in his coffee series to be $\zeta \approx 6$ days. The two procedures outlined in section 3.6.1 and 3.6.2 are examining two completely different series. In section 3.6.1 six typical contracts were woven together and each contract produced approximately 200 returns. If a Taylor- type model was generating these returns one would expect about, $200/\ \zeta = 200/6 = 33$, changes of trend in each subperiod. In section 3.6.2, 28 quasi contracts were woven together, each one producing approximately 41 returns. By similar reasoning one would expect $41/\ \zeta = 41/6 = 7$ changes of trend in each subperiod. These figures have been computed assuming Taylor's estimate of $= 6$ is still valid for the coffee returns in the time period we are examining.

The frequency of trend changes in a given contract may be crucial. It is these very changes in the trend that, as Taylor shows, result in small positive theoretical π_k values. In the long series, the number of such trend changes may be quite high and thus the resulting effect on the r_k values is maybe what we are witnessing in section 3.6.1 (recall U = +1.91 for coffee). In the short series, the number of trend changes is probably low and it is possible that such infrequent changes in trend will not result in any observable effect on the r_k's. But how do we explain the negative U's in Table 3.11? Is it possible that even if a Taylor type model is generating the returns, then weaving such short series together may result in series that have negative theoretical serial correlation coefficients? This is a question we leave for later research.

In conclusion we note that the examination of the quasi contracts, that are always a fixed average time to delivery, produced results completely different to those expected and different to those observed by a similar study on five or six real contracts woven together. These interesting and unexpected results may be due to the fact that the trend-type models proposed by Taylor are not a valid description of each of the quasi contracts. The series of returns that are always a fixed average time to delivery may be negatively correlated at more lags than could be expected under a simple random walk hypothesis. Alternatively, examining such quasi contracts may not be appropriate. The weaving together of so many short series may be inducing the observed negative serial correlations in some way.

Footnotes for Chapter 3

1. Most commodity futures brokers will not allow members of the public to trade in the delivery month contracts unless a substantially larger deposit is placed with them. The risk of not being able to close out a position increases as the final day of the delivery month approaches. The situation of not being able to close out a short position is particularly risky.
2. See Mood (1940) pp367-392.
3. Anderson (1975) suggests that using the above r_k expression is valid for values of k up to about one quarter of the sample size (approximately 20 in our situation).
4. In practice, when n is large (as in this study) the normality condition can be relaxed.
5. See Anderson (1975).
6. A survey of various solutions of the Behrens - Fisher problem and a study of their power characteristrics is given in Scheffe (1970).
7. See Snedecor and Cochran.
8. In testing the stability of means we computed T_1 and T_2 for all pairs of contiguous periods for all commodity series and T_1 and T_2

were always very similar. The ratio T_1/T_2 never moved out of the region 0.98 to 1.02 whether the variances could be considered equal or not.

9. Tests using the T_3 statistic in section 3.3.2 showed that over periods of approximately 84 days, μ_t values were indeed not significantly different from zero.

CHAPTER 4

A STUDY OF THE INTER COMMODITY DISTRIBUTION OF RETURNS

In this chapter we examine the joint distribution of returns of the four soft commodities studied in Chapter 3. The reasons for carrying out such a multivariate study are outlined in section 4.1. The layout of the rest of the chapter is similar in format to that of Chapter 3: the various multivariate procedures are first described and the results and conclusions follow. Many of the tests assume multivariate normality and so the investigation into distributional form is treated first. Tests for multivariate serial correlation and parameter stability follow.

4.1 The need for a multivariate study

Tobin (1958) and others have noted that the assumptions underlying the Markowitz Portfolio Model are that either (i) investors have quadratic utility functions or (ii) returns are multivariate normal. Assumption (i) is quite restrictive and open to question and so interest is usually centred on the possible vailidity of assumption (ii).

All previous empirical studies of returns have examined the question of univariate rather than multivariate normality. Of course if returns are multivariate normal then each component will be univariate normal. If indeed a multivariate normal distribution can explain the returns, then with little or no adjustment the Markowitz Model could be applied to futures markets.

The application of Portfolio Analysis aside, a multivariate study will give information on the complete set of returns rather than the returns on a given single commodity. Previous univariate studies of a variety of price series have resulted in different conclusions for different series, yielding no obvious common conclusion. A multivariate study should produce a more powerful statement on the joint distribution of returns.

In this chapter we study a 4-dimensional set of returns. For reasons outlined in Chapter 3 we divided the data into 14 subperiods. In each subperiod we needed to select a typical contract of each commodity. For simplicity we chose the contracts and subperiods used for the separate univariate studies carried out in Chapter 3.

In Chapter 5 we look at the multivariate distribution of returns of four contracts of the same commodity and it will be useful to contrast it with the results of this study.

4.2 Multivariate procedures and notation

In the sections that follow we describe the non-standard procedures used to examine the returns for multivariate normality, multivariate serial dependence and the stability or otherwise of multivariate population parameters.

Let $\underline{x}_t$ denote the vector (of dimension px1) of returns (logged or otherwise) from day t-1 to day t such that $\underline{x}_t^T = (x_{1t}, x_{2t}, x_{3t}, x_{4t})$ where x_{1t} = returns on cocoa contract, x_{2t} = returns on coffee contract and so on, then:

$$\underline{\bar{x}} = \sum_{t=1}^{n} \underline{x}_t / n = \text{sample mean of } \underline{x}_t,$$

$$S = \sum_{i=1}^{n} \sum_{j=1}^{n} (\underline{x}_i - \underline{\bar{x}})(\underline{x}_j - \underline{\bar{x}})^T / (n - 1) ,$$

$$= \text{sample estimate of variance matrix} ,$$

$\underline{\mu}$ = population mean of returns ,

and V = population variance matrix.

S and V are $p \times p$ matrices and in this study $p = 4$.

4.3 Test of Multivariate Normailty

In this section we briefly outline some tests of multivariate normality. The results of these tests are presented in section 4.8.2.

No previous study has addressed the question of the multivariate distribution of returns on commodity futures prices. There is a vast literature on multivariate analysis and nearly all procedures assume an underlying multivariate normal (hereafter referred to as MVN) distribution. There have been few proposed multivariate distributions other than MVN. Some researchers (eg Malkovich and Afifi (1973)) in examining tests of MVN have used multivariate sets of non normal univariate distributions such as the log normal, uniform and student t distributions.

It is important to note that if the distribution is MVN then the variance matrix describes, completely, the interrelationships between the component variables. If the covariances are zero, one could use separate univariate studies of the marginal distributions to obtain complete information on the joint distribution.

There are an infinite number of ways in which a distribution can be non-normal. This may partly explain the lack of literature on the subject. Cox and Small (1978) note that "while in particular applications very specific kinds of departure from MV normality might be of concern, the departure with the most serious consequences is often the occurrence of appreciable nonlinearity of dependence." In the Cox and Small work interest was centred on this type of non normality and it was for this reason that they examined measures of linearity of regressions. In none of the 14 subperiods examined did we find any evidence of curvature in any of the 6 two-dimensional scatter plots and nowhere did we find correlations higher than 0.55. Accordingly we turned our interest to other measures of departures from MV normality.

4.3.1 Multivariate skewness and kurtosis

Departures from univariate normality are described by the skewness and kurtosis measures b_1 and b_2 already outlined in Chapter 3. Mardia (1970) developed multivariate analogues of these measures.

If $\qquad g_{ij} = (\underline{x}_i - \underline{\bar{x}})^T S^{-1} (\underline{x}_j - \underline{\bar{x}}),$

then
$$b_{1p} = \sum_{i=1}^{n} \sum_{j=1}^{n} g^3_{ij}/n^2,$$
$$b_{2p} = \sum_{i=1}^{n} g^2_{ii}/n ,$$

b_{1p} and b_{2p} are the multivariate skewness and multivariate kurtosis statistics respectively. Mardia (1970) has shown that if $\underline{x}_i$ are MVN then the following functions of b_{1p} and b_{2p} are asymptotically $\chi^2(f)$ and N(0,1) distributed respectively.

$$(1/6)\, n\, b_{1p} \sim \chi^2(f) \qquad \text{with } f = p(p+1)(p+2)/6$$

$$\frac{b2p - p(p+2)}{\sqrt{8\,p(p+2)/n}} \sim N(0,1)$$

4.3.2 Multivariate Normal Plots

Healy (1968) outlines a graphical procedure for the detection of systematic non normality and of outlying values. The Mahalanobis distance d_t of each observation from the sample mean is computed using;

$$d_t = \{ (\underline{x}_t - \underline{\bar{x}})^T S^{-1} (\underline{x}_t - \underline{\bar{x}}) \}^{1/2}$$

If the $\underline{x}_t$ are MVN the d^2_t are $\chi^2(p)$ distributed. The ordered d^2_t are plotted against the expected $\chi^2(p)$ order statistics. If the $\underline{x}_t$ are MVN then the plot should be linear. This is a direct multivariate analogue of the normal order plot. There are problems with computing the expected χ^2 order statistics and in practice Healy (1968) demonstrates that one can use the fact that the cube root of a χ^2 variate is approximately normal. Accordingly the ordered $d^{2/3}_t$ values are plotted against the appropriate expected normal order values.

4.3.3 The W - test for multivariate normality

Royston (1983) developed a very interesting and simple extension of the univariate Shapiro and Wilks test for normality to a test for MVN. We briefly outline the technique.

If $\{ \underline{x}_t \}$, t = 1, 2,...., n, is the set of MVN returns, consider the ith component $\{ x_{it} \}$, t = 1, 2,..., n ordered. Compute the W_i and associated z_i statistic to test univariate normality in the ith component, as described in section 3.2.3. Recall that if the $\{ x_{it} \}$ are univariate normal then z_i is N(0,1) distributed. Large positive values of z_i indicate non normality in the ith component. Consider the set $\{ z_i \}$, $i = 1, 2, \ldots, p$ and define θ_i as follows:

$$\theta_i = \{ \Phi^{-1}(0.5\Phi(-z_i))\}^2 \quad , \quad i = 1, \ldots, p$$

where Φ is the cummulative normal integral. Note that if $\{x_{it}\}$ are normal, each θ_i will be $\chi^2(1)$ distributed. Large values of θ_i indicate non normality in the ith component.

Consider the function, G,

$$G = \sum_{i=1}^{p} \theta_i / p ,$$

and the following two extreme situations. (i) When the components of $\underline{x}_t$ are uncorrelated clearly the W_i, z_i and θ_i will be uncorrelated and thus G will be $\chi^2(p)/p$ distributed. (ii) When the components of $\underline{x}_t$ are perfectly correlated, W_i, z_i and θ_i will be perfectly correlated and G will be $\chi^2(1)$ distributed. For intermediate parent correlations Royston suggests that a natural approximation to the distribution of G is $\chi^2(e)/e$ with e being the "equivalent degrees of freedom". Obviously, e , need not be integral and is estimated using the first two moments of G.

If c_{ij} = correlation between θ_i and θ_j, and

$$\bar{c} = \sum_{j=1}^{p} \sum_{i=1}^{p} c_{ij} \Big/ (p^2 - p) \quad = \text{average of } c_{ij}$$

then

$$e = \frac{p}{1 + (p-1)\bar{c}}$$

Thus one computes H = eG and if $\underline{x}_t$ are MVN then H should be $\chi^2(e)$ distributed.

This all assumes, of course that one knows the correlations c_{ij}.

Royston examined the values of c_{ij} under varying absolute parent correlations within $\underline{x}_t$ in MVN samples. Remarkably, the values of c_{ij} are very small for parent correlations up to 0.7 and so in the study carried out in this chapter (in which the maximum correlation observed was 0.55 and many were very near zero) one could have regarded the H statistic as $\chi^2(4)$ distributed. However in Chapter 5 we encounter distributions with much higher correlations (typically 0.95) and so for uniformity we followed Royston's suggestion and used his proposed method of estimating c_{ij} from the sample correlations, ω_{ij}, by using the function:

$$c_{ij} = (\omega_{ij})^{\lambda} \{1 - (\mu/\nu)\omega_{ij}(1 - \omega_{ij})^{\mu}\},$$

in which $\mu = 0.715$, $\lambda = 5$

and $\nu = 0.21364 + 0.015124(\log(n))^2 - 0.0018034(\log(n))^3$

Royston examined the distribution of H under the null situation of bivariate MVN with parent population up to values of 0.995 and found that the $\chi^2(e)$ approximation was very good. In fact the distribution of H was found to be slightly lighter in the upper tail than the appropriate χ^2 distribution and so tests are fractionally more conservative than one suspects.

For each of the 14 subperiods we used the algorithm provided by Royston (1982b and c) to compute W_i and z_i for each component i = 1, 2, 3 and 4. Each of the six c_{ij} values were computed for the sample correlations using the above function producing values of e. Because e is non integral we used an algorithm due to Narula and Desu (1981) to compute the probabilities associated with values of H from a $\chi^2(e)$ distribution.

Since departure from MVN may occur in lower dimensional[1] space than

$\mathbb{R}^4$ we computed H (and the probability of H under the null) for all of the possible 15 combinations of the four components of $\underline{x}_t$.

4.4 Robust estimation and the detection of outliers

What of distributions failing the MVN tests described in section 4.3? Is there a systematic departure from MVN or are any significant results due to the presence or one or more atypical values? In a given subperiod clearly one could examine the four separate normal order plots and all the 15 possible ordered Mahalanobis distance plots. A systematic departure from linearity in one or more of the 19 plots would suggest non normality. Atypical observations would be highlighted by an otherwise linear plot with one or more outlying points.

A procedure due to Campbell (1980), designed for the robust estimation of variance matrices, provides an extremely useful way of detecting outliers from an otherwise MVN distribution. The sample estimates of V and $\underline{\mu}_t$ in Campbell's procedure are very similar to the classical ones. However each observation $\underline{x}_t$, is given a weight, w_t. Observations coming from the main body of the data are given a reduced weight. The Mahalanobis distance, d_t, plays a central role in deciding which data points are far from the centre of the distribution.

If w_t = weight assigned to observation $\underline{x}_t$,

d_t = Mahalanobis distance of observation $\underline{x}_t$, from weighted mean,

$\overline{\underline{x}}_w$ = weighted sample mean,

S_w = weighted variance estimate,

then

$$\underline{\bar{x}}_w = \sum_{t=1}^{n} \underline{x}_t w_t \Big/ \sum_{t=1}^{n} w_t$$

$$S_w = \sum_{t=1}^{n} w^2_t(\underline{x}_t - \underline{\bar{x}}_w)(\underline{x}_t - \underline{\bar{x}}_w)^T \Big/ \sum_{t=1}^{n} (w^2_t - 1)$$

$$d_t = \{(\underline{x}_t - \underline{x}_w)^T S^{-1}_w (\underline{x}_t - \underline{x}_w)\}^{1/2}$$

$$w_t = \begin{cases} 1 & \text{if } d_t \leqslant d_0 \\ (d_0/d_t)\exp\{-.5(d_t - d_0)^2/d^2_2\} & \text{if } d_t > d_0 \end{cases}$$

with $\quad d_0 = \sqrt{p} + d_1\sqrt{2}$

The value of d_1 determines d_0 and thus the cut off point for the w_t's for what is considered a reasonable distance of $\underline{x}_t$ from $\underline{\bar{x}}_w$. The value of d_2 determines the rate of decrease of the w_t's, associated with outliers, towards zero. Empirical experience leads Campbell to suggest values for d_1 and d_2 of 2.0 and 1.25 respectively.

The solution for $\underline{\bar{x}}_w$, S_w and w_t is iterative. A routine that computes $\underline{x}_t$, S_w and w_t was constructed. Iterations were made conditional on individual components of $\underline{x}_w$ being to within 0.1% of previous values. Following Campbell's suggestions, obsrvations with weights less than 0.3 were designated as "outliers". Simulations using this routine in conjunction with one that generates contaminated MVN samples (see section 4.7) produced estimates of V and $\underline{\mu}$ consistently superior to the classical estimates and always identified the anomalous data. The routine proved extremely useful in the early stages of this research. It was particularly useful in the detection of errors in the original data set.

4.5 Multivariate serial correlation

If the components of the returns vector are independent, then one would think that separate tests of univariate temporal dependence could be merged together in making a joint statement about the multivariate temporal behaviour of $\underline{x}_t$. What if the components of $\underline{x}_t$ are correlated? It would be difficult to make a joint statement from the separate univariate tests. Recall that we did not discover any evidence of any consistent temporal dependence in any of the subperiods examined and we would expect that any joint study of multivariate temporal dependence would yield similarly uninteresting results.

But what if there were any temporally lagged relationships between the returns, $\underline{x}_t$ and $\underline{x}_{t+k}$ of a more complex nature? What if, for example, the cocoa and coffee returns are correlated not only contemporaneously, as they seem to be in the five year period considered, but also, say correlated significantly across different points in time? Cocoa prices going up one day (positive return) could mean that coffee prices will follow the next day. Separate univariate serial correlations analysis could still show the cocoa and coffee returns to be white noise when in fact there exists (possibly) a multivariate temporal pattern.

How does one discover if there is any multivariate temporal dependence in the series? One method would be to examine all the cross correlations of all four commodity returns at all possible lags. This would have been extremely tedious, very time consuming and the author is not sure if any clear conclusion could have been drawn from such a mass of correlation coefficients.

Chitturi (1976) and O'Brien (1980) addressed this very problem and derived measures of multivariate serial correlation (MVSC) together with associated test statistics. Both authors consider the data to be a sequence of p dimensional random variables such that:

$$\underline{x}_t = B\,\underline{x}_{t-1} + \underline{e}_t \quad , t = 2, 3, \ldots\ldots$$

in which B = a $(p \times p)$ matrix of coefficients and $\underline{e}_t$ is a sequence of mutually dependent and identicaly distributed random variables with:

$$E(\underline{e}_t) = \underline{0} \quad \text{and} \quad Var(\underline{e}_t) = V_e$$

One wishes to test the null hypothesis of no MVSC, ie $B = \{0\}$, the matrix of zeros against a general alternative hypothesis of at least one non zero element in B.

Chitturi (1976) considered the sample autocovariance matrices:

$$\Gamma_k = \sum_{t=1}^{n-k} (\underline{x}_t - \overline{\underline{X}})(\underline{x}_{t+k} - \overline{\underline{X}})^T \Big/ (n - k)$$

and showed that under the null hypothesis the sample autocovariance matrices are asymptotically uncorrelated and multivariate normal. He proposed the test statistic t_1.

$$t_1 = \sum_{k=1}^{m} (n - k)\,tr(D_k)$$

where $$D_k = \Gamma_k\,\Gamma_0^{-1}\,\Gamma_k^T\,\Gamma_0^{-1}$$

and showed that t_1 is asymptotically $\chi^2(mp^2)$ distributed.

Although Chutturi (1976) did not explicitly say so, we can assume that since the Γ_k are uncorrelated (under the null hypothesis) then one can consider testing each Γ_k using $t_1(k)$:

$$t_1(k) = (n - k)\mathrm{tr}(D_k)$$

which under the null should be $\chi^2(p^2)$ distributed, asymptotically.

O'Brien (1980) notes that the likelihood ratio statistic given by:

$$t_2(k) = -2\log\lambda = -n\log(C_k)$$

where $$C_k = \left| I_p - ((n-k)/n)^2 D_k \right|$$

is asymptotically $\chi^2(p)$ distributed.

O'Brien also outlines two measures of multivariate serial correlation. The first one, $\tilde{R}^2$, is based on Chitturi's (1976) test statistic and is given by:

$$\tilde{R}^2 = (n-k)^2 \mathrm{tr}(D_k) / pn^2$$

The second one, R^2, is based on O'Brien's test statistic and is given by:

$$R^2 = 1 - (C_k)^{1/p}$$

O'Brien demonstrates that $\tilde{R}^2$ is a function of the sum of the canonical correlation coefficients whereas, R^2 is a function of the product of the canonical correlation coefficients. We will return to these canonical correlations in Chapters 5,6 and 7. O'Brien tested both $t_1(k)$ and $t_2(k)$ using simulations on two dimensional data and reported that both statistics performed equally well with samples larger than n =

50.

Our own simulations study of these test statistics on four dimensional multivariate normal data with and without outliers showed that both $t_1(k)$ and $t_2(k)$ performed well and were insensitive even to extreme outlying values. However for larger lags the $t_2(k)$ statistic rejected the null less frequently than expected for various test sizes.

4.6 Test on Multivariate Parameters

In Chapter 3 we examined the usual sample statistics such as means and variances on each separate univariate commodity returns series. In this chapter we are studying the joint distribution of the set of four dimensional returns. Individually testing the means of the separate components will not be valid if two or more components are correlated. Joint statements and tests are required.

Amongst other things, interest is centred on the stability or otherwise of the multivariate distribution of returns. If Markowitz Portfolio Theory is to be successfully applied to the commodity futures markets it is important that one can use information from the past, such as variances and covariances, to construct portfolios that will be near optimal (efficient) in the future.

All the foregoing tests assume that the data are multivariate normal but Mardia (1979) notes that, broadly speaking, in the presence of non normality the normal theory tests on means are influenced by b_{1p} whereas tests on covariances and correlations are influenced by b_{2p}.

4.6.1 Classical single period multivariate tests on parameters

Two parametrs are of interest in each of the 14 subperiods; the vector of means, $\underline{u}$ and the correlation matrix R

(i) $H_0 : \underline{\mu} = \underline{0}$ V unknown

This is known as the Hotelling one sample T test. This would be consistent with the hypothesis of no persistent trend. Hotelling's one sample T Test is appropriate here. Under the null:

$$T_1 = \frac{(n - \rho)(\underline{\bar{x}} - \underline{0})^T S^{-1} (\underline{\bar{x}} - \underline{0})}{\rho} \sim F(\rho, n-\rho)$$

(ii) $H_0 : R \equiv I\rho$

ie the series are mutually uncorrelated. Box (1949) has shown that under the null:

$$T_2 = -[n - (2\rho + 11)/6] \log|R| \sim \chi^2(\rho(\rho-1)/2)$$

We can use T_2 as a useful measure of the comovement of all the returns.

With $\rho = 4$, we have 4 x 3 /2 = 6 pairwise correlation coefficients in the R matrix. Strictly speaking one cannot test each component of the R matrix as they are interdependent. In fact Elston (1975) provides expressions giving the correlations between correlations.

4.6.2 Stability of multivariate parameters

Do the multivariate parameters change from one time period to another? Recalling the univariate tests in Chapter 3 we note that the individual variances of returns changed frequently from period to period and so we can guess what the answer to the above question is. Nevertheless we will consider in full the stability or otherwise of the parameters by considering the following four hypotheses:

(a) H_{0a} : $\underline{\mu}_1 = \underline{\mu}_2$, $V_1 = V_2$ (test of complete homogeneity)

(b) H_{0b} : $V_1 = V_2$ (test of equal variances/covariances)

(c) H_{0c} : $\underline{\mu}_1 = \underline{\mu}_2$ (no assumption on V_1, V_2)

(d) H_{0d} : $R_1 = R_2$ (no assumption on $\underline{\mu}_1, \underline{\mu}_2, V_1$ or V_2)

where suffix i refers to period i.

The test of H_{0c} is a multivariate extension of the Behrens - Fisher problem and requires the computation of estimates of means and variance matrices via an iterative routine with similarly controversial results. We decided not to carry out test (c).

No known test exists at present for H_{0d}. Other authors have attempted to test corresponding elements of R_1 and R_2 using the Fisher Z_r transform. However as Elston (1975) points out, the elements of R_1 are not independent and so it is difficult to make a joint statement on the stability[2] of R.

And so our concern is reduced to testing H_{0a} and H_{0b}. Mardia (1979) outlines the appropriate likelihood ratio tests:

If S_i = usual unbiased estimate of V_i

S_T = sums of squares and cross products matrix for both periods

= $(n_1 - 1)S_1 + (n_2 - 1)S_2$

S_p = pooled estimate of common variance matrix

= $S_T / (n_1 + n_2 - 2)$

To test H_{0a} and H_{0b} compute

$$T_3 = n\log\left|\frac{S_T}{n}\right| - n_1\log\left|\frac{(n_1 - 1)}{n_1} S_1\right| - n_2\log\left|\frac{(n_2 - 1)}{n_2} S_2\right|$$

$$T_4 = \gamma\,\{(n_1 - 1)\log\left|S_1^{-1}S_p\right| + (n_2 - 1)\log\left|S_2^{-1}S_p\right|\}$$

with $\gamma = 1 - \left[\frac{1}{n_1 - 1} + \frac{1}{n_2 - 2} - \frac{1}{n_1 + n_2 - 2}\right]\kappa$

and $\kappa = (2p^2 + 3p - 1)/(6(p + 1)$

Under the null, T_3 and T_4 are asymptotically χ^2 distributed on $p(p + 3)/2$ and $p(p + 1)$ degrees of freedom respectively.

All the tests described in this section are likelihood ratio tests and were derived assuming a specific distributional form, ie the multivariate normal distribution, and so the comment at the beginning of section 4.6 is relevant.

4.7 Simulating multivariate normal data

In this and later chapters a number of non standard multivariate procedures were examined. Examples are the robust estimation routine due to Campbell (1980) and the testing for multivariate serial correlation due to O'Brien (1980) and Chitturi (1976). It was decided that in such a study it would be useful to be able to generate data sets whose population properties were known. The above non standard procedures could then be applied to such data sets and the sample results compared with the known population values.

Two routines were constructed to simulate multivariate normal observations, with and without outlying values, respectively. The routines are briefly described as follows:

(i) Decide on a population value for $\underline{\mu}$ and V.

(ii) Use the Cholesky decomposition to decompose V into A and A^T such that

$$A \ A^T = V$$

(iii) Generate p independent realisations from the univariate $N \sim (0,1)$ distribution and stack into a px1 vector $\underline{z}$.

(iv) Calculate the required observation $\underline{y}$ using

$$\underline{y} = \underline{\mu} + A \underline{z}$$

Repeat steps (iii) and (iv) until a sample of size n is obtained. It is easy to show that $\underline{y}$ is multivariate normal with mean, $\underline{\mu}$ and variance matix V. Proof:

$$E(\underline{y}) = E(\underline{\mu} + A\underline{z}) = E(\underline{\mu}) + A.E(\underline{z})$$

$$= \underline{\mu}$$

$$Var(\underline{y}) = Var(\underline{\mu} + A\underline{z}) = Var(A\underline{z})$$

$$= E\{(A\underline{z})(A\underline{z})^T\} = A.E(\underline{z}\underline{z}^T).A^T$$

$$= A.I.A^T = A.A^T = V$$

Each component of $\underline{y}$ is a linear combination of normal random variables and is thus normal.

The outlier routine is identical to the above routine except that an additional step is included in each simulation.

$$\underline{y} = \underline{\mu} + A\underline{z} + \underline{\delta}$$

$$\text{where} \quad \underline{\delta} = \begin{cases} \underline{0} & \text{with probability } q \\ \underline{\delta}^* & \text{with probability } 1 - q \end{cases}$$

in which the elements of $\underline{\delta}^*$ are set to five times the standard deviation of a typical component of $\underline{y}$. We can adjust the number of outlying observations by altering q.

These routines proved extremely useful in gaining experience with multivariate analysis.

4.8 Discussion of results of multivariate tests

All the multivariate analyses were carried out on the subperiods using both logged and unlogged returns. Before going into a detailed discussion of each of the sets of results we note here that it was decided for brevity to report only the analysis of the logged data. Recall from the univariate study carried out in Chapter 3 that logging returns tended, if not to normalise otherwise non-normal data, to reduce slightly any positive skew and in some instances also tended to stabilise somewhat the variances. Without going into detail yet, we found similar tendencies in the multivariate analysis. However the difference between the results of the logged and unlogged analyse was small.

4.8.1 The identification of anomalous returns

We consider first the question of outlying observations. Table 4.1 gives a list of the number of outlying observations in each subperiod that were identified by the routine devised by Campbell (1980). Examination of each of the four components of the observed outlier invariably led to the discovery that it was only one of the components that was in fact anomalous.

It is interesting to note that, apart from the first period, the outlier detection routine always converged in less than 17 iterations and typically in only 9 iterations. The first period seems to be a special case. Referring to Figs. 3.2 to 3.5 we see that 16 days before the end of period one the variance of the cocoa and coffee series increased considerably. What we are examining in period one is really two different distributions; those observations before the 86th day and those after. Considering only the first 86 returns, the routine converged in four iterations with no further outliers detected.

Table 4.1

Outlier detection routine (due to Campbell) results

Period	no. of outliers	no. of iterations
1	16	*
2	1	14
3	5	17
4	1	9
5	1	9
6	2	9
7	0	1
8	0	7
9	1	8
10	1	12
11	0	7
12	0	7
13	2	9
14	0	8
Total	14	

(*Routine would not converge until last 16 observations of period were removed)

Apart from the first period there were a total of 14 observations identified as anomalous. This means that in the rest of the 4.5 year period (a span of 1218 - 102 = 1116 days) only 14 atypical returns could be found. Thus only 14/1116 or 1.25% of the returns were considered to be out of the ordinary. Examination of the data in more detail revealed that many of the 14 outlying returns could be attributed to sudden and large changes in the prices of the cocoa and coffee futures. Consider for example the 48th day of period six when there was a simultaneous large drop in the price of the cocoa future (£465 per tonne) and the price of the coffee future (£400 per tonne).

How do these outliers affect the multivariate tests? Is it possible that the presence of a single outlying observation could seriously affect the results of, say, a test for multivariate serial correlation? To answer this question we carried out each test on the complete subperiod and with the outlier(s) removed.

4.8.2 Tests of multivariate normality

We consider first the multivariate kurtosis and skewness measures reported in Table 4.2. There are many instances of extremely large sample values. Eight of the skewness statistics are significant at the 5% level and 5 are signifiant at the 0.1% level. Of the 14 kurtosis statistics, 7 are significant at the 5% level and 5 are significant at the 0.1% level. Not surprisingly those periods with a high skewness statistic also have a high kurtosis statistic. Thus over half of the 14 subperiods yielded significant results and would lead one to the rejection of a hypothesis of multivariate normality.

If, however, the outliers are removed all the statistics are very much reduced in size and only 4 periods would be considered as non multivariate normal. It is interesting to note how much each result is altered by the removal of a few outliers. Consider for example period 9 in which there are 82 returns. Removing one of these returns reduces the skewness measure from 66.79 to 18.67 and the kurtosis measure from 5.53 to 0.98.

We now turn our attention to Tables 4.3a and 4.3b. These Tables have been drawn up using the P(W) values from the Royston's W - test for multivariate normality on all the possible (15) combinations of the four

Table 4.2

Multivariate skewness and kurtosis measures on logged returns

Period	Skewness measure: Complete data	Skewness measure: Outliers removed	Kurtosis measure: Complete data	Kurtosis measure: Outliers removed
1	353.46 c	30.38	29.07 c	2.70 c
2	33.48 a	33.70 a	2.85 b	2.77 c
3	54.90 c	18.85	5.26 c	1.55
4	34.48 a	15.23	2.22 a	1.16
5	23.05	27.17	0.16	-0.26
6	95.18 c	39.50 b	8.42 c	2.42 a
7	22.15	22.15	-1.04	-1.04
8	33.07 a	33.07 a	1.54	1.54
9	66.79 c	18.67	5.53 c	0.98
10	17.72	18.25	1.93	0.78
11	24.37	24.37	0.29	0.29
12	16.56	16.56	-0.49	-0.49
13	142.64 c	19.86	10.12 c	1.61
14	23.80	23.80	0.63	0.63

Skewness measure ~ $\chi^2(20)$ under multivariate normality
Kurtosis measure ~ N(0,1) under multivariate normality

If entry = a then associated value is significant at the 5% level.
If entry = b then associated value is significant at the 1% level.
If entry = c then associated value is significant at the 0.1% level.

Table 4.3a

W - tests of Multivariate Normality on complete data set

Component 1 = Cocoa returns
Component 2 = Coffee returns
Component 3 = Sugar returns
Component 4 = Rubber returns

Significant P(W) values

Components:			All 4	In 3's				In 2's						Individually			
Period	Mul. skew	Mul. kut.	1 2 3 4	1 2 3	2 3 4	1 3 4	1 2 4	1 2	1 3	1 4	2 3	2 4	3 4	1	2	3	4
1	c	c	b	c	-	c	c	c	c	c	-	-	-	c	c	-	-
2	a	b	a	a	a	a	-	-	b	-	b	-	b	-	-	b	-
3	c	c	c	c	c	b	b	c	b	-	c	b	b	-	c	b	-
4	a	a	c	c	c	-	c	c	-	-	c	c	-	-	c	-	-
5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6	c	c	c	c	-	c	c	c	c	c	-	-	-	c	-	-	-
7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8	a	-	-	a	-	a	-	a	a	a	-	-	-	a	-	-	-
9	c	c	a	b	-	b	b	b	b	b	-	-	-	c	-	-	-
10	-	-	-	a	-	-	a	a	a	a	-	-	-	b	-	-	-
11	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
12	-	-	-	-	-	-	-	-	-		-	-	a	-	-	-	a
13	c	c	c	c	c	-	c	c	-	-	c	c	-	-	c	-	-
14	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 4.3b

W - tests of Multivariate Normality on data set with outliers removed

Components:			All 4	In 3's				In 2's						Individually			
	Significant P(W) values																
Period	Mul. skew	Mul. kut.	1 2 3 4	1 2 3	2 3 4	1 3 4	1 2 4	1 2	1 3	1 4	2 3	2 4	3 4	1	2	3	4
1	-	b	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	a	b	-	-	-	-	-	-	a	-	a	-	a	-	-	a	-
3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6	b	a	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8	a	-	-	a	-	a	-	a	a	a	-	-	-	a	-	-	-
9	-	-	-	-	-	-	a	a	a	a	-	-	-	a	-	-	-
10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
11	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
12	-	-	-	-	-	-	-	-	-		-	-	a	-	-	-	a
13	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
14	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

components of the returns vector. It was decided that rather than record the numerical values of the P(W) statistics, a symbol denoting the significance, if any, of the associated test be reported.

Column three of Table 4.3a gives the P(W) value associated with the test of multivariate normality on the complete set of returns. Note that the periods resulting in significant P(W) values almost always also yield significant skewness and kurtosis measures. So we see that the W - test for multivariate normality gives results identical to those obtained using Mardia's multivariate skewness and kurtosis measures.

Columns 4 to 17 of Table 4.3a however are also very interesting. Columns 14 to 17 contain information on the four separate W - tests of univariate normality, columns 8 to 13 contain information on all combinations of two components and columns 4 to 7 contain information on all combinations of three components. It is interesting to note that for any given subperiod, if a particular univariate component of the returns vector is reported as non normal, all combinations of the four returns containing that component will also be reported as non multivariate normal. Consider for example period 13, which results in eight significant P values. Examination of Table 4.3a reveals that it is the second component, coffee, that is causing all the significant results. All combinations containing the coffee component would be considered as non multivariate normal.

This effect is not really surprising. Table 4.4 shows the six inter-commodity correlation coefficients for each subperiod. The maximum correlation coefficient observed was 0.495 in period 3. As noted in section 4.3.3, with correlations less than 0.7 the multivariate W - test is equivalent to simply "adding" the results of the separate univariate W - tests.

At this point it may be useful to consider the values of, e, the effective degrees of freedom used in computing the P(W) statistics in Tables 4.3a and 4.3b. Recall that for highly correlated data e = 1 and for uncorrelated data e = dimension of vector under consideration. The maximum and minimum values of e found in the complete study of the returns were:

(i) With complete four dimensional set:

$e_{max} = 4.000$, $e_{min} = 3.965$

(ii) With all combinations of three dimensional set:

$e_{max} = 3.000$, $e_{min} = 2.976$

(iii) With all combinations of two dimensional set:

$e_{max} = 2.001$, $e_{min} = 1.991$

These values of e reinforce the above remarks.

For tests of multivariate normality, the individual components can therefore be considered as separate, uncorrelated series.

Before moving on to the effect of outliers it is interesting to note the dilution effect in these tests caused by increasing the dimensionality of the vector considered. For example in period ten, in which the P(W) value for the cocoa series is 0.007 (recorded as b in Table 4.3a), the P(W) value for the cocoa and coffee series is 0.012. For the cocoa, coffee and sugar series the value is 0.031 and for all four series the value is 0.061. In this period only the cocoa series is significantly non normal. Notice how in each case the P(W) value is small, resulting (except when considering all four returns) in the rejection of the null hypothesis; but that as more components are added to the cocoa series, the effect of the severe non normality, is gradually diluted. This effect was noticed in all of the subperiods, but although the P(W) value increased as the number of components was increased one generally found that non normality in a single component led to the rejection of the null hypothesis in all combinations containing this component.

We now turn our attention to Table 4.3b. This table gives the results of the tests of multivariate normality on the data after the outliers have been removed. The difference between Tables 4.3a and 4.3b is striking. The number of periods in which we reject the null hypothesis at the 1% level has been reduced to one. The skewness and kurtosis measures result in four periods being significant at the 5% level. The W - test on all the full sets of returns result in no periods in which we reject the null. Examination of all the other combinations of returns shows however that this may be simply a result of the dilution effect. The significant P(W) value for cocoa in period 8 shows up in all of the relevant combinations of two components, two of the relevant combinations of three, but not in the combination of all four.

The only unusual results are in periods 1 and 6 in which we get significant kurtosis and skewness statistics but no significant results in any of the combinations of individual components. These significant results cannot be explained by non normality in any of the individual components and examination of all the 6, two dimensional, scatter diagrams of the returns, does not reveal any obvious anomolies.

In summary then, regarding the distribution of the returns, the multivariate skewness and kurtosis measures in general led to similar conclusions to Royston's W - tests. We also examined many of the 15 x 14 = 210 normal order and multivariate normal order plots and, after the removal of outliers, there was no real evidence of any consistent departures from linearity. It is very interesting that (excluding the first period) the removal of 1.25% of the observations changes the picture markedly, leading to a general acceptance of the hypothesis that the distribution could quite easily be considered as multivariate normal.

Finally, we note that the study using unlogged returns resulted in more subperiods being regarded as non normal. Of the 14 periods, 10 were rejected at the 5% level using the W - test. This compares with 7 of the logged returns series. In general the results were similar and removing outliers greatly reduced the number of significant statistics.

<u>4.8.3 Tests for a persistent multivariate trend</u>

In Table 4.4 are the results of testing the hypothesis of $\underline{\mu} = \underline{0}$ against the general alternative of $\underline{\mu} = \underline{0}$. Considering initially the columns relating to the complete data set we note that only two periods (four and five) produce results that can be considered significant at the 5% level. However, the tests on a number of the other periods produce F statistics that are very near the 5% significance cut off level. For example in period 12 the F statistic is 2.47. This corresponds to a P value of 0.06.

<u>Table 4.4</u>

<u>Testing the hypothesis of no multivariate trend</u>

	Test statistic*	
Period	Complete data	Without outliers
1	1.22	1.99
2	0.96	0.94
3	2.19	2.35
4	2.55 a	2.74 a
5	5.49 b	5.30 b
6	0.74	1.47
7	0.85	0.85
8	2.35	2.35
9	1.02	0.76
10	0.73	0.61
11	0.22	0.22
12	2.47	2.47
13	0.95	0.81
14	1.97	1.97

* Test statistic is ~ F(4,n) under the null hypothesis

These results are similar to those in Tables 3.1 to 3.4 in Chapter 3. The null hypothesis that the vector of returns is identically zero is rejected when one of the relevant tests on one of the components is rejected. It is interesting to note again the diluting effect introduced by considering higher dimensional vectors. In period 12 for example the P value associated with the test on the coffee series is 0.01 but when compounded into a four dimensional vector the P value of the multivariate test becomes 0.06.

There is very little effect from removing outliers. It is known[3] that Hotelling's T Test is overall robust to non normality and here we see that the test seems to be insensitive also to the presence of some extreme outlying observations. Analysing the unlogged returns produced virtually identical results. We conclude that, in most of the subperiods studied the vector of returns could be considered to have a mean of zero.

4.8.4 Inter - commodity correlation coefficients

The complete set of six sample correlation coefficients between the four series of returns for each subperiod is given in Table 4.5. The symbols a, b and c indicate, as usual, if each coefficient can be considered to be significantly different from zero at the appropriate level of significance. Note that these results are obtained by separately testing each coefficient and strictly speaking, cannot be used to make a joint statement on the inter commodity correlation matrix. Table 4.5 also gives the results of the joint test of all coefficients being zero, outlined in section 4.6.1.

Referring to Table 4.5 we see that 26 of the possible 6 x 14 = 84 correlation coefficients are significant at the 5% level. The strongest

Table 4.5

Inter-commodity correlation coefficients

Period	Cocoa / Coffee	Cocoa / Sugar	Cocoa / Rubber	Coffee / Sugar	Coffee / Rubber	Sugar / Rubber	Test* of H_0: $R \equiv I$ Complete Data	Without Outliers
	Pairwise correlation coefficients							
1	0.231a	0.390c	0.163	-0.009	0.084	0.092	24.70c	28.22c
2	0.080	0.027b	0.054	0.163	0.122	0.102	12.76a	11.65
3	0.480c	0.292b	0.403c	0.277a	0.219a	0.253a	46.85c	34.91c
4	0.400c	0.228a	0.388b	0.269a	0.157	0.159	32.56c	26.76c
5	0.262a	0.247a	0.141	0.172	0.271a	0.147	19.68b	27.74b
6	0.495c	-0.185	-0.062	-0.180	-0.112	-0.069	26.91c	13.94a
7	0.155	0.046	0.135	0.072	0.037	0.279a	10.46	10.46
8	0.106	0.144	-0.014	-0.047	-0.056	0.023	3.28	3.28
9	0.172	0.105	0.139	-0.037	-0.113	0.175	8.56	14.80a
10	0.016	0.027	-0.045	0.266a	0.149	0.170	9.77	10.43
11	0.048	0.241a	-0.007	0.152	0.168	-0.012	9.14	9.14
12	0.115	0.213	0.091	0.251a	-0.071	0.261a	16.72a	16.72a
13	0.175	0.379b	0.193	0.273a	0.103	0.120	23.47b	32.99c
14	0.174	0.120	0.162	0.180	0.213a	0.235a	14.28a	14.28a

* Test statistic is $\sim \chi^2(6)$ under null hypothesis

correlations are positive, involve the cocoa and coffee series, but are at the best, weak. There are no significant negative correlation coefficients. One or more significant pairwise correlation coefficient are almost always reflected by a significant result in the joint test.

The correlations, although small, seem quite stable. From period three to period six the cocoa/coffee returns were consistently weakly positively correlated. During this period (early 1976 to mid 1977) the prices of cocoa and coffee futures were rising rapidly.

We have not reported the detailed correlation coefficients for the data with the outliers removed but just make a few remarks on the joint test results given in Table 4.5. Removal of the outliers reduced the magnitude of four of the joint test statistics. Examination of the relevant pairwise correlation coefficients revealed that, not surprisingly, four of the correlations were being exaggerated by the presence of a single outlier and five were being under stated.

We conclude by saying that the set of four returns were weakly jointly positively correlated for the first two years and the last year. The implications this has for applying Portfolio Theory to the futures markets will be discussed in section 4.9.

4.8.5. Stability of multivariate parameters

Table 4.6 gives the results of the tests of the stability of the multivariate parameters as outlined in section 4.6.2.

Considering initially the tests of the equality of the variance matrices, we note that all results are significant at the 0.1% level. This result is not surprising. In Chapter 3, in every period-to-period comparison of the univariate series, we found evidence that at least one of the variances had changed. The results in Table 4.6 and the pictorial

Table 4.6

Stability of multivariate parameters

Period Comparison	Complete homogeneity[1] Complete data	Without outliers	Equality of variance[2] Complete data	Without outliers
1 - 2	140.79	22.47*	145.41	30.77
2 - 3	81.75	66.78	87.63	71.41
3 - 4	31.28	42.74	36.74	46.36
4 - 5	42.62	38.20	47.46	41.46
5 - 6	95.02	90.28	99.97	93.67
6 - 7	45.52	28.42	49.12	32.43
7 - 8	70.03	70.03	74.84	74.84
8 - 9	39.92	33.58	47.23	39.80
9 - 10	46.73	53.11	47.77	54.19
10 - 11	35.03	31.01	36.99	32.97
11 - 12	36.11	36.11	42.14	42.14
12 - 13	41.42	16.09*	42.83	21.23
13 - 14	117.01	83.86	123.61	88.71

1: Complete homogeneity test statistic $\sim \chi^2(14)$ under the null hypothesis
2: Equality of variancetest statistic $\sim \chi^2(10)$ under the null hypothesis

* These are the only values not significant at the 5% level

evidence in Figs. 3.2 to 3.5 simply reinforce these findings. Removal of the outliers reduces somewhat, most of the test statistics, but makes virtually no difference to our conclusions.

The results of the complete homogeneity tests were also predictable. The P values associated with each result have been computed but for brevity are not reported here. In every period to period comparison, the P value of the complete homogeneity test was small (13 were less than 0.001) but not as small as the equality of variance test. This is obviously due to the fact that we are joint hypothesis testing. The results in section 4.8.3 show that the trend values (means) are mostly not significantly different from zero and so it is unlikely that the means $\underline{\mu}_i$ and $\underline{\mu}_{i+1}$ are different in each period to period comparison. What we are observing in the complete homogeneity test are the results of a diluted equality of variance test. Again we note that all the results are significant.

4.8.6 Multivariate serial correlation

The results of the tests for multivariate serial correlation appear in Table 4.7a and 4.7b. Again, for brevity, we simply report the significance or otherwise of each coefficient. Also for brevity we report the results only on O'Brien's test for lags k = 1 to k = 20.

In period one there are seven significant coefficients at the 5% level. In the remaining 13 periods, of the total of 13 x 20 = 260 coefficients, 11 (4.2% of 260) are significant at the 5% level. There is no obvious consistent pattern in the significant coefficients.

Considering Table 4.7b we see that removing the outliers has a most dramatic effect in the number of significant coefficients in period one

Table 4.7a

Significant multivariate serial correlation coefficients on complete data set

Lags = Period	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	a		b	a	c	c		a	b											
2															a					
3									a											
4																				
5																				
6																				
7			a				a													
8			c																	
9					a															
10																				
11																				
12					a															
13				a	a															
14	a			a																

Table 4.7b

Significant multivariate serial correlation coefficients on data set with outliers removed

Lags = Period	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1							a													
2																				
3		b																		
4																				
5																				
6																				
7			a				a													
8			c																	
9					b															
10																				
11																				
12					a															
13				a		a														
14	a			a																

also. The number is reduced from seven to one. The sudden increase in the variance of the cocoa and coffee series towards the end of period one gives rise to the anomalous results in Table 4.3a. Apart from period one, the outliers seem to have very little effect on the test statistic. A total of 11 significant coefficients is very near what we would expect under the null hypothesis (5% of 280 = 14).

Using Chitturi's test statistic yields a similar pattern in the coefficients. The P values associated with the Chitturi's statistics are in general slightly smaller resulting in 16 significant coefficients at the 5% level when testing the data with outliers removed. Analysis on the unlogged returns gave identical results.

We conclude then, that there is no evidence of any type of multivariate temporal dependence. It seems unlikely, then that there are any temporally lagged relationships as described in section 4.5.

4.9 Summary of the inter - commodity distribution of returns

The joint variance of the set of returns varied considerably over the five year period under consideration. This observation alone is justification for the experimental design of this study in which the data were cut into smaller subperiods. Even within these small subperiods we noticed that the variances occasionally changed. It must be stressed that this changing variance witnessed in this study and recorded by many other authors, means that one cannot carry out the usual statistical tests for normality, serial dependence, etc on daily returns in periods longer than 80 days or so.

We conclude then that the returns could be described as being generated from a multivariate normal distribution with the occasional contaminating extreme realization in one or more of the components. The character of these rare extreme returns are varied and not predictable. It does not seem likely that a set of longer-tailed univariate distributions, such as the t or the stable distribution, would explain the returns any better.

There is no evidence that the set of returns has an average that is anything different from zero.

What of the question of temporal dependence? There does not seem to be any. It has occurred to the author to propose a multivariate Taylor type model in which the trend vector ($\underline{\mu}$) would be modelled by a multivariate stochastic process. The study of such a model would require the analysis of very long run series with the associated problem of varying variances. One could of course simply extend Taylor's rescaling technique by using a multivariate exponentially smoothed estimate of the variance matrix. We leave this suggestion for another study.

4.10 Implications for a portfolio analysis model

The successful application of a Markowitz type portfolio model to any investment medium requires, as inputs, estimates (statistical or otherwise) of the expected returns and the riskiness of each asset under consideration. Also required is an estimate of the comovement of asset returns. If one is to use the classical statistical estimates of expected returns (the mean) and risk (the variance) then one would hope that these estimates are meaningful and sufficiently accurate for practical use. If the returns are multivariate normal then the classical estimates of expected return and risk say everything there is to know about the investment set. We have found in our study that the returns are 'almost' multivariate normal.

How meaningful would it be then, to apply the portfolio model to the set of returns? Firstly we found that in most of the periods the average returns could not be considered statistically different from zero. This question of the non significance of average returns must not be over stated. Many other empirical works[4] applying portfolio theory to various investment mediums report returns that are not significantly different from zero. However, investing in the proposed portfolio produced returns that were positive, if not significant, and positive returns with low risks are what we are seeking.

Secondly with respect to the variances, i.e. the risk estimates, there are two points of interest: (i) are the variances the correct instruments to use as risk estimates and (ii) are they accurate as forecasters of future variances?

The variance estimates do not take into account the occasional sudden movements in the prices witnessed in this study. The sudden movements are unpredictable and thus the perceived risk must be higher than that suggested by the classical variance estimates. It is possible that one could bias upwards the classical variance estimates for each component. Such a procedure would require some subjective input as to the likelihood of sudden movements in prices.

Are the classical variance estimates accurate forecasters of future variances? The results outlined in section 4.8.5 demonstrate conclusively that the answer to this question is no. Even ignoring the outlying observations the classical variance estimates do not remain stable.

Summing up, then, we note that a serious draw back to the viability of applying Markowitz Portfolio Theory to the futures markets studied here is the question of the appropriateness of using the usual variance estimates for risk inputs and the observed instability of these estimates from period to period.

However the inter- commodity correlations are small and positive and thus constructing portfolios of futures contracts will afford some risk reduction. Using a technique like the Markowitz procedure will almost certainly result in some diversification but the resulting portfolios are unlikely to be on the efficient frontier. We leave the study of the effectiveness of crudely applying the Markowitz procedure to this data set to another study but report the results of investing in a naive portflio of futures contracts in Chapter 8.

In Chapter 5 we examine the multivariate distribution of returns on all the contracts of a given commodity.

Footnotes for Chapter 4

1. We suspect in this study that a departure from MVN is probably due to departure from univariate normality in one or more or more of the individual components.
2. Simulations carried out by the author have shown that unless the underlying population correlations are higher than 0.7 the individual tests on elements of the R matrices yield meaningful results.
3. See Mardia (1979) p 149.
4. See Watson and Dickinson (1981).

CHAPTER 5

A STUDY OF THE INTRA - CONTRACT DISTRIBUTION OF COMMODITY FUTURES RETURNS

In Chapter 4 we studied a multivariate set of returns in which each component of the returns vector was obtained from a typical contract in one of the four commodity futures markets. In this chapter we examine another type of multivariate distribution, one in which all the components are returns on contracts in the same commodity futures market. In the rubber futures market, for example, there are usually eight different futures contracts which can be traded on any one day. Each contract has a different delivery date. We can construct a vector of returns in which the first component will be the returns on the nearest contract - the one that will reach maturity first, the second component will be the returns on the contract that will reach maturity next and so on. Apart from Dusak (1973), who reports that the returns on different contracts of a given commodity are highly correlated, no other empirical study on such a multivariate distribution appear to have been published.

Such a study may throw light on a number of issues such as:

(i) How the variability of returns on the various contracts are related. There are possibly complex associations between the variability of returns, the trading volume and price expectations for each contract. One would expect the variability of returns in the far contracts to be lower than the variability of returns in the near contracts.

(ii) Whether the inter contract correlations vary over time and the degree of correlation differs across futures markets.

(iii) Whether one can use various futures contracts of the same commodity to construct Markowitz type efficient frontiers.

To answer these questions one must look at the multivariate distribution of returns and the stability or otherwise of the parameters[1].

5.1 Design of the empirical study

There are 7 different contracts one can trade in each day in the cocoa, coffee and sugar futures markets and 8 different contracts in the rubber futures market. The delivery dates of the contracts and the dates when new contracts become available for trading vary from futures market to futures market. Also the time periods in which one could examine the same 7 or 8 contracts is quite small (approximately 42 days in the case of the coffee contracts). For these reasons, it was decided initially to examine only a 4 dimensional set of returns from each market. For consistency with the studies in Chapter 3 and 4 we decided to study the returns over the same 14 subperiods. The 4 contracts from each market were picked from the possible 7 or 8 contracts available each day in the following way.

The first component of the 4 dimensional returns vector was derived from the prices of the contract that was nearest the delivery date. The second component was derived from the prices of the next nearest delivery contract and so on. For reasons mentioned in Chapter 3 it was decided initially not to use prices of futures contracts very near the maturation date and so the first component of the last observation in each subperiod was derived from prices at least one month away from maturity. Details of

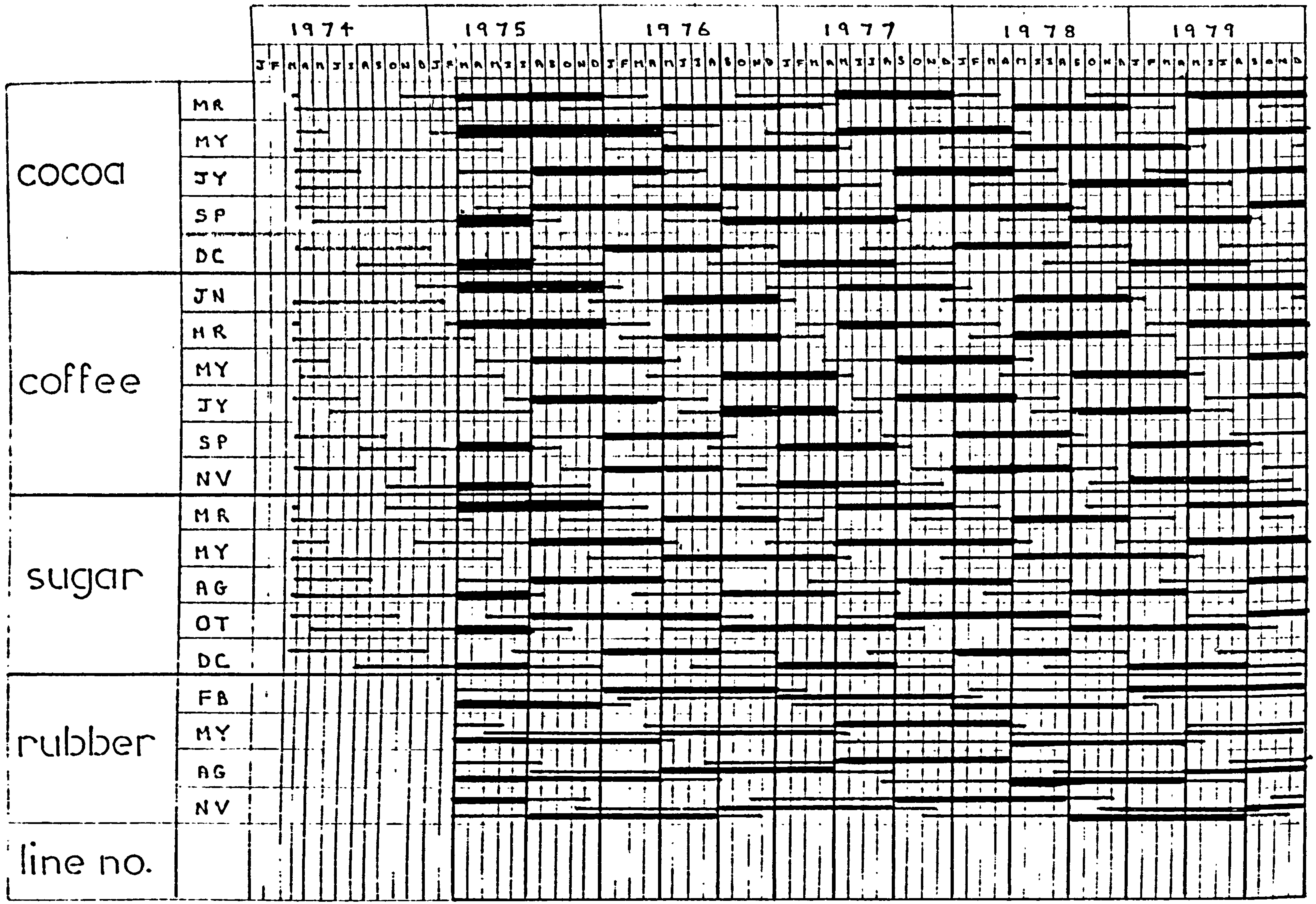

Fig 5.1 Pictorial representation of the sections of contracts used in multivariate study in Chapter 5

the contracts chosen with dates are given in appendix D. A pictorial representation of the portion of each of the contracts used in the study is given in Fig. 5.1.

5.2 Variability of returns on different contracts

In this section we examine the distribution of the variability of returns on futures contracts with different times to maturity. Recall, however from section 3.3.2 that the variances of returns changed considerably from period to period and listing the variances or standard deviations of each of the 4 contract returns for each of the 14 subperiods would not be very useful.

We consider, instead, the variability of returns on each contract relative to the variability of returns of the first contract - the contract with the nearest delivery date. Table 5.1a gives the ratio[2] of the standard deviation of returns of components 2,3 and 4 to the standard deviation of returns on component 1.

The entry of 0.835 in the top left hand corner of Table 5.1a indicates that, in the first period, the second component of the cocoa series had a standard deviation of 83.5% of the standard deviation of the first component. Note that the majority of the entries in Table 5.1a are less than unity reinforcing the prior belief that futures prices of distant contracts are less volatile than the futures prices of near contracts.

From Table 5.1a we can see many instances in which variances decrease monotonically with time to maturity. This can be seen much more clearly by referring to Table 5.1b which has been constructed from Table 5.1a as follows: a zero is placed in a column if a contract has a lower variance than the adjacent (leftward) contract, otherwise place a one[3].

Table 5.1a

Ratios of standard deviations of each contract to nearest contract
(logged returns and complete data)

	Cocoa			Coffee		
Period	2 : 1	3 : 1	4 : 1	2 : 1	3 : 1	4 : 1
1	0.835	0.744	0.715	0.833	0.827	0.828
2	0.964	0.899	0.909	0.963	0.989	0.977
3	1.055	1.066	0.936	0.957	0.950	0.961
4	0.922	0.877	0.882	0.981	0.971	0.979
5	1.018	1.019	1.032	1.029	1.055	1.091
6	1.042	1.069	1.045	1.007	1.038	1.040
7	0.965	0.957	0.956	0.968	0.951	1.002
8	0.891	1.090	0.859	1.007	1.053	1.023
9	1.044	0.941	0.893	1.873	1.127	1.136
10	0.914	0.770	0.762	1.012	1.036	1.001
11	0.971	0.929	0.820	1.059	1.061	1.021
12	0.896	0.868	0.747	1.115	1.185	1.239
13	0.953	0.901	0.886	1.046	1.071	1.238
14	0.972	0.938	0.909	0.944	0.980	1.000

	Sugar			Rubber		
Period	2 : 1	3 : 1	4 : 1	2 : 1	3 : 1	4 : 1
1	0.940	0.913	0.871	0.897	0.859	0.835
2	0.942	0.898	0.886	0.911	0.892	0.822
3	0.954	0.925	0.892	0.850	0.774	0.741
4	0.990	0.908	0.900	0.965	0.927	0.880
5	0.920	0.863	0.820	0.900	0.900	0.855
6	0.955	0.879	0.837	0.892	0.735	0.685
7	0.880	0.849	0.829	0.935	0.925	0.894
8	0.949	0.870	0.814	0.897	0.895	0.901
9	0.955	0.928	0.857	0.823	0.759	0.740
10	0.990	0.975	0.942	0.973	0.919	0.871
11	0.967	0.967	0.907	1.007	0.977	0.967
12	0.833	0.774	0.716	0.907	0.872	0.844
13	0.888	0.823	0.787	0.964	0.899	0.866
14	0.956	0.883	0.842	0.877	0.839	0.844

Table 5.1b

Standard deviation profile across contracts with increasing maturity

	Cocoa			Coffee			Sugar			Rubber		
Period	1-2	2-3	3-4	1-2	2-3	3-4	1-2	2-3	3-4	1-2	2-3	3-4
1	0	0	0	0	0	1	0	0	0	0	0	0
2	0	0	1	0	1	0	0	0	0	0	0	0
3	1	0	0	0	0	1	0	0	0	0	0	0
4	0	0	1	0	0	1	0	0	0	0	0	0
5	1	1	1	1	1	1	0	0	0	0	0	0
6	1	1	0	1	1	1	0	0	0	0	0	0
7	0	0	0	0	0	1	0	0	0	0	0	0
8	0	1	0	1	1	0	0	0	0	0	0	0
9	1	0	0	1	1	1	0	0	0	0	0	0
10	0	0	0	1	1	0	0	0	0	0	0	0
11	0	0	0	1	1	0	0	0	0	1	0	0
12	0	0	0	1	1	1	0	0	0	0	0	0
13	0	0	0	1	1	1	0	0	0	0	0	0
14	0	0	0	0	1	1	0	0	0	0	0	1

0 indicates expected decrease in standard deviation
1 indicates standard deviation profile is opposite to that expected

Table 5.2

Outlier detection routine results

	Cocoa		Coffee		Sugar		Rubber	
Period	no. of outs.	no.of its.	no. of outs.	no. of its.	no. of outs.	no. of its.	no. of outs.	no. of its.
1	3	12	11	10	4	12	6	23
2	1	7	2	9	10	23	2	10
3	13	17	8	10	3	14	14	10
4	2	7	1	9	3	8	1	8
5	2	6	2	8	4	15	4	11
6	1	5	3	9	3	15	4	8
7	0	8	0	9	0	4	2	14
8	4	10	3	13	0	7	0	8
9	6	51	2	8	3	12	5	15
10	1	8	0	9	0	6	1	9
11	0	1	5	24	2	10	4	16
12	3	9	1	8	4	10	1	15
13	0	7	5	11	2	7	0	7
14	0	7	1	9	0	4	1	12

We find that the spread of standard deviations of the sugar contracts is exactly what was expected (there are 14 : 0-0-0 combinations). The rubber series produce almost identical results (12 : 0-0-0 combinations), with only two 1's in the entire set of comparisons. The cocoa series yielded 7 : 0-0-0 combinations 1 : 1-1-1 and the rest were varied. The coffee series has 10 instances in which the second component is less variable than the third and 5 instances in which the distribution of standard deviations is exactly opposite to what one would expect. These results may be due to the anomalies affecting the coffee market in the period examined (1975).

5.3 Multivariate distribution of returns

The sets of intra - contract commodity returns were subjected to the multivariate tests and procedures already applied to the inter - contract commodity returns as outlined in Chapter 4. Testing for no persistent multivariate trend and homogeneity of population parameters produced results that were entirely predictable from the results of the univariate study outlined in Chapter 3. As an example, referring to Table 3.1, period 4, we see that the test on the average univariate return ($H_0 : \mu = 0$) produced a significant result of $T_3 = 2.90$ indicating a significant upward trend in the price of cocoa. This upward trend was of course present in all futures contracts of cocoa in this period and the test on the multivariate trend ($H_0 : \underline{\mu} = \underline{0}$) also produced a significant result. Similarly, the results of testing the stability or otherwise of multivariate variance matrices agreed perfectly with the results of testing the stability of univariate variances.

Because of this duplication it was decided not to report in detail the results of the multivariate tests mentioned above.

5.3.1 Identification of outlying returns

The set of returns was examined using the outlier detection routine devised by Campbell (1980) . The number of outliers detected and the number of iterations required for convergence appear in Tables 5.2.

The numbers of outliers detected in each of the series were: 36, 44, 38 and 45 respectively, which is higher than the number reported in Chapter 4. Also the number of iterations required for convergence (eg 51 in period 9 of the cocoa series) in this study was much higher than that reported in Chapter 4. There appear, then , to be more anomalies when one examines the returns on 4 contracts from the same futures markets than one observes when examining the returns on 4 contracts with each contract being from a different futures market.

It is not possible to classify each and every one of the 163 outliers detected here but it was possible to gain some insight into the general nature of the majority. In each period and with each set of returns a detailed study of the univariate plots in Figs. 3.2 to 3.5, together with all the six 2 dimensional scatter plots and the listings of the Mahalanobis distances led to the classification of each outlier into one of the following three broad categories:

(I) There are a number of instances in which the variance of the returns changes abruptly within a period. As an example, the variance of the rubber series increased towards the end of period 3 and the outlier routine identifies the last 11 observations as being "outliers". If we discount this type of outlier from consideration, then the number of atypical observations is reduced to 26,26,28 and 30 respectively, a total of 110. So 2% of the remaining returns can be considered atypical.

(II) It is clear that a number of the outliers are the result of a simultaneous, sudden change in the level of all of the prices of the futures contracts. Some of these large negative and positive spikes were also observed in section 4.8.1. A total of 31 of the outliers can be attributed to these types of unusual simultaneous price movements.

(III) A close scrutiny of the individual components of the outlying observations that could not be classified as being due to changes in variance or sudden large price moves revealed that many of these remaining outliers were present in contemporaneous pairs. Occasionally one of the four prices used to obtain the returns gets "out of step" with the overall general price movement. Since the set of four returns are highly correlated a small deviation of one return in the wrong direction will result in that observation having a large Mahalanobis distance from the mean of the bulk of the returns. The fact that the outliers tend to be present in pairs indicates that the price that moves out of line one day moves back into line the very next day.

A breakdown of the number and nature of the outliers observed in each of the series is given in Table 5.3. It must be stressed that this classification is not exhaustive and that some outliers could be classified into more than one group. Also throughout we define an outlier as being an observation whose "weight" in the outlier routine is less than 0.30. If we chose a different cut off point the number of outliers would of course be different.

To summarize, approximately 3% of the returns are classified as outliers. Approximately half of these outliers can be attributed to common movements in all of the contract prices and half to an occasional irregular shift of one or more of the prices in relation to the main pattern of price movements. Throughout the rest of this chapter we report results of multivariate analysis of the data with and without the detected outliers.

Table 5.3

Breakdown of outlier type

Type of outlier	Cocoa series	Coffee series	Sugar series	Rubber series	Totals
Change in variance	10	18	10	15	53
Large price price swings	13	6	9	3	31
Price moving out of line	13	20	19	27	79
Total	36	44	38	45	163

5.3.2 Correlations between returns

Dusak (1973) reports that the returns on different contracts from the same commodity futures market are highly correlated. In this section we examine the degree of this inter - contract correlation in more detail.

With 4 contracts there are 6 possible pairwise correlations to consider. These have been computed for all the 14 subperiods. All correlations are very high with most values being near 0.95. The minimum and maximum values observed were 0.8044 and 0.9944 respectively. Rather than report all the 2 x 4 x 6 x 14 = 732 correlation coefficients, we decided to look for a measure of the overall collinearity of the returns.

Recall from section 4.6.1 in which we test the null hypothesis of

$$H_0 : R = I$$

were R is the matrix of correlation coefficients. The test statistic

T_2 is given by:

$$T_2 = -[\, n - (\, 2p + 11\,)/6\,]\, \log|R|.$$

Under the null hypothesis T_2 is $\chi^2(6)$ distributed. One can see that the T_2 statistic could be a useful measure of the comovement of all of the returns. The higher the inter - contract correlations the closer $|R|$ will be to zero. So large correlations will be reflected by large values of T_2.

A second possible measure of the degree of collinearity could be the effective degrees of freedom parameter, e, used in the computation of the W test for multivariate normality reported in section 4.3.2. Small values of e (near unity) indicate high average inter contract correlations. The values of T_2 and e for each period appear in Tables 5.4a and 5.4b.

Tables 5.4a and 5.4b are essentially reporting similar findings. Periods which yield very large T_2 values also yield very low e values. Consider first the results obtained from the data with the outliers removed. The T_2 values for each commodity market are very similar and very high. Similarly the e values are all very low. The average values of the T_2 and e statistics indicate that the contracts with the highest degree of correlation are from the sugar futures market. The cocoa, coffee and rubber futures markets appear to have an almost identical, but lower degree of collinearaity.

We now turn our attention to the effect of the outliers on the measures of collinearity. As a first example consider the two T_2 values corresponding to the cocoa returns in period 1. Note that the T_2 value for the complete data is 849 and for the data with the 3 outlying observations removed is 738. For a second example consider the T_2 values for the rubber series in period 7 in which the values are 690 and 746

Table 5.4a

T measure of collinearity

Period	Cocoa T	Cocoa T*	Coffee T	Coffee T*	Sugar T	Sugar T*	Rubber T	Rubber T*
1	849	738	1251	747	1092	1961	750	727
2	773	764	1105	1111	1149	943	753	768
3	643	581	998	775	861	822	411	528
4	594	738	1130	1057	700	791	562	579
5	995	1036	916	870	736	710	547	621
6	912	805	995	1011	633	706	542	487
7	846	846	682	682	725	725	690	746
8	391	504	608	605	643	643	754	754
9	586	646	365	561	725	749	664	686
10	536	679	777	777	846	846	955	956
11	735	735	619	670	908	846	813	934
12	598	619	524	527	692	745	869	887
13	776	776	841	749	739	761	1029	1029
14	804	804	477	454	987	987	720	714

T = test statistic on R with complete data
T* = test statistic on R with outliers removed

Table 5.4b

e measure of collinearity

Period	Cocoa e	Cocoa e*	Coffee e	Coffee e*	Sugar e	Sugar e*	Rubber e	Rubber e*
1	1.38	1.49	1.16	1.35	1.18	1.18	1.47	1.45
2	1.47	1.48	1.20	1.17	1.17	1.28	1.55	1.50
3	1.45	1.30	1.12	1.14	1.20	1.20	2.16	1.40
4	1.64	1.32	1.09	1.11	1.40	1.30	1.63	1.57
5	1.12	1.10	1.17	1.92	1.37	1.32	1.65	1.52
6	1.17	1.23	1.11	1.09	1.50	1.28	1.73	1.78
7	1.25	1.25	1.40	1.40	1.31	1.31	1.58	1.29
8	2.32	1.81	1.51	1.47	1.49	1.49	1.34	1.34
9	1.61	1.39	2.33	1.56	1.38	1.27	1.55	1.34
10	1.77	1.40	1.31	1.31	1.22	1.22	1.16	1.14
11	1.33	1.33	1.49	1.38	1.18	1.21	1.28	1.15
12	1.59	1.49	1.79	1.73	1.42	1.24	1.24	1.23
13	1.35	1.35	1.32	1.27	1.36	1.30	1.13	1.13
14	1.25	1.25	1.86	1.94	1.13	1.13	1.40	1.40

e = effective degrees of freedom parameter with complete data
e* = effective degrees of freedom parameter with outliers removed

respectively. Why does the removal of a few outliers reduce the average correlation in the first example and increase the average correlation in the second example? The answer is obtained by considering the nature of the outliers present in both examples. In the first example, the outlying observation is of the type II outlined in section 5.3.1 and it is easy to see how such extremes in all components of the returns vector will exaggerate the correlations. In the second example the outlier is of the third type and it is also obvious why such outliers will reduce an otherwise high positive correlation. In order to see more clearly the effect the outliers have on the measures of collinearity we produce Table 5.4c. Table 5.4c has been constructed from Table 5.4a and 5.4b as follows. For each set of returns in each period observe the effect on T2 and e of removing the outliers. If the effect is to reduce the average correlation coefficients, place a zero (0) in the corresponding

Table 5.4c

Effect of removing outliers on collinearity measure

Period	Cocoa		Coffee		Sugar		Rubber	
	T-T*	e-e*	T-T*	e-e*	T-T*	e-e*	T-T*	e-e*
1	0	0	0	0	0	1	0	1
2	0	0	1	0	0	0	1	1
3	0	1	0	0	0	0	1	1
4	1	1	0	0	1	1	1	1
5	1	1	0	0	0	1	1	1
6	0	0	1	1	1	1	0	0
7	-	-	-	-	-	-	1	1
8	1	1	0	1	-	-	-	-
9	1	1	-	-	1	1	1	1
10	1	1	-	-	-	-	1	1
11	-	-	1	1	0	0	1	1
12	1	1	1	1	1	1	1	1
13	-	-	0	1	1	1	-	-
14	-	-	0	0	-	-	0	0

0 = removal of outlier reduces collinearity measure
1 = removal of outlier increases collinearity measure

row and column of Table 5.4c. If removing the outliers increases the average correlation coefficients place a one (1) in the corresponding row and column of Table 5.4c.

From Table 5.4c we see that the reported effects on the average degree of collinearity as measured by the two statistics are very similar. Considering all 4 futures markets there are 44 periods in which at least one outlier was detected. The two measures show different effects of the removal of outliers in only 6 of these 44 periods. If we consider only those periods in which the two measures show similar effects we see from Table 5.4c that there is a predominace of 1's. From this we conclude that overall, the presence of outliers in the data tend to have the effect of reducing the average degree of collinearity.

In Chapters 6,7,8 and 9 we return to the question of inter contract correlation coefficients again and consider in detail the idea of computing a grand average correlation matrix.

5.3.3 Testing the returns for multivariate normality

For each period and each set of returns we computed the multivariate skewness statistic , the multivariate kurtosis statistic and the P(W) value associated with the W - test of multivariate normality. The results appear in Tables 5.5 to 5.8. A count of the number of significant results (at the 5% level) for all series appears in Table 5.9. Referring to

Table 5.5

Multivariate normality tests on the cocoa series (logged data)

	Complete data			Without outliers		
Period	P(W)	Mul. skew	Mul. kut.	P(W)	Mul. skew	Mul. kut.
1	0.000 c	56.67 c	5.17	0.821	39.96 b	3.01 b
2	0.845	36.27 a	3.37 c	0.943	22.08	1.53
3	0.011 a	56.05 c	25.78 c	0.734	26.09	1.27
4	0.915	216.73 c	22.81 c	0.886	32.92 a	2.54 a
5	0.411	28.43	6.90 c	0.413	30.21	1.80
6	0.000 c	97.29 c	6.77 c	0.926	25.65	-0.05
7	0.909	8.55	0.02	0.909	8.55	0.02
8	0.006 b	84.12 c	22.96 c	0.061	35.46 a	3.12 b
9	0.000 c	72.87 c	19.76 c	0.000 c	33.49 a	1.60
10	0.477	756.10 c	36.10 c	0.578	28.14	0.19
11	0.652	15.49	-0.18	0.652	15.49	-0.18
12	0.587	24.72	7.74 c	0.706	27.85	1.24
13	0.654	37.75 b	2.79 b	0.654	37.75 b	2.79 b
14	0.550	13.61	0.84	0.550	13.61	0.84

Table 5.6

Multivariate normality tests on the coffee series (logged data)

	Complete data			Without outliers		
Period	P(W)	Mul. skew	Mul. kut.	P(W)	Mul. skew	Mul. kut.
1	0.000 c	1020.74 c	71.04 c	0.747	32.93 a	1.89
2	0.851	57.92 c	5.81 c	0.909	26.53	1.63
3	0.001 b	80.65 c	11.73 c	0.007 b	28.62	1.31
4	0.000 c	80.46 c	7.93 c	0.098	41.74 c	5.83 c
5	0.755	41.98 c	5.44 c	0.998	13.78	1.40
6	0.306	36.42 a	5.17 c	0.290	25.32	2.17 a
7	0.595	23.27	2.59 b	0.595	23.27	2.59 b
8	0.462	47.73 c	3.97 c	0.252	30.92	1.83
9	0.000	58.10 c	20.29 c	0.420	20.27	0.14
10	0.459	17.51	0.41	0.459	17.51	0.41
11	0.838	19.36	6.64 c	0.778	24.45	3.41 c
12	0.946	20.76	1.58	0.943	26.02	0.74
13	0.000 c	458.53 c	32.10 c	0.217	24.42	2.89 b
14	0.377	69.77 c	6.46 c	0.801	63.33 c	4.92 c

Table 5.7

Multivariate normality tests on the sugar series (logged data)

	Complete data			Without outliers		
Period	P(W)	Mul. skew.	Mul. kut.	P(W)	Mul. skew.	Mul. kut.
1	0.567	50.07 c	8.03 c	0.276	22.31	2.16 a
2	0.000 c	83.18 c	16.42 c	0.764	32.66 a	3.70 c
3	0.004 b	54.28 c	9.15 c	0.740	36.34 a	2.64 b
4	0.830	51.25 c	17.63 c	0.924	27.90	4.22 c
5	0.377	72.99 c	12.05 c	0.505	44.56 c	5.10 c
6	0.083	139.02 c	16.24 c	0.096	32.12 a	4.09 c
7	0.268	26.43	1.33	0.268	26.43	1.33
8	0.102	28.59	2.36 a	0.102	28.59	2.36 a
9	0.512	36.92 a	6.10 c	0.317	32.46 a	4.26 c
10	0.750	51.39 c	2.73 b	0.750	51.39 c	2.73 b
11	0.083	65.03 c	5.00 c	0.904	27.05	0.96
12	0.036 a	549.65 c	31.69 c	0.171	23.25	1.78
13	0.081	98.07 c	14.55 c	0.813	39.93 b	2.98 b
14	0.666	21.44	2.43 a	0.666	21.44	2.43

Table 5.8

Multivariate normality tests on the rubber series (logged data)

	Complete data			Without outliers		
Period	P(W)	Mul. skew.	Mul. kut.	P(W)	Mul. skew.	Mul. kut.
1	0.231	63.27 c	10.14 c	0.232	34.47 a	2.70 b
2	0.867	31.72	4.65 c	0.877	37.85 b	2.87 b
3	0.119	159.02 c	22.66 c	0.349	29.78	2.49 c
4	0.963	30.11	4.42 c	0.959	27.19	2.54 a
5	0.172	49.07 c	15.26 c	0.341	17.03	1.13
6	0.047 a	75.48 c	9.76 c	0.367	16.66	-0.50
7	0.488	31.97 a	11.03 c	0.265	32.74	1.52
8	0.687	19.01	3.49 c	0.687	19.01	3.49 c
9	0.671	23.76	5.41 c	0.779	39.88 b	2.60 b
10	0.625	39.63 b	3.09 c	0.638	19.92	0.52
11	0.844	23.84	11.32 c	0.811	27.73	2.55 a
12	0.036 a	67.95 c	6.66 c	0.047 a	35.14 a	3.47 c
13	0.939	21.41	2.23 a	0.939	21.41	2.23 a
14	0.571	12.16	1.59	0.590	9.22	1.53

Table 5.9

Number of significant (5%) statistics in tests of mulitivariate normality

	Complete data			Outliers removed		
	P(W)	Mul. skew	Mul. kut.	P(W)	Mul. skew	Mul. kut.
Cocoa	5	9	11	1	5	4
Coffee	5	10	12	1	3	6
Sugar	3	11	13	0	6	9
Rubber	2	7	13	1	5	9
Total	15	37	49	3	19	28

Tables 5.5 to 5.9 we make the following observations.

(i) Considering The P(W) values on the complete data sets we note that we obtain significant results in those periods which gave significant results for W - tests of univariate normality. This is of course not surprising. When the univariate W test indicates a significant departure from normality in all 4 components, the joint W - test for multivariate normality also produces a significant result. Removing outlying observations reduces the number of significant P(W) values from a total of 15 to 3.

(ii) There are a large number of significant multivariate skewness results. Of the 4 x 14 = 56 values, 37 are significant at the 5% level. The level of significance of most of the skewness values is striking. In the 4 th period of the cocoa series, for example, the observed value of 216.73 is extremely unlikely under the null hypothesis. The P(W) value for the 4 th period of the cocoa series is however 0.15, There are numerous situations such as this, in which the W - test passes a set of data as multivariate normal but the skewness test rejects multivariate normality. Removal of the outliers predictably reduces the magnitude of skewness measures. However, even with the outliers removed, the number of periods which result in significant skewness statistics is still high (19).

(iii) 49 of the 56 multivariate kurtosis values are significant at the 5% level and 43 of these are significant at the 0.1% level. Removal of the outliers reduces the number of values significant at the 5% level to 28 and at the 0.1% level to 10. With the complete data sets all but 1 of the kurtosis values are positive.

How do we explain these unusual results? Why does one measure of multivariate normality, the skewness or kurtosis value, indicate that the majority of the periods are highly non-multivariate normal, while another measure, the W - test, indicates non-multivariate normality in only a few instances? Even with the outliers removed, 9 of the kurtosis values for the sugar series are significant but none of the P(W) values are. It is also interesting to note that removing the pair of outliers (type 3) from period 7 of the rubber series produced a reduction in the kurtosis value from 11.03 to 1.52 but left the P(W) value and skewness statistic unchanged. There are a number of such instances in which removing outliers results in a drastic drop in the kurtosis value.

One possible explanation for these results is that there may be very complex departures from multivariate normality. It is possible to imagine departures from multivariate normality that would be apparent only when considering the joint distribution but not apparent in the separate univariate distributions[4]. For reasons outlined below we suspect the W - test is insensitive to such complex departures from multivariate normality.

The W - test of multivariate normality produces a statistic that is a mixture of the results of 4 separate tests of univariate normality. The mixing process involves a transformation that uses only correlations between the various component distributions. If the set of returns <u>is</u> multivariate normal then the relationships between the individual components would be completely described by the set of correlation coefficients and the results of the W - test would be valid and meaningful.

The multivariate skewness and kurtosis measures are functions of the Mahalanobis distances and angles of observations from the "centre" of data. It is possible that these measures are sensitive to and thus highlight complex departures from multivariate normality apparent only when considering the joint distribution. If this indeed was the situation, then the correlation matrix would not describe completely the joint distribution and the usefulness of the W - test with our data is thus brought into question.

Finally we observe that in some periods, as in the 13th cocoa series, we find no outliers but still report significant skewness and kurtosis values. The routine computes Mahalanobis distances of each observation form the mean and denotes those observations that are far from the centre as atypical. Could it be that observations resulting in unusual angles are not detected by the outlier routine and it is these observations that are giving rise to the significant skewness and kurtosis values? Unfortunately the literature on the multivariate skewness and kurtosis measures is quite thin and these measures do not appear to be very informative. All one can say is whether a result is significant or not.

5.3.4 Testing returns for multivariate serial correlation

The series were tested for multivariate serial correlation using the routine described in section 4.5. As in Chapter 4, rather than record all the individual P values associated with the multivariate serial correlation coefficients at various lags, we report only those values that could be considered significantly different (at the 5% or level or less) from zero. The analysis was carried out on the logged daily returns and

Table 5.10

Significant Multivariate Serial Correlation Coefficients for Cocoa Series

(a) Complete data set

Lags = / Period	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	R
1	c												a								0.345
2	c			a													a				0.303
3	c	a	a			a	b								b	b					0.435
4	c																				0.358
5	c						a				a										0.433
6	c						b	a	a												0.398
7	c																				0.330
8	c																				0.401
9	c				a		b											b			0.424
10	a																				0.279
11	*	a																			0.272
12	c	c	c		b	b	a	a	b	b	b		a	a							0.441
13	c							a													0.381
14	c									a										a	0.380
Total	13	3	2	1	2	2	5	3	2	2	2	0	2	1	1	1	1	1	0	1	

(b) With outliers removed

Lags = / Period	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	R
1	c	a	a																		0.318
2	b																				0.283
3	c	a																			0.394
4	c					b															0.838
5	c						a						a								0.369
6	c						b	b	b												0.410
7	c																				0.330
8	c		a						a												0.354
9	c																				0.413
10	c																				0.391
11	*	a																			0.272
12	c	a	a																		0.416
13	c							a													0.406
14	c									a										a	0.380
Total	13	4	3	0	0	1	2	2	2	1	0	0	1	0	0	0	0	0	0	1	

(R = value of multivariate serial correlation coefficient at lag 1)

* This coefficient was significant at the 6% level.

Table 5.11

Significant Multivariate Serial Correlation Coefficients for Coffee Series

(a) Complete data set

Lags = Period	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	R
1	c	c	a	a																	0.602
2	c	b			a																0.373
3	c		a																		0.431
4	c																				0.416
5	c																				0.337
6	c	b											a	b		a					0.471
7	c					a							b			b					0.371
8	a							a		a						a		a			0.294
9	c				a																0.400
10	b					a															0.323
11	c																		a	b	0.354
12	c				a	a															0.433
13	c	b							a												0.464
14	c							a								a					0.383
Total	14	4	2	1	3	3	0	2	1	1	0	0	2	1	0	4	0	1	1	1	

(b) With outliers removed

Lags = Period	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	R
1	c																				0.353
2	c	a																			0.377
3	c		a																		0.444
4	c				a														a		0.376
5	c																				0.289
6	c	c																			0.352
7	c						a						b			b					0.371
8	b		a													a					0.330
9	b	a																			0.310
10	b					a															0.323
11	a					a															0.307
12	c																				0.423
13	c	b		a																	0.388
14	c	a																			0.358
Total	14	5	2	1	1	2	1	0	0	0	0	0	1	0	0	2	0	0	1	0	

(R = value of multivariate serial correlation coefficient at lag 1)

Table 5.12

Significant Multivariate Serial Correlation Coefficients for Sugar Series

(a) Complete data set

Lags = Period	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	R
1	c	a	b							a											0.313
2	c	c	c		a					a		c	a								0.394
3	c																				0.357
4	b																				0.330
5	c																				0.385
6	c	a	b	b		a															0.407
7	b			a																	0.319
8	a		a					b	b							a					0.287
9																					0.227
10	b																				0.335
11	b	a													a						0.342
12	b								a												0.307
13	b			a	a	a															0.332
14	c														a						0.360
Total	13	4	4	3	2	2	0	1	2	2	0	1	1	0	2	1	0	0	0	0	

(b) With outliers removed

Lags = Period	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	R
1	b		a							b											0.306
2	c																				0.322
3	a				a	a															0.305
4	a																				0.298
5	b												a								0.323
6	b																				0.346
7	b			a				a													0.319
8	a		a					b	b							a					0.287
9																					0.208
10	b																				0.335
11	b																				0.309
12	b																				0.337
13	a																				0.276
14	c														a						0.360
Total	13	0	2	1	1	1	0	2	1	1	0	0	1	0	1	1	0	0	0	0	

(R = value of multivariate serial correlation coefficient at lag 1)

Table 5.13

Significant Multivariate Serial Correlation Coefficients for Rubber Series

(a) Complete data set

Lags = Period	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	R
1	b	a														a					0.280
2	a																				0.262
3	c	c	c	c	b		a										b			a	0.474
4	c			b									a								0.362
5	c																				0.374
6	b																				0.333
7	c	b																			0.393
8	c																				0.338
9	b																				0.311
10	c												b								0.405
11	c			b			b														0.414
12	*																				0.256
13	c	a																			0.369
14	c	b							b												0.391
Total	13	5	1	3	1	0	2	0	1	0	0	0	2	0	0	1	1	0	0	1	

(b) With outliers removed

Lags = Period	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	R
1	a			a																	0.271
2	b																				0.306
3	c		a			a					a										0.426
4	c																				0.371
5	b																				0.315
6	b																				0.326
7	b				a																0.327
8	b																				0.338
9	*																				0.285
10	c									a			b								0.404
11	c			a											a						0.383
12	b																				0.342
13	c	a																			0.369
14	c	a							a	a											0.388
Total	13	2	1	2	1	1	0	0	1	2	1	0	1	0	1	0	0	0	0	0	

(R = value of multivariate serial correlation coefficient at lag 1)

* This coefficient was significant at the 6% level.

Table 5.14

Count of significant (5%) multivariate serial correlation coefficients

	Cocoa		Coffee		Sugar		Rubber	
Lag	CD	WOR	CD	WOR	CD	WOR	CD	WOR
1	13	13	14	14	13	13	13	13
2	3	4	4	5	4	0	5	2
3	2	3	2	2	4	2	1	1
4	1	0	1	1	3	1	3	2
5	2	0	3	1	2	1	1	1
6	2	1	3	2	2	1	0	1
7	5	2	0	1	0	0	2	0
8	3	2	2	0	1	2	0	0
9	2	2	1	0	2	1	1	1
10	2	1	1	0	2	1	0	2
11	2	0	0	0	0	0	0	1
12	0	0	0	0	1	0	0	0
13	2	1	2	1	1	1	2	1
14	1	0	1	0	0	0	0	0
15	1	0	0	0	2	1	0	1
16	1	0	4	2	1	1	1	0
17	1	0	0	0	0	0	1	0
18	1	0	1	0	0	0	0	0
19	0	0	1	1	0	0	0	0
20	1	1	1	0	0	0	1	0
Total	45	30	41	30	38	25	31	26
R	.372	.366	.404	.357	.335	.309	.354	.347

R = average value of multivariate serial correlation coefficient at lag 1

CD column : complete data set
WOR column : with outliers removed

for brevity again we only considered O'Brien's statistic. The results appear in Tables 5.10 to 5.13. Table 5.14 gives a count of the total number of significant serial correlation coefficients at each lag, for lags 1 to 20 for each series.

The most startling observation to note from Tables 5.10 to 5.14 is the number of significant multivariate serial correlation coefficients of lag 1. Nearly all the coefficients are significant at the 5% level and many are significant at the 1% and the 0.1% level. The number of significant coefficients at lag 2,3 and 4 is also more than one would expect under the null hypothesis. Removing the outliers reduces the magnitude of many of the coefficients and thus the number that can be considered as significant. However, even with the anomalies removed, all but 3 of the 56 coefficients at lag 1 day are significant at the 5% level and 32 are significant at the 0.1% level. Because of these extremely interesting and unexpected results we report the magnitude of the lag 1 coefficients in Tables 5.10 to 5.13 and the average of the coefficients for each series in Table 5.14. The average value of all of the coefficients at lag 1 day is 0.366, which is reduced by removing the outliers to 0.345.

There appears then to be very strong evidence of persistent multivariate serial correlation at lag 1 day. Recall that no evidence of any type of serial dependence was observed when each of the component univariate series were examined in Chapter 3.

5.4 Summary of Chapter 5

In this chapter we looked at the distribution of daily returns on four contracts of differing maturity for the same commodity. In the sugar and rubber series we found that the standard deviation of the returns gradually decreases as one considers contracts with delivery dates further and further into the future. The cocoa and coffee series gave mixed results in this respect.

In the multivariate study we found more outlying observations than expected from the earlier univariate analysis and the anomalies were of 3 main types: (i) instances in which there is a simultaneous change in the variance of the returns; (ii) instances in which there is a single simultaneous large change in the contract prices; and (iii) instances in which one or two contract prices "get out of line" with the general price profile (lasting usually for only one day). Testing for multivariate normality and multivariate serial correlation produced most unexpected results. Many periods showed evidence of multivariate non normality of a possibly complex nature and all periods yielded highly significant multivariate serial correlation coefficients of lag 1.

Is it possible that some, for the moment unknown phenomenon, is giving rise to both these curious sets of results? We investigate this and other questions throughout the rest of this work.

Footnotes for Chapter 5

1. Schrock (1971) outlines theoretically, how by the use of spread positions, an invetstor can construct very low risk portfolios.
2. It is not possible to carry out a formal test for the ratios of of variances using the F test as the returns are not independent.
3. In this study we only compare adjacent contracts.
4. See Kendall and Stuart p 395.

CHAPTER 6

MULTIVARIATE SERIAL CORRELATION IN COMMODITY TIME SERIES: A DETAILED EXAMINATION

In Chapter 5 we reported the discovery of persistent multivariate serial correlation in 4 dimensional sets of returns on contracts in the same futures markets with different delivery dates. As an exercise we also examined sets of returns on 2, 3, 5 and 6 (and 7 and 8 in the case of rubber) contracts from the same futures market. The multivariate serial correlation coefficients of lag 1 were all highly significant. For brevity we do not report these results here.

In this chapter we investigate in more detail the nature of the multivariate temporal dependence. The layout of this chapter is as follows. In section 6.1 we review the properties of O'Briens multivariate serial correlation coefficient (MVSC). In section 6.2 we briefly review canonical correlation analysis and in section 6.3 we show how the MVSC can be decomposed into separate canonical correlation coefficients. The study of up to 6 contract returns requires a different sampling design to that used in Chapter 5 and the new design is described in section 6.4. The results of the analysis are presented in section 6.5. The sample canonical vectors (explained later) are found to be extremely changable. An investigation into the sampling properties of canonical vectors is outlined in section 6.6. Attempts at stabilising sample canonical correlation vectors are reviewed in section 6.7 and the results of applying the procedures to the commodity series are given in section 6.8. A technique that produces grand average canonical vectors over the entire 5 year period is outlined in section 6.9 and the results are given in section 6.10. Concluding remarks are made in section 6.11.

6.1 The properties of O'Brien's multivariate serial correlation coefficient

We list the properties of O'Brien's (1980) multivariate serial correlation coefficient, R, defined in section 4.5:

(i) $0 \leqslant R^2 \leqslant 1$

(ii) In the univariate case, R^2 reduces to the coefficient of determination.

(iii) R^2 has the usual properties of explained variance interpretation.

(iv R^2 can be expressed as the geometric mean of the coefficients of non determination of various linear combinations of the returns:

$$1 - R^2 = \prod_{i=1}^{p} (1 - \lambda_i)^{(1/p)}$$

where $\sqrt{\lambda_i}$, $(i=1,2,..,p)$ are the canonical correlation coefficient between returns on day t and day t - 1.

From property (iv) we see that the multivariate serial correlation coefficient is a simple function of the product of the canonical correlation coefficients. The discovery of significant multivariate serial correlation coefficients implies then that one or more of the canonical correlation coefficients is significantly different from zero.

Associated with each canonical correlation coefficient is a linear combination of returns. The linear combination is obtained using the elements of a canonical variate vector. A study of these canonical variates could prove quite interesting. Before we consider the results of our empirical study, we briefly review canonical analysis.

6.2 Canonical correlation analysis: a synopsis

Canonical correlation analysis involves partitioning a collection of variables, $\underline{x}$, into two sets, an $\underline{x}_1$, set and an $\underline{x}_2$ set. The object is to find linear combinations $\psi_1 = \underline{a}^T\underline{x}_1$ and $\psi_2 = \underline{b}^T \underline{x}_2$ such that ψ_1 and ψ_2 have the largest possible correlation. Such linear combinations can give insight into relationships between the two sets of variables.

Formally: consider a random variable $\underline{x}$ of dimension $((p_1+p_2)x1)$ partitioned as follows:

$$\underline{x} = \begin{bmatrix} \underline{x}_1 \\ \\ \underline{x}_2 \end{bmatrix}$$

where $\underline{x}_1$ is (p_1x1) and $\underline{x}_2$ is (p_2x1). $\underline{x}$ is normally distributed:

$$\underline{x} \quad N \; (\underline{\mu} \; , \; V)$$

with $$\underline{\mu} = \begin{bmatrix} \underline{\mu}_1 \\ \\ \underline{\mu}_2 \end{bmatrix} \qquad V = \begin{bmatrix} V_{11} & V_{12} \\ \\ V_{21} & V_{22} \end{bmatrix}$$

We wish to find vectors $\underline{a}$ $(p_1 x 1)$ and $\underline{b}$ $(p_2 x 1)$ such that subject to the constraints:

$$\underline{a}^T \; V_{11} \; \underline{a} \; = \; 1$$

$$\text{and} \quad \underline{b}^T \; V_{22} \; \underline{b} \; = \; 1$$

$$\text{we maximize} \quad \lambda \; = \; \underline{a}^T \; V_{12} \; \underline{b}$$

It is easy to show that λ and $\underline{a}$ are the eigenvalues and eigenvectors respectively of the matrix A where:

$$A = V^{-1}_{11} \; V_{12} \; V^{-1}_{22} \; V_{21}$$

with $\underline{b}$ given by $$\underline{b} = V^{-1}_{11} \, V_{12} \, \underline{a} / \lambda$$

If the rank of V_{12} is k (< $\min(p_1, p_2)$) so is that of A and there will be k distinct sets of λ , $\underline{a}$, $\underline{b}$. It is usual to consider eigenvalues ranked so that $\lambda_1 > \lambda_2 > \ldots > \lambda_k$. The square roots of the λ_i ,i=1,..k are known as the canonical correlation coefficients and the associated vectors $\underline{a}_i$, $\underline{b}_i$, i = 1,..k the canonical coefficient vectors. It is easy to show that the correlation coefficient between $\underline{a}^T{}_1 \underline{x}_1$ and $\underline{b}^T{}_1 \underline{x}_2$ is $\sqrt{\lambda_1}$. Interest is usually centred on that combination $\underline{a}^T{}_1 \underline{x}_1$, $\underline{b}^T{}_1 \underline{x}_2$ that are maximally correlated.

In an empirical study we use the classical estimates of V_{11}, V_{12} etc and obtain sample estimates : $\hat{\lambda}_i$, $\hat{\underline{a}}_i$, $\hat{\underline{b}}_i$, i=1,...k. Kshirsagor (1972) has derived the sampling distribution of these estimates under the null hypothesis of no correlation between $\underline{x}_1$ and $\underline{x}_2$ (ie V_{12} = 0) but we leave discussion of this subject to section 6.6.

However, it is possible to test the hypothesis that $V_{12} = 0$ using the classic likelihood ratio statistic, L:

$$L = |I - A| = \prod_{i=1}^{k} (1 - \hat{\lambda}_i)$$

which has Wilk's $\Lambda(\rho_1, n-1-\rho_2, \rho_2)$ distribution. Of course we can use Bartlett's approximation and compute the U statistic given by:

$$U = -(m - 1 - (\rho_1 + \rho_2 + 1)/2) \sum_{i=1}^{k} \log(1 - \hat{\lambda}_i),$$

which under the null hypothesis of $\lambda_i = 0$ for all i = 1, 2,.. k, is asymptotically $\chi^2(\rho_1\rho_2)$ distributed.

6.3 Decomposing the multivariate serial correlation coefficient into canonical correlations

Using the notation of section 6.2 we have a sequence of returns $\underline{x}_t$, t = 1, 2,. .. n of dimension $(\rho x 1)$. We consider the vector $\underline{y}_t$ such that

$$\underline{y}_t = \begin{bmatrix} \underline{x}_t \\ \\ \underline{x}_{t-1} \end{bmatrix} \qquad t = 2,3, .. \quad n$$

Assume that $\underline{y}_t$ is normally distributed[1] with

$$\underline{y}_t = \begin{bmatrix} \underline{x}_t \\ \\ \underline{x}_{t-1} \end{bmatrix} \sim N \left[\begin{bmatrix} \underline{\mu}_t \\ \\ \underline{\mu}_{t-1} \end{bmatrix}, \begin{bmatrix} V_{t,t} & V_{t,t-1} \\ \\ V_{t-1,t} & V_{t-1,t-1} \end{bmatrix} \right]$$

Clearly the canonical correlations, $\sqrt{\lambda_i}$ between $\underline{x}_t$ and $\underline{x}_{t-1}$ are obtained from the matrix A:

$$A = V^{-1}_{t,t} V_{t,t-1} V^{-1}_{t-1,t-1} V_{t-1,t}$$

The vectors $\underline{a}_i$ are the associated eigenvectors of A and the vectors $\underline{b}_i$ are obtained using:

$$\underline{b}_i = V^{-1}_{t,t} V_{t,t-1} \underline{a}_i / \lambda_i$$

The multivariate serial correlation coefficient, R is thus given by:

$$1 - R^2 = \prod_{i=1}^{p} (1 - \lambda_i)^{(1/p)}$$

In section 6.5 we report R, along with the associated level of significance, the maximum value of $\sqrt{\lambda_i}$ and corresponding canonical coefficients.

6.4 Sampling design for canonical correlation study

In chapter 5 we examined the logged daily returns on four contracts from the same futures market. As noted earlier on in this chapter we also examined 2, 3, 5, 6, 7 and 8 dimensional sets of returns for various periods. In the desire to discover the exact nature of the multivariate serial correlation of lag 1 that seems to be present we will need to examine the values of λ_i, $\underline{a}_i$ and $\underline{b}_i$ for each period. It is possible that the discovery of some sort of pattern in the vectors

$\underline{a}_1$ and $\underline{b}_1$ could lead to the development of a successful trading rule. As mentioned before interest is usually centred on the coefficients $\underline{a}_1$, $\underline{b}_1$ that are maximally correlated. It is desirable therefore to have estimates of λ_1, $\underline{a}_1$ and $\underline{b}_1$ that are as accurate as possible. For accurate estimates we will need large samples.

There is also the problem of the dimensionalty of the set of returns. How many contracts should we simultaneously examine? It would be desirable to examine as many contracts as are traded. Achieving these two objectives (of having large samples and studying many contract returns) simultaneously is, however, not possible.

As an example it is possible to monitor the same 2 contracts from the rubber market for up to 21 months. With 8 contracts of rubber the maximum period one can study all 8 returns is only 3 months. There are also the arguments outlined in section 3.6.2 on the usefulness of looking at returns on contracts over very long periods.

Taking all these considerations into account it was decided to limit the number of contracts studied to 6. We decided to use the 6 longest running contracts. The period of examination ends when a contract expires (matures). Note that the 6 contracts used in each period are the 6 contracts being examined for univariate serial correlation in section 3.6.2. A diagramatic illustration of the sampling design appears in Fig. 3.10.

By using this sampling design the length of each set of 6 returns will vary for each series. The cocoa series for example are of length 2 months, 3 months, 3 months, 2 months etc and there are 24 periods. In each period the first component of the returns will be the returns on the contract that matures at the end of that period. The second component is the returns on the contract that will be the next to mature and so on. Thus as we proceed from period to period we will be considering a 6 dimensional vector that will represent the returns on the 6 longest running futures contracts traded in a given futures market. Throughout the rest of this chapter we use daily logged returns with complete data sets.

6.5 A Discussion of the multivariate serial correlation and the canonical correlation coefficients

The multivariate serial correlation coefficients at lag 1 and associated P value for each set of returns foreach period are given in Tables 6.1a to 6.1d. Also given in the tables is the root of the maximum eigenvalues ,$\sqrt{\lambda_{max}}$: the correlation coefficient associated with that pair of linear combinations that are maximally correlated.

The most striking observation from the tables is the number of significant R values. All but 12 of the 96 R values are significant at the 5% level and many are significant at the 1% and the 0.1% level. These results confirm the statements made at the beginning of the chapter. There is persistent multivariate serial correlation at a lag of one day. It is interesting to note that the magnitudes of the R values are fairly constant and very similar for each set of returns. The averages and standard deviations of the R values also appear in the tables.

Recall that one can interpret the value of R^2 in the same way as one interprets the coefficient of determination in classical regression. The average value of R^2 for the cocoa series, for example is $R^2 = 0.23$. This implies tht 23% of the variation in the returns on any given day can be explained by the returns on the previous day. This is of course only when we are considering <u>all</u> the contracts in a given futures market. There is no such dependence evident in the individual contract returns.

Consider now the values of $\sqrt{\hat{\lambda}_{max}}$ given in Tables 6.1a to

Table 6.1a   Table 6.1b

Multivariate serial correlation coefficients and maximum canonical correlation coefficents for cocoa and coffee series

Cocoa series

Period	R	P(R)	$\sqrt{\hat{\lambda}_{max}}$
1	.401	.041	.571
2	.593	.000	.822
3	.408	.441	.623
4	.384	.034	.608
5	.520	.000	.733
6	.480	.065	.770
7	.505	.004	.745
8	.488	.017	.671
9	.401	.011	.626
10	.459	.000	.762
11	.537	.002	.752
12	.477	.057	.710
13	.572	.000	.843
14	.401	.013	.667
15	.510	.000	.705
16	.574	.000	.840
17	.496	.011	.683
18	.537	.001	.771
19	.445	.000	.678
20	.510	.000	.673
21	.518	.009	.750
22	.422	.274	.600
23	.540	.001	.729
24	.429	.002	.673
average:	.484		.709

Coffee series

Period	R	P(R)	$\sqrt{\hat{\lambda}_{max}}$
1	.473	.000	.721
2	.703	.000	.927
3	.484	.029	.700
4	.500	.009	.780
5	.579	.000	.826
6	.482	.024	.718
7	.556	.001	.781
8	.069	.000	.909
9	.544	.000	.703
10	.534	.001	.770
11	.543	.001	.752
12	.484	.022	.752
13	.560	.000	.832
14	.509	.016	.789
15	.457	.047	.671
16	.462	.063	.691
17	.451	.162	.594
18	.553	.000	.755
19	.440	.228	.698
20	.435	.184	.640
21	.477	.031	.728
22	.451	.104	.607
23	.438	.201	.629
24	.485	.040	.675
25	.537	.002	.770
26	.418	.303	.646
27	.545	.001	.792
28	.487	.008	.755
29	.825	.011	.985
average:	.499		.744

R = O'Briens mulitivariate serial correlation coefficent at lag 1
P(R) = P value associated with R
$\sqrt{\hat{\lambda}_{max}}$ = maximum correlation between the two linear combinations $\underline{a}^T \underline{x}_t$ and $\underline{b}^T \underline{x}_{t-1}$

Table 6.1c   Table 6.1d

Multivariate serial correlation coefficients and maximum canonical correlation coefficients for the sugar and rubber series

Sugar series

Period	R	P(R)	$\sqrt{\lambda_{max}}$
1	.451	.001	.670
2	.484	.000	.735
3	.540	.000	.770
4	.444	.202	.694
5	.396	.015	.594
6	.524	.006	.730
7	.412	.003	.603
8	.512	.004	.764
9	.433	.200	.660
10	.488	.000	.783
11	.495	.024	.746
12	.386	.035	.566
13	.420	.294	.679
14	.478	.041	.781
15	.397	.022	.622
16	.471	.073	.723
17	.481	.000	.727
18	.469	.045	.746
19	.481	.047	.680
20	.436	.001	.641
21	.495	.033	.774
22	.404	.007	.645
23	.508	.005	.756
24	-	-	-
average:	.461		.700

Rubber series

Period	R	P(R)	$\sqrt{\lambda_{max}}$
1	.447	.002	.640
2	.394	.018	.649
3	.385	.025	.560
4	.445	.001	.670
5	.439	.001	.722
6	.436	.000	.605
7	.419	.000	.712
8	.422	.005	.618
9	.488	.000	.651
10	.460	.000	.613
11	.412	.004	.606
12	.453	.000	.708
13	.431	.002	.596
14	.468	.000	.715
15	.439	.000	.665
16	.460	.000	.606
17	.434	.001	.684
18	.473	.000	.682
19	.427	.001	.665
20	.691	.610	.879
average:	.451		.662

R = O'Briens mulitivariate serial correlation coefficent at lag 1
P(R) = P value associated with R
$\sqrt{\lambda_{max}}$ = maximum correlation between the two linear combinations $\underline{a}^T \underline{x}_t$ and $\underline{b}^T \underline{x}_{t-1}$

Table 6.2

Estimates of a , b components for the 24 periods of cocoa returns

Period		Components 1	2	3	4	5	6	length of vector	sum of comps.
1	a	-123.0	-123.0	-123.0	-123.0	-123.0	-123.0	678.9	19.9
	b	6.8	25.9	-148.1	269.5	-17.7	-49.8	518.2	-17.5
2	a	19.4	-46.6	80.6	114.4	-180.4	28.1	469.4	15.3
	b	-46.6	46.7	-119.8	-72.0	149.6	82.4	519.2	38.3
3	a	10.4	-99.2	309.0	-195.3	-157.9	83.0	854.8	-50.1
	b	2.7	144.4	-381.5	195.1	168.8	-125.7	1018.2	3.8
4	a	37.4	41.7	-222.8	145.9	-63.1	106.4	617.4	-45.5
	b	56.0	-101.0	129.3	-190.5	164.9	-54.5	696.2	-4.3
5	a	-27.6	-140.1	109.9	-178.1	257.1	4.7	717.6	25.9
	b	-70.3	182.9	-73.4	207.1	-288.8	3.0	825.5	-39.5
6	a	56.3	-93.6	30.7	-63.1	221.2	-147.0	611.8	4.6
	b	0.9	50.4	-92.8	6.6	-192.8	229.6	573.2	2.0
7	a	6.2	-55.6	-119.4	293.1	-98.5	-17.7	590.0	8.0
	b	-18.8	-44.0	318.0	-269.9	-54.9	32.6	738.3	-37.0
8	a	-17.2	73.5	-136.3	-6.6	158.4	-76.0	468.1	-4.3
	b	-51.4	39.0	203.5	-93.1	-185.5	101.3	673.9	13.7
9	a	-31.7	219.9	-218.3	297.1	-316.6	51.8	1135.4	2.2
	b	-19.4	129.6	53.1	-268.9	266.7	-160.3	898.0	0.8
10	a	-44.1	242.7	-205.6	-5.4	154.5	-129.9	782.3	12.1
	b	-35.6	-46.7	106.9	-24.7	-148.0	151.1	513.0	3.0
11	a	-5.1	59.9	-90.0	18.8	46.7	-22.0	242.4	8.4
	b	9.7	-22.5	-10.4	-100.2	-22.2	156.5	321.5	10.9
12	a	-46.1	!93.6	132.1	-212.1	-44.1	55.3	583.2	-21.3
	b	20.8	-74.9	-126.7	251.3	-78.6	-15.6	567.8	-23.8

Table 6.2 continued

Period		Components 1	2	3	4	5	6	length of vector	sum of comps.
13	a	12.4	29.5	-36.2	59.4	-73.5	-6.4	217.4	-14.8
	b	30.8	-207.9	222.0	-145.4	61.9	48.9	717.0	10.3
14	a	-10.0	-29.3	255.7	-114.1	-70.3	-21.9	501.3	10.1
	b	-25.2	183.3	-339.8	149.9	54.0	-10.5	762.8	11.7
15	a	-11.8	82.1	21.6	-201.5	-60.5	176.9	554.5	6.9
	b	16.6	-61.3	-30.3	215.4	22.9	-176.1	522.7	-12.8
16	a	-56.1	75.5	-46.4	8.5	-121.4	133.0	440.8	-6.9
	b	-25.6	62.4	-98.2	-8.7	241.0	-188.7	624.5	-17.9
17	a	32.9	5.6	-306.7	-29.4	406.2	-72.2	835.1	36.3
	b	-130.0	24.8	159.3	155.6	-307.8	82.1	859.5	-15.9
18	a	3.6	-241.0	256.8	323.4	-240.6	-65.4	1130.7	36.8
	b	170.1	49.7	-103.9	-62.0	-86.6	18.8	491.2	-13.9
19	a	7.3	182.5	-444.3	339.0	-364.9	275.4	1613.4	-4.9
	b	84.2	-346.4	534.0	-459.9	317.9	-148.9	1891.4	-19.1
20	a	54.4	-75.9	247.7	-383.3	272.2	-108.7	1142.1	6.3
	b	-7.4	93.1	-144.4	276.5	-288.3	61.1	870.8	-9.4
21	a	-33.3	281.0	-220.4	-229L9	405.6	-170.1	1340.3	32.9
	b	87.8	-161.4	-23.8	387.1	-310.8	23.5	994.3	2.6
22	a	47.2	-269.7	377.4	-37.1	-63.8	-35.3	830.5	18.8
	b	7.9	179.2	-386.6	18.2	121.2	38.3	751.3	-21.9
23	a	3.1	-30.0	-140.4	222.8	195.9	-272.2	864.4	-20.9
	b	-66.1	177.8	-22.5	-238.7	38.9	70.1	614.3	-40.5
24	a	21.3	303.4	-258.2	-106.7	-132.5	167.1	989.4	-5.7
	b	39.5	-353.8	606.8	-377.6	208.9	-105.2	1691.7	18.6

6.1d. Not surprisingly the $\sqrt{\hat{\lambda}_{max}}$ values are larger than the corresponding R values. As an example the average $\sqrt{\hat{\lambda}_{max}}$ value for the cocoa series is 0.709. Squaring this value we obtain 0.503. This implies then that it is possible to find a linear comination of returns on any given day, $\underline{b}^T \underline{x}_{t-1}$, that will explain about 50% of the variation of another linear combination of returns on the next day, $\underline{a}^T \underline{x}_{t-1}$. These results are very surprising. How can there by such temporal dependence in the returns and what is causing it?

Some light may be cast on these questions by examining the vectors $\hat{\underline{a}}$ and $\hat{\underline{b}}$; the canonical coefficients. In Table 6.2 we list the estimates of $\underline{a}$ and $\underline{b}$ for each period of the cocoa series.

There is a tremendous variability in the magnitude of the estimates of the elements of the vectors $\underline{a}$ and $\underline{b}$ and it is very difficult to find any consistent pattern. It is possible, however, to make a number of general remarks:

(i) The magnitude of the first elements of $\hat{\underline{a}}$ and $\hat{\underline{b}}$ ($\hat{a}_1$ and $\hat{b}_1$) are invariably smaller than the magnitude of the remaining elements ($\hat{a}_2$ to $\hat{a}_6$ and $\hat{b}_2$ to $\hat{b}_6$).

(ii) Elements with the maximum absolute values tend to be in the middle of the vectors.

(iii) In those situations in which a particular element of $\hat{\underline{a}}$ (say) is large and positive, the corresponding element of $\hat{\underline{b}}$ is usually large and negative.

(iv) Following on from point (iii) we note that the signs of corresponding elements of $\hat{\underline{a}}$ and $\hat{\underline{b}}$ are different more often then they are the same, except that is for the first elements. The frequency of these situations in which the signs of the corresponding elements of $\hat{\underline{a}}$ and $\hat{\underline{b}}$ are different is given in Table 6.2a.

Table 6.2a

Frequency of opposite signs of corresponding elements of $\underline{\hat{a}}$ and $\underline{\hat{b}}$.

Elements:	1	2	3	4	5	6
Observed Frequency	8	19	20	21	21	21
Expected Frequency	12	12	12	12	12	12

Also given in Table 6.2a are the expected frequencies of opposite signs if there were no particular pattern in the signs of the elements.

(v) Column 9 of Table 6.2 gives the "lengths" of the vectors[2] $\underline{\hat{a}}$ and $\underline{\hat{b}}$. These lengths vary considerably from period to period.

(vi) Column 10 of Table 6.2 gives the sums of the elements of each vector. These sums seem remarkably close to zero.

(vii) Examining the sequence of signs of the elements of $\underline{\hat{a}}$ (say) in the order $\hat{a}_1, \hat{a}_2, .., \hat{a}_6$, we notice something very interesting. The signs are certainly not random. There is an excess of +/- and -/+ pairs. With 6 elements there are 5 contiguous + and or - sign pairs to consider. With 24 periods we have 24 x 5 = 120 possible pairs. If the elements of $\underline{\hat{a}}$ were random we would expect 60 +/- and -/+ pairs. The observed number is 87 with only 33 +/+ and -/- pairs. There is a similar excess of +/- and -/+ sequences in the $\underline{\hat{b}}$ vectors.

The estimates $\underline{\hat{a}}$ and $\underline{\hat{b}}$ have been computed for all periods of he the coffee, sugar and rubber series and the general remarks made above apply to all the series. For brevity we do not report the results here.

Although the above general observations are extremely interesting the highly changable values of the elements of the $\hat{\underline{a}}$ and $\hat{\underline{b}}$ vectors make it very difficult to come to any definite conclusions regarding the nature of the overall MV serial correlation present in the series. We pose the question: is the instability of the $\hat{\underline{a}}$ and $\hat{\underline{b}}$ vectors caused by the underlying population values $\underline{a}$ and $\underline{b}$ changing? Or is the instability a characteristic of the estimation procedure? If the population values of $\underline{a}$ and $\underline{b}$ are fixed, two possible reasons for the observed instablity are suggested:

(i) departure from multivariate normality peculiar to these series,

or

(ii) possibly ill - conditioned matrices resulting from highly collinear data.

Obviously we hope that the underlying population $\underline{a}$ and $\underline{b}$ vectors are fairly constant and we need to investigate techniques that will estimate them more accurately. Before we address this problem however it was considered instructive to briefly review the little theoretical work that has been carried out on the sampling distribution of $\hat{\lambda}$, $\hat{\underline{a}}$ and $\hat{\underline{b}}$.

6.6 The sampling distribution of canonical correlations and canonical coefficients. Some simulation results

We consider only the situation in which $p_1 = p_2$. V is the population matrix of the partitioned vector $\underline{x}^T = (\underline{x}^T_1 : \underline{x}^T_2)$ and S is the sample estimate of V:

$$V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \qquad S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}$$

6.6.1 Theoretical sampling distribution of λ , a , b

Kshirsagar (1972) shows that the canonical correlations and canonical vectors can be obtained by considering the matrices D and F given by:

$$S_{12} \; S^{-1}{}_{22} \; S_{21} \; = \; D \; F \; D^T$$

in which F is diagonal with the eigenvalues (ordered) $\hat{\lambda}_1 > \ldots > \hat{\lambda}_p$ corresponding to the squared canonical correlations. The columns of D^{-1} are the canonical vectors $\hat{\underline{a}}$, $\hat{\underline{b}}$. Under the null situation of no correlation between $\underline{x}_1$ and $\underline{x}_2$ (ie V = 0), Kshirsagar has dervied the distribution of λ_i to be as follows:

$$\pi^{p/2} \prod_{i=1}^{p} \left\{ \frac{\psi(n).(1-\hat{\lambda}_i)^{(n-2p-1)/2}}{\sqrt{\lambda_i}\;\;(\psi(p))^2\;\psi(n-p)} \cdot \left\{ \prod_{i>j}^{p} (\hat{\lambda}_i - \hat{\lambda}_j) \right\} d\lambda_i \right\}$$

in which n = sample size and

$$\psi(n) = \Gamma\left[\frac{n+1-i}{2}\right]$$

The distribution of the canonical vectors is given by:

$$\left[\prod_{i=1}^{p} \begin{bmatrix}\psi(p)\\ \psi(n)\end{bmatrix}\right] \cdot \frac{\exp(-0.5.\mathrm{tr}[\;V^{-1}\;D\;D^T\;])}{|D^{-1}|^{(n+p)}\;\sqrt{2^{np}\;\pi^{p^2}\;|V|}} \cdot d(D^{-1})$$

Both these expressions are extremely complex and since they apply only to the null situation have limited usefulness. As Kshirsagar points out the distribution in the non null case are mathematically intractable. In order to gain some insight into the sampling properties of $\hat{\lambda}$, $\underline{\hat{a}}$ and $\underline{\hat{b}}$ it was decided therefore to carry out a simulation study.

6.6.2 Simulation study

Outlined briefly in this section are some of the results of an extensive simulation study of the sampling distribution of $\hat{\lambda}$, $\underline{\hat{a}}$ and $\underline{\hat{b}}$. The estimates of λ, $\underline{a}$ and $\underline{b}$ were obtained from samples of size, n, generated from a multivariate normal distribution of dimension 2ρ. Simulations were carried out using $\rho = 2, 3, 4, 5$ and 6 dimensional vectors and many different variance/covariance structures. For simplicity of exposition we consider here only the results of simulations with $\rho = 2$ dimensions and 4 different covariance structures. It is easy to show that in the study of canonical correlation analysis (see section 6.9.2) that one need only consider the associated partitioned correlation matrix:

$$R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$$

With each R_{ij}, being $(\rho x \rho)$. In the special situations we are interested we obviously set $R_{11} = R_{22}$ with

$$R_{11} = R_{22} = \begin{bmatrix} 1 & \alpha \\ \alpha & 1 \end{bmatrix}$$

Four sets of simulations of samples of size n were carried out using $\alpha =$

0.75, 0.9, 0.95 and 0.99. These four values of α were chosen to illustrate the effect of various degrees of collinearity in the data.

The choice of values for R_{12} was much more difficult. Examination of many of the associated S_{12} matrices used in calculating the results of section 6.5 shows that all the elements of R_{12} are typically small. Many different R_{12} matrices were tried and for brevity we only report here the results using the matrix:

$$R_{12} = \begin{bmatrix} 0.2 & 0.1 \\ 0.1 & 0.0 \end{bmatrix}$$

Two measures of interest associated with the estimates were computed : the bias B(.) and the root mean square error, RMSE(.) defined as follows:

$$B(\hat{\lambda}) = \sum_{j=1}^{m} (\hat{\lambda}_{(j)} - \lambda)/m$$

$$B(\hat{a}_i) = \sum_{j=1}^{m} (\hat{a}_{i(j)} - a_i)/m \quad .$$

$$RMSE(\hat{\lambda}) = \sum_{j=1}^{m} (\hat{\lambda}_{(j)} - \lambda)^2/m$$

$$RMSE(\hat{a}_i) = \sum_{j=1}^{m} (\hat{a}_{i(j)} - a_i)^2/m \quad ,$$

where m = number of samples of size n, i.e. the number of simulations,

λ = maximum population eigenvalue

$\hat{\lambda}_{(j)}$ = j th value of the maximum sample eigenvalue, $j = 1,..,m$

a_i = i th element of population vector $\underline{a}$ associated with maximum eigenvalue,

$\hat{a}_{i(j)}$ = j th estimate of ai, $j = 1,..,m$

We define $B(\hat{b}_i)$ and $RMSE(\hat{b}_i)$ similarly. For comparability with the empirical study of section 6.5 the sample size n was set to 50.

Thus m = 500 samples of size 50 were generated using the routine described in section 4.7. The estimates $\hat{\lambda}_{(j)}, \hat{a}_{i(j)}$ and $\hat{b}_{i(j)}$ were computed for each sample j = 1, to 500. The biases and root mean square erros appear in Table 6.3.

Referring to Table 6.3 we note that the population values of λ, a_i, and b_i, increase as the value of α increases. With regard to the sampling properties of $\hat{\lambda}$, $\hat{\underline{a}}$ and $\hat{\underline{b}}$ note that the bias and the root mean square of each statistic also increases with α. As an example in the very extreme case of $\alpha = 0.99$ when $a_1 = 5.19$ the estimated bias in $\hat{a}_1$ is 0.869 and the RMSE is 2.493. These values are relatively large and thus it is not surprising that the estimated $\hat{\underline{a}}$ and $\hat{\underline{b}}$ vectors in the empirical study in section 6.5 exhibited such marked variability.

Even more interesting, however, is the joint distribution of the individual components of the $\hat{\underline{a}}$ and $\hat{\underline{b}}$ vectors. As an illustrative example, the scatter plots of $\hat{a}_1$, against $\hat{a}_2$ obtained from the first 60 simultations appear in Figs. 6.1 to 6.4. The corresponding population values : a_1 and a_2 are denoted in each diagram by a cross.

When $\alpha = 0.75$, $\hat{a}_1$, and $\hat{a}_2$ are distributed (not uniformly) on an ellipse with negative sloping major axis, as α increases, the elliptical distribution collapses into a narrow distribution around a downward sloping straight line. The diagrams illustrate graphically the increasing variability of each component as α increases. The distribution of $\hat{\underline{b}}$ is similar.

Although here we have only reported the results of a simple 2 dimensional study, examples with other variance\covariance structures and higher dimensions have been examined. In all cases one obtains $\hat{\underline{a}}$ and

Table 6.3

Results of simulation study on the canonical correlation estimates with population correlation matrix R

$$R = \begin{bmatrix} 1.0 & \alpha & .2 & .1 \\ \alpha & 1.0 & .1 & .0 \\ .2 & .1 & 1.0 & \alpha \\ .1 & .0 & \alpha & 1.0 \end{bmatrix}$$

α	=	0.75	0.90	0.95	0.99
Population values	λ	.046	.083	.141	0.579
	a_1	1.244	1.794	2.437	5.19
	a_2	-.364	-.992	-1.666	-4.45
	b_1	1.244	1.794	2.437	5.19
	b_2	-.364	-.992	-1.666	-4.45
Biases	λ	.069	.079	.082	.0386
	a_1	-.326	-.464	-.556	-.8685
	a_2	.051	.235	.315	.7588
	b_1	-.343	-.471	-.650	-.9162
	b_2	.062	.267	.451	.8016
Root mean square errors	λ	.097	.114	.123	.0847
	a_1	.732	1.128	1.519	2.493
	a_2	.870	1.291	1.727	2.676
	b_1	.763	1.113	1.596	2.739
	b_2	.906	1.292	1.782	2.911

(Statistics estimated on 500 simulations of samples of size 50)

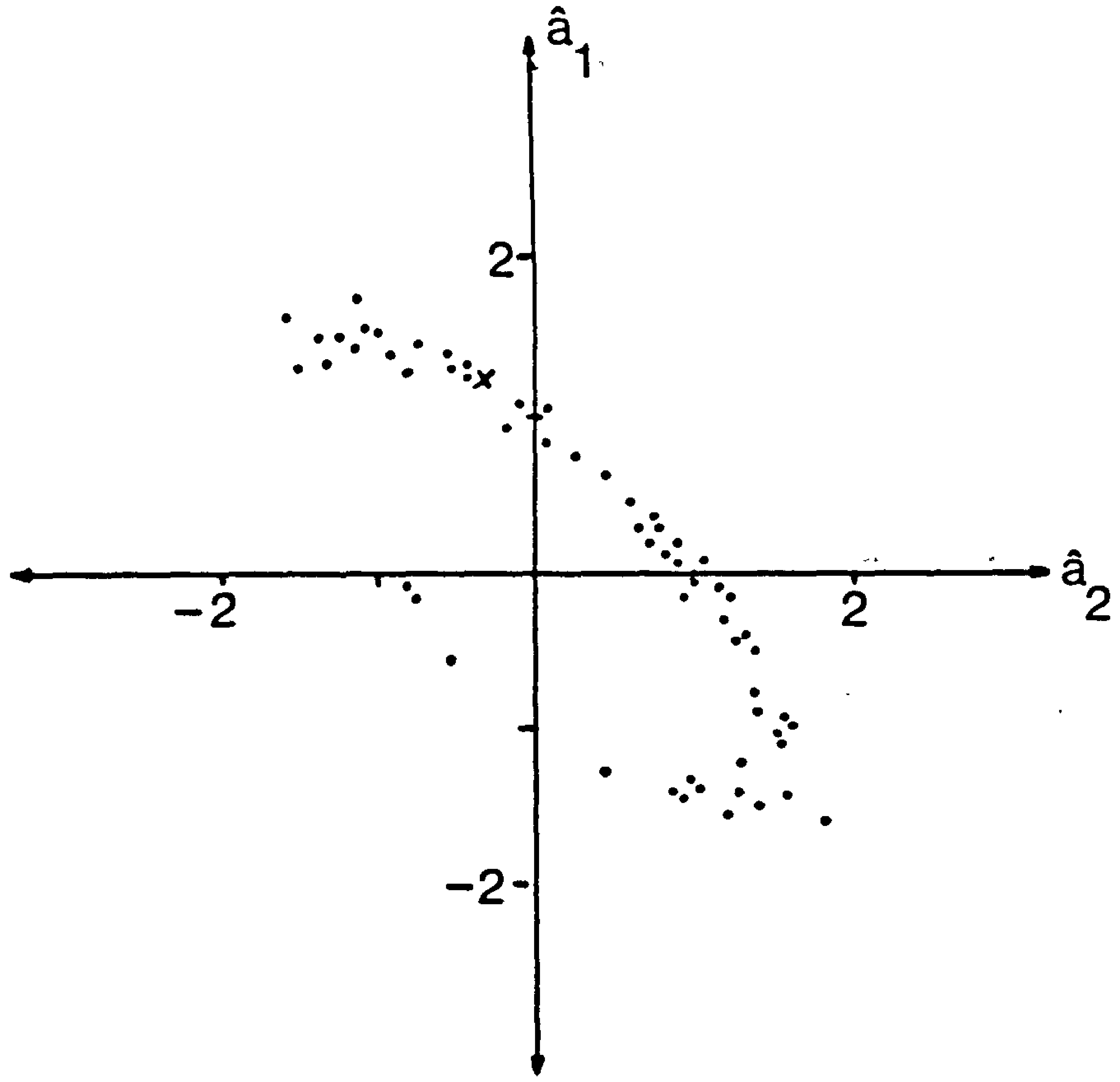

Fig 6.1 Scatter plot of 60 estimates of $\hat{a}_1$ against $\hat{a}_2$ with $\alpha = 0.75$

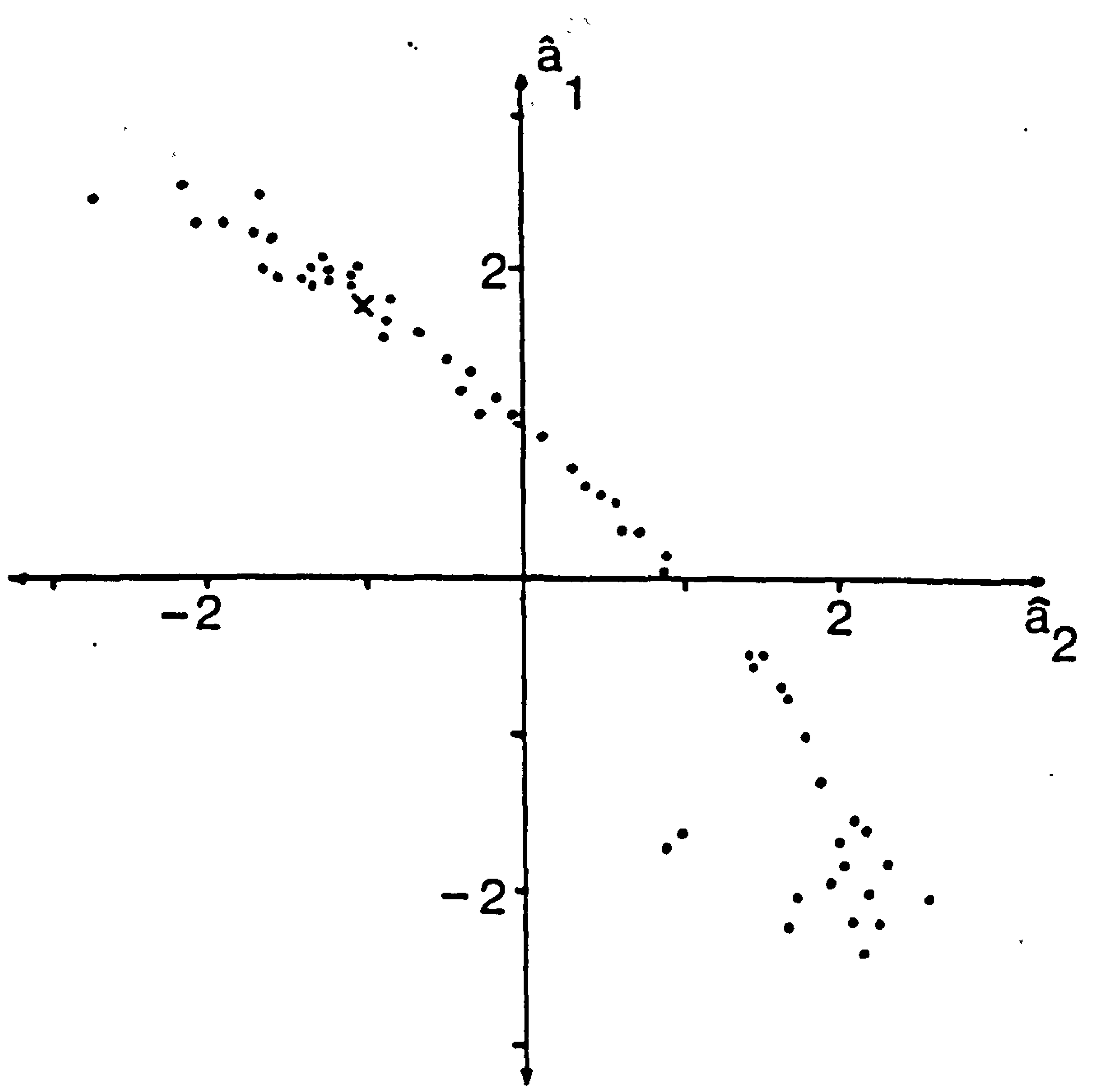

Fig 6.2 Scatter plot of 60 estimates of $\hat{a}_1$ against $\hat{a}_2$ with $\alpha = 0.90$

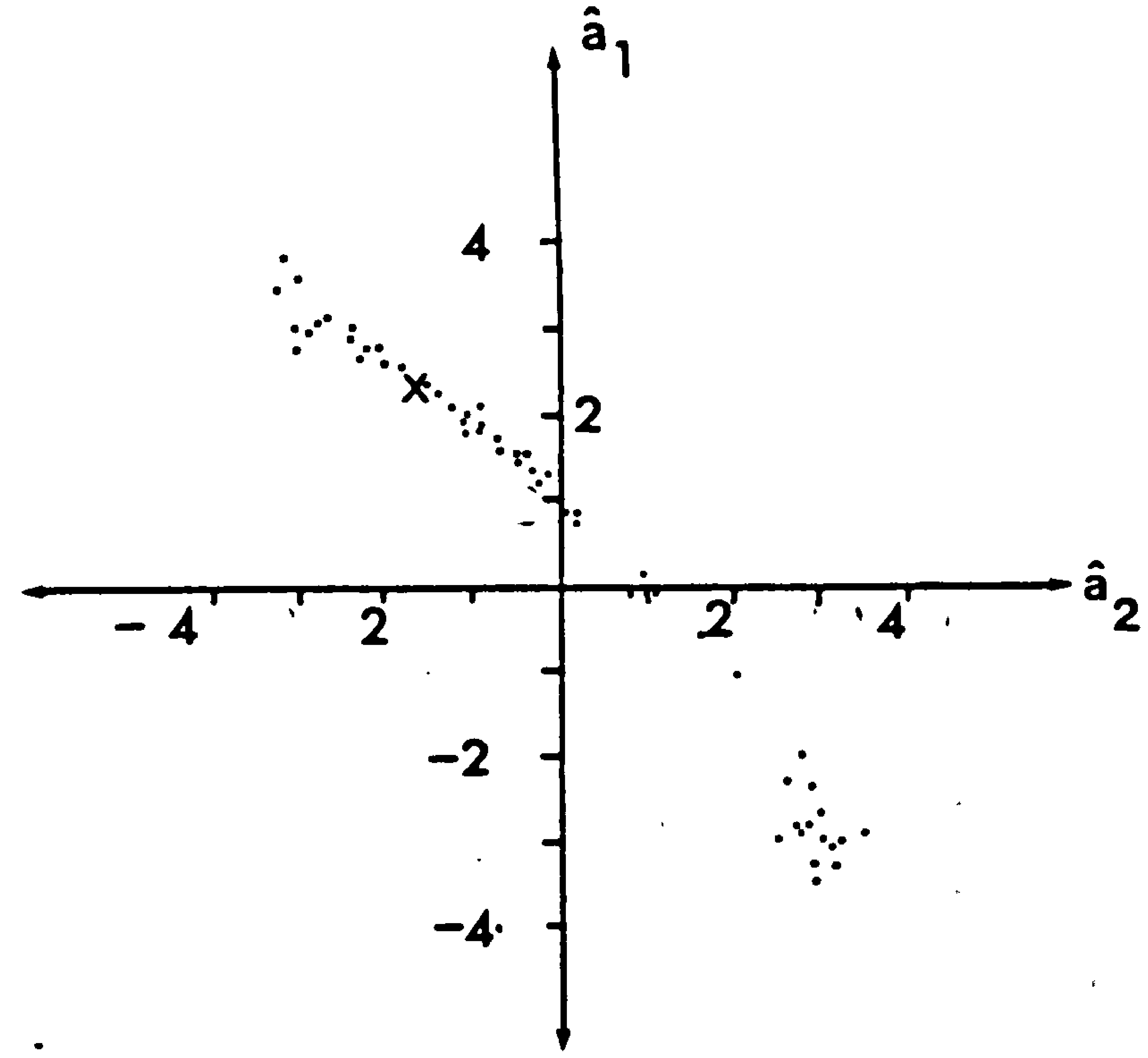

Fig 6.3 Scatter plot of 60 estimates of $\hat{a}_1$ against $\hat{a}_2$ with $\alpha = 0.95$

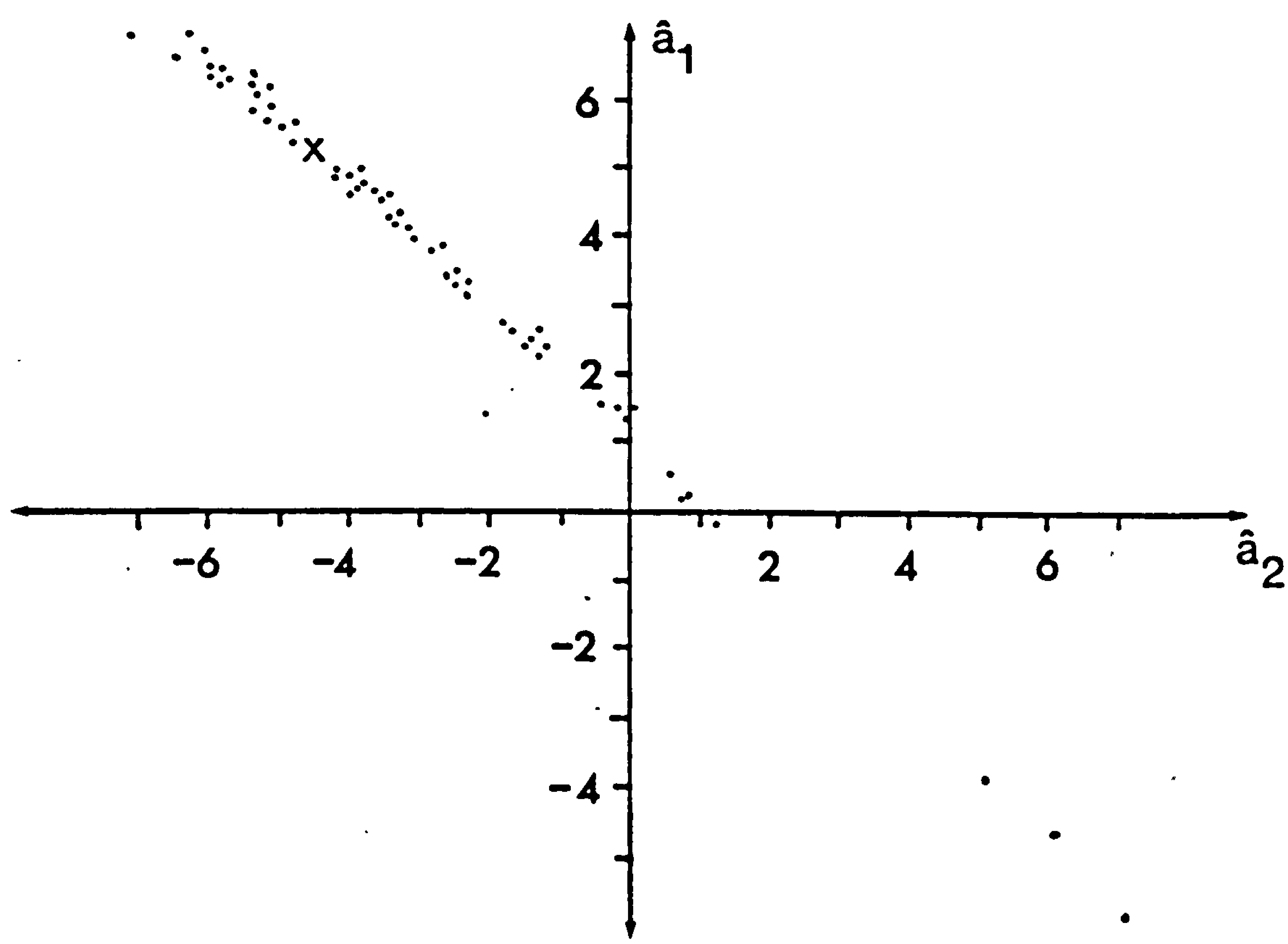

Fig 6.4 Scatter plot of 60 estimates of $\hat{a}_1$ against $\hat{a}_2$ with $\alpha = 0.99$

$\hat{\underline{b}}$ vectors distributed on hyperellipses of varying degrees of eccentricity. The more highly correlated the data, the more eccentric the ellipses. In the cases of extreme collinearity one obtains almost perfect negative correlation between some of the individual components.

6.7 Stabilization of the $\hat{a}$ and $\hat{b}$ vectors

In the empirical study of section 6.5 we report highly variable $\hat{\underline{a}}$ and $\hat{\underline{b}}$ vectors. In the simulation study of section 6.6 we discover that even when the underlying population vectors, $\underline{a}$ and $\underline{b}$, are constant the $\hat{\underline{a}}$ and $\hat{\underline{b}}$ estimates exhibit a very high degree of instability. It is possible then, that in our commodity time series the true values of $\underline{a}$ and $\underline{b}$ are constant but the estimation procedure is giving spurious results?

In the desire to estimate the true values of $\underline{a}$ and $\underline{b}$ in this section we investigate two methods that attempt to produce more stable estimates.

6.7.1 A Canonical Ridge Model

Recall that in the computation of $\hat{\lambda}$, $\hat{\underline{a}}$ and $\hat{\underline{b}}$, the inverses of the matrices S_{11} and S_{22} are used. Vinod (1976) points out that when the variables are highly correlated these matrices tend to be ill-conditioned and their inverses unreliable and very sensitive to small changes in the data. Similar problems of multicollinearity in classical ordinary least squares regression are often encountered. Hoerl and Kennard (1970), in a now famous article, developed a technique known as ridge regression, in which a small arbitrary constant k is added to the diagonals of the near singular (X^TX) matrix. Hoerl and Kennard proved the existence of a ridge estimate that has a lower mean squared error and is "shorter"[3]. The problem still remains in finding that special value of k that will yield the optimal estimate. Hoerl and Kennard recommend plotting the value of each estimate against k for k = 0 to 1.0 and selecting that k which yields "meaningful" values. The literature on ridge regression is extensive.

Vinod (1976) extends the idea of ridge regression to canonical correlation analysis in which we have two near singular matrices; S_{11} and S_{22}. The recommended procedure is to add small positive constants, k_1 and k_2 to the diagonals of S_{11} and S_{22} before inversion. The sample problem thus becomes :

subject to the constraints

$$\underline{a}^{*T} (S_{11} + k_1 I)\underline{a}^* = \underline{b}^{*T} (S_{22} + k_2 I)\underline{b}^* = 1$$

maximize $\lambda^* = \underline{a}^{*T} S_{12} \underline{b}^*$

The ridge estimates, λ^* and $\underline{a}^*$ are obtained by finding the eigenvalues and eigenvectors of

$$A^* = (S_{11} + k_1 I)^{-1} S_{12} (S_{22} + k_2 I)^{-1} S_{21}$$

and $\underline{b}^* = (S_{11} + k_1 I)^{-1} S_{12} \underline{a}^*/ \lambda^*$

It can be shown that the canonical ridge estimates λ^* are less than the classical estimates $\hat{\lambda}$, and that the lengths of the $\underline{a}^*$, $\underline{b}^*$ vectors are less than $\hat{\underline{a}}$ and $\hat{\underline{b}}$. No formal proof of an existence theorem for λ^*, $\underline{a}^*$ and $\underline{b}^*$ has yet been derived.

As Vinod points out it should be possible to plot a ridge trace of the components of $\underline{a}^*$ and $\underline{b}^*$ for various combinations of k_1 and k_2. Instead of using a ridge trace however, Vinod considered three heuristic indicators to guide the user to the optimal combination of k_1 and k_2. These indicators are averages and average sums of squares of estimates[4] and Vinod notes that the addition of values as small as 0.05 stabilize markedly the estimates of $\underline{a}$ and $\underline{b}$.

6.7.2 A Normalized Vector Model

One major source of difficulty in identifying any pattern may be that the lengths of the $\hat{\underline{a}}$ and $\hat{\underline{b}}$ vectors vary considerably from period to period. One way of removing this source of confusion would be to "normalize" the vectors $\hat{\underline{a}}$ and $\hat{\underline{b}}$ by multiplying all components by $1/\hat{\underline{a}}^T\hat{\underline{a}}$ and $1/\hat{\underline{b}}^T\hat{\underline{b}}$ respectively.

It is easy to show that normalizing the vectors in this way does not alter the canonical corelations, $\sqrt{\lambda}$. Only the variance of the resulting canonical variates : $\hat{\underline{a}}^T\underline{x}$ and $\hat{\underline{b}}^T\underline{x}$ are affected. For the moment we are not interested in the canonical variates. In this way then it should be possible to compare the relative magnitudes of each of the elements of $\underline{a}$ and $\underline{b}$ for each period of estimation. For convenience and practical reasons that become obvious in Chapter 7 we will actually normalize the vectors $\hat{\underline{a}}$ and $\hat{\underline{b}}$ by multiplying by $1/\sum_{i=1}^{6}|\hat{a}_i|$ and $1/\sum_{i=1}^{6}|\hat{b}_i|$ respectively.

This alternative approach to the problem of unstable estimates is considerably easier to carry out than the ridge technique. In the following section we report on a simulation study comparing the normalization and ridge techniques.

6.7.3 Simulation study of Ridge verses Normalization approach to canonical coefficient analysis

In section 6.7.2 we note that both in the empirical study of 6.5 and the simulation study of section 6.6.2 the lengths of the $\hat{\underline{a}}$ and $\hat{\underline{b}}$ vectors varied considerably. Since interest is mainly on the relative magnitudes and signs of each of the vectors $\underline{a}$ and $\underline{b}$ we will henceforth consider only normalized vectors (both sample estimates and population values).

Below we outline the simulation experiment.

1. For convenience and continuity we set the population variance/covariance structure to that described in section 6.6.2. The same four values of α were chosen.
2. The constants k_1 and k_2. Vinod tried adding various values of k_1 to S_{11} and k_2 to S_{22}. For brevity we report here the reults of adding the same constant $k = k_1 = k_2$ to both S_{11} and S_{22}. Also for brevity we only report the results for k = 0 (simple normalized classical procedure), k = 0.001, 0.01, 0.05 and 0.1. These were some of the values suggested by Vinod.
3. Samples of size n = 50 were generated and for each value of k, the ridge estimates were calculated using the expressions given in section 6.7.1.
4. The $\underline{a}^*$ and $\underline{b}^*$ estimates were normalized so that

$$\sum_{i=1}^{6} |a^*_i| = \sum_{i=1}^{6} |b^*_i| = 1$$

5. The deviation of each component of $\underline{a}^*$ and $\underline{b}^*$ from the appropriate

normalized population component was computed.

6. This process was repeated for $m = 500$ samples. Different sets of 500 samples were generated for each value of α and for each value of k.

7. The estimated bias and root mean square error of these ridge-normalized estimates are given in Tables 6.4 a-d. These statistics were calculated using identical expressions to the ones given in section 6.7.2 but with normalized vectors in place of unnormalized vectors. Also given in Tables 6.4 a-d are the average lengths and standard deviations of the unnormalized vectors.

6.7.4 Conclusions on study of Ridge vs. Normalization techniques

Consider Tables 6.4 a-d. First we note that, as expected, the lengths and standard deviations of the raw un-normalized vectors are reduced by the addition of a constant, k, to the diagonals of S_{11} and S_{12}. This reduction is much more marked in the very highly collinear populations when $\alpha = 0.99$.

Secondly, note that in all cases the classical estimates of λ (in which $k = 0$) are positively biased. The addition of small k values reduced this bias. The effect of k on the RMSE of the $\hat{\lambda}$ estimates is different for different values of α. In the situations in which the population value of λ is small (i.e. when $\alpha = 0.75, 0.90, 0.95$) there is a marked reduction in the RMSE with the addition of increasingly large k values. In the situation in which λ is large ($\alpha = 0.99$), the RMSE of $\hat{\lambda}$ decreases initially then increases rapidly.

Thirdly consider the estimates of $\underline{a}$ and $\underline{b}$. The picture is a very complex one. We consider first the bias of the $\underline{a}^*$ vector. The a^*_1 component is always negatively biased and the a^*_2 component always positively biased. With all values of α, adding increasingly

Table 6.4a

Simulation study of Ridge verses Normalization approach to canonical correlation analysis

Population values: $\alpha = 0.750$ $\lambda = 0.0459$ $\underline{a}^T = \underline{b}^T = (0.773, -0.227)$

	k	.000	.001	.01	.05	0.10
	λ	.071	.073	.065	.048	.035
	a_1	-.244	-.261	-.265	-.245	-.233
Biases	a_2	.159	.173	.184	.198	.211
	b_1	-.264	-.283	-.260	-.245	-.209
	b_2	.212	.197	.156	.171	.227
	λ	.100	.100	.092	.075	.060
	a_1	.410	.417	.430	.421	.400
RMSE	a_2	.499	.517	.528	.518	.530
	b_1	.424	.447	.426	.413	.385
	b_2	.534	.537	.500	.512	.523
	L	1.845	1.880	1.816	1.647	1.545
	S(L)	1.978	2.018	1.939	1.759	1.636

L = average length of vector, S(L) = standard deviation of length of vector

Table 6.4b

Population values: $\alpha = 0.900$ $\lambda = 0.083$ $\underline{a}^T = \underline{b}^T = (0.644, -0.356)$

	k	.000	.001	.01	.05	0.10
	λ	.076	.076	.056	.026	.000
	a_1	-.174	-.145	-.143	-.133	-.083
Biases	a_2	.249	.223	.222	.307	.273
	b_1	-.176	-.150	-.146	-.109	-.085
	b_2	.235	.231	.252	.264	.317
	λ	.106	.108	.089	.061	.0460
	a_1	.382	.364	.357	.355	.310
RMSE	a_2	.538	.506	.503	.579	.534
	b_1	.387	.351	.355	.301	.321
	b_2	.533	.519	.538	.539	.560
	L	2.812	2.758	2.705	2.099	1.856
	S(L)	3.091	3.028	2.947	2.297	2.018

Table 6.4c

Population values: α = 0 .950 λ = 0.141 $\underline{a}^T = \underline{b}^T$ = (0.594, -0.406)

	k	.000	.001	.01	.05	0.10
	λ	.082	.071	.040	.025	.054
	a_1	-.124	-.103	-.102	-.042	-.001
Biases	a_2	.210	.212	.192	.275	.288
	b_1	-.127	-.107	-.103	-.031	-.004
	b_2	.222	.190	.249	.239	.268
	λ	.118	.112	.087	.061	.071
	a_1	.361	.336	.323	.293	.258
RMSE	a_2	.483	.476	.466	.513	.509
	b_1	.344	.318	.334	.272	.282
	b_2	.504	.463	.518	.482	.486
	L	3.934	3.809	3.634	2.528	2.012
	S(L)	4.324	4.222	9.968	2.801	2.209

Table 6.4d

Population values: α = 0.990 λ = .579 $\underline{a}^T = \underline{b}^T$ = (0.538, -0.462)

	k	.000	.001	.01	.05	0.10
	λ	.0430	.002	.232	.434	.485
	a_1	-.060	-.052	-.048	-.005	-.066
Biases	a_2	.123	.124	.158	.225	.242
	b_1	-.044	-.061	-.017	-.002	-.048
	b_2	.083	.126	.140	.226	.273
	λ	.090	.074	.239	.436	.486
	a_1	.253	.250	.274	.275	.257
RMSE	a_2	.362	.357	.392	.444	.425
	b_1	.227	.265	.236	.276	.283
	b_2	.290	.359	.351	.446	.468
	L	8.636	8.214	5.987	3.301	2.359
	S(L)	9.348	8.913	6.490	3.582	2.540

larger values of k has the effect of reducing the bias on one element (a_1) and increasing the bias on the other element (a_2). These biases are however smaller for larger values of α at all values of k. Exactly similar effects are noted with the $\underline{b}^*$ vector.

Finally we consider the RMSE of the $\underline{a}^*$ and $\underline{b}^*$ vectors. In the cases in which λ is very small (when α = .75 and .90) the additions of k seems to have very little effect on the RMSE of the $\underline{a}^*$ and $\underline{b}^*$ vectors. In the case of α = 0.95 adding k reduces the RMSE of one component of $\underline{a}^*$and $\underline{b}^*$ but increases the RMSE of the other component. With α = 0.99 the RMSE of nearly all components are increased by adding k values.

Many other population variance/covariance structures were used in other simulations with broadly similar results. In populations with large canonical correlations (as witnessed in our empirical study), the biases and RMSE's of the classical (k = 0) normalized $\underline{a}$ and $\underline{b}$ estimates were relatively small and there appeared to be no consistent benefit in using ridge type estimators.

Accordingly in all that follows we report only the classical normalized vectors: $\hat{\underline{a}}\,\hat{\underline{b}}$.

6.8 The normalized canonical coefficients for commodity series

The normalized canonical coefficient estimates associated with the largest canonical correlation have been computed for each period and for each commodity series. For ease of interpretation and clarity of presentation these coefficients have been multiplied by 1000 so the lengths of all vectors are 1000. These values appear in Tables 6.5 to 6.8.

Table 6.5

Normalized estimates of a ,b components for cocoa returns

Period	Vector	Components 1	2	3	4	5	6	sum of comps.
1	A	10	38	-217	397	-274	62	14.5
	B	103	-159	346	-260	34	-95	33.9
2	A	41	-98	172	244	-383	60	32.7
	B	-93	90	-230	-138	288	159	73.8
3	A	12	-115	362	-228	-184	97	-58.6
	B	3	142	-374	192	166	-122	3.8
4	A	-60	-67	361	-235	102	-171	-73.7
	B	-79	145	-185	274	-236	78	-6.1
5	A	-37	-194	153	-247	358	7	36.1
	B	-84	222	-88	251	-349	4	-47.9
6	A	92	-152	50	-102	362	-239	7.5
	B	2	88	-161	12	-335	401	3.4
7	A	10	-93	-201	496	-166	-29	13.5
	B	-24	-59	431	-365	-73	44	-50.2
8	A	-36	157	-290	-13	338	-161	-9.3
	B	-75	58	302	-137	-274	150	20.3
9	A	-27	194	-191	262	-278	46	1.9
	B	-21	144	59	-298	297	-178	0.9
10	A	-55	310	-262	-6	198	-165	15.5
	B	-68	-90	208	-47	-288	295	5.8
11	A	-20	247	-370	77	193	-90	34.6
	B	30	-69	-31	-311	-68	487	33.8
12	A	-78	160	226	-363	-75	95	-36.5
	B	37	-131	-222	443	-137	-26	-41.8
13	A	57	136	-166	273	-337	-28	-67.9
	B	43	-289	310	-202	86	68	14.4
14	A	-19	-57	510	-227	-139	-43	20.1
	B	-32	240	-445	196	71	-13	15.3
15	A	-20	148	39	-362	-108	319	12.4
	B	32	-116	-57	412	44	-336	-24.6
16	A	-126	171	-104	19	-274	302	-15.6
	B	-40	100	-156	-13	386	-301	-28.6
17	A	39	7	-359	-33	476	-84	42.6
	B	-150	29	185	181	-357	96	-18.5
18	A	3	-212	227	286	-212	-57	32.5
	B	346	101	-210	-125	-175	38	-28.3
19	A	4	113	-274	210	-225	171	-3.1
	B	45	-182	282	-242	168	-78	-10.1
20	A	48	-65	217	-335	238	-94	5.5
	B	-8	107	-165	317	-330	70	-10.8
21	A	-24	210	-163	-171	303	-126	24.6
	B	88	-161	-23	389	-312	24	2.6
22	A	57	-324	454	-44	-76	-42	22.6
	B	11	238	-514	24	161	51	-29.1
23	A	4	-34	-161	258	227	-314	-24.2
	B	-107	290	-36	-388	63	114	-65.9
24	A	22	307	-260	-107	-133	169	-5.8
	B	23	-208	359	-222	123	-61	11.0

Table 6.6

Normalized estimates of a, b components for coffee returns

Period	Vector	Components 1	2	3	4	5	6	sum of comps.
1	A	-71	118	162	-191	204	-251	-32.4
	B	-6	-13	-390	383	-77	127	18.2
2	A	-5	132	-450	353	26	-32	20.9
	B	25	-101	381	-381	44	66	32.4
3	A	221	-214	204	-27	-270	61	-27.0
	B	-184	238	144	-177	-137	117	-1.7
4	A	67	-118	-382	406	16	-8	-21.2
	B	61	-233	318	-108	-157	121	-1.3
5	A	-9	133	-234	82	295	-243	21.1
	B	15	30	43	84	-496	330	6.4
6	A	36	-35	-2	-227	456	-240	-15.7
	B	46	-126	-95	339	-285	108	-15.9
7	A	-9	67	-215	232	-275	199	-4.7
	B	-29	-35	205	-254	282	-191	-26.1
8	A	62	-115	39	81	320	-381	5.2
	B	-164	176	-325	222	59	53	18.7
9	A	-46	-27	80	65	339	-440	-31.1
	B	-33	307	-158	-85	-220	192	-1.3
10	A	23	-174	200	-126	-195	281	5.9
	B	-16	175	-390	310	19	-87	6.8
11	A	-37	-356	322	-81	-38	162	-32.7
	B	-43	415	-297	72	-3	-166	-26.5
12	A	98	-335	227	170	-128	-39	-10.1
	B	-88	336	-202	-198	128	45	18.7
13	A	-14	-141	34	91	371	-346	-6.7
	B	-44	169	4	-101	-372	307	-39.0
14	A	-76	-12	235	290	-312	-72	49.4
	B	-11	29	-41	-449	151	316	-7.1
15	A	64	-20	50	-54	-400	409	45.8
	B	19	84	-259	-9	405	-221	15.5
16	A	47	121	13	87	-493	238	12.8
	B	-91	249	36	-244	245	-132	60.6
17	A	-120	494	-80	-205	-70	26	39.7
	B	37	328	-309	106	58	-161	56.4
18	A	-99	110	33	-359	357	-39	0.9
	B	-164	194	-19	249	-316	56	-3.5
19	A	-34	135	-330	120	263	-114	36.6
	B	11	-88	540	-155	-145	-57	101.9
20	A	33	246	-392	224	-43	-60	4.3
	B	-99	-109	407	-299	81	-1	-24.0
21	A	-2	-290	271	-99	201	-133	-54.2
	B	-79	416	-441	4	26	31	-45.3
22	A	29	152	-306	-176	308	27	32.0
	B	-22	-147	111	388	-206	-122	-3.0
23	A	3	-114	247	-146	-233	253	7.0
	B	-2	156	-370	150	177	-142	-34.5
24	A	48	-117	85	254	-371	124	20.2
	B	-167	216	127	-323	157	-7	0.4
25	A	7	-183	22	412	-315	59	-1.7
	B	-56	199	-11	-327	296	-107	-10.7
26	A	-26	75	253	-274	-211	158	-28.1
	B	32	-85	-363	417	-25	75	48.0
27	A	219	-204	112	81	-265	116	56.7
	B	-102	175	-58	-221	328	-111	5.7
28	A	42	-49	-64	-110	423	-308	-69.9
	B	130	-188	-97	192	-251	138	-77.9
29	A	-5	86	145	-365	256	-139	-23.8
	B	55	-204	-47	335	-221	135	49.8

Table 6.7

Normalized estimates of a , b components for sugar returns

Period	Vector	Components 1	2	3	4	5	6	sum of comps.
1	A	79	-148	151	-101	-231	286	33.6
	B	-65	-231	357	-58	131	-153	-25.0
2	A	93	-243	328	-35	75	-223	-7.9
	B	-59	370	-158	-274	62	73	10.7
3	A	0	107	153	-358	-140	240	0.2
	B	83	-188	74	338	-64	-250	-10.4
4	A	127	34	-191	-302	107	238	9.7
	B	-2	86	-53	364	-469	23	-53.1
5	A	-58	43	-164	286	173	-273	4.8
	B	28	-131	152	-25	-334	327	14.4
6	A	-85	82	-352	391	-40	46	39.7
	B	56	-220	284	-147	164	-125	9.0
7	A	-77	-175	132	173	180	-260	-30.2
	B	68	-203	79	22	-257	368	74.3
8	A	-17	-74	413	-408	75	11	-2.4
	B	-39	32	-274	430	-189	34	-8.3
9	A	14	-35	-256	494	-183	15	44.9
	B	29	-176	196	-106	-217	274	-1.8
10	A	1	76	33	-327	395	-166	9.1
	B	-15	177	-73	62	-407	262	2.5
11	A	-12	127	-320	378	-32	-127	11.5
	B	88	-217	266	-137	-129	161	30.5
12	A	74	-23	74	-355	356	-116	7.7
	B	58	-125	78	349	-369	20	.8.7
13	A	55	-25	152	-291	304	-170	23.1
	B	60	6	-178	270	-320	165	0.2
14	A	-77	-1	-84	269	241	-324	20.9
	B	-7	-119	147	-11	-344	368	28.8
15	A	-85	132	-133	30	-254	363	49.3
	B	55	-120	156	-55	285	-326	-8.5
16	A	32	-135	-4	72	-366	388	-14.6
	B	-20	220	-366	-50	288	-53	14.9
17	A	-135	100	-22	-345	301	94	-10.5
	B	204	-363	3	291	-44	-91	-3.4
18	A	11	-96	318	-369	163	-40	-16.1
	B	0	-35	-136	410	-326	89	-3.4
19	A	6	-113	179	120	-386	195	-1.1
	B	-10	32	-87	-231	500	-135	65.1
20	A	-18	78	-370	289	116	-125	-32.6
	B	63	-160	256	-261	-76	182	1.1
21	A	3	-84	167	254	-436	54	-43.7
	B	-21	281	-436	105	136	19	82.1
22	A	-46	162	-116	-341	215	117	-12.2
	B	16	-195	-14	468	-127	-176	-31.6
23	A	30	-137	-1	109	344	-376	-33.8
	B	68	474	-157	-47	-192	-59	82.9

Table 6.8

Normalized estimates of a , b components for rubber returns

Period	Vector	Components 1	2	3	4	5	6	sum of comps.
1	A	-56	35	-250	477	-137	-41	24.1
	B	19	72	134	-493	172	108	11.1
2	A	103	-95	-55	-14	403	-326	11.3
	B	-43	128	-153	168	-265	239	70.7
3	A	-45	-55	200	50	-380	267	34.6
	B	10	32	-184	25	448	-298	31.4
4	A	40	-10	241	166	-497	45	-16.9
	B	-183	62	-291	197	141	124	46.8
5	A	-43	144	-224	300	-232	54	-4.2
	B	-19	46	137	-319	311	-164	-10.7
6	A	27	131	107	-486	110	137	25.2
	B	96	13	-244	354	-249	41	10.8
7	A	47	85	-335	261	121	-149	28.7
	B	-46	-74	413	-225	-140	98	21.2
8	A	-19	3	-28	393	-496	58	-93.4
	B	92	-161	-62	-278	389	14	-10.0
9	A	-61	181	-124	-330	235	67	-35.1
	B	39	-235	210	262	-120	-131	22.8
10	A	45	-203	26	-77	439	-206	21.6
	B	-13	-30	242	-23	-440	248	-21.1
11	A	37	-116	250	-371	218	-4	11.7
	B	65	69	-407	373	-2	-82	13.4
12	A	-80	54	-70	256	-332	205	28.8
	B	10	-96	207	-212	272	-199	-21.0
13	A	-21	-2	-101	145	355	-372	0.1
	B	16	33	75	-47	-433	394	35.2
14	A	-33	7	196	-60	-414	286	-21.8
	B	-20	105	-229	143	243	-257	-18.0
15	A	-30	36	117	216	-468	131	-0.6
	B	16	47	-213	-133	436	-151	-0.3
16	A	-10	-203	74	217	-282	211	5.2
	B	-1	153	-140	-317	344	-42	-6.3
17	A	-20	12	16	-174	468	-308	-9.6
	B	11	-44	50	260	-455	177	-2.3
18	A	9	-60	83	-118	398	-329	-21.6
	B	-5	-47	64	26	-441	414	8.6
19	A	-9	60	-297	428	-22	-180	-23.5
	B	11	66	383	-392	6	-141	-69.1
20	A	-29	-286	300	-7	166	-208	-67.7
	B	-42	-75	342	39	136	-363	33.9

6.8.1 Remark on canonical coefficient estimates

Referring to Tables 6.5 to 6.8, we note that there is still a tremendous variation from period to period in the individual componenets of $\hat{\underline{a}}$ and $\hat{\underline{b}}$ vectors. In all there are 96 periods x 2 vectors x 6 components = 1152 individual estimates to consider and although the picture is a very complex one some general patterns can be observed.

(i) The sums of the components of each vector in each period are very near zero. The positive components seem to 'cancel out' the negative components.

(ii) The values of $|\hat{a}_i|$ and $|\hat{b}_i|$ range from 1 to 540.

(iii) Considering the absolute values of $\hat{a}_i$ and $\hat{b}_i$ only, we note that the first component is usually the smallest.

(iv) The $\hat{\underline{a}}$ component with the largest absolute value tends to be 'further down' the vector : the 3rd, 4th or 5th element. In each period one of the six absolute values, $\hat{a}_1, \ldots\ldots, \hat{a}_6$ is maximal. Table 6.9 gives a record of the frequency of each component containing the maximum value. Consider for example the rubber series. We see that in 11 of the 20 periods the maximum component of $\hat{\underline{a}}$ was the 5th element : a_5.

Note that the frequency distributions are not uniform. The first component is maximal only once. For cocoa, coffee and rubber the most frequent maximal component is the fifth, for sugar it is the fourth. We can make similar remarks relating to the $\underline{b}$ vector.

Recall that the $\underline{a}$ and $\underline{b}$ vectors give us the canonical variates : $\underline{a}^T\underline{x}_t$ and $\underline{b}^T\underline{x}_{t-1}$ that are maximally correlated. We can see now that the contributions to these variates from the individual returns in

Table 6.9

Frequency of maximum absolute components within $\hat{a}$, $\hat{b}$ vectors (obtained from Tables 6.5 - 6.8)

	no. of periods	$\hat{a}$ vector						$\hat{b}$ vecter						no. of matched pairs	no. of matched pairs with opposite sign
Component		1	2	3	4	5	6	1	2	3	4	5	6		
Cocoa	24	0	2	6	6	8	2	1	0	9	6	5	3	9	9
Coffee	29	0	4	3	6	11	5	0	5	8	9	7	2	14	13
Suger	23	0	0	3	9	5	6	0	3	5	6	5	4	11	11
Rubber	20	0	0	2	6	11	1	0	0	2	6	10	2	16	16

Table 6.10

Frequnecy of $\hat{a}$,$\hat{b}$ components with opposite signs (obtained from Tables 6.5 - 6.8)

Components:	1	2	3	4	5	6	Expected Freq.	No. of Periods
Cocoa	8	19	20	21	21	21	12.0	24
Coffee	17	22	22	25	24	26	14.5	29
Suger	16	15	15	19	19	18	11.5	23
Rubber	14	11	14	17	19	14	10.0	20

$\underline{x}_t$ and $\underline{x}_{t-1}$ are not uniformly spread. It appears that the 3rd, 4th and 5th components of the 6 dimensional returns vectors are mostly responsible for the observed multivariate dependence. We return to this point again later.

(v) Still on the subject of the component with the largest value it is interesting to note the 'pairing' of such components in the $\hat{\underline{a}}$ and $\hat{\underline{b}}$ vectors. Table 6.9 gives the number of periods in which the maximum absolute value of the elements of $\hat{\underline{a}}$ and $\hat{\underline{b}}$ vectors occur in the same component. Also reported in Table 6.9 is the number of these pairs in which the signs of the relevent maximum a_i and b_i are opposite. In the case of rubber for example in 16 of the 20 periods when a particular component of $\hat{\underline{a}}$ is maximal, the corresponding component of $\hat{\underline{b}}$ is also maximal. In all 16 cases the associated $\hat{a}_i$, $\hat{b}_i$ pairs had opposite signs. Reference to Table 6.9 shows that for the other three series roughly half of the periods resulted in this pairing of maximal components and in all but 1 case the signs of the relevent $\hat{a}_i$, $\hat{b}_i$ pair were opposite.

(vi) Following on from (v) we constructed Table 6.10 which lists the frequencies of occurrences in which the signs of corresponding elements of $\hat{\underline{a}}$ and $\hat{\underline{b}}$ are different for each element, not just the maximum ones. Also given in Table 6.10 is the expected frequency of these occurrences if the signs of corresponding elements of $\hat{\underline{a}}$ and $\hat{\underline{b}}$ were assumed to be random.

Consider as an example the 20 rubber periods. We would expect, under a hypothesis of random signs that half (i.e. 10) of the periods would result in corresponding elements of $\hat{\underline{a}}$ and $\hat{\underline{b}}$ to be different. However we see that in 19 periods the signs of the 5th elements $\hat{a}_5$ and $\hat{b}_5$ were different. Table 6.10 shows that in all but one of the 24 situations considered, the observed frequencies of opposite signs far exceeds that expected[5]. We note also that the frequency of different $\hat{a}_i$, $\hat{b}_i$ signs is lowest in the case of the 1st componenet.

(vii) Finally, as originally mentioned in section 6.5, we notice that the sequence of the signs of the elements within a given vector do not appear to be random. If one considers the elements of $\underline{\hat{a}}$ (say) in the sequence $\hat{a}_1,..,\hat{a}_6$ there appears to be a predominance of sign switching. For example in the situations in which $\hat{a}_2$ (say) is positive it is more likely that $\hat{a}_1$ and $\hat{a}_3$ will be negative than positive. This phenomenon is neatly summed up in Table 6.11. In this table the frequency of contiguous sign combinations is given. Consider as an example the rubber series. In each period we have 6 elements of the $\underline{\hat{a}}$ vector resulting in 5 contiguous + and or - sign pairs. With 20 periods we have 20 x 5 = 100 pairs of signs. If the signs were random then one would expect half (i.e. 50) of the cases to result in a + following a - or a - following a +. Similarly we would expect half (i.e. 50) of the cases to result in +/+ or -/- combinations. We see from Table 6.11 that the observed frequencies are 25 and 75 respectively. The likelihood of this outcome (if the signs are really random) can be computed using:

$$Z = \frac{\hat{p} - 0.5}{\sqrt{\dfrac{0.5 \times 0.5}{n}}} ,$$

where n = number of pairs of signs (= 100 for rubber) and $\hat{p}$ = proportion of pairs resulting in +/- or -/+ (= 0.75 for rubber). If the signs are random Z is N ~ (0, 1). We see from Table 6.11 that in all cases there is a very significant excess of +/- or -/+ sequences over +/+ or -/- sequences in both the $\underline{\hat{a}}$ and $\underline{\hat{b}}$ vectors.

We sum up then by saying that although the individual elements of the normalized $\underline{\hat{a}}$, $\underline{\hat{b}}$ vectors vary considerably from period to period we can observe some very interesting and unexpected patterns in the estimates.

Table 6.11

Frequency of contiguous sign combinations within $\hat{a}$ and $\hat{b}$ vectors (obtained from Tables 6.5 - 6.8)

		Sign combinations +/+ or -/-	+/- or -/+	Z value
Cocoa	a	33	87	4.93
	b	33	87	4.93
	E	60	60	
Coffee	a	47	98	4.24
	b	43	102	4.90
	E	72.5	72.5	
Sugar	a	30	85	5.13
	b	28	87	5.50
	E	57.5	57.5	
Rubber	a	25	75	5.00
	b	38	62	2.40
	E	50	50	

E = expected frequency under null hypothesis of random sign allocation

Z should be N(0,1) distributed under the null hypothesis of random sign distribution

The larger (both positive and negative) elements tend to be in the 3rd, 4th and 5th components and the smallest elements are invariably in the 1st component. The sums of the components are all very near zero.

One can possibly best sum up the typical (most frequent) configuration of the signs of the $\hat{\underline{a}}$ and $\hat{\underline{b}}$ estimates in the following

$\hat{\underline{a}}^T \longrightarrow$	$+ a_1$	$- a_2$	$+ a_3$	$- a_4$	$+ a_5$	$- a_6$
$\hat{\underline{b}}^T \longrightarrow$	$- b_1$	$+ b_2$	$- b_3$	$+ b_4$	$- b_5$	$+ b_6$

Adjacent elements within $\hat{\underline{a}}$ or $\hat{\underline{b}}$ tend to have opposite signs as do corresponding elements in $\hat{\underline{a}}$ and $\hat{\underline{b}}$.

What is also particularly interesting is that the above properties are exhibited by <u>all</u> of the commodity series. Whatever phenomenon is giving rise to these results must be common to all 4 futures markets. Before considering what phenomenon in the futures markets could be giving rise to these results we review the possibility of arriving at single, 'grand average' $\hat{\underline{a}}$ and $\hat{\underline{b}}$ estimates.

6.9 An average canonical coefficient vector

We see from section 6.8 that although the normalized $\hat{\underline{a}}$ and $\hat{\underline{b}}$ vectors vary from period to period there is a discernable pattern in the distribution of the individual elements. Simulations similar to the ones described in section 6.7.4 but with 6 dimensional vectors have shown that with fixed population vectors $\underline{a}$, $\underline{b}$ and highly collinear data similarly varied $\hat{\underline{a}}$ and $\hat{\underline{b}}$ estimates are obtained. It is possible then that with our commodity series there exist fixed 'population' $\underline{a}$ and $\underline{b}$ vectors, but that the highly collinear nature of the data and the relatively small sample sizes are very likely giving rise to extremely varied estimates. In this section we attempt to estimate these assumed constant population $\underline{a}$ and $\underline{b}$ vectors.

6.9.1 Aggregate variance/covariance matrices

Throughout this and the next sectin we will refer to the coffee series as an example. Recall that there are 29 periods of estimation. Each period contains approximately 42 daily returns on 6 contracts. In total there are 1189 daily returns.

In Table 6.6 there are 29 estimates of λ, $\underline{a}$ and $\underline{b}$. Each estimate is computed from expressions involving the variance/covariance matrix estimates; S_{11}, S_{22} and S_{12}. There are 78 elements making up these 3 matrices. These 78 elements are computed from sums of squares and cross products from the original sample. The original sample is typically 42 observations of 6 returns. Thus we are using 42 x 6 = 252 returns to compute the 78 elements of S_{11}, S_{22} and S_{12} and hence $\hat{\lambda}$ $\hat{\underline{a}}$ and $\hat{\underline{b}}$. With such a low ratio of data to parameter estimates we could not expect the resulting $\hat{\underline{a}}$ and $\hat{\underline{b}}$ vectors to be very accurate. Considering the paucity of data and the fact that it is highly collinear it is not surprising that the observed $\hat{\underline{a}}$, $\hat{\underline{b}}$ vectors are highly varied.

If we assume that there exist fixed constant population $\underline{a}$ and $\underline{b}$ vectors we could obtain much more accurate estimates by using larger samples. One possible approach would be to estimate S_{11}, S_{22} and S_{12} over the entire 5 year period and this could be achieved as follows.

Treat each of the 29 periods as separate samples of data each containing about 42 returns. Although in each period we are examining slightly different contracts we can consider the 6 dimensional set of returns in the manner as we did in section 3.6.2. Each element of the vector will represent the returns on a contract that is always on average some fixed 'distance' from maturity. It would be possible then to estimate each element of S_{11}, S_{22} and S_{12} using all 1189 returns. The 1st element of S_{11} for example would be computed from the sums of squares obtained by considering all 1189 returns on the contract that was always 'nearest' to maturity.

However recall from section 3.3.2 and 4.8.5 that the variances and covariances of returns vary tremendously over the 5 year period. The ratio of maximum variance to minimum variance in the coffee series for example was found to be 900. If we computed the S matrices as suggested above the contribution to each element of S_{11} etc in periods of low variability would be completely swamped by the contributions from periods of high variability.

The discovery of persistent multivariate serial correlation implies that the elements of the V_{12} matrix are not all identically zero. The accurate estimation of the elements of this matrix is a crucial point in the search for the underlying $\underline{a}$ and $\underline{b}$ vectors. The estimates (36 in all) are all measures of association between the vector $\underline{x}_t$ and $\underline{x}_{t-1}$. Unfortunately these estimates will also vary with the underlying variance of the series. What is needed is a procedure that is invariant to the changing variability. An obvious candidate is the correlation matrix R.

6.9.2 Canonical correlation analysis with correlation matrices

Mardia (1979) shows that canonical correlation analysis using the correlation matrix leads to essentially the same results as canonical correlation analysis using the variance/covariance matrix. If V and R are the population variance/covariance and correlation matrices respectively partitioned as in section 6.6.2 it is easy to show that:

$$\lambda = \lambda_R$$

$$\underline{a} = [\mathrm{diag}\ (V_{11})^{-1/2}]\ \underline{a}_R$$

$$\underline{b} = [\mathrm{diag}\ (V_{22})^{-1/2}]\ \underline{b}_R$$

where λ, $\underline{a}$ and $\underline{b}$ are the canonical correlations and vectors of coefficients obtained from the V matrix and λ_R, $\underline{a}_R$ and $\underline{b}_R$ are the corresponding parameters obtained from the R matrix.

So we see that the canonical correlation coefficients using V and R are identical and the coefficient vectors are simple linear transforms of one another.

Note that with the data we have, the study in section 5.2.1 has shown that the elements of diag (V_{11}) and diag (V_{22}) are all very similar within each period[6] but vary markedly from period to period. For the purposes of this study then we will consider the diagonal elements of V_{11} and V_{22} to be identical and so:

$$\underline{a} = c_a.I\ \underline{a}_R$$

$$\underline{b} = c_b.I\ \underline{b}_R$$

where c_a and c_b are constants. Of course we are only interested in normalized vectors and so the constants in the above expression are reduced to unity. Thus the results of canonical correlation analysis of the series we are examining using correlation matrices will yield <u>identical</u> results to an analysis using the variance/covariance matrices.

6.9.3 An aggregate correlation matrix

If we are considering using aggregate correlation matrices to find accurate estimates of a and b we will need to estimate the 36 elements of R_{12}, 15 elements of R_{11} and 15 elements of R_{22}. Each of these 66 estimates will obviously need to be derived from the 29 individual estimates obtained from each period.

Donner and Rosner (1980) addressed the problem of finding common correlation coefficients from $k \geqslant 2$ independent samples. They showed that of the 4 methods they examined (some fairly sophisticated) a standard score method equivalent to finding a weighted average of the simple correlation coefficients proved generally superior. This weighted averaging procedure was shown to be particularly good when the population correlation coefficients were very small.

Donner and Rosner considered the estimation of a single correlation coefficient. In this section we extend their ideas to jointly estimating all 66 correlation coefficients in the R matrix. We believe that the procedure suggested by Donner and Rosner is particularly suited to our problem for two reasons:

(i) We strongly suspect that the elements of R_{12} are very small but not necessarily zero. The accurate estimation of these elements is crucial to the accurate estimation of a and b.

(ii) As noted in section 4.6.2 the author, using many simulations, has found that sample estimates of the elements of correlation matrices can be considered to be virtually independent, provided the parent correlations are not large ($\leqslant$ 0.7). In brief, the weighted average procedure for estimating the crucial R matrix will, have the similar optimal properties noted by Donner and Rosner when studying simple correlation coefficients. Accordingly the pooled estimate of the components $(R)_{ij}$ $(i,j = 1,12)$ of the correlation matrix R were computed using:

$$(R)_{ij} = \frac{\sum_{k=1}^{29} (r_k)_{ij} (n_k - 1)}{\sum_{k=1}^{29} (n_k - 1)} \qquad \ldots\ldots\ldots\ldots\ldots\ldots 7.9.3$$

where n_k = no. of returns in period k,

$(r_k)_{ij}$ = i, j th element of correlation matrix estimated in period k.

The resulting R matrix is partitioned as in section 6.6.2 and the pooled, aggregate or grand average $\hat{\lambda}_R$, $\hat{\underline{a}}_R$, $\hat{\underline{b}}_R$ estimates derived. The results appear in Table 6.12.

6.9.4 The pooled canonical coefficient vectors

We make the following remarks on the values presented in Table 6.12.

The pooled λ values

The pooled canonical correlation coefficients are virtually identical all being approximately 0.450. This is slightly smaller than the estimates obtained for most of the 96 individual periods.

In retrospect this situation of lower sample values of canonical correlations is entirely what one would expect. This phenomenon can be

Table 6.12

Average normalized $\hat{a}$ and $\hat{b}$ estimates from pooled R matrices

Components:		1	2	3	4	5	6	sum of +ve comps	sum of -ve comps	sum of all comps
Cocoa	a	+16	+19	-326	+394	-143	-24	+507	-498	+14
	b	+21	-119	+305	-379	+141	+35	+502	-498	+4
Coffee	a	-34	+10	+117	-308	+370	-161	+497	-503	-6
	b	+20	-29	-101	+357	-370	+123	+500	-500	0
Sugar	a	+2	-56	+147	+11	-443	+341	+501	-499	+2
	b	-14	+33	-143	-7	+460	-343	+493	-507	-14
Rubber	a	+27	-89	+132	-269	+343	-140	+502	-498	+4
	b	-9	+54	-119	+292	-372	+154	+500	-500	0

Average $\hat{\lambda}_{max}$ estimates from pooled R matrices:

Cocoa	Coffee	Sugar	Rubber
0.453	0.463	0.441	0.463

explained by considering initially the pooling of R matrices just over the first two periods. In period one, the estimation of the elements of R_{12}, R_{22} and R_{11} result in one special pair of linear combinations $\hat{\underline{a}}_{(1)}$, $\hat{\underline{b}}_{(1)}$ that yield a maximum canonical correlation $\hat{\lambda}_{(1)}$ between canonical variates $\hat{\psi}_{1(1)} = \hat{\underline{a}}^T_{(1)}\underline{x}_t$ and $\hat{\psi}_{2(1)} = \hat{\underline{b}}^T_{(1)}\underline{x}_{t-1}$. Any pair of linear combinations even slightly different from these will result in a lower correlation coefficient. In period two the process is repeated and we arrive at a new pair of optimal linear combinations $\hat{\underline{a}}_{(2)}$ and $\hat{\underline{b}}_{(2)}$ that result in a maximal correlation, $\hat{\lambda}_{(2)}$. The averaging of the R matrices over these 2 periods will result in new pair of pooled vectors: $\hat{\underline{a}}_{(p)}$, $\hat{\underline{b}}_{(p)}$ that will be somewhere between $\hat{\underline{a}}_{(1)}$, $\hat{\underline{b}}_{(1)}$ and $\hat{\underline{a}}_{(2)}$, $\hat{\underline{b}}_{(2)}$. Since the pooled vectors will not be exactly equal to these found in period one or two, the correlation coefficients between $\hat{\underline{a}}^T_{(p)}\underline{x}$ and $\hat{\underline{b}}^T_{(p)}\underline{x}$ must be slightly less than the maximum correlation found in period one and less than the maximum correlation found in period two. Thus the averaging of R matrices over two periods must result in a lower average maximum canonical coefficient. It is not surprising then that extending the averaging process to over 20 periods or more results in a correlation coefficient as low as 0.450.

Comparing absolute values of $\hat{a}$ and $\hat{b}$ elements

Recall the general pattern observed in the distribution of the absolute values of the elements of all 96 individual $\hat{\underline{a}}$, $\hat{\underline{b}}$ vectors noted in section 6.8.1 .. The distribution of the absolute values of the pooled elements is exactly what we would have predicted. The profiles of the values is as follows:

With the exception of sugar, the smallest elements are in component one. As we consider successive components the elements gradually increase in size reaching a maximum at the 4th or 5th component and then fall off to a medium value in the 6th component. If the elements of appropriate pairs of $\hat{\underline{a}}$ and $\hat{\underline{b}}$ vectors are ranked there is an almost perfect correspondence in ranks. Not only do largest components match but also second largest components etc down to smallest components.

The Pattern of signs within and across vectors

With only one exception (1st elements of cocoa vectors) all the signs of corresponding elements of $\hat{\underline{a}}$ and $\hat{\underline{b}}$ vectors are opposite. The configuration of + and - signs within each vector is also exactly what was expected. There are only 7 instances out of a possible (8 x 5) = 40 in which the signs of two adjacent elements are identical; 33 of the 40 sign configurations are of the +/- or -/+ type.

Sums of Element Values

It is interesting to consider the three sums reported in Table 6.12.

(i) The sum of all the positive elements of a given vector are all very near +500.

(ii) The sum of all the negative elements of a given vector are all very near -500.

(iii) The sum of all the elements of a given vector are near zero.

Summary of pooled vector results

We sum up by saying that, in the light of the general pattern in the individual estimates noted in section 6.8.1, the average estimates given in Table 6.12 seem to be what we would expect. We will thus consider that these pooled vector estimates represent the underlying population vectors that we have been seeking.

We make one final remark on these average vectors. There is a remarkable similarity in the $\hat{\underline{a}}$ and $\hat{\underline{b}}$ vectors. This similarity is investigated and exploited further in Chapters 7 and 8.

6.10 Interpretation of canonical coefficient vectors

We now address the question of what these canonical coefficient vectors tell us about the observed persistent multivariate temporal dependence found in all the series.

The $\hat{\underline{a}}$, $\hat{\underline{b}}$ vectors can be used to compute a linear combination of returns on day t, $\hat{\psi}_1 = \hat{\underline{a}}^T \underline{x}_t$ and a linear combination of returns on day t-1, $\hat{\psi}_2 = \hat{\underline{b}}^T \underline{x}_{t-1}$ such that the correlation between $\hat{\psi}_1$ and $\hat{\psi}_2$ is maximum and positive. The $\hat{\psi}_1$, $\hat{\psi}_2$ series are special mixes of returns on day t and t-1 and we can refer to them as portfolios. The contribution to each portfolio from the $\underline{x}_t$, $\underline{x}_{t-1}$ series are given by the elements of $\hat{\underline{a}}$ and $\hat{\underline{b}}$ respectively. If we consider as an example the pooled rubber vectors:

$\hat{\underline{a}}^T$ =	+27	-89	+132	-269	+343	-140
$\hat{\underline{b}}^T$ =	-9	+54	-119	+292	-372	+154

The $\hat{\psi}_1$ portfolio is made up by summing 2.7%, 13.2% and 34.3% of the returns on contracts 1, 3 and 5 (a total of 50.2%), and subracting 8.9%, 26.9% and 14.0% of the returns on contracts 2, 4 and 6 (a total of 49.8%). The $\hat{\psi}_2$ series is made up by summing a similar collection of contract returns (with reverse signs) on day t-1. It is easy to see that if the correlation between the $\hat{\psi}_1$ and $\hat{\psi}_2$ series is large and positive the correlation between the $\hat{\psi}_1$ and $-\hat{\psi}_2$ series would be large and negative.

Considering the ψ_1 and $-\psi_2$ series is much more informative, since they are virtually identical. In the case of the rubber series we see that the maximum contribution to the $\hat{\psi}_1$ and $-\hat{\psi}_2$ series is contract number 5. The contract with the next highest contribution is contract number 4 and so on. It appears then that the observed multivariate temporal dependence can be explained by large and significantly negatively serially correlated linear combinations of returns (portfolios). Most of this negative serial correlation can be attributed to the 3rd, 4th and 5th contracts. It is interesting to note that no such negative serial correlation was discovered in any of the individual contracts when examined univariately (see section 3.6.2).

What phenomenon is giving rise to this negative serial correlation of lag 1 day witnessed in these complex linear combinations?

We consider two possible explanations:

(i) In section 5.3.2 we reported that the correlation between returns on different contracts in the same futures market are very high, typically +0.95. The returns and prices of all 6 contracts move together very closely. However there may be days when 1 or 2 of the contract prices lags behind the overall general movement of the rest of the market. This "getting out of line" with the overall price distribution would obviously not last long and trading would ensure that all prices were back to the norm within, say, at most one day. These small perturbations in the prices and hence returns may not be picked up in a univariate statistical examination of each of the 6 series. However when considering all 6 series together such anomalies may "stand out" quite clearly in a statistical sense. What we are suggesting here then is that the futures markets are not perfectly efficient. This model of market inefficiency is investigated in more detail in Chapter 9.

(ii) A market inefficiency of a different nature lies behind our second suggested model to explain the observed temporal dependence. This model is concerned with the overall pattern of the spread of prices of all the contracts in a given futures market.

It is well known that in periods of plentiful supply of the spot commodity there is a contango in which the 'near' prices are lower than the 'far' prices. In periods of short supply the reverse situation,a backwardation, is more usual; the prices of the 'near' contracts are higher than the prices of the 'far' contracts. In any period, the supply and demand of the spot and future commodity together with the current interest rates and storage charges should determine the relevant price profile.

From day to day information on factors that determine the price profile will arrive at the market. If the markets disseminate this information efficiently the change from one price profile to another should be gradual and smooth. If the markets are not perfectly efficient at disseminating this information, it is possible that some of the contract prices are not adjusted to their 'correct' level immediately.

If one or more prices were to 'get out of line' in the manner suggested in (i) or (ii) above and 'get back into line' the <u>very</u> next day, this could explain the observed negative serial correlations between certain linear combinations of returns. Furthermore from the discussion on the magnitudes of the $\hat{\underline{a}}$ and $\hat{\underline{b}}$ vectors it would seem that prices that 'get out of line' more frequently are the'far' ones (3rd, 4th and 5th contracts). This may be a reasonable explanation since the volume traded in these far contracts is relatively low.

6.11 Conclusions of Chapter 6

In this chapter we investigated the nature of the multivariate serial correlation found in the joint distribution of the returns on contracts in the same commodity futures market. It was demonstrated that most of the correlation could be explained by certain linear combinations of returns. This chapter was concerned mainly with the attempt to estimate accurately these special linear combinations. The highly collinear data gave rise to very changeable estimates but experiments with the use of a ridge - regression procedure was shown not produce any significant improvements. Although the estimates were very changeablethere was a very clear pattern in the individual components and this was discussed at length. Finally a grand averaging procedure was used to produce estimates for these linear combinations over the entire 5 year period. The general pattern in all of the individual components and the pattern in the grand average estimates suggests the possibility of a specific type of multivariate inefficiency. This multivariate inefficiency could to be mostly due to small price perturbations in the middle and far contracts.

The discovery of this multivariate inefficiency and the special nature of the linear combinations prompted the author into investigating the possibility of constructing various trading rules. These trading rules are outlined in Chapters 7 and 8. A model of the multivariate inefficiency is considered in Chapter 9.

Footnotes for Chapter 6

1. We know from Chapter 5 that the assumption of multivariate normality is suspect but hope that the procedure is robust to the type of departures from normality witnessed here.

2. We define the length L of the vector $\underline{\hat{a}}$ as:

$$L = \sum_{i=1}^{6} |\hat{a}_i|$$

3. Hoerl and Kennard (1970) and Vinod (1976) define the length of a vector $\underline{\hat{a}}$ as:

$$L = \underline{\hat{a}}^T \underline{\hat{a}}$$

4. In the calculation of his indicators, Vinod (1976) states that he is using five sets of data. Unfortunately these five sets were overlapping periods and thus are not independent. Accordingly one must treat his results with scepticism.

5. All these results have been tested using the standard test of proportions of H_0: $p = 0.5$ against H_1: $p \neq 0.5$, where p = proportion of pairs resulting in different signs. All except the first component of cocoa series result in rejection of the null hypothesis at the 0.1% level. See 6.8.1 (vii) for more details.

6 For the sugar and rubber series the standard deviations tended to decrease monotonically as one considers contracts with longer andd longer times to maturity.

CHAPTER 7

TIME DEPENDENT PORTFOLIOS

In Chapter 6 we noted that it is possible to find constructs $\hat{\psi}_1 = \hat{\underline{a}}^T\underline{x}_t$ and $-\hat{\psi}_2 = -\hat{\underline{b}}^T\underline{x}_{t-1}$ that are significantly negatively correlated and that $\hat{\underline{a}}$ is nearly always very similar to $-\hat{\underline{b}}$ (in each separate period or in the grand aggregate estimate). If this really is the case then it should be possible to find a vector $\hat{\underline{c}}$ such that $\hat{\underline{c}}^T\underline{x}_t$ and $-\hat{\underline{c}}^T\underline{x}_{t-1}$ are negatively correlated and in which $\hat{\underline{c}}$ will be somewhere "between" $\hat{\underline{a}}$ and $-\hat{\underline{b}}$. , i.e.

$$\hat{\underline{c}} \simeq \hat{\underline{a}} \simeq -\hat{\underline{b}}$$

If this were possible then the vector $\hat{\underline{c}}$ will denote a portfolio of returns that will exhibit negative serial correlation at lag 1. The portfolio $\hat{\underline{c}}$ will delineate a new asset, one made of various proportions of long and short positions in commodity futures contracts.

Classical portfolio theory begins by assuming that the set of returns is a set of temporally independent time series with fixed mean and variance. The portfolios delineated by Markowitz procedures are efficient in the mean/variance sense but will still result in temporally independent series (having the lowest variance possible, at the given average returns).

The purpose of this chapter is to construct portfolios which, irrespective of expected returns or risk, will exhibit a temporal dependence (negative correlation). Constructing such temporally dependent portfolios has not previously been reported in the literature and leads on to the development of multivariate trading rules which are dealt with in Chapter 8.

The layout of this chapter is as follows. In section 7.1 we show that the temporally dependent portfolios we are seeking are the eigenvectors of certain matrices. In section 7.2 we show that the resulting eigenvectors are always real and in section 7.3 we discuss their sampling properties. The portfolios delineated by the procedure on the data set are discussed in section 7.4. In section 7.5 we give a brief report on some of the other statistical properties of the resultant portfolios. In section 7.6 we consider portfolios produced using the grand averaging pooling technique. Concluding remarks are made in section 7.7

7.1 Derivation of the $\underline{c}$ vector

Using the notation of sections 6.3 and 6.2 we consider partitioned vector $\underline{y}_t$ consisting of the returns of day t, $\underline{x}_t$ and day t-1, $\underline{x}_{t-1}$; t = 2, . .. n. Assume that $\underline{y}_t$ is multivariate normal of dimension $2p$.

$$\underline{y}_t = \begin{bmatrix} \underline{x}_t \\ \\ \underline{x}_{t-1} \end{bmatrix} \sim N \left[\begin{bmatrix} \underline{\mu}_t \\ \\ \underline{\mu}_{t-1} \end{bmatrix} , \begin{bmatrix} V_{t,t} & V_{t,t-1} \\ \\ V_{t-1,t} & V_{t-1,t-1} \end{bmatrix} \right]$$

For reasons outlined in Chapter 5 we can assume that $V_{t,t} = V_{t-1,t-1}$. For simplicity of exposition denote $V_{t,t}$ by V_{11} and $V_{t-1,t-1}$ by V_{22}.

We wish to find $\underline{c}$ such that the linear combination $\underline{c}^T\underline{x}_t$ will be maximally negatively correlated with $\underline{c}^T\underline{x}_{t-1}$. The correlation coefficient between $\underline{c}^T\underline{x}_t$ and $\underline{c}^T\underline{x}_{t-1}$ is given by γ, where:

$$\gamma = \frac{\mathrm{Cov}[\ \underline{c}^T\ \underline{x}_t,\ \underline{c}^T\ \underline{x}_{t-1}\ ]}{\sqrt{\mathrm{Var}[\ \underline{c}^T\ \underline{x}_t\ ].\mathrm{Var}[\ \underline{c}^T\ \underline{x}_{t-1}\ ]}}$$

$$\gamma = \frac{\underline{c}^T\ V_{12}\ \underline{c}}{\underline{c}^T\ V_{11}\ \underline{c}}$$

It is possible to choose $\underline{c}$ such that $\underline{c}^T\ V_{11}\ \underline{c} = 1$ and so

$$\gamma = \underline{c}^T\ V_{12}\ \underline{c} \quad \ldots\ldots\ldots\ldots \quad (7.1)$$

The problem formally is:

Minimize $\qquad \gamma = \underline{c}^T\ V_{12}\ \underline{c}$

Subject to the constraint

$$\underline{c}^T\ V_{11}\ \underline{c} = 1 \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (7.2)$$

We form the Lagrangian

$$Z = \underline{c}^T\ V_{12}\ \underline{c} - \lambda\ (\underline{c}^T\ V_{11}\ \underline{c}\ -1)$$

differentiate with respect to $\underline{c}$ and equate to zero

$$\frac{\partial Z}{\partial \underline{c}} = V_{12}\ \underline{c} + V^T{}_{12}\ \underline{c} - \lambda\ V_{11}\ \underline{c} - \lambda\ V^T{}_{11}\ \underline{c}$$

$$= \quad (V_{12} + V_{21})\,\underline{c} \quad - \quad 2\,\lambda\,V_{11}\,\underline{c} \quad = \underline{0}$$

$$= \quad [(V_{12} + V_{21}) \quad - 2\,\lambda\,V_{11}]\,\underline{c} \quad = \quad \underline{0} \qquad \ldots\ldots (7.3)$$

Pre multiply by V^{-1}_{11} (V_{11} is non singular)

$$[V^{-1}_{11}(V_{12} + V_{21}) \quad - \quad 2\,\lambda\,I]\,\underline{c} \quad = \quad \underline{0} \qquad \ldots\ldots (7.4)$$

Thus 2λ and $\underline{c}$ will be the eigenvalues and eigenvectors of the matrix

$$D \quad = \quad [\ V^{-1}_{11}(V_{12} + V_{21})\ ] \qquad \ldots\ldots (7.5)$$

The correlation coefficient r can be obtained by pre multiplication of (7.3) by $\underline{c}^T$:

$$\underline{c}^T\,[V_{12} + V_{21} - \ 2\,\lambda\,V_{11}]\,\underline{c} \quad = \quad \underline{c}^T\,\underline{0}$$

$$\underline{c}^T\,V_{12}\underline{c} \quad + \quad \underline{c}^T\,V_{21}\,\underline{c} \quad - \quad 2\lambda\,\underline{c}^T\,V_{11}\,\underline{c} \quad = \quad 0$$

$$2\,\underline{c}^T\,V_{12}\,\underline{c} \quad - \quad 2\,\lambda \quad = \quad 0$$

$$2\,r \quad - \quad 2\,\lambda = \quad 0$$

therefore $\qquad r = \lambda \quad = \quad 1/2$ eigenvalue of D

If the rank of matrix D is k ($1 \leqslant k \leqslant p$) there will be k eigenvalues and eigenvectors. The eigenvalue of most interest will be the minimum one (maximum negative correlation). Accordingly we will compute the minimum eigenvalue ($2r_{min}$) and associated eigenvector ($\underline{c}$).

7.2 Proof that γ and $\underline{c}$ are real

In this section we prove that the eigenvalues and eigenvectors of D in (7.5) are real. For ease of exposition we express D as:

$$D = A.B$$

in which $A = V^{-1}_{11}$ and $B = V_{12} + V_{21}$

Note that A and B are symmetric, but that A.B is not necessarily symmetric. A standard result in the theory of matrices states that the non zero eigenvalues of AB and BA are the same and have the same multiplicity. If $\underline{c}$ is a non-trivial eigenvector of AB for an eigenvalue $2\gamma = 0$, then $\nu = B\underline{c}$ is a non trivial eigenvector of BA. If possible let a non trivial eigenvalue of AB be 2γ with eigenvector $\underline{c}$ and let $\underline{c}$ be be such that:

$$\underline{c} = \underline{x} + i\underline{y}, \qquad 2\gamma = a + i.b$$

with $\qquad i = \sqrt{-1}$

Then the eigenvector and eigenvalue of BA will be $B\underline{c}$ and 2γ respectively, with:

$$B\underline{c} = B\,\underline{x} + i\,B\,\underline{y}$$

and $\qquad (BA)(B\underline{c}) = 2\gamma(B\underline{c})$

$$(BAB)\underline{c} = 2\tau(B\underline{c})$$

$$(BAB)(\underline{x} + i\,\underline{y}) = (a + i\,b)\,(B\,\underline{x} + i\,B\,\underline{y})$$

$$BAB\underline{x} + i(BAB\underline{y}) = aB\underline{x} + iaB\underline{y} + ibB\underline{x} - bB\underline{y}$$

Equating real and imaginary parts

$$BAB\underline{x} = aB\underline{x} - bB\underline{y} \quad \ldots\ldots \; (7.6)$$

$$BAB\underline{y} = bB\underline{x} + aB\underline{y} \quad \ldots\ldots \; (7.7)$$

Premultiplying (7.6) by $\underline{y}^T$ and (7.7) by $\underline{x}^T$

$$\underline{y}^T BAB\underline{x} = a\underline{y}^T B\underline{x} - b\underline{y}^T B\underline{y} \quad \ldots\ldots\ldots \; (7.8)$$

$$\underline{x}^T BAB\underline{y} = b\underline{x}^T B\underline{x} + a\underline{x}^T B\underline{y} \quad \ldots\ldots\ldots \; (7.9)$$

Subtracting (7.9) from (7.8) yields:

$$0 = -\,b\,[\underline{y}^T B\underline{y} + \underline{x}^T B\underline{x}\,]$$

Since $\underline{y}$, $\underline{x}$ are non trivial b must be zero. Hence $\underline{c}$ is real and so is 2τ. And so the vector $\underline{c}$ and eigenvalue, 2τ of A in (7.5) are real.

7.3 Estimation of τ and c

To estimate the temporally dependent portfolios, $\underline{c}$, and the degree of temporal dependence, τ, (the serial correlation coefficient) one may simply substitute the classical sample estimates S_{11}, and S_{12} in expression (7.5) and solve for the eigenvectors and eigenvalues respectively. Note that in all that follows we are only interested in the minimum τ value and associated $\underline{c}$ vector and so without loss of generality we will refer always to τ as half the minimum eigenvalue of D in (7.5).

7.3.1 Sampling properties of τ and c

We noted in section 6.6.1 that the distribution of the canonical vectors and canonical correlations were extremely complex in the null case and (to date) mathematically intractable in the non null case. The author is unaware of any published work relating to the sampling distribution of eigenvectors and eigenvalues of matrices of the type in (7.5). We strongly suspect that the distributions will be complex. Similarly there exists no published work on tests of significance of the eigenvalues of matrices of the type in (7.5). Accordingly it was decided, in this study, to leave aside the question of the sampling distribution and tests of significance of $\hat{\tau}$ and $\hat{\underline{c}}$.

7.3.2 Estimating the variance/covariance matrix

In section 7.1 we made the assumption that $V_{t,t} = V_{t-1,t-1}$, $V_{t,t}$ being estimated for t = 1 to n - 1 and $V_{t-1,t-1}$, being estimated for t = 2 to n. An initial study showed that the two estimates were virtually identical so the question arose as to which estimate to use in the full study. It was decided to use a third estimate, one which uses squares and sums of squares spanning the entire set of observations from t = 1 to n. All three estimates yielded almost identical results for $\hat{\tau}$ and $\hat{\underline{c}}$.

7.3.3 The problems of highly collinear data

Note that the matrix D in (7.5) involves the inverse of V_{11}. In the estimation of τ and $\underline{c}$ then, this will involve the inversion of S_{11}, the sample estimate of V_{11}. As noted in section 6.7.1, with highly collinear data, the inverses tend to be unreliable and very sensitive to small changes in the data. We suspect therefore that we will encounter similar problems of instability in the estimation of $\underline{c}$. In particular, the lengths of $\hat{\underline{c}}$ will probably vary considerably.

The question of finding methods to reduce such instability in canonical vector estimates was treated in section 6.7.4. It was found that the normalization of vectors proved equally as good as using the ridge regression technique. We have not carried out a simulation study into the usefulness or otherwise of applying the ridge technique to this problem and instead simply present the normalized $\hat{\underline{c}}$ vectors.

Note that in Chapter 6 we studied logged returns. In this chapter we are interested in "real" portfolios of returns and so we now revert to considering unlogged returns. It should however be noted that the study was carried out on logged data as well with almost identical results.

7.4 The sample results

7.4.1 The serial correlation coefficients

The sample values of the serial correlation coefficients, $\hat{\tau}$, associated with the portfolio $\hat{\underline{c}}^T \underline{x}_t$ appear in Tables 7.1 to 7.4. Also given in Tables 7.1 to 7.4 are the sample values of the maximal canonical correlation coefficients1, $\hat{\lambda}_{max}$, associated with the linear combination[1]:

Table 7.1

Cocoa series

Period	$\hat{\lambda}_{max}$	$\hat{r}$
1	.573	-.442
2	.841	-.589
3	.623	-.586
4	.602	-.512
5	.741	-.691
6	.783	-.684
7	.736	-.673
8	.697	-.615
9	.638	-.526
10	.768	-.665
11	.772	-.598
12	.717	-.521
13	.854	-.648
14	.673	-.607
15	.724	-.696
16	.845	-.733
17	.681	-.598
18	.768	-.607
19	.683	-.565
20	.674	-.599
21	.761	-.659
22	.597	-.567
23	.732	-.625
24	.676	-.482

Table 7.2

Coffee series

Period	$\hat{\lambda}_{max}$	$\hat{r}$
1	.720	-.716
2	.942	-.825
3	.695	-.611
4	.782	-.620
5	.832	-.692
6	.714	-.621
7	.777	-.672
8	.882	-.614
9	.711	-.561
10	.780	-.678
11	.750	-.593
12	.752	-.750
13	.824	-.726
14	.774	-.695
15	.707	-.694
16	.712	-.569
17	.602	-.530
18	.751	-.620
19	.693	-.406
20	.623	-.589
21	.696	-.516
22	.615	-.553
23	.629	-.610
24	.657	-.640
25	.759	-.716
26	.620	-.532
27	.788	-.718
28	.744	-.627
29	.984	-.911

$\hat{\lambda}_{max}$ = maximum correlation between $\underline{a}^T\underline{x}_t$ and $\underline{b}^T\underline{x}_{t-1}$ series

$\hat{r}$ = minimum correlation between $\underline{c}^T\underline{x}_t$ and $\underline{c}^T\underline{x}_{t-1}$ series

Table 7.3

Sugar series

Period	$\hat{\lambda}_{max}$	$\hat{\gamma}$
1	.668	-.592
2	.738	-.552
3	.770	-.765
4	.668	-.496
5	.596	-.544
6	.744	-.631
7	.597	-.571
8	.760	-.711
9	.667	-.561
10	.785	-.624
11	.759	-.589
12	.555	-.537
13	.682	-.478
14	.779	-.698
15	.609	-.599
16	.726	-.572
17	.708	-.566
18	.744	-.588
19	.681	-.634
20	.655	-.559
21	.774	-.589
22	.641	-.577
23	.788	-.591
24	-	-

Table 7.4

Rubber series

Period	$\hat{\lambda}_{max}$	$\hat{\gamma}$
1	.642	-.555
2	.640	-.554
3	.562	-.544
4	.665	-.563
5	.712	-.602
6	.610	-.487
7	.726	-.632
8	.620	-.533
9	.659	-.616
10	.616	-.517
11	.612	-.528
12	.712	-.658
13	.606	-.554
14	.712	-.607
15	.665	-.643
16	.609	-.579
17	.686	-.664
18	.692	-.651
19	.665	-.526
20	.879	-.853

$\hat{\lambda}_{max}$ = maximum correlation between $\underline{a}^T\underline{x}_t$ and $\underline{b}^T\underline{x}_{t-1}$ series

$\hat{\gamma}$ = minimum correlation between $\underline{c}^T\underline{x}_t$ and $\underline{c}^T\underline{x}_{t-1}$ series

$\hat{\underline{a}}^T\underline{x}_t$ and $\hat{\underline{b}}^T\underline{x}_{t-1}$. Note in all cases the $\hat{r}$ values are negative. Note also that in each period $|\hat{r}| < \hat{\lambda}_{max}$. This is what one would have expected. The canonical correlation coefficient is the maximal correlation possible between any two linear combinations of the returns : $\underline{x}_t$, $\underline{x}_{t-1}$. By constraining the linear combinations to be of the form $\hat{\underline{c}}^T\underline{x}_t$ and $\hat{\underline{c}}^T\underline{x}_{t-1}$ it is not suprising that one obtains results that are slightly "off" the maximum possible value. What is remarkable about these results is the "closeness" of the two sets of correlations. In many periods the $\hat{r}$ values are very near the maximum possible values set by $\hat{\lambda}_{max}$.

We have shown then that in each of the 96 periods examined it is possible to find a portfolio of returns that exhibit large and significant (not necessarily in the statistical sense) negative serial correlation coefficients of lag 1.

7.4.2 The composition of the $\hat{\underline{c}}$ vector

Tables 7.5 to 7.8 list the components of the sample estimates of the $\underline{c}$ vector for each period. As before, for ease of exposition, the absolute values of the components are normalized to sum to 1000. We make the following observations.

(i) Comparing the values of the components of the $\hat{\underline{c}}$ vectors with the corresponding values of the components of the $\hat{\underline{a}}$ and $\hat{\underline{b}}$ vectors in Tables 6.5 to 6.8, we see that the $\hat{\underline{c}}$ vector is "somewhere between" $\hat{\underline{a}}$ and $\hat{\underline{b}}$.

(ii) The components that are the largest (in absolute terms) are in the far contracts. Table 7.9 gives a frequency distribution of the maximal component for each series. This table is very similar to the

Table 7.5

Normalized estimates of $\hat{c}$ components for cocoa returns

Period	Components 1	2	3	4	5	6	sum of comps.	length of vector
1	-6	-54	327	-384	182	-48	17	3.17
2	14	-16	172	293	-254	-251	-42	3.86
3	-4	-91	365	-249	-169	121	-27	2.56
4	-34	139	-303	254	-136	134	54	3.43
5	22	-160	56	-279	455	-28	66	1.39
6	-21	82	-88	84	-392	333	-2	2.67
7	42	-23	-376	477	1	-81	40	1.93
8	24	43	-352	-7	426	-148	-14	2.27
9	9	-27	141	-356	342	-125	-16	2.52
10	7	-205	245	-26	-271	246	-4	3.23
11	-16	178	-195	164	138	-308	-39	3.98
12	-29	23	-298	446	35	-170	7	2.83
13	6	-191	268	-301	148	87	17	6.35
14	2	-132	456	-184	-190	36	-12	1.78
15	24	-135	-51	396	67	-327	-26	2.61
16	31	-34	-31	9	464	-431	8	1.79
17	-6	110	-306	144	249	-185	6	3.72
18	-5	183	-492	305	-2	14	3	4.58
19	1	-147	279	-234	217	-123	-7	3.21
20	-182	301	-98	163	-234	21	-29	3.23
21	29	-139	29	411	-373	19	-24	2.95
22	13	-299	498	2	-170	-18	26	1.52
23	16	-9	176	-470	296	-33	-24	3.24
24	-25	-220	334	-136	173	-112	14	2.46

Table 7.6

Normalized estimates of c components for coffee returns

Period	Components 1	2	3	4	5	6	sum of comps.	length of vector
1	-22	8	282	-178	200	-311	-21	2.85
2	-55	96	-173	114	293	-269	6	11.59
3	214	-430	198	76	-20	-62	-24	2.70
4	-34	-24	334	-372	159	-77	-14	2.06
5	29	-88	189	-37	-380	276	-11	1.48
6	-15	-18	236	-414	264	-53	0	2.17
7	45	179	-124	-370	32	249	11	3.75
8	-55	13	271	-413	217	-32	1	3.00
9	-8	210	-109	-279	-94	300	20	1.88
10	17	-118	252	-222	-160	232	1	6.07
11	-12	351	158	-215	-262	-2	18	3.75
12	-90	371	-252	-152	66	68	11	2.30
13	57	-226	106	95	247	-269	10	2.82
14	34	18	-278	-234	183	253	-24	3.46
15	28	-57	141	-67	-382	326	-11	1.59
16	-46	26	116	-294	366	-151	17	2.20
17	22	80	-243	321	-247	87	20	2.94
18	2	2	44	-354	436	-162	-32	3.05
19	16	40	-432	264	169	-80	-23	2.24
20	41	162	-405	299	-34	-58	5	3.02
21	13	261	-202	-279	235	-9	19	1.90
22	-29	-141	230	264	-227	-109	-12	2.89
23	-2	-140	321	-78	-280	180	1	2.57
24	-7	54	-151	352	-348	87	-13	1.60
25	17	-148	-99	470	-251	15	4	3.79
26	26	52	-233	300	-268	121	-2	1.60
27	105	-154	120	185	-334	102	24	3.84
28	-70	98	128	-277	276	-151	4	3.72
29	18	24	-297	406	-195	60	16	1.53

Table 7.7

Normalized estimates of c components for sugar returns

Period	Components 1	2	3	4	5	6	sum of comps.	length of vector
1	204	-35	-56	-214	-177	315	37	3.28
2	-107	-15	102	398	-353	-24	1	6.83
3	-74	322	-130	-279	-12	183	10	2.19
4	-10	-166	123	370	-248	-83	-14	3.72
5	-22	82	-203	203	210	-279	-9	2.33
6	-104	322	-394	151	26	3	4	4.14
7	-41	-31	41	-31	439	-417	-40	1.73
8	-9	-66	406	-425	93	1	0	2.51
9	7	191	-419	304	19	-60	42	1.82
10	4	-57	147	-292	353	-147	8	1.97
11	-24	132	-292	359	-28	-167	-20	1.76
12	-18	-45	0	398	-436	104	3	2.21
13	-26	-91	122	-71	386	-304	16	3.60
14	68	-55	24	-108	-311	433	51	1.88
15	-84	178	-178	13	-233	313	9	2.87
16	-26	187	-209	-61	316	-201	6	6.29
17	-70	146	68	-292	-134	291	9	2.41
18	44	-42	143	-458	299	14	0	2.56
19	-1	-59	129	213	-450	149	-19	1.72
20	-23	134	-312	259	105	-166	-3	3.71
21	-16	168	-269	-114	347	-85	31	9.08
22	33	-242	139	329	-117	-141	1	2.35
23	-20	-19	-37	-150	504	-270	8	2.07

Table 7.8

Normalized estimates of c components for rubber returns

Period	Components 1	2	3	4	5	6	sum of comps.	length of vector
1	58	129	110	-506	11	186	-12	2.11
2	62	-114	87	-120	337	-280	-28	3.69
3	-13	-111	263	10	-363	239	25	1.83
4	77	-12	331	-223	-262	94	5	3.01
5	28	-53	172	-327	302	-118	4	2.86
6	36	-60	5	365	-460	74	-40	2.99
7	69	41	-380	250	138	-121	-3	2.58
8	-64	119	172	137	-453	55	-34	2.07
9	89	-324	83	376	-11	-117	96	2.29
10	-79	186	-38	83	-390	223	-15	2.74
11	-32	50	-341	440	4	-132	-11	3.42
12	63	-134	136	-180	292	-196	-19	3.63
13	26	8	-80	201	-414	271	12	2.89
14	-3	96	-299	116	291	-196	5	2.05
15	28	-9	-210	-122	468	-162	-7	2.41
16	-21	203	-117	-279	299	-81	4	5.83
17	-15	46	-49	-205	453	-231	-1	2.90
18	4	10	-35	39	439	-472	-15	2.08
19	3	-92	424	-390	64	-26	-17	3.22
20	-20	386	-353	-125	62	53	3	2.61

Table 7.9

Frequency of maximal absolute $\hat{c}$ components
(obtained from Tables 7.5 - 7.8)

		Component					
	no. of periods	1	2	3	4	5	6
Cocoa series	24	0	1	8	9	5	1
Coffee series	29	0	3	4	1	8	3
Sugar series	23	0	1	3	7	9	4
Rubber series	20	0	1	4	4	10	1

Table 7.10

Frequnecy of contiguous sign combinations within $\hat{c}$ vector
(obtained from Tables 7.5 - 7.8)

		Sign combinations +/+ or -/-	+/+ or -/-	Z value
Cocoa series	c	23	97	6.76
	E	60	60	
Coffee series	c	47	98	4.24
	E	72	72	
Sugar series	c	37	83	4.20
	E	60	60	
Rubber series	c	24	76	5.20
	E	50	50	

E = Expected frequency if sign combinations are random corresponding summary Table 6.9. The 4th and 5th components are the largest in more than 50% of the periods.

(iii) As was noted in section 6.8.1 in the analysis of the $\hat{\underline{a}}$, $\hat{\underline{b}}$ vectors, there is an excessive switching in the sequence of signs of the components within any given $\hat{\underline{c}}$ vector. Table 7.10 gives a summary of the total number of sign sequences (in pairs) along with the associated test statistic Z. Recall that under the hypothesis of random sign sequences Z should be N (0,1) distributed. Note that all the Z values are highly significant.

(iv) Note that the sum of all the $\hat{\underline{c}}$ components in each period is always very near zero. Thus we observe that in each period the temporally dependent portfolio delineated is a nearly perfect spread. As an example consider the 6th period of the cocoa series:

$$\underline{c}^T = (-21, 82, -88, 84, -392, 333)$$

$$\sum_{i=1}^{6} c_i = -2$$

The designated portfolio in this case would consist of long positions in the 2nd, 4th and 6th contracts and short positions in the 1st, 3rd and 5th contracts. The relative numbers of each contract are given by the absolute values of the components. In this case 21 short in contract 1, 82 long in contract 2 and so on. The total long position is 499 contracts and the total short position 501 contracts. The overall net position is given by the sum of the components, i.e. -2 ; or 2 short. A perfectly spread portfolio[2] would have an overall net position of zero. This portfolio of contracts over 6 delivery dates is an extremely complex multivariate spread and although not a perfect spread the risk associated with it will be very small indeed. We return to this point later in section 7.5.1.

Reference to Tables 7.5 to 7.8 reveals that although some of the individual contract positions are as large as 470 the overall position of each portfolio is very near zero.

It is an extremely interesting result that the temporally dependent portfolios that "fall out" of this analysis invariably constitute near perfect (though complex) spreads. We return to this point again in Chapter 9 when discussing the possible mechanisms giving rise to the observed multivariate dependence.

(v) In the above we have outlined the observed general pattern in the $\hat{\underline{c}}$ vectors. The pattern is very similar to that found when studying the $\hat{\underline{a}}$ and $\hat{\underline{b}}$ vectors. However, although these patterns are discernible and very similar across commodity series, there is still a great deal of variation in the estimates of particular $\hat{\underline{c}}$ components from period to period. As in the case of the variation observed in the $\hat{\underline{a}}$ and $\hat{\underline{b}}$ vectors we suggest two possible explanations:

(a) Whatever the mechanism that is causing the observed multivariate serial correlation (MVSC), its nature and its effect on the set of multivariate returns is changing from period to period. As a result the maximally temporally dependent portfolios $\hat{\underline{c}}^T\underline{x}_t$ will also change from period to period. In some periods more weight will be put on the 4th contract and in others more weight will be put on the 5th contract and so on. In short, one possible explanation for the observed variation in $\hat{\underline{c}}$ is that the real vector $\underline{c}$ is varying.

(b) The mechanism that is giving rise to the observed MVSC is the result of some persistent phenomenon in the commodity futures markets. The underlying MVSC is thus also a constant [as appears to be the case when examining the correlation coefficients]. Accordingly the $\underline{c}$ vector should also be fixed and constant. If this were the case then the observed highly variable components of $\hat{\underline{c}}$ must be explained by sampling variation. This certainly could be the case. Recall the extreme variations in the $\hat{\underline{a}}$, $\hat{\underline{b}}$ vectors observed. As noted the estimation problem associated with the inversion of near singular matrices yields highly variable results. We return to this suggestion again in section 7.6.

7.5 A statistical examination of the ex post temporally dependent portfolios

It was decided to examine the 96 sets of time series, $\hat{\underline{c}}^T\underline{x}_t$ delineated by the methods of section 7.2. We know that the series will exhibit negative serial correlation but it will also be interesting to discover if (i) the serial correlation is caused by one or two outliers, (ii) if the series have temporal dependencies at lags other than one day.

Rather than list the results of all the usual statistical tests on all 96 series we summarize our findings in the following brief statements.

7.5.1 Distribution of returns

Some of the distributions of returns appear to be normal and certainly nearly all were symmetrical. This is not a surprising result since, as reported in Chapter 5, on the multivariate study of individual contract returns, the joint distribution could be considered as being generated from a multivariate normal process. Not surprisingly, there also appeared occasional large spikes and/or sudden chages in the variability of returns. The large spikes in the series could usually (but not always - see Chapter 9) be explained by large spikes in the joint distribution of individual contract returns.

The Means of the returns were all not significantly different from zero.

The standard deviations of the returns were all very much smaller than the standard deviations of returns of a contemporaneous individual typical contract. Table 7.11 lists the standard deviations of returns on the $\hat{\underline{c}}^T \underline{x}_t$ series, the standard deviations of returns on the contract that is nearest maturity and the ratio of these two measures. Note how small the ratios are. The range of values of the ratio are 0.047 to .365 but most are typically 0.100. Recall from section 7.4 that these series are returns on near perfect spread portfolios. Table 7.11 demonstrates the much larger risks when investing in a given contract as compared with a spread portfolio.

7.5.2 A further investigation into the nature of the temporal dependence

If the returns series, $\hat{\underline{c}}^T \underline{x}_t$ exhibit a temporal dependence of a more complex nature than first order autocorrelation, it may be possible to fit a Box - Jenkins type model. The autocorrelation function (ACF) and the partial auto- correlation function (PACF) of each of the 96 series were examined and almost without exception they had the following pattern:

(i) The first ACF coefficient was negative and highly significant. The remaining ACF coefficients were nearly always insignificant and distributed randomly positive or negative about the value zero.

Table 7.11

Standard deviations of returns on near contract and $\hat{c}^T x_t$ portfolio

Cocoa series

Period	near ct.	portfolio	ratio
1	1102.	208.	0.189
2	1979.	158.	0.080
3	1493.	119.	0.079
4	1025.	183.	0.178
5	991.	181.	0.183
6	3143.	300.	0.096
7	1847.	332.	0.180
8	2685.	529.	0.197
9	4477.	364.	0.081
10	10270.	662.	0.064
11	7586.	1521.	0.291
12	8032.	776.	0.097
13	6828.	820.	0.120
14	5390.	584.	0.108
15	5107.	604.	0.118
16	3964.	861.	0.217
17	3448.	327.	0.095
18	3033.	243.	0.080
19	2499.	231.	0.092
20	2706.	722.	0.267
21	2511.	314.	0.125
22	2966.	421.	0.142
23	2821.	350.	0.124
24	2412.	225.	0.093

Coffee series

Period	near ct.	portfolio	ratio
1	511.	102.	0.200
2	3998.	348.	0.087
3	1168.	261.	0.223
4	759.	84.	0.111
5	1090.	120.	0.110
6	1114.	127.	0.114
7	3702.	374.	0.101
8	4952.	247.	0.050
9	2385.	326.	0.137
10	3290.	333.	0.101
11	8314.	1004.	0.121
12	12226.	1362.	0.111
13	13239.	1172.	0.088
14	14279.	2358.	0.165
15	10068.	1817.	0.180
16	7622.	903.	0.119
17	6201.	790.	0.127
18	3496.	930.	0.266
19	2944.	634.	0.215
20	6788.	589.	0.087
21	5753.	631.	0.110
22	4590.	554.	0.121
23	4042.	424.	0.105
24	2195.	391.	0.178
25	2471.	383.	0.155
26	6155.	831.	0.135
27	1650.	386.	0.234
28	2340.	420.	0.180
29	2585.	387.	0.150

Table 7.11 continued

Standard deviations of returns on near contract and $\hat{c}^T x_t$ portfolio

Sugar series

Period	near ct.	portfolio	ratio
1	1097.	400.	0.365
2	1076.	85.	0.079
3	392.	48.	0.121
4	307.	32.	0.105
5	301.	23.	0.075
6	296.	28.	0.094
7	497.	93.	0.188
8	472.	60.	0.127
9	446.	53.	0.119
10	228.	44.	0.192
11	305.	29.	0.096
12	188.	28.	0.149
13	243.	38.	0.157
14	164.	47.	0.285
15	152.	26.	0.171
16	140.	19.	0.139
17	157.	28.	0.176
18	184.	20.	0.108
19	275.	18.	0.067
20	138.	12.	0.083
21	188.	14.	0.075
22	138.	24.	0.171
23	335.	33.	0.098
24	382.	-	-

Rubber series

Period	near ct.	portfolio	ratio
1	40.	8.	0.201
2	80.	15.	0.184
3	58.	11.	0.197
4	43.	10.	0.234
5	107.	28.	0.265
6	111.	32.	0.287
7	147.	24.	0.161
8	86.	29.	0.333
9	97.	21.	0.212
10	66.	13.	0.189
11	57.	11.	0.199
12	66.	7.	0.106
13	76.	9.	0.118
14	83.	7.	0.087
15	84.	7.	0.088
16	86.	11.	0.128
17	101.	13.	0.125
18	108.	9.	0.087
19	72.	8.	0.117
20	92.	7.	0.073

(ii) The first PACF coefficient, by definition, is equal to the first ACF coefficient and so was found to be negative and significant. The second PACF coefficient was found to be negative and smaller than the first PACF coefficient. The third PACF coefficient was found to be negative and smaller than the second PACF coefficient and so on. The absolute value of the PACF coefficients decayed with increasing lag.

This pattern of ACF and PACF coefficients is consistent with a simple moving average process of order 1. Fitting 96 simple moving average models produced 96 estimates of the first order parameter. These estimates were in the range 0.50 to 0.70 .

So a tentative model of the series of returns is:

$$\underline{\hat{c}}^T \underline{x}_t = e_t - 0.65\, e_{t-1} \quad \ldots\ldots (7.10)$$

e_t = random disturbance, $E(e_t) = 0$, $V(e_t) = \sigma^2$, $E(e_t\, e_{t+i}) = 0$ for all i.

Obviously one could spend considerable time on a further examination of Box Jenkins type models to explain the returns. We leave such a more detailed discussion to another study. We will return to the model (7.10) later.

7.5.3 The influence of anomalies in the returns

It is well known that the presence of a single anomalous value can seriously affect the value of the first order serial correlation coefficient. How much of the negative serial correlation is due to some of the large spikes found on close examination of the series?

All 96 series were examined pictorially for large spikes. Those series that were found to contain anomalous observations were examined again with the value of the outlier reduced and "brought into line" with the rest of the series. In all cases the value of the first ACF and PACF coefficients were reduced slightly, but still remained significant. The overall patterns of the ACF's and the PACF's remained unchanged. As suspected, the values of the corresponding first order moving average parameters were also slightly reduced.

So although the anomalies tend to exaggerate the first order serial correlation, the "adjusted" series were found to still exhibit significant temporal dependence. It is interesting to compare these results with those of section 5.3.4. In section 5.3.4 highly significant multivariate serial correlation coefficients of lag 1 day are reduced (but still remain significant) by the removal of anomalous values.

7.5.4 Runs tests on the series _ a trading rule idea

A series that exhibits negative serial correlation of lag 1 day should also produce significant values in a runs test. Runs tests on all 96 series were carried out, and produced highly significant results. In every series there were too many runs above and below the median value. Adjusting the anomalies found in some series yielded identical results (because the runs test is robust to outliers).

It is interesting to look at a particular $\hat{\underline{c}}^{T}\underline{x}_t$ series. As an example consider a plot of the rubber series from the 15th period in Fig. 7.1. In this series there are 64 observations; 32 above the median and 32 below the median. Using the expression given in section 3.2.1 we see that

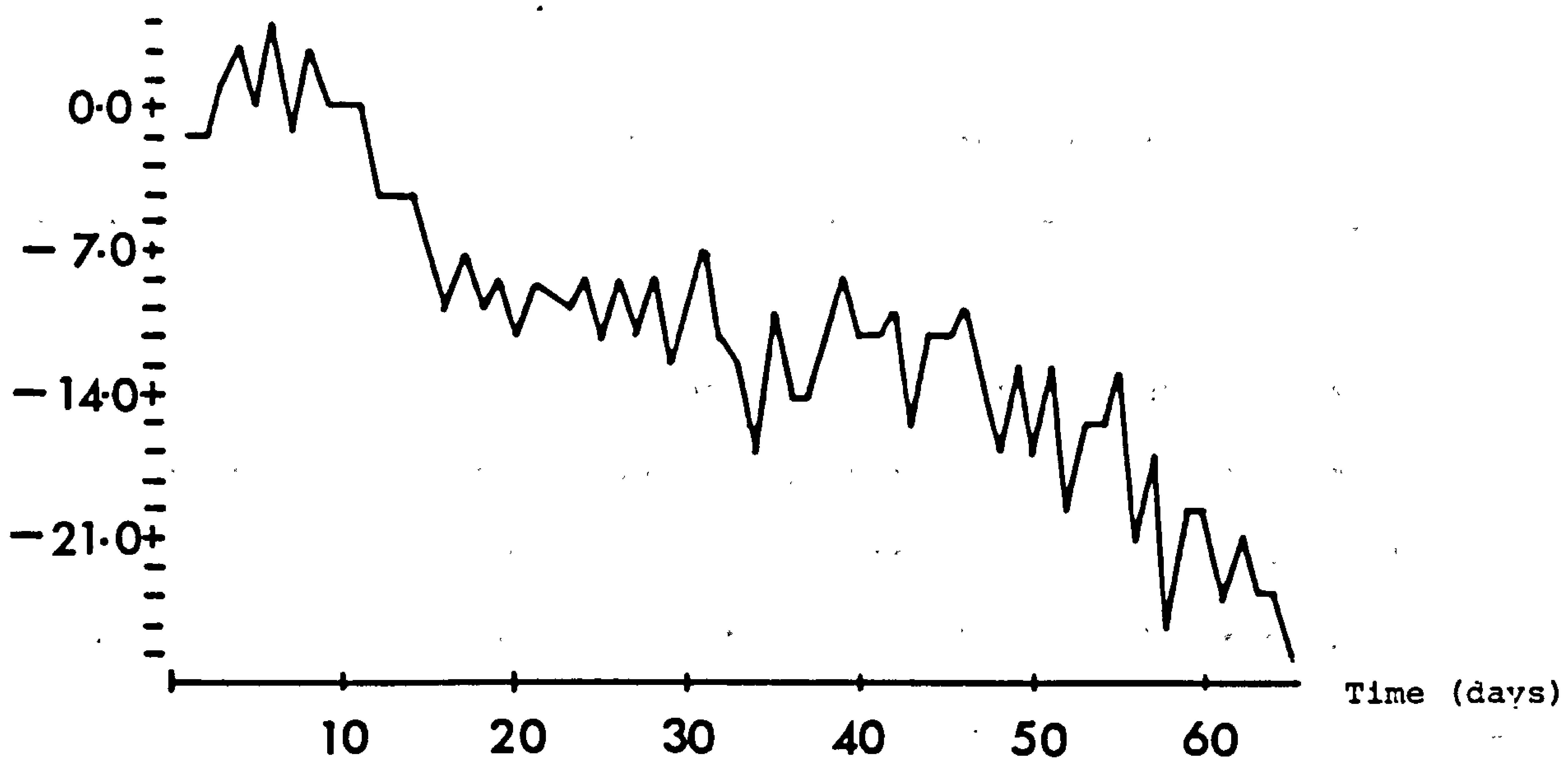

Fig 7.1 Prices of ex post portfolio in the 15th rubber series

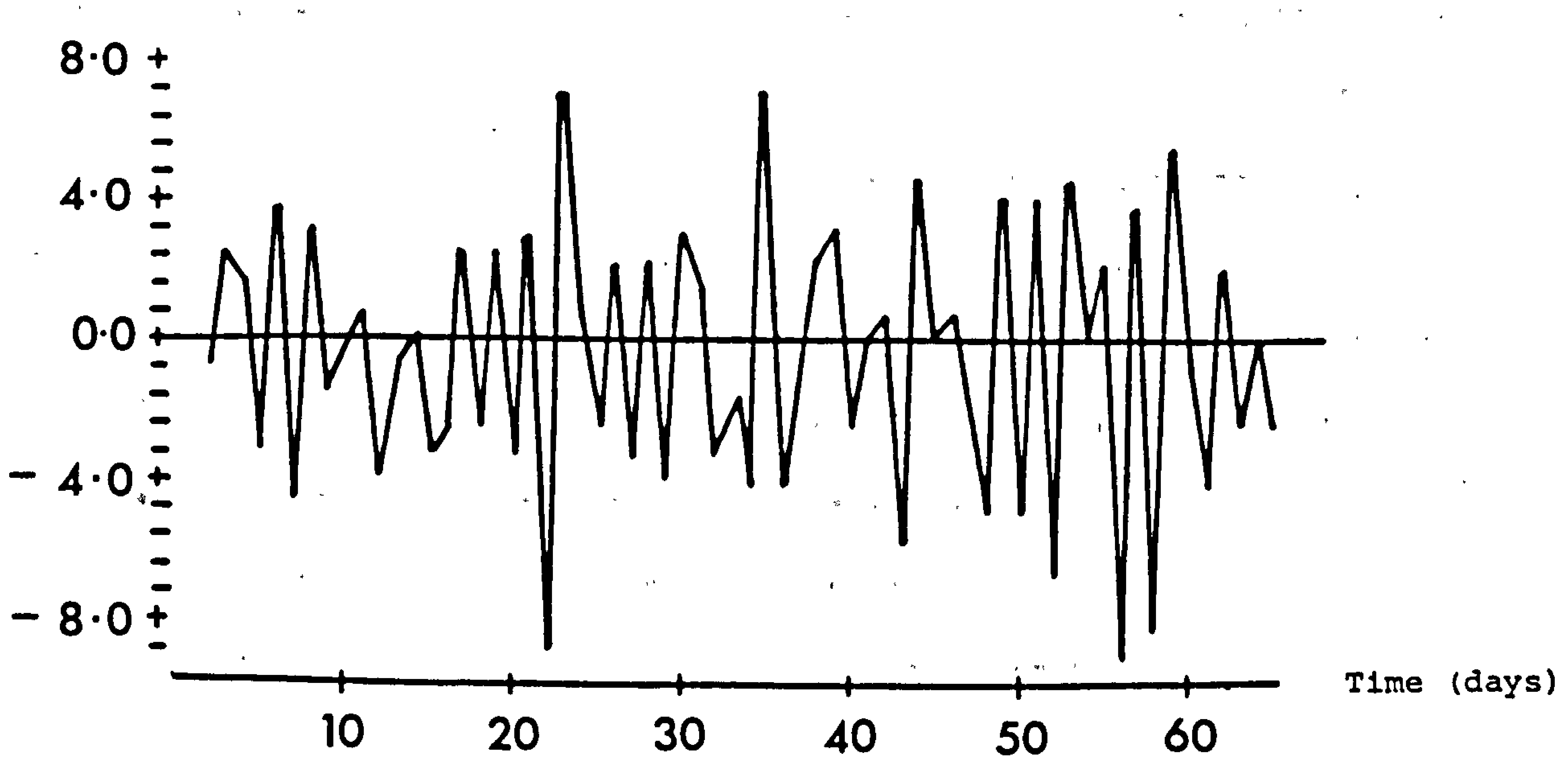

Fig 7.2 Returns of ex post portfolio in the 15th rubber series

the expected number of runs above and below the median is 33. The observed number of runs[3] is 47. There is an excessive reversalling in the returns.

It is even more interesting if we consider the series of "prices" that resulted in this particular returns series. See Fig. 7.2. Here we are considering the prices of a new asset; a portfolio of long and short positions in commodity futures contracts of various maturity dates. We see that the "price" of this asset is fluctuating in a non random manner. In particular, if the "price" falls (rises) from day i - 1 to day i, it is more likely to rise (fall) from day i to day i + 1 than fall (rise).

7.6 A pooled $\hat{\underline{c}}$ vector

We see from section 7.4.2 that although the normalized $\hat{\underline{c}}$ vectors vary from period to period there is a discernible pattern in the distribution of the individual components. Tables 7.9 and 7.10 give a brief description of this distribution. In section 7.4.2 we suggest that the observed variation in the components of $\hat{\underline{c}}$ could be simply due to sampling variation and problems associated with the inversion of nearly singular S_{11} matrices.

Is it possible to find one grand average $\hat{\underline{c}}$ vector that will be most temporally dependent throughout the entire 5 year period? If the mechanism giving rise to the observed MVSC is in fact due to some constant phenomenon then the estimation of the one overall temporally dependent portfolio $\hat{\underline{c}}^T \underline{x}_t$ should be possible. In order to calculate such an average $\hat{\underline{c}}$ vector we will need to estimate average V_{11} and V_{12} matrices. In the case of the coffee series, for example, we will need to average 29 individual estimates of V_{11} and V_{12}. As noted in section 6.9 in the computation of a grand average $\underline{a}$, $\underline{b}$ vector, the tremendous variation in variances and covariances over time will mean that estimates made in periods of high variation will completely swamp estimates made in periods of low variation. What is needed is a procedure that is variance free. The obvious step is to consider the possibility of estimating $\underline{c}$ from correlation matrices R_{11} and R_{12}.

7.6.1 Temporally dependent portfolios using correlation matrices

Below we show that by considering an appropriate transformation of the returns vector, $\underline{x}_t$, it is possible to compute $\underline{c}$ from the partitioned correlation matrix.

Using the notation of section 7.1, consider a non singular transformation of the returns vector $\underline{x}_t$ as follows:

$$\underline{x}^*_t = U^T \underline{x}_t \quad , \quad \underline{x}^*_{t-1} = U^T \underline{x}_{t-1}$$

$$\text{Vor } (\underline{x}^*_t) = \text{Vor } (U^T \underline{x}_t)$$

or

$$V^*_{11} = U^T V_{11} U$$

and

$$\text{Cov } (\underline{x}^*_t, \underline{x}^*_{t-1}) = \text{Cov } (U^T \underline{x}_t, U^T \underline{x}_{t-1})$$

or

$$V^*_{12} = U^T V_{12} U$$

In order to find the vector $\underline{c}^*$ such that $\underline{c}^{*T} \underline{x}_t$ and $\underline{c}^{*T}\underline{x}_{t-1}$ are maximally negatively correlated, we find the eigenvalues and eigenvectors of A^*,

$$A^* = V^{*-1}_{11} (V^*_{12} + V^*_{21})$$

$$= [U^T V_{11} U]^{-1} [U^T V_{12} U + U^T V_{21} U]$$

$$= U^{-1} V^{-1}_{11} (U^T)^{-1} U^T (V_{12} + V_{21}) U$$

$$A^* = U^{-1} V^{-1}_{11} (V_{12} + V_{21}) U \qquad \text{. . . . . (7.10)}$$

It is a standard result that the eigenvalues of A and A^* are identical and further than if $\underline{c}$ is an eigenvector of A for $2\,\Upsilon$ then $\underline{c} = (U^T)^{-1}\,\underline{c}$ is an eigenvector of A^* for $2\,\Upsilon$. If we let $U = [\text{diag } V_{11}]^{-1/2}$, then

$$V^*_{11} = [\text{diag } V_{11}]^{-1/2} \; V_{11} \; [\text{diag } V_{11}]^{-1/2}$$

$$\left.\begin{array}{lll} \text{i.e.} & V^*_{11} = & R_{11} \\ \text{and} & V^*_{12} = & R_{12} \end{array}\right\} \quad \text{matrices of correlation coefficients}$$

So the temporal correlation coefficient (Υ = 1/2 eigenvalue) obtained using the V and R matrices are identical. The $\underline{c}$ vector can be computed from the $\underline{c}^*$ vector (obtained using the correlation matrices) as follows:

$$\underline{c} = U^T \underline{c}^* = [\text{diag } V_{11}]^{-1/2}\, \underline{c}_* \qquad \text{. . . . . (7.11)}$$

Each component of $\underline{c}$ will be the component of $\underline{c}^*$ divided by the standard deviation of the return on that component. Note that with this particular transformation, A^* can be expressed in terms of the R matrices:

$$A^* = R^{-1}_{11} (R_{12} + R_{21}) \qquad \ldots\ldots (7.12)$$

7.6.2 Computing the pooled R matrix

Recall from section 6.9.3 the remarks made concerning the computation of average correlation coefficients. For reasons noted in that section we decided to estimate average or pooled correlation matrices R_{11} and R_{12} over the entire 5 year period using the technique of Donner and Rosser. In the case of coffee, for example, 29 individual estimates of each of the 15 components of R_{11} and 36 components of R_{12} are computed. The pooling is carried out using the expression (6.9.3).

7.6.3 Computing various pooled R matrices

Computing one pooled estimate of R_{11} and R_{12} and the substitution into (7.12) will yield one grand estimate of $\underline{c}^*$ for the entire 5 year period. However this single result will still not inform us as to the stability or otherwise of the population parameters. It was decided to produce a number of pooled estimates of $\underline{c}^*$. Accordingly we produced estimates obtained by pooling R values over period 1 and 2, then another estimate by pooling over periods 1, 2 and 3, and so on until finally we obtain an estimate by pooling over all (29 in the case of coffee series) periods.

7.6.4 The pooled serial correlation coefficients

The various pooled estimates of, r, the serial correlation coefficients are given in Tables 7.12 to 7.15. They are all negative. Note that over the initial 4 or 5 periods there is some small variation in the values but that the estimates rapidly stablize to a value of around - 0.455. The final values are identical to or very slightly smaller than the corresponding pooled estimates of the canonical correlations presented in Table 6.12.

It would appear then, that almost all of the MVSC witnessed in the series could be explained, not necessarily by the complex considerations of $\underline{a}$ and $\underline{b}$ vectors outlined in Chapter 6 but by a single vector $\underline{c}$.

7.6.5 The pooled estimates of c and c*

Note that $\underline{c}$ given by (7.11) in terms of $\underline{c}^*$ requires the diagonal components of $(V_{11})^{-1/2}$, the standard deviations of the returns on the 6 components of $\underline{x}_t$. These standard deviations $(\sigma_1,\sigma_2,\ldots,\sigma_6)$ vary tremendously from period to period and so the question arises as to what values we should use. For consistency we will of course consider both $\underline{c}$ and $\underline{c}^*$ after normalization. What effect does the division by each standard deviation in expression (7.11) have? If all the σ_i's were identical it is easy to see that each $c_i = c^*_i$, $i = 1,\ldots,6$. Recall the study of returns on the four contracts that are nearest maturity in section 5.2.1. We observed that in most (but not all) periods the standard deviations were very similar in magnitude but that generally they tended to be smaller the more distant the maturity date was, i.e. $(\sigma_1>\sigma_2>\ldots>\sigma_6)$. Equation (7.11) implies that the components of $\underline{c}$ associated with the more distant contracts (eg c_5, c_6) will be larger then the corresponding components of $\underline{c}^*$, (c^*_5, c^*_6). By a similar reasoning the values of the components of $\underline{c}$ associated with the contracts that are near maturity (e.g. c_1, c_2) will be smaller than

Table 7.12

Normalized estimates of τ values and c components for cocoa returns using cumulative average correlation matrices

Period	Components 1	2	3	4	5	6	sum of comps.	τ
1	9	78	−382	338	151	42	66	−0.442
2	−72	122	−350	366	−7	−83	−24	−0.400
3	−30	64	−323	372	59	−153	−11	−0.400
4	−31	99	−343	390	−119	17	13	−0.388
5	−22	136	−390	367	−85	0	6	−0.478
6	−25	152	−384	347	−88	4	6	−0.472
7	−10	113	−379	391	−93	−14	8	−0.475
8	−11	115	−389	389	−75	−22	7	−0.477
9	−10	112	−387	391	−87	−13	6	−0.475
10	−10	117	−390	386	−84	−14	5	−0.474
11	−12	119	−377	383	−101	−9	3	−0.475
12	−13	107	−369	395	−109	−8	3	−0.473
13	−3	92	−345	405	−151	5	3	−0.474
14	−4	136	−431	365	−32	−32	2	−0.472
15	2	111	−399	386	−49	−53	−2	−0.469
16	−9	116	−313	387	−128	−48	5	−0.455
17	−7	114	−324	389	−102	−65	5	−0.456
18	−7	119	−337	384	−80	−73	6	−0.456
19	−6	119	−337	383	−99	−55	5	−0.454
20	−9	118	−319	385	−129	−40	6	−0.454
21	−8	109	−304	395	−154	−31	7	−0.453
22	−9	110	−302	393	−160	−26	6	−0.453
23	−9	109	−299	394	−163	−26	6	−0.454
24	−8	111	−306	392	−160	−23	6	−0.452

Table 7.13

Normalized estimates of τ values and c components for coffee returns using cumulative average correlation matrices

Period	Components 1	2	3	4	5	6	sum of comps.	τ
1	-26	7	288	-179	186	-314	-38	-0.716
2	-15	-16	300	-148	191	-330	-18	-0.654
3	26	30	249	68	244	383	14	-0.556
4	18	62	267	145	224	283	17	-0.526
5	32	1	181	173	312	302	15	-0.519
6	-29	-3	-181	-226	-310	252	13	-0.503
7	-28	-4	203	-250	291	-224	-12	-0.504
8	-28	-5	-211	-261	-283	-212	-12	-0.499
9	-22	-20	-221	-254	-272	-210	-13	-0.487
10	29	-2	-191	245	-302	231	10	-0.484
11	-28	-1	-184	-236	-309	-241	-11	-0.483
12	-29	-14	-181	-260	-301	-215	-8	-0.475
13	-27	-6	-171	-237	-319	-240	-8	-0.466
14	-36	-152	-100	-450	-15	-247	-2	-0.455
15	17	-13	89	-91	-394	395	3	-0.475
16	-32	-55	-102	-7	-445	-360	-1	-0.462
17	-38	-72	-18	-231	-426	-214	-3	-0.456
18	27	-3	-124	285	-370	190	5	-0.477
19	27	4	-145	319	-353	153	5	-0.452
20	-27	-11	-153	-324	-345	-140	-4	-0.444
21	-25	-14	-158	-333	-340	-130	-4	-0.437
22	25	21	-160	332	-338	125	5	-0.427
23	26	12	-165	353	-332	112	6	-0.425
24	24	16	-162	356	-336	107	5	-0.440
25	-23	-7	-163	-372	-335	-101	-5	-0.439
26	23	10	-165	366	-332	103	5	-0.443
27	22	8	-138	338	-359	134	5	-0.451
28	-22	-7	-138	324	-360	149	-4	-0.458
29	-22	-4	139	-335	359	-141	4	-0.461

Table 7.14

Normalized estimates of [illegible] values and c components for sugar returns using cumulative average correlation matrices

Period	Components 1	2	3	4	5	6	sum of comps.	[illegible]
1	286	-39	-71	-191	-152	260	93	-0.592
2	279	-198	184	-268	-17	54	34	-0.561
3	332	-206	97	-150	-130	85	28	-0.538
4	69	-280	190	204	38	-218	3	-0.508
5	-102	249	-103	-208	-89	249	-4	-0.491
6	-206	337	-224	94	-75	65	-9	-0.470
7	35	-5	134	-408	-74	344	26	-0.444
8	-9	-24	-160	347	133	-327	-22	-0.438
9	-2	-103	-334	-342	-47	-171	-15	-0.440
10	1	-47	152	-327	349	-124	4	-0.450
11	1	-55	164	-327	337	-117	3	-0.447
12	2	-22	113	-350	387	-126	4	-0.452
13	2	-20	108	-340	391	-138	3	-0.445
14	-11	9	75	-322	416	-167	0	-0.431
15	0	1	75	-326	424	-174	0	-0.416
16	-8	12	64	-309	424	-183	0	-0.414
17	-10	15	55	-291	429	-200	-2	-0.416
18	6	10	66	*315	-423	*180	2	-0.412
19	4	-14	-62	331	-424	164	-1	-0.415
20	1	-5	-73	332	-422	167	-0	-0.414
21	3	-11	-68	327	-422	169	-2	-0.413
22	3	-12	-65	336	-425	160	-3	-0.411
23	5	-15	-62	331	-424	163	-2	-0.412
24	7	-20	-56	330	-425	162	-2	-0.412

Table 7.15

Normalized estimates of τ values and c components for rubber returns using cumulative average correlation matrices

Period	Components 1	2	3	4	5	6	sum of comps.	τ
1	83	148	102	-482	1	184	36	-0.555
2	-49	88	-161	299	-303	99	-27	-0.468
3	63	-52	3	-204	457	-221	46	-0.458
4	-38	17	133	72	-479	260	-35	-0.403
5	-24	69	-138	274	-351	144	-26	-0.470
6	-14	44	-129	330	-372	111	-30	-0.460
7	-19	70	-151	278	-336	146	-12	-0.474
8	-24	73	-137	300	-345	122	-11	-0.466
9	-25	71	-130	305	-354	116	-17	-0.463
10	-27	80	-112	281	-368	131	-15	-0.462
11	-24	-80	-127	-288	-355	-126	-12	-0.455
12	-26	87	-134	274	-344	134	-9	-0.454
13	-24	82	-133	279	-347	135	-8	-0.454
14	-22	80	-129	275	-351	142	-5	-0.453
15	-22	79	-121	272	-360	145	-7	-0.453
16	22	-81	119	-269	361	-148	4	-0.452
17	20	-74	109	-269	373	-154	5	-0.454
18	20	-74	102	-263	379	-161	3	-0.453
19	21	-70	105	-269	376	-159	4	-0.453
20	20	-69	102	-267	380	-162	4	-0.451

the corresponding componenets of $\underline{c}^*$ (e.g. c^*_1, c^*_2). Equation (7.11) affords a gentle rescalling of $\underline{c}^*$ to $\underline{c}$.

However the comparison of some estimates of $\underline{c}^*$ and $\underline{c}$ from a number of sub samples studies shows the difference to be remarkably small. Because of this slight difference and the ambiguity as to what values of σ_1, σ_2 etc to choose we simply report the various pooled $\underline{c}^*$ vectors computed from (7.12) and bear in mind that the far components should probably be slightly larger and that the near components should be slightly smaller. See Tables 7.12 to 7.15 for the $\underline{c}^*$ estimates.

Referring to Tables 7.12 to 7.15 one sees that, as with the serial correlation coefficients, $\hat{r}$, there is an initial instability in the pattern of the individual components of $\underline{c}^*$. However after the first 6 or 7 periods the pooling seems to produce remarkably stable $\underline{c}^*$ vectors. The final estimates of $\underline{c}^*$ are almost identical to the grand average estimates of $\hat{\underline{a}}$ and $\hat{\underline{b}}$ vectors produced in section 6.9.

What are we to conclude from these results? It certainly looks as if it is possible that there is an underlying constant $\underline{c}$ vector. The vector is slightly different for each future series but the overall picture is very much the same. The components with the highest contribution to the portfolio of returns are the 4th and 5th in the case of the coffee, sugar and rubber series and the 3rd and 4th in the cocoa series. In 4 series the grand average pooled results are almost a perfect multivariate spread.

7.7 Conclusions of Chapter 7

The discovery of persistent multivariate serial correlation outlined in Chapter 5, its special nature investigated in Chapter 6 and the discovery, in this chapter, of a particular portfolio $\underline{c}$ that could explain most of the observed temporal dependence suggests a possible model for the multivariate distribution of prices. This model is further investigated in Chapter 9.

In Chapter 8 we consider the application of a number of trading rules based on these temporally dependent portfolios.

Footnotes for Chapter 7

1. These $\hat{\lambda}_{max}$ values have been computed on the unlogged returns.
2. We are using the term perfect spread here to mean a net overall neutral position. The resulting spread may of course not be perfect in the sense of zero risk.
3. The $P(Z_r)$ value associated with this result is 0.004.

CHAPTER 8

TRADING RULES

In this chapter we examine the application of three trading rules designed to exploit the observed multivariate serial correlation. The layout of this chapter is as follows. In section 8.1 we outline the results of a study into the temporal dependence of ex ante portfolios. In section 8.2 the general idea of a scheme that could exploit excessive price reversaling is introduced. In section 8.3 the first trading rule (rule 1) is outlined along with a discussion of the results obtained by applying the rule to all four series. In section 8.4 a slightly more sophisticated rule (rule 2), one which is designed to produce a smoother flow of returns, is described. In section 8.5 we consider the impact of transaction costs on these rules and in section 8.6 we explicitly incorporate such costs into our final trading rule, rule 3. In section 8.7 we compare the returns from applying these rules to the returns received by investing in single commodity futures contracts and in the British Stock Market over the same period. In section 8.8 we consider the application of simulationeously applying rule 3 to all four series. The practical limitations of applying these rules is dealt with in section 8.9 and concluding comments are made in secion 8.10.

8.1 An ex ante analysis of temporally dependent portfolios

In Chapter 7 we demonstrated that it was possible to find in each period ex post portfolios of commodity futures contracts that exhibited negative serial correlation but that the portfolios were not the same from period to period. A trading rule designed to exploit negatively serially correlated returns series must be able to select ex ante portfolios that are still negatively serially correlated. How can this be achieved? We consider three possibilities:

(i) Use a grand average estimate based on all previous periods. If we believe the underlying population $\underline{c}$ vector is fixed, each additional period will provide information on the vector. Using the grand average technique of section 7.7.4 one could pool all the estimates of the R matrices from period 1 to i in order to obtain an estimate of $\underline{c}$ for period i + 1.

(ii) If the underlying $\underline{c}$ vector is changing, then the grand average technique suggested in (i) above will not be appropriate. Averaging the R matrices over all periods will not yield valid estimates for $\underline{c}$ in period i + 1. It may be more useful to consider a weighted averaging technique, a procedure that will place more weight on the most recent estimates. The author has experimented with exponentially smoothed estimates of correlation matrices and has encountered problems, particularly in arriving at positive definite estimates.

(iii) Although we see from Tables 7.5 to 7.8 that the $\hat{\underline{c}}$ estimates vary from period to period, we could let $\hat{\underline{c}}_i$ be the predictor of $\underline{c}_{i+1}$. If the underlying $\underline{c}$ vector is changing from period to period this procedure will surely be optimal[1].

It was decided to find which of the schemes mentioned above would yield superior ex ante estimates of $\underline{c}$. A measure of the effectiveness of these estimates will be the degree of negative serial correlation witnessed in period i + 1 using the estimates derived in period i. Accordingly we used methods (i) and (iii) above to compute estimates of

$\underline{c}$ in period i. The serial correlation coefficients of the resulting portfolio of returns $\hat{\underline{c}}^T{}_i\underline{x}$ in period (i + 1) were computed using estimates of $\underline{c}$ from period i. The method that yields the ex ante portfolios with the maximum negative serial correlation will be superior. In Tables 8.1 to 8.4 we list the results of this study. Also given in these tables are the maximum negative serial correlation coefficients that one could have obtained ex post.

Referring to Tables 8.1 to 8.4 we note that, with the exception of one period in the coffee series, all (i.e. 92) ex ante portfolios exhibit negative serial correlation coefficients using methods (i) and (iii). This is an extremely encouraging result. It appears then that it is possible to contruct ex ante portfolios that will in future periods exhibit the hoped for excesive price reversaling.

We now turn our attention to the question as to which method produces superior results. Consider the ex ante coefficients in Tables 8.1 to 8.4. In each period an asterisk appears in either column 2 or 3. The asterisk denotes the particular technique which produced the largest negative serial correlation. The results for the cocoa, sugar and rubber series are very similar. The ratio of the total number of periods in which method (i) is superior to method (iii) is almost exactly 2 to 1 in each of the 3 series. For the coffee series the result is almost exactly the opposite. These results are mirrored in the average ex ante serial correlation coefficients. For the cocoa, sugar and rubber series the average ex ante serial correlation coefficient is more negative using method (i) and for the coffee series the average using method (iii) is more negative. It must be pointed out, however, that the difference produced by the two methods is very small indeed.

For comparison, the averages of the ex post serial correlation coefficients are also given in Tables 8.1 to 8.4. In all 4 series the ex

Table 8.1

Correlations between $\hat{c}^T x_t$ and $\hat{c}^T x_{t-1}$ for cocoa series

Period	Ex post minimum	Ex ante method (iii)	Ex ante method (i)
2	-.589	-.037*	+.110
3	-.586	-.323*	-.105
4	-.512	-.261	-.268*
5	-.691	-.538	-.566*
6	-.684	-.344*	-.281
7	-.673	-.110	-.443*
8	-.615	-.415	-.475*
9	-.526	-.396*	-.283
10	-.665	-.223	-.297*
11	-.598	-.539*	-.499
12	-.521	-.209	-.303*
13	-.648	-.109	-.444*
14	-.607	-.042	-.387*
15	-.696	-.442	-.470*
16	-.733	-.515	-.536*
17	-.598	-.428	-.454*
18	-.607	-.372	-.395*
19	-.565	-.370*	-.205
20	-.597	-.511*	-.447
21	-.659	-.087	-.293*
22	-.567	-.336	-.399*
23	-.625	-.331	-.568*
24	-.482	-.229	-.329*
averages:	-.611	-.313	-.362
frequency of *'s :		7	16

* indicates when a particular method produces the smallest correlation

Table 8.2

Correlation between $\hat{c}^T x_t$ and $\hat{c}^T x_{t-1}$ for coffee series

Period	Ex post minimum	Ex ante method (iii)	Ex ante method (i)
2	-.825	-.373*	+.448
3	-.611	-.422*	-.201
4	-.620	-.008	-.085*
5	-.692	-.522*	-.381
6	-.621	-.455*	-.324
7	-.672	-.603*	-.510
8	-.614	-.210	-.323*
9	-.561	-.355*	-.343
10	-.678	-.412*	-.382
11	-.593	-.414	-.431*
12	-.750	-.318*	-.229
13	-.726	-.383*	-.273
14	-.695	-.340*	-.264
15	-.694	-.252*	-.185
16	-.569	-.468*	-.424
17	-.530	-.464*	-.363
18	-.620	-.523*	-.422
19	-.406	-.347*	-.307
20	-.589	-.443*	-.261
21	-.516	-.295	-.403*
22	-.553	-.431*	-.251
23	-.610	-.212	-.297*
24	-.640	-.253	-.613*
25	-.716	-.279	-.407*
26	-.532	-.278*	-.163
27	-.718	-.453	-.487*
28	-.627	-.378	-.554*
29	-.911	-.124	-.386*
averages:	-.639	-.331	-.315
frequency of *'s :		18	10

TABLE 8.3

Correlation between $\hat{c}^T x_t$ and $\hat{c}^T x_{t-1}$ for sugar series

Period	Ex post minimum	Ex ante method (iii)	Ex ante method (i)
2	-.552	-.338*	-.201
3	-.765	-.085*	-.043
4	-.496	-.428*	-.025
5	-.544	-.313	-.322*
6	-.631	-.412*	-.366
7	-.571	-.381	-.433*
8	-.711	-.342	-.355*
9	-.561	-.425*	-.365
10	-.624	-.318	-.475*
11	-.589	-.229	-.258*
12	-.537	-.223	-.430*
13	-.478	-.299	-.324*
14	-.698	-.463*	-.217
15	-.599	-.287*	-.235
16	-.572	-.273	-.436*
17	-.566	-.420	-.454*
18	-.588	-.003	-.035*
19	-.634	-.460	-.539*
20	-.559	-.327	-.364*
21	-.589	-.432*	-.345
22	-.577	-.319	-.388*
23	-.591	-.061	-.462*
24	-	-.075	-.349*
averages:	-.592	-.301	-.323
frequency of *'s :		8	15

Table 8.4

Correlation between $\hat{c}^T x_t$ and $\hat{c}^T x_{t-1}$ for rubber series

Period	Expost minimum	Ex ante method (iii)	Ex ante method (i)
2	-.554	-.152	-.194*
3	-.544	-.344	-.360*
4	-.563	-.311*	-.127
5	-.602	-.209	-.489*
6	-.487	-.311	-.345*
7	-.632	-.439	-.510*
8	-.533	-.335	-.420*
9	-.616	-.429	-.465*
10	-.517	-.386	-.426*
11	-.528	-.235*	-.170
12	-.658	-.243	-.387*
13	-.554	-.372	-.511*
14	-.607	-.443*	-.324
15	-.643	-.458*	-.455
16	-.579	-.252	-.336*
17	-.664	-.564*	-.561
18	-.651	-.434*	-.294
19	-.526	-.330	-.414*
20	-.853	-.279	-.279*
averages:	-.595	-.343	-.372
frequency of *'s:		6	13

ante portfolios produced using methods (i) or (iii) have a much reduced degree of negative serial correlation: typically about half the negative serial correlation that could have been obtained ex post.

8.2 A reverse filter rule

How can one exploit the negatively serially correlated return series obtained in all the sub periods examined? One possibility would be to investigate the modelling of these series using Box - Jenkins techniques. If the ex ante portfolios could be explained by simple moving average processes originally mentioned in section 7.5.2, then it may be possible to forecast the price of the series at some future time horizon and trade accordingly.

However a trading rule that is possibly much simpler to consider would be the use of a reverse filter rule. One such rule could be : if the price of the portfolio moves up a large amount, say x%, sell and close out the next day, if the price falls a large amount, say x% buy and close out the next day. Loeb (1979) noticed a small degree of negative serial correlation in some of the simple spread series he examined and experimented with this type of reverse filter rule. His results were disappointing, with the rule only invoking one trade about every two years. In the next section we set out in detail the first of three trading rules used in this study.

8.3 Trading rule 1

To outline the rationale and the results of the trading rule it is helpful to consider a particular sub-period in detail. To this end we present as an example the ex ante portfolio prices of the 17th rubber series plotted in Fig. 8.1. Below we explain how this series is arrived at.

Recall from section 8.1 that it was found that of the two suggested methods for calculating $\hat{\underline{c}}$ (the one using the average correlation matrices of all the previous periods) seemed to produced optimal results[2]. Accordingly the $\hat{\underline{c}}$ vector used in constructing the ex ante portfolio in our example was computed using the average correlation matrices of all the previous 16 periods. The individual components of this $\hat{\underline{c}}$ vector appear in row 16 of Table 7.15. Note that, as in all periods, the sum of the $\hat{\underline{c}}$ components is near zero, representing a near perfect spread. The sum of all the terms corresponding to long positions is very near +500 and the sum of all the terms corresponding to short positions is very near -500. It is thus possible to think of the portfolio as a complex, multiple spread of 500 long and 500 short positions. These numbers are very large and it would be much more instructive to consider the portfolio in terms of a typical single spread, i.e. one of 1 long and 1 short position. Accordingly the $\hat{\underline{c}}$ vectors are normalized so that $\sum_{i=1}^{6} |\hat{c}_i| = 2$. This proves to be particularly useful later on when we have to take into account deposit considerations. The time series plot of the 62 daily prices of the ex ante portfolio is given in Fig. 8.1. The units are in £'s per unit spread (1 net long and 1 net short position).

Note that more than 50% of the prices in Fig. 8.1 are negative. These represent situations in which the contracts that are held in short positions have on average, a higher price than the contracts that are held in long positions. The use of a % type reverse filter rule is thus impracticable. Instead we consider an absolute price change type reverse filter rule.

Consider the returns on the above portfolio. The price differences, are plotted in Fig. 8.2. We construct a fixed band above and below the

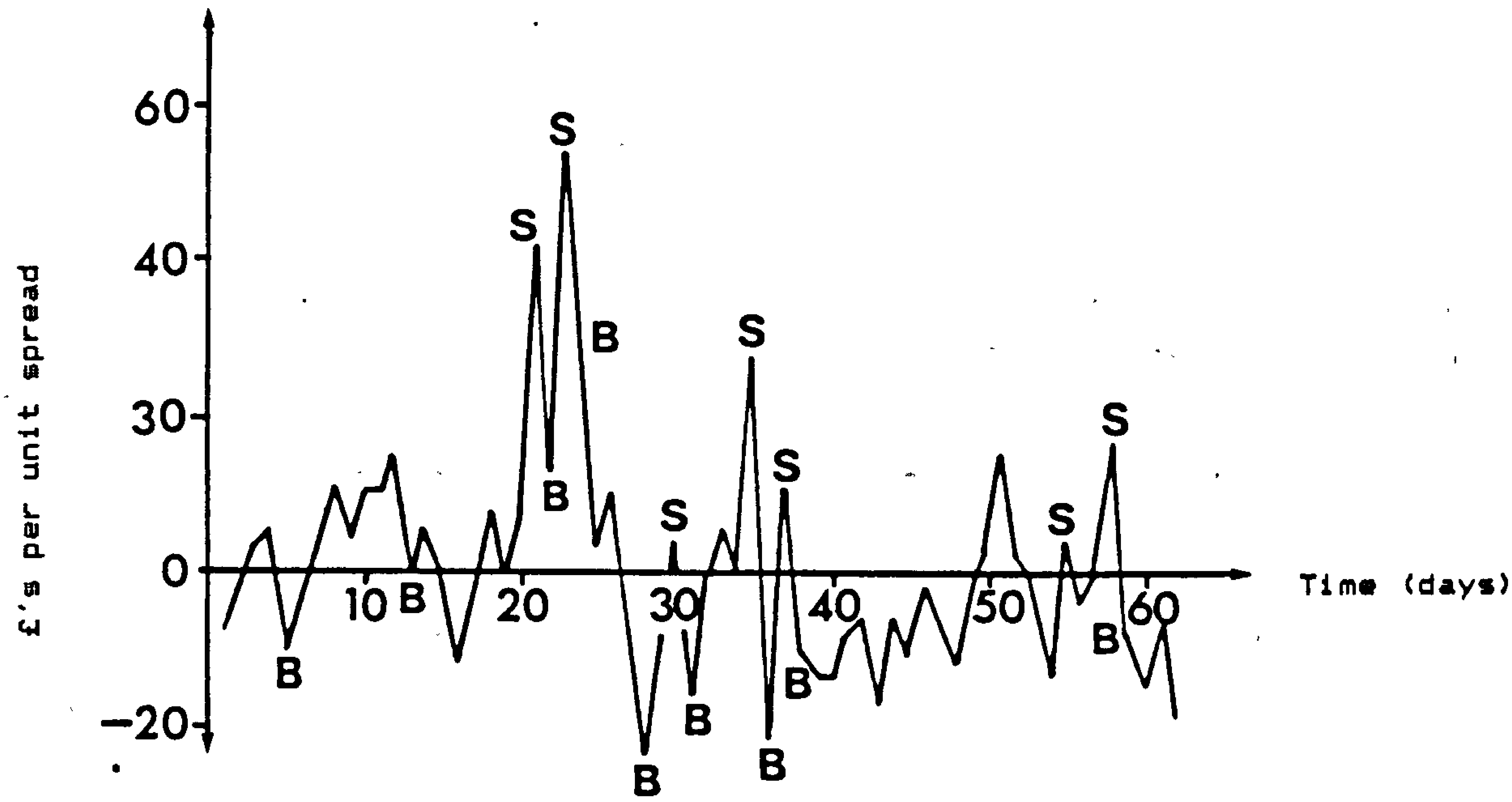

Fig 8.1 Prices of ex ante portfolio in the 17th rubber series

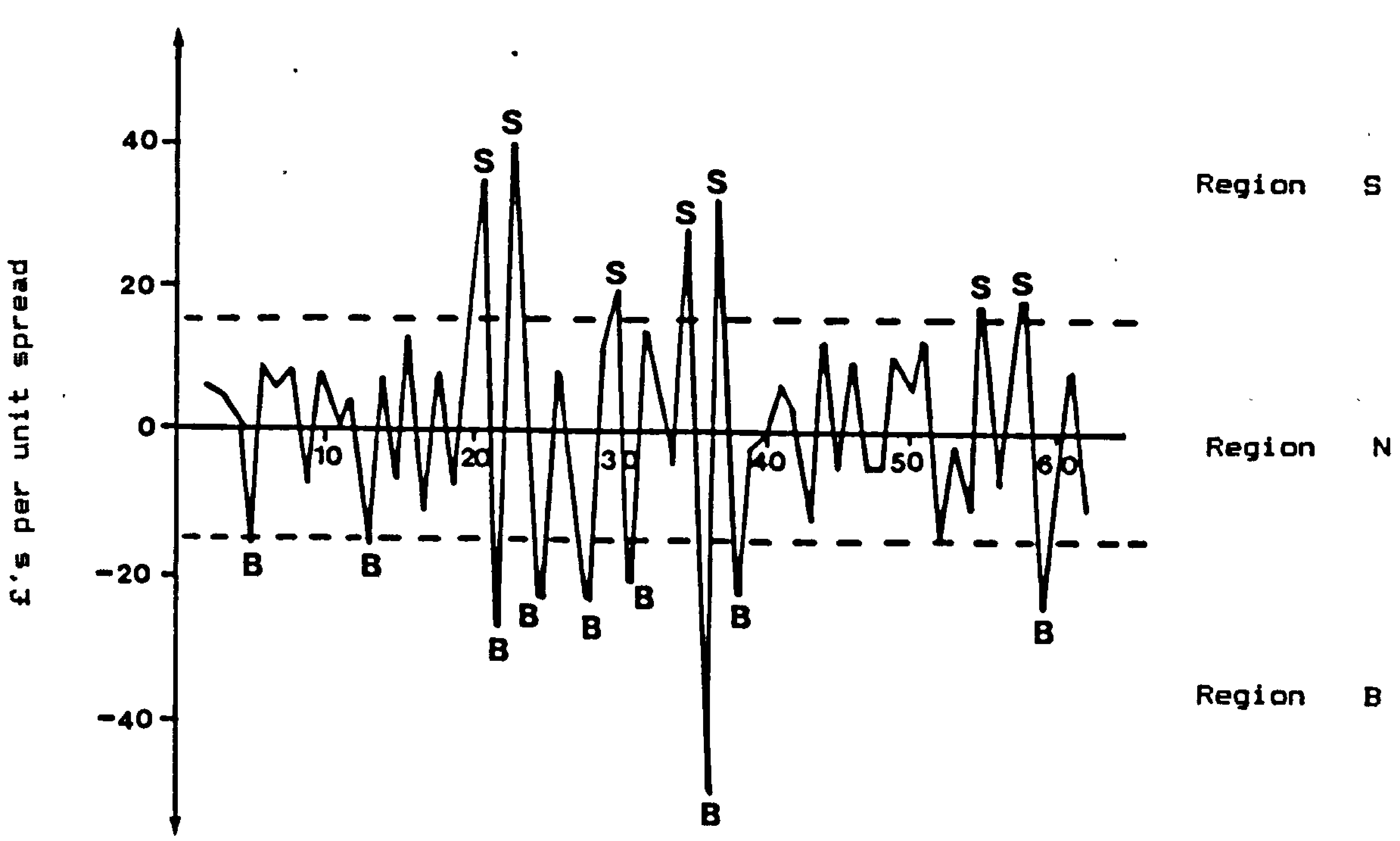

Fig 8.2 Returns of ex ante portfolio in the 17th rubber series

Table 8.5 - Trading Rule 1 on 17th Rubber Series

Width of N band in Fig. 8.2 = ± £15.12 = ± 1.4 x St. dev of returns in in 16th period

No. of days into period 17 signal occurs	Signal (B=buy,S=sell)	Profit received on closing the following day in £'s
5	B	8.84
13	B	7.28
21	S	28.16
22	B	40.98
23	S	23.38
24	B	-27.74
28	B	10.72
30	S	20.84
31	B	14.48
35	S	49.46
36	B	32.78
37	S	22.14
38	B	-2.72
55	S	7.86
58	S	25.14
59	B	-6.58
		£ 255.00

Total no. of buy signals = 9

Total no. of sell signals = 7

Total no. of signals = 16

Total profit on period = £255.00

Average profit per trade = £15.94

zero return line. This band encompasses a region we designate as region N (for null). The rule is as follows. If the price of the portfolio rises significantly, the returns will leave region N and enter the region S and we interpret this as a sell signal [denoted S in Figs. 8.1 and 8.2]. The portfolio is sold at that price and the position is closed out the very next day, whatever happens to the price, by a purchase. Using similar reasoning, if the price falls significantly the returns will enter region B and we interpret this as a buy signal [denoted B in Figs. 8.1 and 8.2]. The portfolio is bought at that price and the position is closed out the very next day by a sale, whatever happens to the price. Note that if the returns plot swings from region S straight into region B our rule dictates that the original position (sold short) be closed out and a further, opposite position be entered by another buy transaction.

In the use of this rule on the data set given it should be noted that we are making three major assumptions:

(i) One can get transactions at the close of day prices on <u>that</u> day.

(ii) Transactions are executed at the one price - the seller prices used in this study.

(iii) There are no transaction costs involved.

Later in this chapter we attempt to bring transaction costs into account. We return to the problems associated with assumption (i) and (ii) in section 8.9.

Table 8.5 gives in more detail the results of applying the trading rule to the 17th rubber series. A profit of £255 or £15.96 per trade over the short period of 62 days is very encouraging, particularly when one considers that at that time the deposit recommended by the ICCH for traders in rubber futures dealing in spreads was £300 per spread. Before we go on to consider the results for the entire 5 year period, it is important to note why the band width of region N was set to ± £15.12.

It was noticed in studying many of the ex post and ex ante returns series that a good deal of the negative serial correlation could be attributed to the larger swings. The reverse filter rule was constructed to exploit these large swings. The width of the region N should be set in such a way as to produce the maximum number of profitable signals, i.e. signals that result in large and opposite returns the very next day. If the band is set too narrow, many signals will occur with the possible execution of many loss-making trades. If the band width is set too wide, the frequency of signals would be so low as to make the trading rule inpractical. It was decided therefore that a useful technique would be to set the band to a suitable multiple of the standard deviation of the returns. The question then arises "what is the standard deviation of the returns"? We cannot use the ex post value for period 17 and so an obvious and reasonable candidate is the standard deviation of the returns of the most recent series. In this particular example we set the band width to a value of 1.4 times the standard deviation of the returns estimated from period 16.

We now consider the results of applying the trading rule to all of the 19 rubber series. Note that the first period of 58 days is used as an initialization period - one in which the first $\hat{\underline{c}}$ vector and the first standard deviation of returns are computed. In each subsequent period, i, the $\hat{\underline{c}}_i$ are computed using the averages of the correlation matrices of all the previous i - 1 periods. Also in each period, i, the width of region N is set to 1.4 times the standard deviation of the returns in the most recent period, i - 1. A brief summary of the profits in each period appear in Table 8.6 and we make the following observations:

Table 8.6

Trading Rule 1 on complete rubber series with scale factor set to 1.4

Period	St.dev. of returns	no.of days	no.of buy sigs	no.of sell sigs	no.of trades	no.of days per trade	profit	profit per trade
1	24.00	58	-	-	-	-	-	-
2	26.73	64	5	4	9	7.1	62.50	6.94
3	15.99	65	2	0	2	37.5	-4.24	-2.12
4	15.99	62	5	5	10	6.2	35.88	3.56
5	31.95	63	12	12	24	2.6	100.86	4.20
6	28.98	65	2	4	6	10.8	33.10	5.52
7	69.51	65	16	12	28	2.3	983.78	35.14
8	30.66	61	0	0	0	-	0.00	-
9	33.57	64	7	5	12	5.3	331.46	27.62
10	19.95	63	1	1	2	31.5	93.84	46.92
11	14.19	65	1	2	3	21.7	-23.76	-7.92
12	7.80	61	0	1	1	61.0	-2.44	-2.44
13	14.31	62	12	11	23	2.7	127.18	5.53
14	12.48	15	6	4	10	6.5	116.52	11.65
15	9.36	65	1	2	3	21.7	44.92	14.97
16	10.80	61	7	4	11	5.5	61.80	5.62
17	16.26	62	9	7	16	3.9	255.00	15.94
18	10.02	65	1	0	1	65.0	19.62	19.62
19	18.75	65	11	9	20	3.3	194.00	9.70
20	21.27	17	1	1	2	8.5	51.84	25.92
Totals			99	84	183		2481.86	

Average no. of days/trade = 6.34   Average profit/trade = £13.56

Average no. of trades/month = 3.33

St. dev. of no. of trades/month = 3.32

Monthly returns average = £45.13

Monthly returns st. dev. = £85.48

Coefficient of monthly variation = 1.89

(i) A total profit of £2481.86 with 183 trades, or £13.56 per trade is encouraging.

(ii) Consider all 19 periods. All but 4 result in positive profits. We return to the question of the statistical significance of these profits in section 8.3.5.

(iii) The numbers of buy (99) and sell (84) signals in each period are approximately equal. This result is not significantly different from a 50 to 50 ratio.

(iv) The number of signals in each period varies considerably. A maximum of 28 signals occurred in period 7. No signals occurred in period 8. For a useful trading rule one would like a fairly constant rate of trading opportunities. Trading every 2 or 3 days in period 7 followed by a wait of 61 days without a trade in period 8 does not seem very satisfactory. Examination of column 2 of Table 8.6 reveals that a possible cause of the large fluctuations in the number of trading signals per period is the variation in the standard deviation of returns from period to period. In period 6 the standard deviation was found to be £28.98. Multiplying this by 1.4 we set the width of band N in period 7 to be ± £40.57. In period 7 the returns turned out to be more than twice as varied with a standard deviation of £69.51. Not surprisingly many (28) trading signals were observed resulting in a large profit of £983.78. In the next period, period 8, the band width is set to ± 1.40 x 69.51 = ± £97.31 a very wide band. After period 8 the standard deviation of returns had fallen back to £30.66, with the net result being that no trading signals were generated. However, although there are a number of such instances in which large changes in the variation of returns produce 'too many' or too few trades, the overall profits are very encouraging.

8.3.1 Monthly returns to trading rule 1

In section 8.7 we will be comparing the returns to our trading rule with returns from the British Stock Market. There is an immediate difficulty in comparing two such streams of returns. In many empirical studies of the stock market the standard measure is the monthly return. One considers buying a stock on the first day of the month and selling the stock on the last day of the month, recording the return as a percentage of the original investment. Throughout the month the investor has a given sum of money completely committed to the stock. This sum of money is at risk for the entire month and the return constitutes one payment at the end of each month.

For an individual using trading rule 1, in rubber say, a deposit of £300, approximately, is placed in an account and occasionally (in our example on average once every 6 or 7 days) a trade is invoked. Only at these times is the individual exposed to risk and the returns thus constitute a stream of profits over unequal intervals of time.

In order to compare the trading rule with the more standard assests it was decided to present the trading rule returns as monthly returns. This was done by accumulating profits generated by the rule over each month and reporting these as a single sum. In some months, say, up to 5 or so trades would thus produce a single profit payment that we consider to be received on the last day of the month. Although the risks associated with the rule and investing in the stock market are not comparable one can simply look at the two streams of monthly returns. Accordingly in all that follows we report, along with other information, the means and the standard deviations of the monthly returns in £'s per unit spread traded. Note that the entire period covers 55 months.

8.3.2 Deposits required for spread trading

At this point it is also useful to consider the deposits recommended by the ICCH for traders in spreads. Although there has been a good deal of variation throughout the period, typical deposits required for spread trading were:

Cocoa : £450, Coffee : £400, Sugar : £500 and Rubber : £300

These are the minimum sums required to hold simultaneously 1 long and 1 short position.

8.3.3 Adjusting the number of trade signals. The scale factor

In the above example we had defined the region N to be set rigidly to $\pm$ 1.4 times the standard deviation of returns (estimated from the previous period). In the entire rubber series 183 trading signals were observed. An obvious way to alter the number of trading signals would be to alter the width of the band N. Reducing the width (choosing a smaller multiplier) will probably increase the number of signals, increasing the width (choosing a larger multiplier) will probably decrease the number of signals. In all that follows we will now refer to this multiplier as the scale factor.

The rule was applied to all 4 series with many values of the scale factor. For brevity we present the results using scale factors of 0.6, 1.0, 1.4, 1.8, 2.2, 2.6 and 3.0 in Tables 8.7 to 8.10.

Table 8.7

Trading rule 1 applied to complete cocoa series

Scale	no. of trades	profit (£'s)	no.of days/ trades	profit/ trade (£'s)	Days/month μ_d	Days/month σ_d	Monthly returns mean (£'s)	Monthly returns st. dev. (£'s)	Monthly returns coeff of var. (ν)
.6	424	9435	2.74	22.25	7.7	3.5	171.55	276.28	1.61c
1.0	249	8486	4.7	34.08	4.5	3.3	154.29	263.87	1.71c
1.4	153	6300	7.6	41.18	2.8	2.7	114.55	230.39	2.09c
1.8	96	4831	12.1	50.32	1.8	2.0	87.84	208.25	2.37b
2.2	65	3927	17.9	60.42	1.2	1.7	71.40	210.37	2.95a
2.6	41	4239	28.3	103.39	0.8	1.3	77.07	181.40	2.35b
3.0	33	3972	35.2	120.36	0.6	1.1	72.22	182.88	2.53b

μ_d = average number of trades per month
σ_d = st. dev. of number of trades per month

Test of significant of monthly returns : small ν are significant

If ν < 3.71 returns are sig. diff. from zero at 5% level (a)
If ν < 2.79 returns are sig. diff. from zero at 1% level (b)
If ν < 2.12 returns are sig. diff. from zero at 0.1% level(c)

Table 8.8

Trading rule 1 applied to complete coffee series

Scale	no. of trades	profit (£'s)	no.of days/ trades	profit/ trade (£'s)	Days/month μ_d	Days/month σ_d	Monthly returns mean (£'s)	Monthly returns st. dev. (£'s)	Monthly returns coeff of var. (ν)
.6	459	6034	2.53	13.15	8.4	3.6	109.71	189.83	1.73c
1.0	290	4469	4.00	15.41	5.3	3.5	81.25	125.19	1.54c
1.4	184	4025	6.30	21.88	3.4	2.8	73.18	126.28	1.73c
1.8	125	3304	9.3	26.43	2.3	2.5	60.00	117.81	1.96c
2.2	70	2508	16.6	35.83	1.3	1.5	45.60	110.21	2.42b
2.6	40	1994	29.0	49.85	0.7	1.1	36.25	100.11	2.76b
3.0	28	1505	41.4	53.75	0.5	0.9	27.36	93.91	3.43a

Table 8.9

Trading rule 1 applied to complete sugar series

Scale	no. of trades	profit (£'s)	no. of days/ trade	profit/ trade (£'s)	Days/month μ_d	Days/month σ_d	Monthly returns mean (£'s)	Monthly returns st. dev (£'s)	coeff of var. (ν)
.6	423	3401	2.7	8.04	7.7	3.2	61.84	82.76	1.34c
1.0	236	2633	4.9	11.16	4.3	2.8	47.87	89.48	1.87c
1.4	129	1791	9.0	13.88	2.4	2.2	32.56	82.22	2.52b
1.8	67	1346	17.3	20.08	1.2	1.7	24.46	80.89	3.28a
2.2	40	805	29.0	20.12	0.7	1.3	14.63	41.39	2.83a
2.6	30	417	38.7	13.90	0.6	1.2	7.58	28.51	3.76
3.0	23	348	50.4	15.13	0.4	1.1	6.33	28.06	4.43

Table 8.10

Trading rule 1 applied to complete rubber series

Scale	no. of trades	profit (£'s)	no. of days/ trade	profit/ trade (£'s)	Days/month μ_d	Days/month σ_d	Returns Monthly mean (£'s)	Returns Monthly st. dev (£'s)	coeff of var. (ν)
.6	457	3541	2.5	7.74	8.3	3.4	64.38	92.69	1.44c
1.0	278	2866	4.2	10.31	5.1	3.5	52.12	88.33	1.69c
1.4	183	2482	6.3	13.56	3.3	3.2	45.12	85.48	1.89c
1.8	119	2191	9.8	18.41	2.2	2.8	39.83	80.41	2.02c
2.2	78	1553	14.9	19.91	1.4	2.0	28.23	56.81	2.01c
2.6	52	1088	22.3	20.92	1.0	1.5	19.78	47.47	2.40b
3.0	36	904	32.2	25.11	0.7	1.2	16.44	49.82	3.03a

8.3.4 General discussion of results of returns to trading rule 1 on all four series

Referring to Tables 8.7 to 8.10 we note that, not surprisingly, the profits obtained in each series are not exactly identical. However the dependence of the total returns and the number of trading signals on the value of the scale parameter are remarkably similar across all series. We make the following remarks on the general patterns observed:

(i) For each value of the scale parameter we note that the number of trades invoked in each series is very similar.

(ii) As expected, small values of the scale parameter produce a large number of trades. Increasingly the scale parameter reduces the number of trades.

(iii) With the scale parameter set to 0.6 the average number of days between trades is approximately 2.6. With the largest value of the scale parameter set to 3.0, the average number of days between trades increases to about 45.

(iv) Large positive profits result with small scale values. Increasing the scale value reduces the number of trades <u>and</u> reduces the total profit.

(v) Although the total profits and the number of trades decrease with increasing scale values, the average profit per trade is small for small scale values and increases with increasing scale values (to a spectacular £120.36 for cocoa series).

(vi) The mean monthly returns and the standard deviation of monthly returns are high for small scale values and decrease with increasing scale values.

(vii) The rate of the decrease in the monthly means and standard deviations of the returns as the scale value increases is however not identical. The coefficient of variation (v = standard deviation / mean) is an indicator of these different rates of dependence on the scale value. Coefficients of variation are small for small scale values and large for large scale values. From this we conclude that increasing the scale value reduces the mean monthly return proportionately more than the standard deviation of the monthly returns.

The frequency of trading signals and the association between the returns and the scale values are remarkably similar across all series. We now discuss in more detail the magnitude of returns generated by rule 1 in each series.

Without exception, for each scale value, the ranking of the returns as measured by total profit, profit per trade, or profit per month are as follows:

cocoa > coffee > rubber > sugar

Over the whole period of 55 months the cocoa series gave the most spectacular total returns of £9,435, and the sugar series gave the least total returns, £3,541. When one considers that the initial deposits required to achieve these profits are £450 and £500 respectively indicating profits of 2,000% and 700%, the rule seems to be a very successful one.

As early as the original works by Alexander (1960) and Houthakker (1959), authors have been reporting large positive returns by the application of simple x% filter type rules to commodity futures markets. Recall however that in these early works only certain values of the filter level produced positive profits. Many filter values produced negative profits and nowhere has the statistical significance of the returns been tested. Note that in our study no negative total returns were found at all, with any value of the scale factor.

Finally, an important consideration in the application of any trading strategy is the level of risk. Usually high returns are associated with high risk. There is no exception in the returns to rule 1. Consider as an example the monthly returns on the cocoa series with the scale value set to 0.6. Although an investor would be receiving an average of £171.25 per month for a deposit of £450, the monthly standard deviation of returns is very high : £276.28. Consider the, not unlikely, possibility of two successive months in which the returns are 1 standard deviations below the mean (i.e. -£242.87). In such situations the deposit of £450 would be wiped out. In reality then, one would need to maintain deposits larger than the minimum sums recommended by the ICCH[3].

8.3.5 Statistical significance of returns to rule 1

The total profits and profits per trade to rule 1 on first sight seem extremely encouraging across all commodity series. Of course with any trading rule it is possible to end up with positive profits through sheer chance. What is needed is a test of significance. If the rule works, the returns should on average be greater than zero. An obvious and simple way to investigate this is to examine all the individual returns produced in each series. Consider in more detail, for example, the returns of the rule on the rubber series with a scale value of 1.0. There were n = 278 trades, an average of £10.31 per trade and the standard deviation of returns (not reported in Table 8.10 for brevity) of £31.48. We carry out the classic two sided t test:

$$t = \frac{10.31 - 0}{\sqrt{\dfrac{31.48}{278}}} = +30.64$$

a highly significant result.

This test was carried out on all series with all scale values and with the exception of the returns on two high scale values on the sugar series all returns proved highly significant.

An equivalent test can be performed on the monthly returns using the coefficients of variation as follows:

$$t = \frac{\bar{\mu} - 0}{\sigma/\sqrt{n}} = \frac{\sqrt{n}}{\nu}$$

or equivalently $\nu = \dfrac{\sqrt{n}}{t}$, with $n = 55$

Thus small ν statistics in Tables 8.7 to 8.10 correspond to large t statistics and significant values are : 3.71, 2.79 and 2.12 at 5%, 1% and 0.1% levels respectively.

Observe from the tables that with small scale values all results are significant at the 0.1% level, at medium scale values some are significant only at the 1% level and all but 2 results at the high scale values are significant at the 5% level.

This then confirms, statistically, that the returns to rule 1 (in the absence of transaction costs) are significantly greater than zero. The chance of receiving these very large positive returns with a series of prices that follows the classic random walk is less than 1 in a 1000.

8.4 Trading rule 2

We now consider a slightly different rule in which the standard deviation of returns within a period is continuously updated using the iterative relation

$$\hat{\sigma}^2_t = \alpha \hat{\xi}^2_t + (1 - \alpha) \hat{\sigma}^2_{t-1}$$

where $\hat{\xi}_t = \hat{\underline{c}}^T \underline{x}_t$ = return on day t

$\hat{\sigma}_t$ = estimate of standard deviation of returns on day t

α = a smoothing constant ϵ (0,1)

In this relation we obviously assume $E(\hat{\xi}_t) = 0$. It is easy to show, that if the returns are generated from a process with a constant standard deviation σ, then:

$$E(\hat{\sigma}^2_t) = \sigma^2 \qquad \text{as } t \longrightarrow \infty$$

This new procedure is one in which each day a new estimate of the standard deviation is derived as an exponentially smoothed average of the previous days' estimates. Using this new technique, when the standard deviation of returns increases (decreases) the band width for the trading rule will be automatically increased (decreased). The rate of response to changes in the standard deviations will be set by the value of α. The results of applying this new procedure (trading rule 2, with α = .2) to

Table 8.11

Trading Rule 2 on all rubber prices with scale factor set to 1.4 and smoothing constant set to $\alpha = 0.2$

Period	no. of days	no.of buy sigs.	no. of sell sigs.	no. of trading sigs.	no. of days between trades	profit in £'s	profit per trade in £'s
1	58	-	-	-	-	-	-
2	64	6	4	10	6.4	23.92	2.39
3	65	5	3	8	8.1	76.42	9.55
4	62	6	5	11	5.6	63.92	5.81
5	63	5	4	9	7.0	148.08	16.45
6	65	6	6	12	5.4	209.78	17.48
7	65	8	5	13	5.0	631.34	48.56
8	61	5	5	10	6.1	177.94	17.79
9	64	4	7	11	5.8	370.26	29.11
10	63	3	3	6	10.5	159.08	26.51
11	65	5	8	13	5.0	28.36	2.18
12	61	3	5	8	7.6	49.60	6.20
13	62	7	8	15	4.1	138.06	9.20
14	65	6	6	12	5.4	123.68	10.31
15	65	6	4	10	6.5	86.62	8.66
16	61	5	4	9	6.8	59.12	6.57
17	62	5	6	11	5.6	229.26	20.84
18	65	4	6	10	6.5	83.26	8.33
19	65	8	7	15	4.3	181.72	12.11
20	17	1	3	4	4.3	72.44	18.11
Totals		98	99	197		£2862.78	

Average no. of days between trades = 5.92

Average no.of trades per month = 3.58

Average profit per trade = £14.53

Monthly return average = £52.05

St. dev. of no. of trades per month = 1.46

Monthly return st. dev. = £63.22

Coefficent of variation = 1.21

the entire rubber series are given in Table 8.11.

We now compare the results obtained using rule 1 and rule 2 on the rubber series by reference to Table 8.6 and 8.11.

(i) The total profits over the entire 5 years are similar although rule 2 produces slightly (£381) more than rule 1.

(ii) The two rules generate remarkably similar numbers of trading signals.

(iii) The profits per trades are similar but again rule 2 generates slightly superior results (£14.52 verses £13.56).

(iv) The number of trades in each period is, on average, almost the same for both procedures. However, as expected, the variation in the number of trades from period to period is very much smaller using rule 2. (A comparison of the number of trades generated by both technique in periods 7 and 8 highlight this feature very well).

(v) Point (iv) is reinforced when one considers the number of trades per month. While numbers of trades per month are similar, the standard deviation of the number of trades per month for rule 2 is half the corresponding standard deviation for rule 1. The stream of trading signals is "twice as smooth" using rule 2.

(vi) The monthly average return using rule 2 is higher than the monthly average return using rule 1.

(vii) The standard deviation of monthly returns using rule 2 is lower than the standard deviation of monthly returns using rule 1.

(viii) The coefficient of variation of monthly returns is lower using rule 2.

In every respect then, rule 2 seems superior to rule 1. The overall profit is larger, the variability of returns is lower and most importantly, rule 2 seems to generate a much smoother stream of trading opportunities. Tables 8.12 to 8.15 contain the results of applying rule 2 to all 4 series using the following values of α : 0.02 (low), 0.14 (medium) and 0.26 (high).

In the tables 8.12 to 8.15 and many of the remaining tables in this chapter we use the following abreviations:

n tds.	:	total number of trades over all periods
days	:	average number of days between trades
pft	:	total profit in £'s received per unit spread
pft/td	:	average profit in £'s per trade per unit spread
days/mnth	:	the two entries correspond to the average number of trades per month and the standard deviation of the number of trades per month respecitively
mon ave.	:	average monthly return in £'s from trading rule
st.dv.	:	standard deviation of monthly returns in £'s from trading rule
c of var	:	coefficient of variation of monthly returns from trading rule

8.4.1 General discussion of returns to rule 2

In this section we discuss (a) the effects of the smoothing constant, α, on the returns to rule 2 and, (b) the differences (if any) between the results of rule 2 and rule 1.

There is a tremendous quantity of information in Table 8.12 to 8.15 and in order to address (a) and (b) above we look in detail at each of the statistics : n tds, pft etc etc and for ease of reference we use the

Table 8.12

Trading rule 2 on cocoa series with low, medium and high α values

α = 0.02

scale	n tds	pft	days	pft/td	days/mnth		mon ave	st.dv.	c of var.
.6	394	9625.	2.94	24.43	7.16	3.15	175.00	282.45	1.61c
1.0	194	7021.	5.98	36.19	3.53	2.54	127.66	250.40	1.96c
1.4	103	4817.	11.26	46.77	1.87	1.88	87.58	215.35	2.46b
1.8	61	5043.	19.02	82.67	1.11	1.36	91.69	193.71	2.11c
2.2	38	4195.	30.53	110.39	0.69	1.10	76.27	183.66	2.41b
2.6	29	4320.	40.00	148.98	0.53	0.96	78.55	182.48	2.32b
3.0	18	3615.	64.44	200.82	0.33	0.67	65.72	177.77	2.70b

α = 0.14

scale	n tds	pft	days	pft/td	days/mnth		mon ave.	st.dv.	c of var.
.6	478	11082.	2.43	23.18	8.69	2.62	201.49	320.07	1.59c
1.0	302	9313.	3.84	30.84	5.49	2.51	169.33	250.10	1.48c
1.4	174	6062.	6.67	34.84	3.16	1.69	110.21	183.64	1.67c
1.8	96	4371.	12.08	45.54	1.75	1.31	79.48	198.31	2.50b
2.2	57	4910.	20.35	86.14	1.04	1.02	89.27	201.18	2.25b
2.6	34	4436.	34.12	130.47	0.62	0.85	80.66	197.15	2.44b
3.0	19	4131.	61.05	217.41	0.35	0.55	75.10	180.08	2.40b

α = 0.26

scale	n tds	pft	days	pft/td	days/mnth		mon ave.	st.dv.	c of var.
.6	502	10045.	2.31	21.60	9.13	2.25	197.19	334.88	1.70c
1.0	334	9399.	3.47	28.14	6.07	2.19	170.89	255.95	1.50c
1.4	195	6929.	5.95	35.53	3.55	1.50	125.98	194.22	1.54c
1.8	125	7042.	9.28	56.34	2.27	1.16	128.04	191.57	1.50c
2.2	70	5489.	16.57	78.41	1.27	0.95	99.79	192.30	1.93c
2.6	44	4222.	26.36	95.95	0.80	0.76	76.76	188.42	2.45b
3.0	28	3820.	41.35	136.42	0.51	0.74	69.45	185.85	2.68b

See page 240 for legend

Table 8.13

Trading rule 2 on coffee series with low, medium and high α values

$\alpha = 0.02$

scale	n tds	pft	days	pft/td	days/mnth		mon ave	st.dv.	c of var.
.6	414	6672.	2.80	16.11	7.53	3.18	121.30	203.28	1.68c
1.0	232	4301.	5.00	18.54	4.22	2.79	87.20	115.81	1.48c
1.4	140	2636.	8.29	18.83	2.55	2.18	47.93	112.44	2.35b
1.8	79	2695.	14.68	34.11	1.44	1.60	48.99	115.66	2.36b
2.2	40	1926.	29.00	48.15	0.73	0.97	35.02	99.35	2.84a
2.6	28	1520.	41.43	54.28	0.51	0.79	27.63	98.54	3.57a
3.0	17	1222.	68.24	71.87	0.31	0.66	22.22	91.96	4.14a

$\alpha = 0.14$

scale	n tds	pft	days	pft/td	days/mnth		mon ave.	st.dv.	c of var.
.6	483	7334.	2.40	15.18	8.78	2.33	133.34	198.80	1.49c
1.0	292	5038.	3.97	17.25	5.31	1.97	91.60	183.01	2.00c
1.4	178	3579.	6.52	20.11	3.24	1.56	65.08	105.66	1.62c
1.8	106	2775.	10.94	26.18	1.93	1.25	50.45	98.27	1.95c
2.2	59	2159.	19.66	36.59	1.07	0.88	39.26	92.71	2.36b
2.6	34	1749.	34.12	51.44	0.62	0.65	31.80	96.98	3.05a
3.0	23	1668.	50.43	72.51	0.42	0.53	30.32	97.18	3.21a

$\alpha = 0.26$

scale	n tds	pft	days	pft/td	days/mnth		mon ave.	st.dv.	c of var.
.6	514	7014.	2.26	13.65	9.35	2.07	127.53	202.25	1.59c
1.0	321	5560.	3.61	17.32	5.84	1.64	101.10	182.83	1.81c
1.4	204	4814.	5.69	23.60	3.71	1.42	87.54	176.66	2.02c
1.8	130	2552.	8.92	19.63	2.36	1.21	46.40	93.31	2.01c
2.2	71	2110.	16.34	29.72	1.29	0.76	38.36	98.40	2.56b
2.6	49	2029.	23.67	41.41	0.89	0.63	36.90	99.03	2.68b
3.0	37	1905.	31.35	51.49	0.67	0.61	34.64	99.53	2.87a

Table 8.14

Trading rule 2 on sugar series with low, medium and high α values

$\alpha = 0.02$

scale	n tds	pft	days	pft/td	days/mnth		mon ave	st.dv.	c of var.
.6	400	3128.	2.90	7.82	7.27	3.17	56.86	84.28	1.48c
1.0	215	2566.	5.40	11.94	3.91	2.80	46.66	92.25	1.98c
1.4	108	1481.	10.74	13.72	1.96	1.82	26.93	64.37	2.39b
1.8	58	1302.	20.00	22.44	1.05	1.43	23.67	80.20	3.39a
2.2	41	780.	28.29	19.02	0.75	1.28	14.18	40.90	2.89a
2.6	23	532.	50.43	23.12	0.42	0.83	9.67	28.69	2.97a
3.0	15	384.	77.33	25.61	0.27	0.62	6.98	28.29	4.05

$\alpha = 0.14$

scale	n tds	pft	days	pft/td	days/mnth		mon ave.	st.dv.	c of var.
.6	491	3363.	2.36	6.85	8.93	2.49	61.15	93.18	1.52c
1.0	318	3203.	3.65	10.07	5.78	2.05	58.23	71.68	1.23c
1.4	182	2508.	6.37	13.78	3.31	1.72	45.60	80.99	1.78c
1.8	102	1533.	11.37	15.02	1.85	1.33	27.86	57.56	2.07c
2.2	58	957.	20.00	16.50	1.05	1.11	17.40	32.70	1.88c
2.6	33	746.	35.15	22.61	0.60	0.76	13.56	35.16	2.59b
3.0	19	526.	61.05	27.70	0.35	0.52	9.57	32.03	3.35a

$\alpha = 0.26$

scale	n tds	pft	days	pft/td	days/mnth		mon ave.	st.dv.	c of var.
.6	511	3503.	2.27	6.85	9.29	2.31	63.69	98.67	1.55c
1.0	347	2983.	3.34	8.60	6.31	1.85	54.24	74.15	1.37c
1.4	203	2742.	5.71	13.51	3.69	1.46	49.85	81.42	1.63c
1.8	125	1855.	9.28	14.84	2.27	1.21	33.73	58.35	1.73c
2.2	79	1231.	14.68	15.59	1.44	1.00	22.39	35.68	1.59c
2.6	51	905.	22.75	17.74	0.93	0.79	16.45	34.35	2.09c
3.0	26	661.	44.62	25.43	0.47	0.54	12.02	32.31	2.69b

Table 8.15

Trading rule 2 on rubber series with low, medium and high α values

$\alpha = 0.02$

scale	n tds	pft	days	pft/td	days/mnth		mon ave	st.dv.	c of var.
.6	446	3611.	2.60	8.10	8.11	2.98	65.66	93.79	1.43c
1.0	272	3297.	4.26	12.12	4.95	2.75	59.95	87.04	1.45c
1.4	142	2347.	8.17	16.53	2.58	2.02	42.68	75.98	1.78c
1.8	74	1501.	15.68	20.29	1.35	1.53	27.30	57.99	2.12b
2.2	46	1248.	25.22	27.14	0.84	1.12	22.69	49.83	2.20b
2.6	29	715.	40.00	24.64	0.53	0.86	12.99	44.48	3.42a
3.0	19	587.	61.05	30.87	0.35	0.70	10.66	45.19	4.24

$\alpha = 0.14$

scale	n tds	pft	days	pft/td	days/mnth		mon ave.	st.dv.	c of var.
.6	509	3578.	2.28	7.03	9.25	2.30	65.05	79.04	1.22c
1.0	326	3326.	3.56	10.20	5.93	1.95	60.46	89.31	1.48c
1.4	183	2691.	6.34	14.70	3.33	1.49	48.92	63.66	1.30c
1.8	109	2138.	10.64	19.61	1.98	1.25	38.87	59.07	1.52c
2.2	61	1201.	19.02	19.69	1.11	0.81	21.84	46.36	2.12b
2.6	35	880.	33.14	25.14	0.64	0.65	16.00	44.31	2.77b
3.0	22	730.	52.73	33.19	0.40	0.49	13.27	40.57	3.06a

$\alpha = 0.26$

scale	n tds	pft	days	pft/td	days/mnth		mon ave.	st.dv.	c of var.
.6	532	3669.	2.18	6.90	9.67	2.06	66.71	78.01	1.17c
1.0	348	3436.	3.33	9.87	6.33	1.74	62.47	90.91	1.46c
1.4	210	2939.	5.52	13.99	3.82	1.40	53.43	62.83	1.18c
1.8	127	2337.	9.13	18.40	2.31	1.22	42.49	58.51	1.38c
2.2	86	1657.	13.49	19.27	1.56	1.03	30.13	46.70	1.55c
2.6	48	1112.	24.17	23.17	0.87	0.75	20.22	45.00	2.23b
3.0	31	976.	37.42	31.50	0.56	0.60	17.75	43.98	2.48b

notation, $n(\alpha_i, \gamma_j)$, say to refer to the number of trades on rule 2 with the smoothing constant set to α_i and scale parameter set to γ_j. We define $\alpha_1 = 0.02$, $\alpha_2 = 0.14$ and $\alpha_3 = 0.26$. Similarly, we define $\gamma_1 = 0.06$, ... $\gamma_7 = 3.0$. Also, when referring to rule 1 results, we can, without confusion, refer to, say, the number of trades as $n(\gamma_j)$.

The effects of the various α's on the results to rule 2 and the differences between rules 1 and 2 are remarkably similar across commodity series. Accordingly, in what follows (except in certain situations) we are referring to the results of all 4 sets of returns.

(1) Number of trades n ( )

(a) In each series and at every combination of α and γ, $n(\alpha, \gamma)$ are remarkably similar across series. The fixing of α and γ produces almost identical numbers of trades in the cocoa, coffee, sugar and rubber series. As expected, for each value of α, increasing γ decreases $n(\alpha, \gamma)$. Without exception, larger values of α generate more trades, i.e.

$$n(\alpha_1, \gamma) < n(\alpha_2, \gamma) < n(\alpha_3, \gamma)$$

for all γ. At small γ values the maximum difference is quite large (approximately 100 with $\gamma = 0.6$) and decreases with large γ (approximately 10 with $\gamma = 3.0$).

(b) We observe that rule 2 does not always produce more trades than rule 1. In some instances this is so, in others not. It is interesting to note that the rate of change of $n(\alpha, \gamma)$ and $n(\gamma)$ with respect to γ is different. Whatever the value of, α, the decrease in the number of trades caused by increasing γ is more marked with rule 2 than with rule 1. With rule 2 the number of trades generated is more sensitive to the scale parameter setting.

(2) Total profits P ( )

(a) As observed with rule 1, total profits are high for low γ and decrease as γ increases. At most γ values, the ranking of profits across α values is:

$$P(\alpha_1,\gamma) < P(\alpha_2,\gamma) < P(\alpha_3,\gamma)$$

With some values of γ, in the cocoa and coffee series, this ranking is reversed. By considering the profits across commodity series we note that in general the ranking is:

$$\text{cocoa} > \text{coffee} > \text{rubber} > \text{sugar}$$

at each combination of α and γ.

(b) Total profits to rules 1 and 2 are similar at each value of γ.

(3) Profits per trade Pt ( )

(a) Pt (α,γ) increases as γ increases at all values of α. The ranking of the profits per trade is identical to the ranking of the total profits given in 2 (a) with cocoa yielding the highest value at £217 and sugar yielding the lowest at £7. The rankings of profit per trade over increasing α values is given generally by:

$$Pt(\alpha_1,\gamma) > Pt(\alpha_2,\gamma) > Pt(\alpha_3,\gamma)$$

The difference between Pt (α_1,γ) and Pt (α_3,γ) is small at low γ values and large at large γ values. We deduce from this that although n(α,γ) and P(α,γ) both increase with increasing α, the ratio: Pt(α,γ) decreases.

(b) At small γ values, the profit per trdes to rule 1 and 2 are similar. At larger γ values rule 2 results are generally superior to rule 1 results.

(4) Monthly returns means and standard deviation : M ( ) and S ( )

The average monthly returns are simply the total profits divided by 55 and so the pattern of variation of values is exactly duplicated those of the total profits given in (2) above.

(a) As γ increases, the standard deviation of monthly returns decreases. There is no obvious consistent pattern in the effects of α on S (α,γ) except that they are all very similar across γ. The cocoa S(γ) values are the largest and the sugar and rubber values, the lowest.

(b) The S (α,γ) and S (γ) values seem to be similar.

(5) Coefficients of variation of month returns ν( )

(a) As γ increases, ν(α,γ) increases. With the exception of the rubber series (in which ν(α_1,γ) > ν(α_2,γ) > ν(α_3,γ) for almost all γ), there does not seem to be any general pattern across α values. Nearly all n(α,γ) values are significant at the 1% level and many are significant at the 0.1% level. The values of n(α,γ) are, in general remarkably similar across series. The rubber series produced the smallest values with $\nu(0.26, 0.6) = 1.17$

(b) The $\nu(\alpha,\gamma)$ and $\nu(\gamma)$ values appear to be quite similar.

(6) Variation of streams of returns : $\mu_d(\ )$, $\sigma_d(\ )$

The average number of trades per month is simply the total number of trades divided by 55 and so remarks in (1) above apply. As γ increases, $\mu_d(\alpha,\gamma)$ decreases.

(a) The standard deviation of the number of trades per month, $\sigma_d(\alpha,\gamma)$ decreases as γ increases. In almost every series and at all γ values:

$$\sigma_d(\alpha_1,\gamma) > \sigma_d(\alpha_2,\gamma) > \sigma_d(\alpha_3,\gamma)$$

It appears then that high α's smooth out the numbers of trades per month more than low α's.

(b) Without exception the $\sigma_d(\alpha,\gamma)$ values are smaller than the $\sigma_d(\gamma)$ values. As expected, the variation in the number of trades per month is lower using rule 2.

8.4.2 Concluding remarks on trading rule 2

By considering Tables 8.12 to 8.15 along with the above summary we note that the effects of the two parameters in the application of trading rule 2 is complex but consistent across all commodity series. We can generate a large number of trading signals by selecting high smoothing constants and small scale parameters. Such trades will be frequent and result in relatively small returns per trade on average. Alternatively, one could set α to a small value and γ to a large value to generate few but very profitable trades.

The values of α and γ which one would wish to set in a practical situation would be governed by the cost of transactions which we have so far ignored. If trading costs are small, we could achieve high profits by generating a large number of trades. If trading costs are high it would be desirable to generate a small number of more profitable trades. In section 8.5 we now consider the question of transaction costs

8.5 Transaction costs

For private investors in the futures markets, transaction costs will depend on the broker and to a certain extent on what volume of trading is involved. At the time of writing the author has noted that the commissions quoted for trades on the London soft futures markets vary considerably from broker to broker. In order to arrive at some sort of typical commissions that would have been charged over our period, it was decided to refer to the ICCH Procedural Manuals for the years 1975 to 1979. However, studying the relevant sections in each manual reveals a confusing array of different commissions for each commodity futures market. There are, for example, 27 different possible commissions listed for trading in the cocoa futures market. Which commission one is charged depends on what status one has within the relevant market association. The Cocoa Terminal Market Association for example has trade members, broker members, and city home members. The lowest commissions are charged for trades between trade members and the highest are for trades between non members.

Rather than examine the trading rule results with all the possible rates of commissions we decided to investigate the profits received by 3 types of individual; ones required to pay, (i) the minimum commission, (ii) a medium commission and (iii) the maximum commission. Although there was some variation in these commissions charges throughout the 5 year period covered, the values chosen to be typical for each futures market are given in Table 8.16. These are the commissions charged per spread per round turn, (i.e. 1 buy followed by 1 sell or vice versa).

Table 8.16

Commissions in £'s per spread per round turn

	Commissions		
	Low	Medium	High
Cocoa	8.00	20.00	32.00
Coffee	6.00	20.00	28.00
Sugar	9.00	20.00	36.00
Rubber	9.00	20.00	39.00

8.5.1 Rule 2 with transaction costs

Trading rules 1 and 2 with the inclusion of transaction costs were applied to the data sets with many values of α and γ. However since the overall profits to rule 1 and 2 are similar (except for large γ) it was decided for brevity to report results for rule 2 only. Also rather than present returns for many values of α we consider results for only one value; the medium value of $\alpha = 0.14$. Also for brevity we decided not to report the statistics relating to the flow of trading signals as these are independent of commissions and are already presented in Tables 8.12 to 8.15.

The results appear in Tables 8.17 to 8.20. We make the following observations:

<u>(i) Low commissions</u>

In all series the profit per trade, as expected, is small for low γ values and increases with γ. In all but two instances ($\gamma = .6$ in sugar and rubber series) the profits per trade are positive. The ranking of total profits across series is as in section 8.4.1. The sugar profits are quite low with only £4.78 per trade at $\gamma = 1.4$ and the cocoa profits are high at £26.84 at $\gamma = 1.4$. Many of the cocoa and coffee returns are significant at the 1% level. However, although in the rubber and sugar series with γ greater than 0.6 the total returns are positive, only 2 rubber and 1 sugar results are significant. Note the monthly returns and the standard deviation of monthly returns. The sugar and rubber returns have relatively low standard deviations of £30 to £70 per month whereas the cocoa and coffee standard deviations are typically 2 to 5 times these values.

Thus with low commissions, positive profits were still possible. When one considers the margin requirements, a profit of £7258 in the cocoa series is very large indeed. Although the sugar and rubber results are disappointing in that the (small) commissions remove a good deal of the profits recorded in section 8.4.1, the returns with middle to high γ values are still quite large.

Table 8.17

Trading rule 2 with α = 0.14 applied to the cocoa series with low, medium and high commission charges

Commission level	scale	profit	profit /trade	Monthly returns means (£'s)	st. dev. (£'s)	coef. of var
	.6	7258.	15.18	131.96	321.83	2.44 b
	1.0	6897.	22.84	125.40	249.13	1.99 c
	1.4	4670.	26.84	84.91	184.18	2.17 b
Low	1.8	3603.	37.54	65.52	196.26	3.00 a
	2.2	4454.	78.14	80.98	196.14	2.42 b
	2.6	4164.	122.47	75.71	193.34	2.55 b
	3.0	3979.	209.41	72.34	177.55	2.45 b
	.6	1522.	3.18	27.67	326.98	11.82
	1.0	3273.	10.84	59.51	250.69	4.21
	1.4	2582.	14.84	46.94	186.83	3.98
Medium	1.8	2451.	25.54	44.57	194.21	4.36
	2.2	3770.	66.14	68.54	188.98	2.76 b
	2.6	3756.	110.47	68.29	187.94	2.75 b
	3.0	3751.	197.41	68.20	173.90	2.55 b
	.6	-4214.	-8.82	-76.62	335.00	-
	1.0	-351.	-1.16	-6.38	255.80	-
	1.4	494.	2.84	8.98	191.60	21.34
High	1.8	1299.	13.54	23.63	193.41	8.19
	2.2	3086.	54.14	56.10	182.37	3.25 a
	2.6	3348.	98.47	60.87	182.96	3.01 a
	3.0	3523.	185.41	64.05	170.43	2.66 b

Table 8.18

Trading rule 2 with $\alpha = 0.14$ applied to the coffee series with low, medium and high commission charges

Commission level	scale	profit	profit /trade	Monthly returns means (£'s)	st. dev. (£'s)	coef. of var
Low	.6	4436.	9.18	80.65	201.22	2.49 b
	1.0	3286.	11.25	59.75	183.85	3.08 a
	1.4	2511.	14.11	45.66	106.23	2.33 b
	1.8	2139.	20.18	38.89	97.35	2.50 b
	2.2	1805.	30.59	32.82	90.94	2.77 b
	2.6	1545.	45.44	28.09	95.78	3.41 a
	3.0	1530.	66.51	27.81	96.04	3.45 a
Med	.6	-2326.	-4.82	-42.29	210.40	-
	1.0	-802.	-2.75	-15.58	188.69	-
	1.4	19.	0.11	0.35	110.69	318.11
	1.8	655.	6.18	11.91	97.41	8.18
	2.2	979.	16.59	17.80	87.92	4.94
	2.6	1069.	31.44	19.44	93.55	4.81
	3.0	1208.	52.51	21.96	93.76	4.27
High	.6	-6190.	-12.82	-112.55	217.67	-
	1.0	-3138.	-10.75	-57.05	193.17	-
	1.4	-1405.	-7.89	-25.54	115.17	-
	1.8	-193.	-1.82	-3.51	98.83	-
	2.2	507.	8.59	9.22	86.94	9.43
	2.6	797.	23.44	14.49	92.66	6.39
	3.0	1024.	44.51	18.61	92.70	4.98

Table 8.19

Trading rule 2 with α = 0.14 applied to the sugar series with low, medium and high commission charges

Commission level	scale	profit	profit /trade	Monthly returns means (£'s)	st. dev. (£'s)	coef of var
	.6	-1056.	-2.15	-19.20	84.94	-
	1.0	341.	1.07	6.19	68.39	11.04
	1.4	870.	4.78	15.82	74.77	4.73
Low	1.8	615.	6.02	11.17	51.18	4.58
	2.2	435.	7.50	7.91	27.70	3.50 a
	2.6	449.	13.61	8.16	31.44	3.85
	3.0	355.	18.70	6.46	29.59	4.58
	.6	-6457.	-13.15	-117.40	82.40	-
	1.0	-3157.	-9.93	-57.41	71.00	-
	1.4	-1132.	-6.22	-20.58	71.11	-
Medium	1.8	-507.	-4.98	-9.23	46.44	-
	2.2	-203.	-3.50	-3.69	25.90	-
	2.6	86.	2.61	1.56	28.50	18.22
	3.0	146.	7.70	2.66	27.39	10.29
	.6	-14896.	-29.15	-260.23	94.11	-
	1.0	-8245.	-25.93	-149.92	85.95	-
	1.4	-4044.	-22.22	-73.52	74.55	-
High	1.8	-2139.	-20.98	-38.90	47.36	-
	2.2	-1131.	-19.50	-20.56	32.64	-
	2.6	-442.	-13.39	-8.04	28.39	-
	3.0	-158.	-8.30	-2.87	26.14	-

Table 8.20

Trading rule 2 with α = 0.14 applied to the rubber series with low, medium and high commission charges

Commission level	scale	profit	profit /trade	Monthly returns: means (£'s)	Monthly returns: st. dev. (£'s)	Monthly returns: coef of var
Low	.6	-1003.	-1.97	-18.24	78.01	-
	1.0	392.	1.20	7.12	85.35	11.99
	1.4	1044.	5.70	18.97	59.53	3.14 a
	1.8	1157.	10.61	21.04	53.73	2.55 b
	2.2	652.	10.69	11.86	44.15	3.72
	2.6	565.	16.14	10.27	42.68	4.16
	3.0	532.	24.19	9.67	38.98	4.03
Medium	.6	-6602.	-12.97	-120.04	83.94	-
	1.0	-3194.	-9.80	-58.08	85.30	-
	1.4	-969.	-5,30	-17.63	58.42	-
	1.8	-42.	-0.39	-0.76	49.98	-
	2.2	-19.	-0.31	-0.34	42.99	-
	2.6	180.	5.14	3.27	41.73	12.75
	3.0	290.	13.19	5.27	37.67	7.14
High	.6	-16273	-31.97	-295.88	108.19	-
	1.0	-9388.	-28.80	-170.70	97.13	-
	1.4	-4446.	-24.30	-80.85	66.73	-
	1.8	-2113.	-19.39	-38.42	52.18	-
	2.2	-1178.	-19.31	-21.41	45.23	-
	2.6	-485.	-13.86	-8.82	42.92	-
	3.0	-128.	-5.81	-2.33	37.21	-

(ii) Medium commissions

The relative ranking of profits and profits per trade are as before. However the increased commissions produce significantly different returns. At the low γ values, very large losses result. With γ set to 0.6, losses of the order of £6000 are realised in the sugar and rubber series. However even at the low γ values, when many non profitable trades are executed, the cocoa series still produces positive (though not statistically significant) returns. Increasing, γ, and reducing the number of many of the non profit making trades reduces the losses in the sugar and rubber series and similarly increases the profits in the cocoa and coffee series. At large γ's, the sugar and rubber profits are positive, though small. Only 3 values are significant; the cocoa returns at γ values greater than 2.7.

(iii) High commissions

Very large and significant losses are realised with high commission charges. A loss of £14,313 occurs with γ = 0.6 in the sugar series. No γ values produces a positive profit in the sugar and rubber series. At γ = 0.6 and 1.0 even the cocoa returns are negative. At γ values higher than 1.0 one still observes significantly positive large profits in the cocoa series.

8.5.2 Concluding remarks on rules 1 and 2 with transaction costs

The extremely large profits reported to rules 1 and 2 reported in sections 8.3 and 8.4 are drastically altered by the inclusion of commission charges. As predicted, the setting of low γ values resulted in many trades and thus induced large transaction costs. However, considering trades between members in which low commissions are charged the profits to rules 1 and 2 are still very large indeed, especially on the cocoa and coffee series. The medium and high charges imposed on non members would have removed most profits received and actually resulted in extremely large losses in the sugar and rubber series. However even with the highest charges the rules would have produced positive profits in the cocoa series and it seems possible that a private individual may have been able to realise large positive returns over this period.

8.6 Trading rule 3

The large positive profits reported in section 8.3 and 8.4 are much reduced by the inclusion of transaction costs. Although the negatively serially correlated portfolios are generating trades that are on average highly significant and positive, the cost of each transaction in many instances is reducing the resulting profits of each trade and in the sugar and rubber series resulting in losses. This may be due to the slightly unrealistic nature of rules 1 and 2.

In rules 1 and 2 a trade is invoked if the price change of the portfolio moves out of a region (band N in Fig. 8.2) set by the standard deviation of returns of the series. With non zero transaction costs this may be slightly impractical. It is not very useful to construct, ex ante, a portfolio that exhibits negative serial correlation if the price movements are not large enough to cover transaction costs. Recall from Tables 8.5 that there is tremendous variation in the standard deviation of the portfolio returns over the whole period. In periods in which the standard deviations of the returns on the portfolios are very low, rules 1 and 2 still continue to generate trading signals even though most of the prices never move enough on the subsequent days to cover the cost of the trades.

We now consider a possibly much more practicable trading rule; rule 3. In rule 3 the band width is set as in rule 2 to a multiple of the standard deviation of the returns of the portfolio. This standard deviation is the exponentially smoothed estimate and is updated every day. However, in order to prevent trades that seem likely to result in profits less than transaction costs we add an additional constraint. If on day, t, a portfolio return moves into the region B or S in Fig. 8.2, a trading signal is given <u>only</u> if the magnitude of the returns have been greater than the expected transaction cost.

Rule 3 was applied to each data set with many different values of the scale parameter, γ, and smoothing constant, α. For brevity we report only those results in which γ = .6, 1.0, 1.4, 1.8, 2.2, 2.6 and 3.0 and with α = 0.14. The results appear in Tables 8.21 to 8.24.

8.6.1 Results of trading rule 3

For ease of reference and in order to compare results of various scale and commission values and to compare rule 3 with rule 2, we now use a slightly different notation. When considering the number of trades,for example, we refer to n (c,γ,k) in which k has the value 2 for rule 2 and 3 for rule 3. The c parameter has the three values : 1, for low commissions, 2 for medium commissions and 3 for high commissions. The scale parameter, γ, as before takes the values γ_1, = .06, γ_2 = 1.0 etc.

A comparison of the results in Tables 8.21 to 8.24 with the results in Table 8.17 to 8.20 yields a startlingly consistent pattern across all series. In what follows we shall (a) consider the effects of the various

Table 8.21

Trading rule 3 with $\alpha = 0.14$ applied to the cocoa series

scale	n tds	pft	days	pft/td	days/mnth		mon ave	st.dv.	c of var.		Commission level
.6	477	7284.	2.43	15.27	8.67	2.61	132.43	321.66	2.43	b	
1.0	302	6897.	3.84	22.84	5.49	2.51	125.40	249.13	1.99	c	
1.4	174	4670.	6.67	26.84	3.16	1.69	84.91	184.18	2.17	b	
1.8	96	3603.	12.08	37.54	1.75	1.31	65.52	196.26	3.00	a	LOW
2.2	57	4454.	20.35	78.14	1.04	1.02	80.98	196.14	2.42	b	
2.6	34	4164.	34.12	122.47	0.62	0.85	75.71	193.34	2.55	b	
3.0	19	3979.	61.05	209.41	0.35	0.55	72.34	177.55	2.45	b	
.6	420	2659.	2.76	6.33	7.64	2.90	48.34	318.13	6.58		
1.0	283	3529.	4.10	12.47	5.15	2.49	64.17	248.69	3.88		
1.4	171	2619.	6.78	15.32	3.11	1.69	47.63	186.45	3.91		
1.8	96	2451.	12.08	25.54	1.75	1.31	44.57	194.21	4.36		MEDIUM
2.2	57	3770.	20.35	66.14	1.04	1.02	68.54	188.98	2.76	b	
2.6	34	3756.	34.12	110.47	0.62	0.85	68.29	187.94	2.75	b	
3.0	19	3751.	61.05	197.41	0.35	0.55	68.20	173.90	2.55	b	
.6	328	629.	3.54	1.92	5.96	3.37	11.44	303.73	26.54		
1.0	239	1216.	4.85	5.09	4.35	2.71	22.11	247.26	11.18		
1.4	148	1214.	7.84	8.20	2.69	1.76	22.07	187.75	8.51		
1.8	87	1600.	13.33	18.39	1.58	1.27	29.09	193.00	6.63		HIGH
2.2	55	3148.	21.09	57.24	1.00	1.04	57.24	181.91	3.18	a	
2.6	34	3348.	34.12	98.47	0.62	0.85	60.87	182.96	3.01	a	
3.0	19	3523.	61.05	185.41	0.35	0.55	64.05	170.43	2.66	b	

Table 8.22

Trading rule 3 with α = 0.14 applied to the coffee series

scale	n tds	pft	days	pft/td	days/mnth		mon ave	st.dv.	c of var.		Commission level
.6	442	4563.	2.62	10.32	8.04	2.63	82.97	200.04	2.41	b	
1.0	282	3302.	4.11	11.71	5.13	1.93	60.04	183.76	3.06	a	
1.4	178	2511.	6.52	14.11	3.24	1.56	45.66	106.23	2.33	b	
1.8	106	2139.	10.94	20.18	1.93	1.25	38.89	97.35	2.50	b	LOW
2.2	59	1805.	19.66	30.59	1.07	0.88	32.82	90.94	2.77	b	
2.6	34	1545.	34.12	45.44	0.62	0.65	28.09	95.78	3.41	a	
3.0	23	1530.	50.43	66.51	0.42	0.53	27.81	96.04	3.45	a	
.6	290	810.	4.00	2.79	5.27	3.63	14.73	188.79	12.81		
1.0	206	553.	5.63	2.69	3.75	2.36	10.06	180.55	17.94		
1.4	132	769.	8.79	5.83	2.40	1.56	13.99	104.07	7.44		
1.8	83	1037.	13.98	12.49	1.51	1.17	18.85	94.38	5.01		MEDIUM
2.2	54	1066.	21.48	19.74	0.98	0.89	19.38	87.07	4.49		
2.6	32	1112.	36.25	34.75	0.58	0.63	20.22	93.01	4.60		
3.0	22	1236.	52.73	56.19	0.40	0.53	22.48	93.45	4.16		
.6	232	-320.	5.00	0.99	4.22	3.65	-4.19	182.82	-		
1.0	169	-274.	6.86	-1.62	3.07	2.38	-4.98	177.24	-		
1.4	113	69.	10.27	0.61	2.05	1.64	1.26	103.21	81.92		
1.8	76	484.	15.26	6.37	1.38	1.18	8.80	93.18	10.59		HIGH
2.2	52	634.	22.31	12.18	0.95	0.89	11.52	85.73	7.44		
2.6	31	847.	37.42	27.31	0.56	0.60	15.39	91.90	5.97		
3.0	21	1051.	55.24	50.04	0.38	0.49	19.10	92.35	4.83		

Table 8.23

Trading rule 3 with α = 0.14 applied to sugar series

scale	n tds	pft	days	pft/td	days/mnth		mon ave	st.dv.	c of var.	Commission level
.6	412	-556.	2.82	-1.35	7.49	2.79	-10.12	82.13	-	
1.0	297	450.	3.91	1.51	5.40	2.18	8.17	67.61	8.27	
1.4	176	909.	6.59	5.16	3.20	1.75	16.52	74.54	4.51	
1.8	100	618.	11.60	6.18	1.82	1.35	11.23	51.16	4.56	LOW
2.2	58	435.	20.00	7.50	1.05	1.11	7.91	27.70	3.50 a	
2.6	33	449.	35.15	13.61	0.60	0.76	8.16	31.44	3.85	
3.0	19	355.	61.05	18.70	0.35	0.52	6.46	29.59	4.58	
.6	199	-1792.	5.83	-9.01	3.62	3.03	-32.59	70.65	-	
1.0	166	-1132.	6.99	-6.82	3.02	2.34	-20.59	62.18	-	
1.4	122	-395.	9.51	-3.24	2.22	1.80	-7.19	66.09	-	
1.8	85	-310.	13.65	-3.65	1.55	1.34	-5.65	44.80	-	MEDIUM
2.2	50	-95.	23.20	-1.90	0.91	1.01	-1.73	24.05	-	
2.6	31	117.	37.42	3.78	0.56	0.76	2.13	28.50	-	
3.0	18	166.	64.44	9.21	0.33	0.51	3.01	27.22	-	
.6	90	-1602.	12.89	-17.80	1.64	2.46	-29.13	59.32	-	
1.0	78	-1143.	14.87	-14.65	1.42	2.04	-20.78	52.92	-	
1.4	60	-581.	19.33	-9.69	1.09	1.62	-10.57	47.96	-	
1.8	42	-453.	27.62	-10.80	0.76	1.14	-8.25	30.27	-	HIGH
2.2	31	-387.	37.42	-12.49	0.56	0.86	7.04	20.03	-	
2.6	20	-73.	58.00	-3.67	0.36	0.68	-1.34	24.52	-	
3.0	13	38.	89.23	2.89	0.24	0.47	0.68	23.73	-	

Table 8.24

Trading rule 3 with α = 0.14 applied to rubber series

scale	n tds	pft	days	pft/td	days/mnth		mon ave	st.dv.	c of var.		Commission level
.6	434	-372.	2.67	-0.86	7.89	2.41	-6.77	73.85	-		
1.0	320	408.	3.63	1.27	5.82	1.94	7.41	85.31	11.51		
1.4	182	1028.	6.37	5.65	3.31	1.50	18.69	59.65	3.19	a	
1.8	109	1157.	10.64	10.61	1.98	1.25	21.04	53.73	2.55	b	LOW
2.2	61	652.	19.02	10.69	1.11	0.81	11.86	44.15	3.72	a	
2.6	35	565.	33.14	16.14	0.64	0.65	10.27	42.68	4.16		
3.0	22	532.	52.73	24.19	0.40	0.49	9.67	38.98	4.03		
.6	238	-1504.	4.87	-6.32	4.33	3.20	-27.34	64.68	-		
1.0	198	-873.	5.86	-4.41	3.60	2.49	-15.88	74.95	-		
1.4	135	-254.	8.59	-1.88	2.45	1.56	-4.63	52.05	-		
1.8	97	56.	11.96	0.58	1.76	1.26	1.02	48.99	47.89		MEDIUM
2.2	59	12.	19.66	0.20	1.07	0.81	0.22	42.76	195.78		
2.6	35	180.	33.14	5.14	0.64	0.65	3.27	41.73	12.75		
3.0	22	290.	52.73	13.19	0.40	0.49	5.27	37.67	7.14		
.6	94	-1715.	12.34	-18.25	1.71	2.62	-31.18	59.51	-		
1.0	85	-1445.	13.65	-17.00	1.55	2.21	-26.27	63.57	-		
1.4	65	-942.	17.85	-14.49	1.18	1.58	-17.13	45.89	-		
1.8	51	-505.	22.75	-9.90	0.93	1.30	-9.18	42.75	-		HIGH
2.2	33	-347.	35.15	-10.51	0.60	0.85	-6.31	41.38	-		
2.6	21	-104.	55.24	-4.95	0.38	0.56	-1.89	40.44	-		
3.0	15	48.	77.33	3.18	0.27	0.45	0.87	35.97	41.52		

γ and c values on results to rule 3 and (b) compare corresponding results across rules 2 and rules 3. We consider in detail each of the four statistics (i) number of trades, (ii) total profits, (iii) profits per trade and (iv) the coefficients of variation of monthly returns.

<u>(i) Number of trades n (, ,)</u>

(a) Without exception and at each γ level the number of trades induced by rule 3 at various c levels is ranked as follows:

$$n(1, \gamma, 3) \geqslant n(2, \gamma, 3) \geqslant n(3, \gamma, 3)$$

The difference between $n(1, \gamma, 3)$ and $n(3, \gamma, 3)$ is maximal at low γ values (eg in the rubber series at $\gamma = .6$ the difference is 434 - 94 = 340). At larger γ values the differences are very much reduced (it is zero above $\gamma = 2.2$ in cocoa series). This is exactly what we would have expected. At low scale values the commission constraint in rule 3 is dominant and increasing the commissions will result in significantly fewer trading signals. At larger scale values the commission constraint becomes less influential and in the cocoa series is completely redundant at γ above 2.2.

(b) Again without exception and at all c and γ levels

$$n(c, \gamma, 3) \leqslant n(c, \gamma, 2),$$

with the equality holding at high γ and low c values in all series. With hindsight this result was only to be expected, as rule 3 is rule 2 with an additional constraint. The number of trades to rule 3 must be less than or equal to the number of trades to rule 2, other things being equal.

(ii) Total Profits P (, ,)

(a) All the cocoa and all but two of the coffee profits are positive. In the sugar and rubber series at the medium and high commission values, the profits are negative. However at the value of γ = 3.0, both sugar and rubber series produce small but positive profits. All the medium and high commission level profits increase with increasing γ. In the cocoa and coffee series with low commissions this trend is reversed. At every value of γ and in all series :

$$P(1, \gamma, 3) > P(2, \gamma, 3) > P(3, \gamma, 3).$$

The difference between P (1 , γ, 3) and P (3 , γ, 3) is most marked at low γ values (eg in cocoa series a difference of £7284 - £679 = £6655 at γ = 0.6 is realized) and least marked at high γ values.

(b) Without exception, at every value of γ and c and across all series:

$$P(c, \gamma, 3) \geqslant P(c, \gamma, 2),$$

the difference being most marked at high c and low γ values. A loss of £16773 at γ = .6 in the rubber series in rule 2 is reduced to one of only £1715 in rule 3. In the cocoa series with γ = .6, a loss of £4214 with rule 2 can be compared to a profit of £629 with rule 3.

At higher γ values the differences between the profits to rules 2 and 3 becomes less and less (in the cocoa series the total profits to rules 2 and 3 are equal at medium to high γ values). It is interesting to look at this phenomenon in more detail.

With low γ and high c values, many expensive trades that would be invoked by rule 2 are being blocked in rule 3. At high γ and/or low c values few and relatively cheap trades are invoked by rule 2 and these would also be invoked by rule 3 resulting in similar profits.

The very interesting (but not immediately obvious) result to note here is that the inclusion of an additional commission constraint in rule 3 does not appear to be blocking any potentially profitable trades. The constraint seems to work as designed; many non profitable trades are blocked and no profitable trades are overlooked.

(iii) Profits per trade Pt (, ,)

(a) The profits per trade are spectacular in the cocoa series even when considering high transaction costs where values range from £1.97 to £185.41 per trade. The coffee series also produces mostly positive profits per trade. The sugar and rubber series at the low commission rates produce positive values but these become negative as one considers higher commissions. The patterns in the profit per trade figures over various γ and c values is identical across all series and can be summed by the following equalities:

$$Pt\ (c, \gamma_1, 3) < Pt\ (c, \gamma_2, 3) < \ldots\ldots\ldots < Pt\ (c, \gamma_7, 3)$$

at all c, and

$$Pt\ (1, \gamma, 3) > Pt\ (2, \gamma, 3) > Pt\ (3, \gamma, 3)$$

at all γ.

With hindsight neither of these results is surprising. At a given c, increasing γ reduces the number of non profitable trades resulting in an overall larger profit per trade. At a given γ, value, increasing c also has the effect of reducing the number of non profitable trades but each trade invoked is more expensive with the resultant reduction in the overall profits per trade.

(b) Similarly the pattern in the difference between rules 2 and 3 is identical across all series and at all c and γ values:

$$Pt\ (c, \gamma, 3) \geqslant Pt\ (c, \gamma, 2)$$

The difference between the results from the two rules is most marked at low γ and high c values, particularly in the sugar and rubber series. For example in the sugar series at the high c value and $\gamma = 0.6$ an average loss of £260.23 per trade from rule 2 is altered to an average loss per trade of only £17.80. These large differences at high c and low γ values are obviously due to the dominance of the commission constraint and the remarks made about the total profits apply directly here.

(iv) Coefficients of variation $\nu(\ ,\ ,\ )$

(a) In the cocoa and coffee series all of the coefficients are signficant at the low commission values. At the medium to high commission levels, the cocoa results are significant above $\gamma = 2.2$. None of the sugar or rubber results prove significant. The pattern in the variation of coefficients over various γ and c values is not as consistent as the patterns observed in the other statistics we have considered thus far. However, it is still possible to discern a general pattern and this is best summed up by the inequalities :

(i) $\nu(c, \gamma_1, 3) > \nu(c, \gamma_2, 3)) \ldots > \nu(c, \gamma_7, 3)$

over most c values, and

(ii) $\nu(1, \gamma, 3) < \nu(2, \gamma, 3) < \nu(3, \gamma, 3)$

over all γ values.

(b) The coefficients of variation to rule 3 are equal to or less than the corresponding coefficients in rule 2, ie

$$\nu(c, \gamma, 3) < \nu(c, \gamma, 2)$$

at all c and γ.

From this we can conclude that rule 3 is relatively less risky than rule 2.

8.6.2 Concluding remarks on trading rule 3

It appears then that the introduction of a commission constraint in rule 3 is successful. Rule 3 is superior to rule 2 on every count:

(i) The number trades generated by rule 3 is less than the number of trades generated by rule 2.

(ii) Each of the trades is on average more profitable with rule 3 than with rule 2.

(iii) The total profits to rule 3 are larger than the total profits to rule 2.

(iv) The relative riskiness of the stream of returns to rule 3 is lower than the relative riskiness of returns to rule 2.

All of the cocoa and some of the coffee series profits are large, positive, and highly statistically significant. At low commission rates large positive profits could have been obtained in all series. For members of the relevant trade associations who enjoy low transaction costs the returns would have been very large indeed.

8.7 Comparing rule 3 returns to other assets

So far in this chapter we have looked at the stream of returns to various trading rules involving complex spread portfolios. Rule 3 seemed to be the most successful rule and with hindsight probably a much more realistic one to use in practice. It will be interesting to compare these returns with the returns obtained in other investments over the same period i.e. (i) long positions in each futures markets and (ii) a portfoliio stocks in the UK market.

8.7.1 Returns on net long positions

Using rule 3 on a particular series is essentially watching the price of a very complex six dimensional portfolio every day. Keeping close track of such a spread over time results, every now and then, in the generation of a trading opportunity. The returns from these trades are lumped together each month and considered to be one monthly return. Consider now an alternative strategy. At the begining of each month go long one futures contract in each market. At the end of each month close out the position by a sale. Repeat this process for each of the 55 months being considered. Brokers usually require identical deposits and charge

Table 8.25

Returns (£'s) on long positions in contracts one delivery period from maturity

Commission level		Profit	Monthly mean	Monthly st. dev.	Coeff of var.
	Cocoa	20815	378	1812	4.79
	Coffee	11335	206	1275	6.19
Low	Sugar	-3248	-59	584	-
	Rubber	967	18	489	27.80
	Portfolio	7467	136	610	4.49
	Cocoa	20155	366	1812	4.94
	Coffee	10565	192	1275	6.64
Medium	Sugar	-3853	-70	584	-
	Rubber	362	7	489	74.2
	Portfolio	6807	124	610	4.93
	Cocoa	19495	354	1812	5.11
	Coffee	10125	184	1275	6.93
High	Sugar	-4733	-86	584	-
	Rubber	-683	-12	489	-
	Portfolio	6349	116	610	5.25

identical commissions on net long positions and on spreads. Thus we can use the same sums given in section 8.3.2 and 8.5 to support our alternative stream of trades. The results appear in Table 8.25 and we make the following observations on the returns to long positions:

(i) The cocoa and coffee profits and monthly means are very large indeed. Over the period concerned the prices of cocoa and coffee rose sharply. The rubber series produces a small profit that is reduced to a small loss at higher commission levels. The sugar prices fell over the period and results in large losses.

(ii) The monthly standard deviations are very large. As noted in section 7.5.1 the standard deviation of the underlying futures contract returns are between 10 to 30 times the standard deviation of the returns of the specially constructed portfolios. Comparing Tables 8.25 with Tables 8.21

to 8.24 we note that indeed the returns on the net long contracts are also between 10 to 30 times as variable as the returns obtained using rule 3.

(iii) The coefficients of variation are so large that no return is statistically significant. This just confirms what was originally reported in section 3.3.2 that there are no significant trends.

We conclude then that although the returns to an investor in long cocoa and coffee contracts were large, the risks borne by such investors were very large indeed. The returns to rule 3, although lower on average than net long positions, were very much less risky.

8.7.2 Returns on the FT index

We now compare the returns from rule 3 to the returns one would have received by investing in a market portfoliio on the British Stock Market. Returns on the FT index (including dividends) were supplied by the London Business School. The deposits required to maintain a spread position in each of the futures markets were approximately £400 and so in order to compare like with like we examined the returns on £400 invested in the FT index over the 55 months in question. The question of commission charges is a difficult one as charges in the stock market vary with the size of transactions. As a rough figure we assumed transaction costs were 1% per round turn. Two strategies were examined. (i) Buy and hold for 55 months and (ii) buy and sell the FT index at the beginning and the end of each month. The returns on each stratergy were : (i) P = £307.00, M = £5.58, S = £18.91 and ν = 3.39; (ii) P = £95.04, M = £1.73, M = £18.91 and ν = 10.93.

As expected the returns to both strategies are very low, as are the standard deviation of returns. The relative riskiness of the returns, as measured by the coefficients of variation, is certainly higher than many of the corresponding measures of risk of returns obtained from rule 3.

It appears then that over the period considered applying trading rule 3 in the futures markets would have resulted in returns in the risk- adjusted sense that were higher than returns witnessed in the stock market.

8.8 Simultaneous application of rule 3 to all series

It is interesting to examine the joint distribution of the monthly returns produced by applying rule 3 to each of the series. In particular we are specifically interested in the correlations between returns. All 4 x 3/2 = 6 correlation coefficients were computed from returns generated by rule 3 on many α, γ and c values. All correlations were very similar and below we report a typical correlation matrix generated with α set to 0.14, γ set to 1.0 and low commission rates:

	cocoa	coffee	sugar	rubber
cocoa				
coffee	0.181			
sugar	-0.129	-0.135		
rubber	0.047	0.010	0.057	
FT	0.045	0.193	-0.058	-0.039

For comparative purposes we also report the correlations between the returns on the FT index and the 4 sets of trading rule returns.

Note that all correlations are small and that none are significantly different from zero. This result is interesting. The returns to rule 3 arise out of some, as yet, unexplained phenomenon that gives rise to persistent complex multivariate serial correlation. The phenomenon appears to be consistent within each commodity series. The zero correlations between each stream of returns suggests that the phenomenon is not common across commodity series. We are however only for the moment studying monthly returns and we have not yet investigated the possibility of joint daily multivariate serial correlation existing across commodity series.

However this lack of correlation between monthly returns to rule 3 does suggest the possibility of constructing portfolios that would have risks even lower than those outlined in section 8.6. It is well known that given a multivariate set of returns, provided the components are not perfectly positively correlated, there is scope for Markowitz-type diversification. The correlations reported above are near zero and some are negative and accordingly we constructed a naive portfolio, one in which the returns each month are simply the sum of the 4 sets of returns obtained from applying rule 3 to each series. The total margin required for such a portfolio would be £1650.

The mechanics of such a scheme would require the continuous simultaneous monitoring of 4 different portfolios, one from each futures market. On one day a trade would be executed in, say, the sugar market, on another day a trade may be executed in the coffee market and so on. On some days all 4 portfolios would be traded. The results of these trades are given in Table 8.26. In order to compare returns on this naive strategy with returns on each individual series it is necessary to correct for the larger commissions.

Table 8.26

Returns (in £'s)using rule 3 simultaneously on all 4 futures series

Commission level	scale	no. of trades	profit	monthly mean	monthly st. dev.	coeff. of var.	
	.6	1765	10918	198.51	440.00	2.22	b
	1.0	901	11056	201.02	346.00	1.72	c
	1.4	710	9117	165.78	244.40	1.47	c
Low	1.8	411	7517	136.67	241.20	1.76	c
	2.2	235	7346	133.56	218.40	1.64	c
	2.6	136	6723	122.24	218.40	1.79	c
	3.0	83	6396	116.29	208.80	1.80	c
	.6	1147	173	3.14	416.00	132.48	
	1.0	853	2077	37.77	332.00	8.79	
	1.4	560	2739	49.80	241.60	4.85	
Medium	1.8	361	3234	58.80	232.40	3.95	
	2.2	220	4753	86.42	210.8	2.44	b
	2.6	132	5165	93.92	209.6	2.23	b
	3.0	81	5443	98.96	200.40	2.03	c
	.6	744	-2918	-53.06	396.0	-	
	1.0	571	-1645	-30.99	333.2	-	
	1.4	386	-240	-4.37	245.6	-	
High	1.8	256	1125	20.46	231.6	11.32	
	2.2	111	3048	55.42	211.6	3.82	
	2.6	106	4017	73.04	207.6	2.84	a
	3.0	68	4659	84.70	193.2	2.28	b
FT index	*	55	392	7.13	77.99	10.94	
	**	1	1283	23.33	77.99	3.34	a

* Investing £1650 in FT index over 55 months with 1% turn round commission.
** Buy and hold £1650 in FT index fo entire period.

Only 3 of the total profits are negative. The very worst situation arises with a loss of £2918 with $\gamma = 0.6$ and high commission charges. The most impressive profits are £11056 with $\gamma = 1.0$ at the lowest commission level. All of the coefficients of variation at the low c level are significant and those above $\gamma = 2.2$ are significant even at the highest c level. These coefficients of variation of the returns to the naive strategy are, as expected, lower than many of the coefficients of variation of the individual returns.

Finally, for comparison purposes we also report in Table 8.26, information relating to the returns obtained by investing an equivalent sum, £1650, in the FT index over that period. Note that the coefficients of variation are larger than many of the ones realised from the naive portfolio.

We summarize then by saying tht at low commission levels, large and relatively low-risk returns were possible from the simultaneous application of rule 3 to all 4 futures markets. Even at the highest commission levels and with $\gamma = 2.6$, say, a stream of (on average positive) profits would have been obtained. At this value of γ, a profit of £4017 would have been realised. This represents an annualized return of 21% (if one is to use the deposit as an albeit inaccurate measure of initial investment). This profit would have been realised by the execution of 106 trades, approximately 1 every 11 days.

8.9 Limitations of trading rules

In the application of our final and most realistic trading scheme, rule 3 we explicitly incorporate the transaction costs as part of the trading strategy. However the author is still aware of 3 possible drawbacks to the practicalities of applying these rules. They relate to assumptions about (i) the prices; (ii) the possibility of getting executions at closing prices; and (iii) placing spreads in six contracts simultaneously.

8.9.1 Closing seller prices

All the prices used in this study were closing seller prices on each day (30,450 prices in all). In practice the execution of our rule requires the simultaneous buying and selling of 6 different contracts. It is well known that even in very liquid markets there is always a buyer-seller spread and such spreads will reduce the profits to our trading rules. The buyer-seller spread could be considered to be essentially an additional transaction cost and one could attempt to incorporate this into our rules by increasing such costs. However it is not really that simple. The buyer-seller spreads are not usually the same on the near and far contracts. It is more likely that a lower volume of trading in the far contracts will result in larger buyer-seller spreads and so a simple increase in transaction costs would not be the appropriate course of action.

The ICCH did supply all the closing buyer prices on each contract and each day and the author intends in time to edit these and incorporate them into this study. Nevertheless the author is still reasonably confident that even with the inclusion of buyer-seller spreads (provided they are not excessively large) the profits in the cocoa and coffee series at the low commission levels would be positive and significant.

8.9.2 Execution at close of day

In our trading rules, a trading signal is generated if the price of a special protfolio moves out of a certain band. Using these rules we have assumed that it is possible to get executions at the prices given at the close of each day. This may not be realistic. Once the market is closed no more trades are possible and one may be considering a slightly unrealistic situation of a trader waiting until a few moments before the close of day for a trading signal.

However, if it is not possible in reality to trade at the close of day prices it may be that prices in the final hours of trading are not that far from the closing prices. Indeed, it is perfectly feasible that the rule could be successfully applied not just to end of day prices but to, say end of hour prices. It may be that the returns to a rule which continuously monitors the prices of all 6 contracts minute by minute may result in significantly larger profits than those reported here.

8.9.3 Spreads of 6 contracts

A major assumption of all our trading rules has been that a trader can simultaneously buy and sell different quantities of 6 contracts to form complex spread positions. It is of course possible for an individual, through a broker, to buy a simple 2 dimensional spread across any pair of contracts. In fact there is evidence (Schrock(1971)) that a significant proportion of the trades in the American futures markets are of this sort. However the simultaneous purchase and sale of 6 contracts in the proportions delineated by procedures outlined in chapter 8 may not be possible in practice.

The author has, however, experimented with applying the above trading rules to portfolios that are much less complex than the ones used in this study. The revised portfolios contained very much fewer positions, often only in 3 or 4 different futures contracts. These constrained portfolios still exhibited ex ante multivariate serial correlation and produced equally encouraging profits. We leave the discussion of this research to another work.

8.10 Concluding remarks on Chapter 8

In this chapter we outlined three simple trading rules that were designed to exploit the observed persistent multivariate temporal dependence in commodity futures prices. The trading rules proved very profitable on all series with zero or low commission charges. Even with the very largest transaction costs the rules still proved profitable in the cocoa and coffee series. With low transaction costs the stream of returns to these rules proved not only to be positive, but highly statistically significant. Furthermore it appears that investing in such a scheme would have resulted in a sequence of returns that would have yielded, in a risk-adjusted sense, returns superior to those obtained by investing in a basket of common stocks.

Footnotes for Chapter 8

1. This is equivalent to using an exponential prcedure with the smoothing constant set to 1.0.
2. Note that when the study was repeated with method (iii) , type $\underline{\hat{c}}$ estimates, the results were remarkably similar.
3. However, it is possible to consider a slightly more sophisticated

and ambitious strategy in which the investor would increase his trading unit to 2 or 3 spreads in those times when there had been a run of positive returns. Such a strategy may result in even larger returns (and risks) than those reported here.

CHAPTER 9

MULTIVARIATE MODELS OF COMMODITY FUTURES PRICES

In this chapter we propose two models of the multivariate distribution of prices that could explain much of the observed significant multivariate serial correlation of lags of 1 day.

Recall from the univariate study of returns in Chapter 3 that, apart from the coffee series, there was no evidence of significant serial correlation at a lag of 1 day. However as reported in Chapter 5 there is evidence of a persistent and highly significant multivariate serial correlation at a lag of 1 day in all series. This was examined in more detail in Chapter 6 and in Chapter 7 led to the development of special linear combinations of returns that seemed to explain almost all of the observed multivariate serial correlation.

In section 5.3.1 we outlined the results of applying the robust estimation and outlier detection routine to the distribution of returns on 4 typical futures contracts from the same market. Many more outliers were detected than in a similar study on 4 contracts, with each contract being from a different futures market.

Half of the observed outliers appeared in contemporaneous pairs and corresponded to situations in which one of the 4 contract prices moved out of line for one day only to get back into line the very next day.

Although only 79 out of a total of 30450 returns could be attributed to these infrequent deviations of contract returns from the norm, the discovery did suggest a possible cause for the observed MVSC.

It may be that there are many such instances of prices occasionally deviating from the main price profile. If these deviations are small compared to the overall movements in daily returns they may not be detectable even by the outlier routine.

Consider an albeit unrealistic example of the prices of all 6 contracts moving from day to day up and down in perfect unison. For simplicity assume that the variability of all the returns are identical. Consider a day when events in the futures market dictate that all prices should rise by, say 30 units. On that day the contract returns should all be +30 units. Consider the situation however in which for some reason the price of the 6th contract only rises 20 units. The returns on the 6th contract that day will be only +20 units. Assume also that the next day the 1st 5 contracts rise a further 30 units. If the 6th contract is to get back into line its price must rise by 40 units. Thus one small price disturbance has caused two anomalous returns, one below and one above the average returns of most of the contracts on each day. If such disturbances of prices are small relative to the overall movements of returns then it is unlikely that any anomalous temporal patterm will be observed by a univariate examination of each series.

Consider however the distribution of returns to an individual trading a spread of 1 long position in each of the 1st, 3rd and 5th contracts and 1 short position in each of the 2nd, 4th and 6th contracts. If the prices move perfectly in unison the returns on the spread will be exactly zero each and every day. However, if as described above, the price of the 6th contract lags behind by 10 units on one day the returns on the spread will be +10 units. The very next day the returns on the spread will be -10 units.

Price disturbances in any of the contracts will always give rise to such switching of spread returns. Furthermore even though the disturbances may be small relative to the variation of returns within each contract series, the resulting positive and negative disturbances in the spread returns compared to the overall variability of spread returns can be very large indeed. It is for this reason that we now propose the first of two models of the multivariate distribution of returns.

9.1 Model 1 A simple model of the multivariate distribution of daily returns

The first model that we propose to explain the multivariate distribution of prices is as follows:

$$\begin{aligned} \underline{P}_t &= \underline{P}_{t-1} + \underline{r}_t \\ \underline{r}_t &= \underline{x}_t + \underline{d}_t - \theta\, \underline{d}_{t-1} \end{aligned} \qquad \ldots\ldots (9.1)$$

In which:

(i) $\underline{P}_t$ is the (p dimensional) vector of prices at the close of trading on day t.

(ii) $\underline{r}_t$ is the vector of returns on day t.

(iii) $\underline{x}_t$ is a vector of random variables that represents the majority of the returns on day t. We define $E(\underline{x}_t) = \underline{0}$ and $Var(\underline{x}_t) = V_x$, With V_x being a symmetric positive definite matrix. Furthermore, all the co-movement of returns in the series is explained by the off diagonal terms of V_x. Accordingly V_x is nearly singular with many off diagonal terms being near 0.95. Also $\underline{x}_t$ and $\underline{x}_{t-k}$ are independent for all k = 1, 2, ..

(iv) $\underline{d}_t$ is a vector of disturbance terms. The i th component of $\underline{d}_t$, d_{it} takes on the values $\{\Delta_i, 0, -\Delta_i\}$ with probabilities $\{p_i/2, (1-p_i), p_i/2\}$ respectively. Each component of $\underline{d}_t$ is indpendent of the other components and furthermore pairs : ($\underline{d}_t$ and $\underline{d}_{t-k}$) and ($\underline{x}_t$, $\underline{d}_{t-1}$) are independent for all t = 1, 2, .. and k = 0, 1, 2, ...

It is easy to show that

$$E\ (\underline{d}_t) = 0 \quad , \quad Var\ (\underline{d}_t) = diag\ \{\Delta^2_i p_i\}$$
$$= V_d$$

(v) θ is a (ρ x ρ) matrix of constants and in model 1 we shall assume all off diagonal terms are zero, i.e.

$$\theta = diag\ \{\ \theta_i\ \}$$

In this way we model exactly the situation outlined in the introduction to this chapter. On most days the returns are simply identical to $\underline{x}_t$. Occasionally the i th component of $\underline{r}_t$ will be disturbed by a small amount; Δ_i or $-\Delta_i$. This will occur with a probability of p_i. The next day a certain fraction, θ_i, of this disturbance in the returns will be removed. The disturbances to the returns and subsequent prices are completely at random and the rates are set by the parameters p_i, i = 1, to ρ. A priori we would assume pi be small. The magnitude of the disturbances will be governed by the parameters Δ_i, i = i to ρ and the degree of correction of each disturbance on the following days is set by θ_i, i = 1 to ρ.

We now investigate the properties of such a model and show analytically that the presence of small disturbance described by this model can give rise to MVSC.

9.2 Derivation of the MVSC coefficient and temporally dependent portfolios of model 1

Consider the model given by expression 9.1. We now derive the expected returns, the variance of returns on day ,t, and the covariance of returns between day t and day t-1.

$$\underline{r}_t = \underline{x}_t + \underline{d}_t - \theta\, \underline{d}_{t-1}$$

$$E(\underline{r}_t) = E(\underline{x}_t) + E(\underline{d}_t) - \theta\, E(\underline{d}_{t-1})$$

$$E(\underline{r}_t) = \underline{0} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (9.2)$$

$$\mathrm{Var}(\underline{r}_t) = \mathrm{Var}(\underline{x}_t) + \mathrm{Var}(\underline{d}_t) + \theta\, \mathrm{Var}(\underline{d}_{t-1})\, \theta^T$$

$$V_r = V_x + V_d + \theta\, V_d\, \theta^T \quad \ldots\ldots\ldots\ldots\ldots\ldots (9.3)$$

In model 1 we set $\theta = \mathrm{diag}\{\theta_i\}$ in which all $\theta_i > 0$
and therefore (9.3) reduces to

$$V_r = V_x + \mathrm{diag}\,[(1 + \theta^2_i)\, \Delta^2_i\, p_i] \quad \ldots\ldots\ldots\ldots (9.4)$$

$$\mathrm{Cov}(\underline{r}_t, \underline{r}_{t-k}) = \mathrm{Cov}\,[(\underline{x}_t + \underline{d}_t - \theta\, \underline{d}_t), (\underline{x}_{t-k} + \underline{d}_{t-k} - \theta\, d_{t-k-1})$$

$$= \left.\begin{array}{l} -\theta\, \mathrm{Var}(\underline{d}_t) \\ 0 \end{array}\right\} \quad \begin{array}{l} \text{for } k = 1 \\ \text{for } k = 2, 3, \ldots \end{array}$$

Thus

$$\mathrm{Cov}(\underline{r}_t, \underline{r}_{t-1}) = -\theta\, V_d \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (9.5)$$

In model 1 this reduces to

$$\mathrm{Cov}(\underline{r}_t, \underline{r}_{t-1}) = \mathrm{diag}\{-\theta_i\, \Delta^2_i\, p_i\} \quad \ldots\ldots\ldots\ldots (9.6)$$

In model 1, then, the variance of returns matrix is represented by equation 9.4 and is simply the variance matrix of the $\underline{x}_t$ process with additional (small) terms of $(1 + \theta^2_i)\Delta^2_i p_i$ along the diagonal. These small additions will have the effect of increasing, slightly, the variance of each component of $\underline{r}_t$ over the variance of the corresponding component of the $\underline{x}_t$ process. In addition to this, the correlation coefficients between each pair of components of $\underline{r}_t$ will be slightly lower than the correlation coefficients of corresponding pairs of the $\underline{x}_t$ process.

The covariance between returns on day t and day t - 1 is defined purely in terms of the parameters θ_i, Δ_i and p_i . The covariance matrix in expression 9.5 is diagonal. Thus the i th component of $\underline{r}_t$ is negatively correlated with the i th component of $\underline{r}_{t-1}$ and not correlated with any of the other components of $\underline{r}_{t-1}$. Assuming model 1 is a true description of the process generating the returns, we now derive the temporally correlated portfolios, $\underline{c}$, referred to in Chapters 7 and 8. Substituting the variance and covariance matrices derived in 9.4 and 9.6 into equation 7.5 we have:

$$A = [V_x + \text{diag}\,[(1 + \theta^2_i)\Delta^2_i p_i]^{-1}\ [\ -2\ \text{diag}\ [\theta_i\ \Delta^2_i\ p_i]\ ]. \ . \ . (9.7)$$

Interest is centred on the minimum eigenvalue (2τ) and corresponding eigenvector $\underline{c}$ of A. If we divide 9.7 by -2 we have

$$A^* = [V_x + \text{diag}\,[(1 + \theta^2_i)\Delta^2_i p_i]^{-1}\ [\ \text{diag}\ [\theta_i\ \Delta^2_i\ p_i]\ ]\ . \ (9.8)$$

and interest is now centred on the maximum eigenvalue $(-\tau)$ and associated eigenvector $(\underline{c})$ of A^*. The $\underline{c}$ vector will give the special portfolio $\underline{c}^T\ \underline{r}_t$ that will be maximally negatively correlated with

$\underline{c}^T \underline{r}_{t-1}$. The serial correlation coefficient will be, $-\tau$. It is interesting to note that if we express A^* in (9.8) as

$$A^* = B^{-1} C$$

in which
$$B = V_x + \text{diag}\,[(1 + \theta^2_i)\Delta^2_i p_i]$$

and
$$C = \text{diag}\,(\theta_i \Delta^2_i p_i)$$

then since B and C are positive definite, all the eigenvalues of A^* are real and positive . Thus the maximum eigenvalue, $-\tau$, will be positive and so the correlation between $\underline{c}^T \underline{r}_t$ and $\underline{c}^T \underline{r}_{t-1}$, τ , will always be negative.

The analytic derivation of the eigenvalues and eigenvectors of A^* in (9.8) in terms of θ_i, Δ_i and p_i appears to be intractable and so the author experimented with various θ_i, Δ_i and p_i values in an attempt to duplicate the empirical temporal portfolios outlined in Chapter 7.

However the problem is more complex than finding θ_i, Δ_i and p_i values that produce $\underline{c}$ and τ similar to the values observed in practice. For consistency the variances and covariances given by (9.4) and (9.6) must also be reasonably similar to the ones observed in practice. We defer the treatment of this problem to section 9.3 but note in passing that in the experimentation process the maximum negative correlation coefficient, τ , that was produced was -0.49. Recall from section 7.6.4 that using the grand average correlation matrices the average temporal correlation appeared in all 4 series to be approximately -0.45. This led the author to investigate, in more detail, the properties of the eigenvalues of, A^*, under certain simplifying assumptions.

9.2.1 A theoretical lower bound on the degree of serial correlation in multivariate portfolios

Consider, again, the matrix A^* in (9.8). It is simpler if we turn our interest to the inverse of A^*, $A^{**} = (A^*)^{-1}$. It is well known that if $(-\tau)$ and $\underline{c}$ are the corresponding eigenvalues and eigenvectors of A^* then $(-1/\tau)$ and $\underline{c}$ are corresponding eigenvalues and eigenvectors of A^{**}, where,

$$A^{**} = \text{diag}\left[\frac{1}{\theta_i \Delta^2_i\ p_i}\right] V_x + \text{diag}\left[\frac{(1+\theta^2_i)}{\theta_i}\right] \quad \ldots\ldots (9.9)$$

So the minimum eigenvalues $(-1/\tau)$ of A^{**} will correspond to the maximum egienvalues $(-\tau)$ of A^*.

We now make three assumptions; (i) each θ_i is equal to a constant correction factor θ_∞, (ii) each disturbance component Δ_i is equal to a constant value Δ_∞ and (iii) all p_i values are identical and equal to p_∞. These values substituted in (9.9) affords a simplification:

$$A^{**} = \left[\frac{1}{\theta_\infty \Delta^2_\infty\ p_\infty}\right] . V_x + \left[\frac{1+\theta^2_\infty}{\theta_\infty}\right] I \quad \ldots\ldots (9.10)$$

With these simplifications the eigenvalues of A^{**}, $(-1/\tau)$, can easily be expressed in terms of the eigenvalues of V_x, λ_x as follows:

$$-\frac{1}{\tau} = \frac{1}{\theta_\infty \Delta^2_\infty\ p_\infty} . \lambda_x + \frac{1+\theta^2_\infty}{\theta_\infty}$$

$$\text{or} \quad (-r) \quad = \quad \frac{+1}{\left[\dfrac{1}{\theta_{\varnothing}\Delta^2_{\varnothing}\, p_{\varnothing}}\right] \lambda_* \quad + \left[\dfrac{1+\theta^2_{\varnothing}}{\theta_{\varnothing}}\right]} \quad \ldots . \, (9.11)$$

The value of $(-r)$ of interest is the maximum value and this will be given in terms of the minimum eigenvalue of V_* in (9.11). The V_* matrix is positive definite and so all, even the smallest, λ_*, will be positive. However in our models V_* will be nearly singular and thus the smallest eigenvalue, λ_*, will be near zero. If we substitute the lower bound of such, λ_*, i.e. zero in (9.11) we have

$$r \quad = \quad \frac{-\theta_{\varnothing}}{1+\theta^2_{\varnothing}} \quad \ldots . \, (9.12)$$

Note that in this situation the magnitude of the temporal dependence is now dependent only on, $\theta_{\varnothing}$, the degree of correction of price disturbances and independent of the magnitude ($\Delta_{\varnothing}$) and frequency ($p_{\varnothing}$) of disturbances.

A priori we would expect $\theta_{\varnothing}$ to lie in the range (0, 1) and it is interesting to note that if we substitute $\theta_{\varnothing} = 1$ (corresponding to model in which a price disturbance is corrected perfectly the very next day) into (9.12) we obtain $r = -0.5$.

Thus if the multivariate set of highly collinear prices were generated according to model 1, we see that MVSC of lag 1 day would be observed. Furthermore with certain simplifying assumptions relating to the parameters of the model it is possible to show that the maximum degree of negative serial correlation possible would be -0.5. It is extremely interesting to note that the maximal negative serial correlation observed when using the grand average technique over the entire period was in fact -0.461.

9.3 The Grand Average R_{11} and R_{12} matrices

At this stage we will consider how one might estimate the parameters of model 1 that would best explain the observed variance/covariance matrices and resulting eigenvalues and eigenvectors.
Obviously we would wish to find θ_i, Δ_i, p_i and V_u that result in V_r and $Cov(\underline{r}_t, \underline{r}_{t-1})$ matrices that are as near as possible to the observed ones. However here again we encounter the problem present throughout this study, that of fluctuating variances. Model 1 assumes a constant variance structure.

It is possible of course to extend model 1 to one of fluctuating variance and one could use either of the variance processes suggested by Taylor (1980). We will instead approach the problem in a much cruder way, and simply assume that variances remain fairly constant within the short time periods examined and remodelthe process in terms of correlation matrices. Although the variances vary considerably over time we are assuming the correlation structure is constant. In fact this very assumption underlies the rationale of calculating the grand average correlation matrices used in Chapters 6 and 7.

Using the notation of Chapters 6 and 7 and the variance/covariance matrices given in (9.4) and (9.6) we arrive at the following correlation matrices for model 1.

$$R_{11} \quad = \quad \{\mathrm{diag}\ (V_r)\}^{-1/2}\ V_r\ \{\mathrm{diag}\ (V_r)\}^{-1/2} \qquad \ldots\ldots(9.13)$$

and

$$R_{12} \quad = \quad \{\mathrm{diag}\ (V_r)\}^{-1/2}\mathrm{Cov}(\underline{r}_t,\ \underline{r}_{t-1})\ \{\mathrm{diag}\ (V_r)\}^{-1/2} \ldots\ldots(9.14)$$

Note that in (9.14) all the off diagonal terms of R_{12} are zero and so we would expect the sample R_{12} matrix to contain negative elements along the diagonal term and all off diagonal elements to be zero. It was considered for reasons argued in section 6.9.3 that a good estimate of R_{11} and R_{12} for each series would be the grand average matrices and accordingly we now present these in Tables 9.1 to 9.4

Referring to Tables 9.1 to 9.4 we notice that the sample R_{12} matrices are not exactly as expected, i.e. purely diagonal. Although many of the diagonal terms are negative, many are smaller in magnitude than the off diagonal terms. However there appears to be a general pattern common to each R_{12} matrix. We make the following remarks on this general pattern:

(i) With the exception of two elements, all the elements in the first column are positive.

(ii) Apart from the terms in column 1, most of the elements are negative.

(iii) In each row as one spans the elements from column 1 to column 6, the elements tend to decrease in value.

These results are interesting and certainly unexpected. It appears then that the correlations between elements of $\underline{r}_t$ and $\underline{r}_{t-1}$ is more complex than that proposed by model 1. The asymmetry of R_{12} is difficult to explain.

Consider the 1st row and the 1st column of the rubber series, as an

Table 9.1

Grand average R_{11} and R_{12} matrices for the cocoa series

$r_{i,t}$ = returns on ith component on day t

R_{11} =

	$r_{1,t}$	$r_{2,t}$	$r_{3,t}$	$r_{4,t}$	$r_{5,t}$	$r_{6,t}$
$r_{1,t}$	1.000					
$r_{2,t}$	.888	1.000				
$r_{3,t}$	.843	.954	1.000			
$r_{4,t}$	.812	.930	.957	1.000		
$r_{5,t}$	.798	.909	.942	.957	1.000	
$r_{6,t}$	.782	.889	.923	.936	.954	1.000

R_{12} =

	$r_{1,t}$	$r_{2,t}$	$r_{3,t}$	$r_{4,t}$	$r_{5,t}$	$r_{6,t}$
$r_{1,t-1}$	.021	-.014	-.036	-.037	-.034	-.036
$r_{2,t-1}$	.043	-.022	-.029	-.032	-.018	-.026
$r_{3,t-1}$	.059	.003	-.030	-.018	-.006	-.013
$r_{4,t-1}$	.044	-.007	-.025	-.046	-.019	-.025
$r_{5,t-1}$	.056	.004	-.018	-.022	-.031	-.025
$r_{6,t-1}$	.060	.007	-.014	-.016	-.016	-.042

Table 9.2

Grand average R_{11} and R_{12} matrices for the coffee series

R_{11} =

	$r_{1,t}$	$r_{2,t}$	$r_{3,t}$	$r_{4,t}$	$r_{5,t}$	$r_{6,t}$
$r_{1,t}$	1.000					
$r_{2,t}$	.811	1.000				
$r_{3,t}$	.773	.946	1.000			
$r_{4,t}$	.752	.936	.959	1.000		
$r_{5,t}$	.740	.912	.937	.971	1.000	
$r_{6,t}$	.719	.884	.914	.948	.960	1.000

R_{12} =

	$r_{1,t}$	$r_{2,t}$	$r_{3,t}$	$r_{4,t}$	$r_{5,t}$	$r_{6,t}$
$r_{1,t-1}$	.005	.008	-.001	-.021	-.013	-.004
$r_{2,t-1}$	.018	-.011	-.004	-.016	-.008	-.007
$r_{3,t-1}$	.021	-.003	-.027	-.019	-.008	-.001
$r_{4,t-1}$	.023	.009	-.002	-.018	.001	.013
$r_{5,t-1}$	.032	.013	.001	-.010	-.012	.015
$r_{6,t-1}$	.027	.008	.002	-.014	-.010	-.013

Table 9.3

Grand average R_{11} and R_{12} matrices for the sugar series

		$r_{1,t}$	$r_{2,t}$	$r_{3,t}$	$r_{4,t}$	$r_{5,t}$	$r_{6,t}$
	$r_{1,t}$	1.000					
	$r_{2,t}$	.804	1.000				
R_{11} =	$r_{3,t}$	.783	.978	1.000			
	$r_{4,t}$	.768	.955	.975	1.000		
	$r_{5,t}$	.746	.934	.959	.978	1.000	
	$r_{6,t}$	.726	.918	.943	.951	.967	1.000

		$r_{1,t}$	$r_{2,t}$	$r_{3,t}$	$r_{4,t}$	$r_{5,t}$	$r_{6,t}$
	$r_{1,t-1}$	-.030	-.027	-.045	-.046	-.050	-.062
	$r_{2,t-1}$	.079	-.001	-.015	-.013	-.020	-.030
R_{12} =	$r_{3,t-1}$	.076	.002	-.018	-.010	-.018	-.027
	$r_{4,t-1}$	.068	-.005	-.019	-.020	-.024	-.028
	$r_{5,t-1}$	.059	-.014	-.026	-.020	-.040	-.040
	$r_{6,t-1}$	.056	-.017	-.028	-.018	-.033	-.057

Table 9.4

Grand average R_{11} and R_{12} matrices for the rubber series

		$r_{1,t}$	$r_{2,t}$	$r_{3,t}$	$r_{4,t}$	$r_{5,t}$	$r_{6,t}$
	$r_{1,t}$	1.000					
	$r_{2,t}$	.815	1.000				
R_{11} =	$r_{3,t}$	.799	.951	1.000			
	$r_{4,t}$	.779	.930	.974	1.000		
	$r_{5,t}$	.763	.908	.954	.967	1.000	
	$r_{6,t}$	.741	.878	.916	.933	.942	1.000

		$r_{1,t}$	$r_{2,t}$	$r_{3,t}$	$r_{4,t}$	$r_{5,t}$	$r_{6,t}$
	$r_{1,t-1}$	-.068	-.045	-.057	-.064	-.070	-.078
	$r_{2,t-1}$	.021	-.039	-.038	-.042	-.058	-.072
R_{12} =	$r_{3,t-1}$	.023	-.007	-.035	-.037	-.053	-.066
	$r_{4,t-1}$	.026	-.010	-.030	-.051	-.060	-.070
	$r_{5,t-1}$	.017	-.011	-.034	-.046	-.080	-.074
	$r_{6,t-1}$	.028	-.009	-.021	-.034	-.053	-.091

example. In particular we consider the elements:

$$(R_{12})_{1,6} = -0.078 \quad \text{and} \quad (R_{12})_{6,1} = +0.028$$

From this we infer that the components $r_{1,t}$, and $r_{6,t-1}$ are negatively correlated whereas componenets $r_{6,t-1}$ and $r_{1,t}$ are positively correlated. This effect is observed in all 4 series.

It is interesting to consider also the four R_{11} matrices. Note that all the correlations are high and positive and that as one considers components that are further apart the correlations decrease. This seems a reasonable result. The degree of comovement of contract prices (and hence returns) may be directly proportional to the distance in days between maturity dates. However, what perhaps is unexpected is the relatively low values of the coefficients in the first column of each R_{11} matrix. It appears that there is a high degree of correlation amongst the contract returns that are not immediately due for delivery; contracts 2 through 6, but that this very high degree of correlation is not shared with the contract that is due for immediate delivery; contract 1. The above observations suggest that the distribution of returns on contract 1 is subtely different from the distribution of the returns on all the other contracts and furthermore that the nature of the association between the entire set of returns on day t and day t-1 is much more complex than that explained by model 1.

9.4 Model 2

In this section we consider a more complex model of the multivariate distribution of returns that could explain, not only the observed MVSC but also the observed R_{11} and R_{12} matrices reported in section 9.3.

Consider again the model of returns defined in section 9.1:

$$\underline{r}_t = \underline{x}_t + \underline{d}_t - \theta\, \underline{d}_{t-1}$$

In this model 1 we assumed θ to be a diagonal matrix. In this way, the i th component of $\underline{r}_t$, $(r_{i,t})$ would be correlated only with the i th component of $\underline{r}_{t-1}$, $(r_{i,t-1})$ through θ_i, i.e.

$$r_{i,t} = x_{i,t} + d_{i,t} - \theta_i\, d_{i,t-1}$$

If we now revert to a more general model in which θ is a square matrix of non zero constants, $\{\theta_{ij}\}$ the association between $\underline{r}_t$ and $\underline{r}_{t-1}$ becomes more complex. In this model, model 2, the ith component of $\underline{r}_t$ is now given by:

$$r_{i,t} = x_{i,t} + d_{i,t} - \sum_{j=1}^{p} \theta_{ij}\, d_{j,t-1}$$

In this model, $r_{i,t}$ would experience not only corrections from disturbances to $r_{i,t-1}$ but also corrections from disturbances to $r_{k,t-1}$, through $\theta_{i,k}$. This may be reasonble in the sense that a random shock to the i th contract price may have an impact the next day on the prices of the other contracts. The θ_{ij} will represent the influence of a disturbance to the j th contract on day t-1 on the returns in the ith contract on day t.

Using arguments similar to section 9.3 we derive the theoretical

R_{11} and R_{12} matrices of model 2.

$$R_{11} = \{\text{diag}\,(V_r)\}^{-1/2}\, V_r\, \{\text{diag}\,(V_r)\}^{1-/2}$$

$$R_{12} = -\,\{\text{diag}\,(V_r)\}^{-1/2}\, \theta\, V_d\, \{\text{diag}\,(V_r)\}^{-1/2}$$

with $V_r = V_x + V_d + \theta\, V_d\, \theta^T$

. . . .(9.15)

and $V_d = \text{diag}\,\{\Delta^2_i\, p_i\}$

Note that each element of R_{11} and R_{12} in (9.15) is a complex function of O_{ij}, Δ_i, p_i, $i = 1,..,p$ and the elements of V_x.

9.4.1 Estimation of parameter values for Model 2

The estimation of the values of the parameters in Model 2 presents a problem. There are 21 elements of V_x, 36 elements of θ and 6 $(\Delta^2_i p_i)$ terms[3]. How does one find the values of all 63 parameters? As an attempt at estimation we considered minimising a function of the form:

$$F = \sum_{i>j} [(R_{11})_{i,j} - (\widehat{R}_{11})_{i,j}]^2 + \sum_{i,j} [(R_{12})_{i,j} - (\widehat{R}_{12})_{i,j}]^2 \quad . . \ (9.16)$$

where $(R_{IJ})_{i,j}$ = theoretical elements of R_{IJ} matrix given by (9.15)

and $(\widehat{R}_{IJ})_{i,j}$ = sample values of elements of grand average R_{IJ} matrix given in Tables 9.1 to 9.4 .

In this way it was hoped that one would be able to find parameter values that resulted in theoretical R_{11} and R_{12} matrices that were as 'near' as possible the observed matrices. The minimization is of course subject to the constraints:

$$|V_x| > 0$$

$$|V_r| > 0$$

$$\begin{vmatrix} R_{11} & R_{12} \\ R_{21} & R_{11} \end{vmatrix} > 0 \qquad (9.17)$$

The minimisation of such a function requires the analytic derivation of the differential of F and the constraints with respect to each of the 63 variables. Accordingly a program was constructed that computed the values of (i) F given by (9.16) , (ii) the constraints given by (9.17), (iii) the rates of change of F and the constraints with respect to each parameter (iv) the R_{11} and R_{12} matrices and (v) the eigenvalues and eigenvectors of A given by (7.12). The program called the NAG minimisation routine: E04VBF.

Many problems have been encountered with this procedure. No single global minimum could be found and those local minima that were found were 'ill conditioned', a small shift in one of the 63 parameter values alters the values of F drastically. Altering the range constraints on the parameters invariably produced different local minima. The computer processor time for a typical minimisation run was 3 to 4 hours.

In an attempt to simplify the problem and possibly speed up the minimisation routine, many simplifying assumptions on the parameter structures were tried. Examples of some of the simplifications are given below.

(i) Constraining all the correlation coefficients defining the $\underline{x}_t$

process to be modelled by

$$(r_x)_{i,j} = a\,(\omega)^{|(i-j+1)|}$$

with a, $\omega < 1$

In this way the correlation between near components of $\underline{x}_t$ is higher than the correlation between distant components.

(ii) Constraining all the standard deviations of the $\underline{x}_t$ process to decrease monotonically according to the relation:

$$(\sigma_x)_i = (\sigma_x)_1 - (\sigma_c)i$$

in which σ_c = change in standard deviation of $\underline{x}_t$ vector from component to component.

(iii) Many simplifications on the structure of the θ matrix were tried.

None of these simplifications has yet led to a satisfactory solution and we leave this problem as the subject for further research.

9.4.2 Some suggested parameter values for model 2

Although attempts at finding global minima have yet to prove fruitful, some degree of success has been obtained by the simple substitution of trial parameter values into (9.15). With 63 variables there is obviously an enormous number of possibilities to consider. However the detailed empirical examination of the multivariate distribution outlined in Chapter 5 led one to experiment with certain combinations of parameter values. Many such combinations produced theoretical R_{11} and R_{12} matrices that were very near those given in Table 9.1 to 9.4.

In Table 9.5 we present one example of such a set of parameter values along with the resulting theoretical τ and $\underline{c}$ values. We briefly outline the rationale behind choosing this particular set.

(i) Recall the remarks on the components of the $\underline{c}$ vectors made in section 7.6.5. We suspect that the major contributing factor to the multivariate temporal dependence comes from the middle to far contracts. With this in mind we choose $(\Delta^2{}_i p_i)$ values to be scaled accordingly.

(ii) In section 5.2.1 we noted that the standard deviation of returns generally decreased as one considered contracts further and further from maturity. The standard deviations of the $\underline{x}_t$ process were thus set to decrease monotonically.

(iii) In section 5.3.2 we noted that all the cross contract correlations were very high. The correlation coefficients of the $\underline{x}_t$ process were thus set to near unity.

(iv) Considering the grand average R_{12} matrices in Table 9.1 to 9.4 we note that the correlation between each component of $\underline{r}_t$ and $\underline{r}_{t-1}$ is extremely complex and in an attempt to reproduce a typical R_{12} matrix we set θ to be as in Table 9.5. For simplicity the diagonal terms are set to 1.0 and all off diagonal terms (except the 1st column) were set to 0.3. The elements in the 1st column were set to -1.0.

The substitution of these parameter values in (9.15) result in R_{11} and R_{12} matrices remarkably near those given in Table 9.1 to 9.4. Furthermore the resulting theoretical MVSC, τ, and the $\underline{c}$ vector are consistent with those of the grand average estimates given in Tables 7.12 to 7.15.

It must be pointed out that very many different sets of parameter values have been tried and the dependence of R_{12} and R_{11} matrices on the individual values is very complex indeed. Increasing the disturbance term parameters $\Delta^2{}_i p_i$ tends to reduce the elements of R_{11} and increase the

Table 9.5

A set of parameter values for model 2 with theoretical results

σ_x:	1.00	0.96	0.92	0.88	0.84	0.80

	1.000					
	0.950	1.000				
	0.912	0.980	1.000			
Cor_x:	0.876	0.951	0.980	1.000		
	0.840	0.922	0.951	0.980	1.000	
	0.807	0.894	0.922	0.951	0.980	1.000

$\Delta^2{}_i p_i$:	0.0204	0.0424	0.0562	0.0697	0.0467	0.0294

	1.0	0.3	0.3	0.3	0.3	0.3
	-1.0	1.0	0.3	0.3	0.3	0.3
	-1.0	0.3	1.0	0.3	0.3	0.3
θ	-1.0	0.3	0.3	1.0	0.3	0.3
	-1.0	0.3	0.3	0.3	1.0	0.3
	-1.0	0.3	0.3	0.3	0.3	1.0

In the above σ_x = the standard deviation of the $\underline{x}_t$ process and Cor_x = the correlation structure of the $\underline{x}_t$ process.

The substituting these into equatuion (9.15) yields:

	1.000					
	0.875	1.000				
R_{11}:	0.829	0.912	1.000			
	0.783	0.872	0.887	1.000		
	0.761	0.860	0.875	0.887	1.000	
	0.737	0.847	0.862	0.874	0.912	1.000

	-0.019	-0.012	-0.016	-0.021	-0.015	-0.010
	0.019	-0.041	-0.016	-0.021	-0.015	-0.010
	0.020	-0.012	-0.056	-0.021	-0.015	-0.010
R_{12}:	0.020	-0.013	-0.017	-0.073	-0.016	-0.011
	0.022	-0.014	-0.018	-0.023	-0.056	-0.011
	0.023	-0.014	-0.020	-0.025	-0.018	-0.040

							Sum of components
$\underline{c}^T$:	-13	103	-248	320	-237	79	4

τ = -0.443

elements of R_{12}. The τ and $\underline{c}$ values are very sensitive to the specification of θ. The relationship between the parameter values and the resulting τ, $\underline{c}$ values is left as the subject for further research.

9.5 Implications of models for multivariate distributions of commodity futures returns

It is likely that processes of the type described in this chapter will have complex multivariate distributions. If the $\underline{x}_t$ process in model 1 or 2 is multivariate normal then disturbances $\underline{d}_t$ and corrections $\theta\underline{d}_{t-1}$ will obviously result in multivariate non - normal distributions. In section 5.3.3 we investigated the multivariate distribution of returns on four contracts from the same futures market. Highly significant multivariate skewness and multivariate kurtosis statistics were reported. Removing outliers reduced the magnitude of the statistics but many remained significant. Thus if the returns are generated by processes described by the above models this may explain not only the observed multivariate temporal behaviour but also the observed extreme departure from multivariate normality.

The mathematical treatment of the multivariate skewness and kurtosis of model 1 and 2 returns appears to be intractable and we leave this as an interesting area of research.

9.6 Conclusions of Chapter 9

In this chapter we outlined two models of the multivariate distribution of commodity futures prices. These models involve small perturbations in individual components of the price vector. We have demonstrated that processes generated from these models would exhibit a multivariate temporal dependence consistent with that observed throughout this study and could also explain the significant multivariate skewness and kurtosis values. In the first model, under certain limiting conditions, the maximum degree of negative serial correlation possible in multivariate spread portfolios is 0.5. A second model was introduced in an attempt to explain the complex nature of the observed R_{12} matrices. Although we have yet to estimate the parameters of this second model, a trial set of parameters produced theoretical results that were very near those observed in practice.

CHAPTER 10

SUMMARY

In this study we examined in detail the univariate and multivariate distribution of daily returns of futures contracts in the four major London soft commodity futures markets. The univariate study revealed that the distributions were highly nonstationary in variance and only one series (rubber) could be described as approximately normally distributed. An investigation into the univariate temporal dependence of long series of contracts woven together produced evidence of long term negative serial correlation.

In the multivariate analysis, three relatively recent procedures (the multivariate serial correlation, the multivariate extension of the W - test for normality and a multivariate outlier detection routine) were investigated. The four dimensional vector of returns with one component from each futures markets can be viewed as being generated from a serially independent multivariate normal process with non - constant variance/covariance structure with occasional contaminating extreme realisations.

Examining the multivariate distribution in which all the components are returns on contracts in the same futures market however produced different and very unexpected results. Significant departures from multivariate normality were witnessed and highly significant multivariate serial correlation coefficients of lag one day were discovered. The departures from multivariate normality could not be completely explained by the presence of obvious anomalous "spikes" in the plots and the sensitivity of the multivariate W - test was brought into question.

The multivariate temporal dependence was shown to be due to the correlation between certain linear combinations of returns. The near singularity of the sample variance/covariance matrices encountered, caused estimation problems and an extensive study into the possibility of using enhanced (ridge - regression type) estimators was carried out, unfortunately with little success.

In an attempt to obtain average linear combination estimates over the entire time period studied (five years) the idea of estimating grand average correlation matrices was investigated. The resulting grand average linear combination estimates were remarkably similar across all four series of returns and led to the discovery that most of the observed multivariate serial correlation could be explained by certain complex multivariate spread portfolios.

Three multivariate trading rules of various degrees of sophistication were devised to exploit the observed temporal behaviour and when applied to all four series produced positive and statistically significant returns. The introduction of non zero transaction costs reduced returns but still produced positive profits in the cocoa and coffee series. There are a number of practical limitations to the implementation of these trading rules, not the least being the purchase of complex six dimensional spread portfolios. Also the Author is aware that the large returns in the cocoa and coffee series may be peculiar to those markets in the time period examined.

Finally, models of multivariate processes that could explain both the observed departure from multivariate normality and the significant multivariate serial correlation coefficients are presented. In one model we show analytically that under certain simplifying assumptions the multivariate spread portfolios have a theoretical serial correlation consistent with that observed.

REFERENCES

Alexander, S.S. , 1961, "Price Movements in Speculative Marktes: Trends or Random Walks." Industrial Management Review (May), 2, pp. 7-26.

Anderson, O.D. , 1975, "Time Series Analysis and Forecasting: the Box - Jenkins approach." Butterworth Press, London.

Bear, R.M. , 1972, "Margin Levels and the Behaviour of Futures Prices." Journal of Financial and Quantitative Analysis (September), pp. 1907-1930.

Blattberg, R. and N. Gonedes, 1974, "A Comparison of the Stable and Student Distributions as Statistical Models for Stock Prices." Journal of Business, 47, pp. 244-288.

Bachelier, L, 1900, "Theory of Speculation," reprinted in : P.Cootner, ed., 1964, The Random Character of Stock-market Prices (MIT Press, Boston).

Box, G.E.P., 1949, "A General Distribution Theory for a class of Likelihood Criteria." Biometrika, Vol. 36, pp. 317-346.

Campbell, N.A., 1980, "Robust Procedures in Multivariate Analysis I. Robust Covariance Estimation." Applied Statistics, Vol. 29, No.3, pp 231-237.

Cargill, T.F. and Rausser, G.C., 1975, "Temporal Price Behaviour in Commodity Futures Markets." Journal of Finance, Vol. XXX, No.4, pp. 1043-1053.

Chitturi, R.V., 1976, "Distribution of Multivariate White Noise Autocorrelations." Journal of the American Statistical Association, Vol. 71, pp. 223-226.

Chu, K., 1978, "Short-run Forecasting of Commodity Prices: An Application of Autoregressive Moving Average Models." Internaltional Monetary Fund Staff Papers, Vol 25, pp. 90-111.

Cox, C.C., 1976, "Futures Trading and Market Information." Journal of Political Economy, Vol. 84, No.6, pp.1215-1237.

Cox, D.R. and Small, N.J.H. 1978, "Testing Multivariate Normality." Biometrika, Vol. 65, No. 2, pp. 263-272.

Donner, A. and Rosner, B., 1980, "On Inferences Concerning a Common Correlation Coefficient." Applied Statistics, Vol. 29, No.1, pp. 69-76.

Dusak, K., 1973, "Futures Trading and Investor Returns: An Investigationjof Commodity Market Risk Premiums." Journal of Political Economy, Vol. 81, No.6, pp. 1387-1406.

Elston, R.C., 1975, "On Correlations Between Correlations." Biometrika, Vol. 62, No.1, pp.133-147.

Fama, E.F., 1970, "The Foundation of Finance." Basic Books Inc., Publishers, New York.

Fama, E.F. and Roll, R., 1968, "Some Properties of Symmetric Stable Distributions." Journal of the American Statistical Association, Vol. 63, (September), pp. 817-836.

Healy, M.J.R., 1968, "Multivariate Normal Plotting." Applied Statistics, Vol. 17, pp. 157-161.

Hoerl, A.E. and Kennard, R.W., 1970, "Ridge Regression: Biased Estimation for Nonorthogonal Problems." Technometrics, Vol. 12, No.1, pp.55-82.

Holthausen, D.M. and Hughes, J.S., 1978, "Commodity Returns and Capital Asset Pricing." Financial Management, (Summer), pp37-44.

Houthakker, H., 1959, "Can Speculators Forecast Prices?" ,Review of Economics and Statistics, Vol. 39, No.2, pp. 143-151.

Kendall, M.G., 1953, "The Analysis of Economic Time-Series, Part I: Prices," Journal of The Royal Statistical Society, Vol. 96, pp. 11-25.

Kendall, M.G. and Stuart, A., 1976, "The Advanced Theory of Statistics Volume 3." Third edition, Charles Griffin and Co. ltd.

Kshirsagar, A.M., 1972, "Multivariate Analysis." Marcell Dekker, New York.

Labys, W.C., 1976, "Naivety and Optimality in Commodity Price Forecasting." Primary Commodity Prices, Rotterdam University Press.

Labys, W.C. and Perrin, Y., 1976, "Multivariate Analysis of Price Aspects of Commodity Stabilization." Weltivirtschaftliches Archiv, Vol. 112, pp.556-564.

Larson, A.B., 1960, "Measurement of a Random Process in Futures Prices." Food research Institute Studies, Vol. 1, No.3, pp. 313-324.

Leuthold, R.M., 1972, "Random Walk and Price Trends: The Live Cattle Futures Market." Journal of Finance, Vol. XXVII, No. 3, pp. 879-888.

Loeb, M., 1979, "The Behaviour of Spot and Future Prices," Unpublished Ph.D. Thesis, London Business School.

Malkovich, J.F. and Afifi, A.A., 1973, "On Tests For Multivariate Normality." Journal of the American Statistical Association, Vol. 68, No. 341, pp. 176-179.

Mandelbrot, B., 1963, "The Variation of Certain Speculative Prices." Journal of Business, Vol. 36, No. 4, pp. 394-419.

Mardia, K.V., 1970, "Measures of Multivariate Skewness and Kurtosis with applications." Biometrka, Vol. 57, No. 3, pp. 519-530.

Mardia, K.V., Kent, J.T. and Bibby, J.M., 1979, "Multivariate Analysis." Academic Press.

Markowitz, H., 1952, "Portfolio Selection." Journal of Finance, (March), pp. 71-91.

Martell, T.F. and Philippatos, G.C., 1974, "Adaptive Information and Dependence on Commodity Markets." Journal of Finance, Vol. 29, No.2, pp.493-498.

Mood, A.M., 1940, "The Distribution Theory of Runs." Annals of Mathematics and Statistics, Vol. 11, pp.367-392.

Narula, S.C. and Desu, M.M., 1981, "Algorithm AS170: Computation of Probability and Non-centrality Parameter of a Non-central Chi-squared Distribution." Applied Statistics, Vol. 30, No.3, pp.349-352.

O' Brien, P.C., 1980, "A Likelihood Test for Multivariate Serial Correlation." Biometrika, Vol. 67, No. 3, pp. 531-537.

Osborne, M.F.M., 1959, "Brownian motion in the Stock Market," Operations Research, Vol.10, pp. 145-173.

Pearson, E.S., D'Agostino, R.B. and Bowen, K.C., 1977, "Tests For Departures From Normality: Comparison Of Powers." Biometrika, Vol. 64, pp.231-246.

Pinches, G.C., 1977, "Discussion." Journal of Finance, Vol. XXIX, No.2, pp.105-120.

Praetz, P.D., 1972, "The Distribution of Share Price Changes." Journal of Business, Vol.45, pp.49-55.

____________, 1975, "Testing the Efficient Markets Theory on the Sydney Wool Futures Exchange." Australian Economic Papers, Vol. 14, pp.240-249.

____________, 1976a, "On the Methodology of testing for independence in futures prices: A Comment." Journal of Finance, Vol.31, (June) pp. 977-979.

____________, 1976b, "Rates of Return on Filter Tests." Journal of Finance, Vol.31, (March) pp.71-75.

____________, 1978, "Distribution of Stock Market Returns: 1958-1973." Australian Journal of Management, Vol.3, No.1, pp. 81-90.

____________, 1979, "A General Test of a Filter Effect." Journal of Finance and Quantitative Analysis, Vol. XIV, No.2, (June), pp.385-394.

Press, S.J., 1967, "A Compund Events Model for Security Prices." Journal of Business, Vol.40, pp.317-335.

Royston, J.P., 1982a, "An Extension of Shapiro and Wilk's W-test for normality to large samples." Applied Statisitcs, Vol.31, pp.115-124.

____________, 1982b, "Algorithm AS181: The W-test for Normality." Applied Statistics, Vol.31, pp.176-180.

____________, 1982c, "Algorithm AS177: Expected Normal Order Statistics (Exact and Approximate)." Applied Statistics, Vol.31, pp.161-165.

Royston, J.P., 1983, "Some Techniques for assessing Multivariate Normality based on the Shapiro - Wilk's W." Applied Statistics, Vol.32, No.2, pp. 121-133.

Sarhan, A.E. and Greenberg, B.G., 1956, "Estimation of Location and Scale Parameters by Order Statistics from Single and Double Censored Samples. Part I." Annuls of Mathematical Statistics, Vol. 27, pp.427-451.

Scheffe, H., 1970, Practical Solutions of the Behrens-Fisher Problem." Journal of American Statistical Association, Vol.65, pp.1501.

Schrock, N.W., 1971, "The Theory of Asset Choice: Simultaneous Holding of Short and Long Positions in the futures markets." Journal of Political Economy, Vol. 79, No. 2, pp. 270-293.

Shapiro, S.S. and Wilk, M.B., 1965, "An Analysis of Variance Test for Normality (Complete Samples)." Biometrika, Vol.57, pp.591-611.

Snedecor, G.W. and Cochran, W.G., 1967, "Statistical Methods." Sixth edition. University Press.

Smidt, S., 1965, "A Test of the Serial Independence of Price Changes in Soybean Futures." Food Research Institute Studies, Supplement 5, pp.117-136.

Stevenson, R.A., and Bear, R.M., 1976, "Commodity Futures: Trends of Random Walks." Journal of Finance, Vol.XXV, (March), pp.233-255.

Taylor, S.J., 1980, "Conjectured Models for Trends in Financial Prices, Tests and Forecasts." Journal of Royal Statistical Society, Series A, Vol.143, pp.338-362.

__________, 1983, "Trading Rules for Investors in Apparantly Inefficient Markets." Published in "Futures Marktes. Modelling, Managing and Monitoring Futures Trading." Edited by M.E. Streit, Basil Blackwell Publishers Ltd., pp. 165-198.

Taylor, S.J. and Kingsman, B.G., 1978, "Non-stationarity in Sugar Prices." Journal of the Operational Research Society, Vol.29, pp.971-980.

____________________, 1979, "An Analysis of the Variance and Distribution of Commodity Price Changes." Australian Journal of Management, Vol.4, No.2, pp.135-149.

Tschoegl, A.E., 1978, "Weak-Form Efficiency In the Gold Market." Working Parer, Massachusetss Institute of Technology.

Tobin, J., 1958, "Liquidity Preference as Behaviour Towards Risk." Review of Economic Studies, (Feb), 25, pp.65-85.

Vinod, H.D., 1976, "Canonical Ridge and Econometrics of Joint Production." Journal of Econometrics, Vol.4, pp.147-166.

APPENDIX A

THE DATA SET

In this appendix we describe the data set, how the numerous errors were identified and corrected and the final rearrangement into a chronological file that was much easier to access in subsequent analysis.

A.1. Description of the data set

The International Commodity Clearing House (ICCH) made available a magnetic tape containing daily information on all the futures contracts of cocoa, coffee, sugar from the 1 st. January, 1974 to the 31 st December, 1979 inclusive and on the rubber series from the 13 th March 1975 to the 31 st December, 1979 inclusive. The data set contained information on the following prices: open, high, low,closing buyer and closing seller for each contract. The set contained 5,120,000 prices in all

Virtually all previous statistical empirical research on the commodity futures markets use daily, weekly or monthly closing prices and so it was decided to use the closing seller price as the one price representing that contract on each day. Accordingly a smaller subfile (a random access file) was constructed containing only the daily closing seller prices. This new file contained 43,022 records. Each record contained information on the price, the date, the commodity code and the particular contract.

Each record had the following format:

1, 06047, MY8, 195500,

where,

1	=	commodity code (1 = cocoa)
06047	=	date (6th day of 4th month 1977)
MY8	=	delivery month of contract (MAY 1978)
195500	=	closing seller price (£1955 per tonne)

So, on the 6 th April, 1977 a contract (10 tonnes) of cocoa for delivery at the end of May, 1977 could have been purchased (at the close of day) at a price of £1955 per tonne. All the cocoa prices (from 13 th March,1974 to 31 st December,1979) were at the begining of the file followed by all the coffee prices followed by all the sugar prices and finally all the rubber prices. A description of the contracts on the ICCH tape is given in Table A.1.

The total number of contracts is = 37 + 42 + 35 + 27 = 141 , the length of each contract varies from 14 months for the coffee series to 24 months for the rubber series. The original tape also contained information on the 'current month' contracts, e.g. January, 1977 Sugar. These current months are simply duplicates of the nearest delivery months (in the above example it would be March, 1977 Sugar) and so were not copied into the new abreviated data set.

Table A.1

Complete contracts on ICCH data tape

Cocoa contracts

	Delivery Month				
Year	March	May	July	September	December
1974	/	/	/	/	/
1975	/	/	/	/	/
1976	/	/	/	/	/
1977	/	/	/	/	/
1978	/	/	/	/	/
1979	/	/	/	/	/
1980	/	/	/	/	/
1981	/	/	x	x	x

Total number of contracts = 37

Coffee contracts

	Delivery Month					
Year	March	May	July	September	November	December
1974	x	/	/	/	/	/
1975	/	/	/	/	/	/
1976	/	/	/	/	/	/
1977	/	/	/	/	/	/
1978	/	/	/	/	/	/
1979	/	/	/	/	/	/
1980	/	/	/	/	/	/
1981	/	x	x	x	x	x

Total number of contracts = 42

Sugar contracts

	Delivery Month				
Year	March	May	August	October	December
1974	/	/	/	/	/
1975	/	/	/	/	/
1976	/	/	/	/	/
1977	/	/	/	/	/
1978	/	/	/	/	/
1979	/	/	/	/	/
1980	/	/	/	/	/

Total number of contracts = 35

Table A.1 continued

Rubber contracts

	Delivery Month			
Year	February	May	August	November
1974	x	x	x	x
1975	x	/	/	/
1976	/	/	/	/
1977	/	/	/	/
1978	/	/	/	/
1979	/	/	/	/
1980	/	/	/	/
1981	/	/	/	/

Total number of contracts = 27

A.2 Identification and removal of errors

With a file containing 43,022 prices it seemed likely that some errors would be present. Before any serious statistical work was to be carried out, the data had to be examined in detail for anomalies.

A.2.1. Missing trading days

A routine that generated trading day codes was constructed. A second routine checked the entire file for missing observations. There were no missing trading days.

A.2.2. Missing Significant Digits

Listing all the prices of a few randomly chosen contracts showed that some large obvious errors were present in the data. As an example consider the recorded prices of the March, 1978 cocoa contract for the period 16 th Decmber, 1976 to 23 rd December, 1976 inclusive.

date	recorded price
16126	155000
17126	158000
20126	60500
21126	61000
22126	162000
23126	162000

Obviously on the 20 th and 21 st December, the ICCH operator did not record the most significant digit. Clearly the prices on these dates were 160500 and 161000 respectively. It was decided that in order to disclose all of the errors of this type a time series plot of each contract had to be made. In Figs. A.1 and A.2 are two typical examples of initial plots of the unedited data set.

The large spikes and abrupt changes in the level of the series are clearly due to errors of the type mentioned above. The sugar prices were particularly bad in this respect,especially in the earlier half of the time period considered. Consider for example the October 1975 contract on the days 20 th November 1974 to 3 rd December 1974.

Fig A.1 Time series plot of prices of unedited sugar August 1974 contract

Fig A.2 Time series plot of prices of unedited sugar March 1975 contract

date	contract	recorded price	real price
20114	OT5	7050	57050
21114	OT5	3000	53000
22114	OT5	0	50000
25114	OT5	8000	48000
26114	OT5	6000	46000
27114	OT5	5000	45000
28114	OT5	9525	39525
29114	OT5	2000	42000
02124	OT5	0	40000
03124	OT5	8000	38000

A.2.3. Incorrect significant digits

Occasionally an incorrect significant digit was recorded, eg on 2 nd January 1977 the September 1978 cocoa contract was recorded as 296000 (£2960 per tonne). It should have been 196000 (£1960 per tonne). Whenever an error was discovered the real closing seller price for that day was obtained from past copies of The Financial Times and/or records on Microfiche at the ICCH. As the errors in each contract were corrected new plots were made and new errors were disclosed. The process was repeated until all 141 plots were free from obvious anomolous spikes and changes in level.

In all there were 5725 such mistakes corrected, most of them in the sugar series. In Table A.2 a breakdown of the number of errors is given.

Table A.2

	No. of prices	No. of errors	% Errors
Cocoa	10841	924	8.5%
Coffee	10351	522	5.0%
Sugar	11089	4268	38.5%
Rubber	10738	11	0.1%
Total	43022	5725	13.0%

The number of errors, their special nature, and the size of the data file meant that the use of a text editor such as SOS on the DEC10 would have been very time consuming and the computer workspace required prohibitive. As a result a suite of editing programmes specially designed to search and correct the file was created. The programmes proved extremely useful, being both fast and requiring very little workspace.

A.3 Rearrangement of data into final form

Before the final sample of contracts (outlined in chapter 3) was decided on, many different combinations of multiples of contracts were examined. To make this task easier the data file was rearranged so that each record contained the prices of all contracts of all 4 futures markets. In the new rearranged data set there is only one record per day with a total of 1218 records in all. A routine was then constructed that could read the prices of any number of specified contracts in a given record and any number of consecutive records. This final data reaarangement made the subsequent univariate and multivariate analysis particularly simple. A pictorial representation of the entire data set is given in Fig. A.3.

A.4 Computer routines

All computations and data rearrangements were carried out on The City of London Polytechnic's DEC10 and DEC20 computers. All programmes were written in FORTRAN 77. In Chapters 4 through to 9 we outline work involving computations requiring the inversion of matrices and the solution of eigen value problems. In Chapter 9 the minimization of a 63 variable function with non linear constraints in the variables was involved.

These computations were carried out using the NAG library routines. Routine F01ABF was used to invert positive definite matrices, F01AKF,F01APF and F02AQF were used to find the eigenvalues and eigenvectors of general matrices . E04VBF was used in an attemp to solve the minimization problem. G05DDF was used to generate standard normal random variables in simulation routines. All computations were carried out at the single precision level.

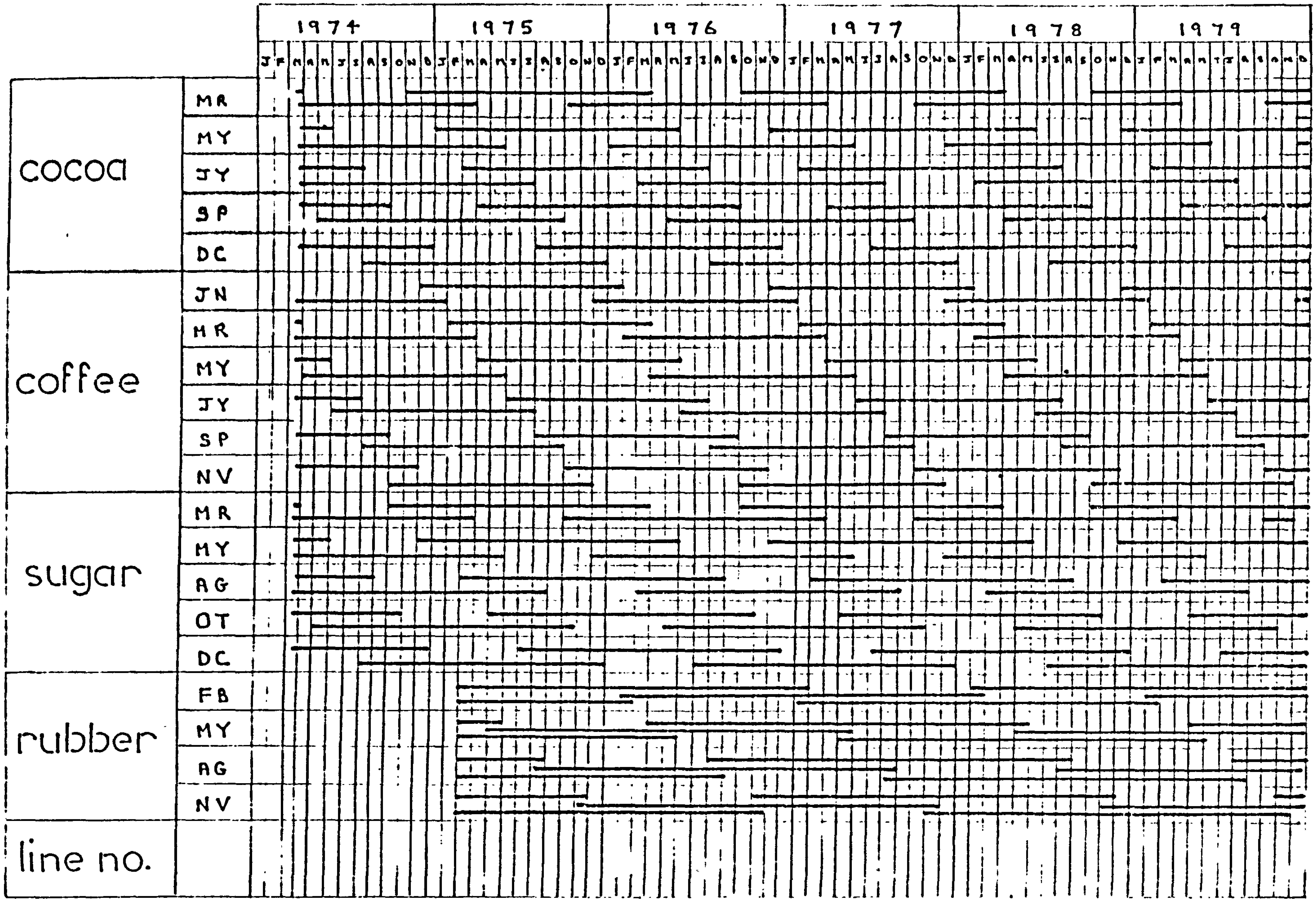

Fig A.3 Pictorial represention of entire ICCH data set

APPENDIX B

Sections of contracts studied in Chapters 3 and 4

Period	No. of days	Cocoa contract	Coffee contract	Sugar contract	Rubber contract	Dates from - to
1	102	Dec. 75	Nov. 75	Mar. 76	May 76	7/03/75 - 31/07/75
2	107	May 76	May 76	May 76	Aug. 76	1/08/75 - 31/12/75
3	85	Sep. 76	Sep. 76	Oct. 76	Feb. 77	2/01/76 - 30/04/76
4	86	Dec. 76	Jan. 77	Mar. 77	May 77	3/05/76 - 31/08/76
5	87	Jul. 77	May 77	May 77	Aug. 77	1/09/76 - 31/12/76
6	83	Sep. 77	Sep. 77	Oct. 77	Feb. 78	4/01/77 - 29/04/77
7	86	Dec. 77	Jan. 78	Mar. 78	May 78	2/05/77 - 31/08/77
8	86	May 78	May 78	May 78	Aug. 78	1/09/77 - 30/12/77
9	83	Sep. 78	Sep. 78	Oct. 78	Feb. 79	3/01/78 - 28/04/78
10	87	Dec. 78	Jan. 79	Mar. 79	May 79	2/05/78 - 31/08/78
11	85	May 79	May 79	May 79	Aug. 79	1/09/78 - 29/12/78
12	84	Sep. 79	Sep. 79	Oct. 79	Feb. 80	2/01/79 - 30/04/79
13	87	Dec. 79	Jan. 80	Mar. 80	May 80	1/05/79 - 31/08/79
14	83	May 80	May 80	May 80	Aug. 80	3/09/79 - 28/12/79

APPENDIX C

Sections of contracts woven together in study of long term serial correlation in section 3.6.1

Series	Contracts	Dates
Cocoa	Mar. 76	7/03/75 - 31/12/75
	Mar. 77	2/01/76 - 31/12/76
	Mar. 78	4/01/77 - 30/12/77
	Mar. 79	3/01/78 - 29/12/78
	Mar. 80	2/01/79 - 28/12/79
Coffee	Mar. 76	7/03/75 - 30/01/76
	Jan. 77	2/02/76 - 30/11/76
	Nov. 77	1/12/76 - 30/09/77
	Sep. 78	3/10/77 - 31/07/78
	Jul. 79	1/08/78 - 31/05/79
	May 80	1/06/79 - 28/12/79
Sugar	Mar. 76	7/03/75 - 31/12/75
	Mar. 76	2/01/76 - 31/12/76
	Mar. 78	4/01/77 - 30/12/77
	Mar. 79	3/01/78 - 29/12/78
	Mar. 80	2/01/79 - 28/12/79
Rubber	Feb. 76	7/03/75 - 28/11/75
	Feb. 77	28/11/75 - 30/11/76
	Feb. 78	30/11/76 - 30/11/77
	Feb. 79	30/11/77 - 30/11/78
	Feb. 80	30/11/78 - 28/12/79

APPENDIX D

Sections of cocoa contracts studied in Chapter 5

Period	No. of days	first contract	second contract	third contract	fourth contract	Dates from - to
1	102	Sep. 75	Dec. 75	Mar. 76	May 76	7/03/75 - 31/07/75
2	107	Mar. 76	May 76	Jul. 76	Sep. 76	1/08/75 - 31/12/75
3	85	May 76	Jul. 76	Sep. 76	Dec. 76	2/01/76 - 30/04/76
4	86	Sep. 76	Dec. 76	Mar. 77	May 77	3/05/76 - 31/08/76
5	87	Mar. 77	May 77	Jul. 77	Sep. 77	1/09/76 - 31/12/76
6	83	May 77	Jul. 77	Sep. 77	Dec. 77	4/01/77 - 29/04/77
7	86	Sep. 77	Dec. 77	Mar. 78	May 78	2/05/77 - 31/08/77
8	86	Mar. 78	May 78	Jul. 78	Sep. 78	1/09/77 - 30/12/77
9	83	May 78	Jul. 78	Sep. 78	Dec. 78	3/01/78 - 28/04/78
10	87	Sep. 78	Dec. 78	Mar. 79	May 79	2/05/78 - 31/08/78
11	85	Mar. 79	May 79	Jul. 79	Sep. 79	1/09/78 - 29/12/78
12	84	May 79	Jul. 79	Sep. 79	Dec. 79	2/01/79 - 30/04/79
13	87	Sep. 79	Dec. 79	Mar. 80	May 80	1/05/79 - 31/08/79
14	83	Mar. 80	May 80	Jul. 80	Sep. 80	3/09/79 - 28/12/79

APPENDIX D (continued)

Sections of coffee contracts studied in Chapter 5

Period	No. of days	first contract	second contract	third contract	fourth contract	Dates from - to
1	102	Sep. 75	Nov. 75	Jan. 76	Mar. 76	7/03/75 - 31/07/75
2	107	Jan. 76	Mar. 76	May 76	Jul. 76	1/08/75 - 31/12/75
3	85	May 76	Jul. 76	Sep. 76	Nov. 76	2/01/76 - 30/04/76
4	86	Sep. 76	Nov. 76	Jan. 78	Mar. 78	3/05/76 - 31/08/76
5	87	Jan. 77	Mar. 77	May 77	Jul. 77	1/09/76 - 31/12/76
6	83	May 77	Jul. 77	Sep. 77	Nov. 77	4/01/77 - 29/04/77
7	86	Sep. 77	Nov. 77	Jan. 78	Mar. 78	2/05/77 - 31/08/77
8	86	Jan. 78	Mar. 78	May 78	Jul. 78	1/09/77 - 30/12/77
9	83	May 78	Jul. 78	Sep. 78	Nov. 78	3/01/78 - 28/04/78
10	87	Sep. 78	Nov. 78	Jan. 79	Mar. 79	2/05/78 - 31/08/78
11	85	Jan. 79	Mar. 79	May 79	Jul. 79	1/09/78 - 29/12/78
12	84	May 79	Jul. 79	Sep. 79	Nov. 79	2/01/79 - 30/04/79
13	87	Sep. 79	Nov. 79	Jan. 80	Mar. 80	1/05/79 - 31/08/79
14	83	Jan. 80	Mar. 80	May 80	Jul. 80	3/09/79 - 28/12/79

APPENDIX D (continued)

Sections of sugar contracts studied in Chapter 5

Period	No. of days	first contract	second contract	third contract	fourth contract	Dates from - to
1	102	Aug. 75	Oct. 75	Dec. 75	Mar. 76	7/03/75 - 31/07/75
2	107	Mar. 76	May 76	Aug. 76	Oct. 76	1/08/75 - 31/12/75
3	85	May 76	Aug. 76	Oct. 76	Dec. 76	2/01/76 - 30/04/76
4	86	Oct. 76	Dec. 76	Mar. 77	May 77	3/05/76 - 31/08/76
5	87	Mar. 77	May 77	Aug. 77	Oct. 77	1/09/76 - 31/12/76
6	83	May 77	Aug. 77	Oct. 77	Dec. 77	4/01/77 - 29/04/77
7	86	Oct. 77	Dec. 77	Mar. 78	May 78	2/05/77 - 31/08/77
8	86	Mar. 78	May 78	Aug. 78	Oct. 78	1/09/77 - 30/12/77
9	83	May 78	Aug. 78	Oct. 78	Dec. 78	3/01/78 - 28/04/78
10	87	Oct. 78	Dec. 78	Mar. 79	May 79	2/05/78 - 31/08/78
11	85	Mar. 79	May 79	Aug. 79	Oct. 79	1/09/78 - 29/12/78
12	84	May 79	Aug. 79	Oct. 79	Dec. 79	2/01/79 - 30/04/79
13	87	Oct. 79	Dec. 79	Mar. 80	May 80	1/05/79 - 31/08/79
14	83	Mar. 80	May 80	Aug. 80	Oct. 80	3/09/79 - 28/12/79

APPENDIX D (continued)

Sections of rubber contracts studied in Chapter 5

Period	No. of days	first contract	second contract	third contract	fourth contract	Dates from - to
1	102	Nov. 75	Feb. 76	May 76	Aug. 76	7/03/75 - 31/07/75
2	107	Feb. 76	May 76	Aug. 76	Nov. 76	1/08/75 - 31/12/75
3	85	May 76	Aug. 76	Nov. 76	Feb. 77	2/01/76 - 30/04/76
4	86	Nov. 76	Feb. 77	May 77	Aug. 77	3/05/76 - 31/08/76
5	87	Feb. 77	May 77	Aug. 77	Nov. 77	1/09/76 - 31/12/76
6	83	May 77	Aug. 77	Nov. 77	Feb. 78	4/01/77 - 29/04/77
7	86	Nov. 77	Feb. 78	May 78	Aug 78	2/05/77 - 31/08/77
8	86	Feb. 78	May 78	Aug. 78	Nov. 78	1/09/77 - 30/12/77
9	83	May 78	Aug. 78	Nov. 78	Feb. 79	3/01/78 - 28/04/78
10	87	Nov. 78	Feb. 79	May 79	Aug. 79	2/05/78 - 31/08/78
11	85	Feb. 79	May 79	Aug. 79	Nov. 79	1/09/78 - 29/12/78
12	84	May 79	Aug. 79	Nov. 79	Feb. 80	2/01/79 - 30/04/79
13	87	Nov. 79	Feb. 80	May 80	Aug. 80	1/05/79 - 31/08/79
14	83	Feb. 80	May 80	Aug. 80	Nov. 80	3/09/79 - 28/12/79