A G R I C U L T U R A L

E C O N O M I C S

R E S E A R C H   U N I T

Lincoln College

QUANTITATIVE TECHNIQUES FOR FORECASTING

A REVIEW WITH APPLICATIONS TO NEW ZEALAND WOOL PRICES FOR 1974-5

by

Joan Rodgers

Research Report No. **69**

1974

THE AGRICULTURAL ECONOMICS RESEARCH UNIT

THE UNIT was established in 1962 at Lincoln College, University of Canterbury, with an annual grant from the Department of Scientific and Industrial Research. This general grant has been supplemented by grants from commercial and other oganisations for specific research projects within New Zealand and overseas.

The Unit has on hand a long-term programme of research in the fields of agricultural production, marketing and policy, resource economics, and the economics of location and transportation. The results of these research studies are published as Research Reports as projects are completed. In addition, technical papers, discussion papers and reprints of papers published or delivered elsewhere are available on request. For a list of previous publications see inside back cover.

RESEARCH STAFF : 1974

Director

W. O. McCarthy, M.Agr.Sc., Ph.D. (Iowa)

Deputy Director

P. D. Chudleigh, B.Sc., Ph.D. (N.S.W.)

Senior Research Economist

G. W. Kitson, M.Hort.Sc.

Research Economists

T. I. Ambler, B.A. (Cant.), B.C.A. (Victoria), A.N.Z.I.M.

R. J. Gillespie, B.Agr.Sc. (Trinity Coll.)

J. D. Gough, B.Sc. (Cant.), M.Com. (Hons.) (Cant.)

University Lecturing Staff

A. T. G. McArthur, B.Sc. (Agr.) (Lond.), M.Agr.Sc.

J. L. Rodgers, B.Ag.Ec. (U.N.E.), Dip.Inf.Proc. (Q'ld)

R. J. Brodie, B.Sc. (Cant.), M.A. (Otago)

P R E F A C E

Since its inception the Research Unit has been involved to a significant extent in econometric model building, sometimes with the objective of estimating prices or quantities of agricultural commodities.

Thus for example over the last two years we have contracted to estimate the price of New Zealand wool for the ensuing season. Initially this led us to review available techniques and in particular to explore further the Box-Jenkins and leading indicator approaches.

Joan Rodgers and Les Woods have been associated with much of this work. We now consider it appropriate to draw together our literature reviews and working papers, add some empirical material and publish for a wider audience.

Owen McCarthy
Director

CONTENTS

1. INTRODUCTION

The econometrician faces two types of forecasting problem, which may be termed conditional forecasting and unconditional forecasting.

In conditional forecasting the econometrician must answer "what .... if?" questions. For example, "what will be the level of demand for commodity x, if its price is $y per kilogram and if consumer incomes average $z per capita?". An econometric model, relating quantity demanded to price and income, can be used to answer questions of this type.

Unconditional forecasting requires answering "what ....?" questions. For example, "what will be next year's demand for commodity x?". An econometric model is not always suitable for solving this sort of problem, since, before the model can be used to predict demand, the researcher may first have to predict the price of commodity x and average consumer incomes one year hence. In many cases this is as difficult as predicting demand itself. In other words, the problem is transformed from one of predicting values of the endogenous variable (quantity demanded) to one of predicting values of the exogenous variables (price and income).[1]

This paper discusses a number of quantitative techniques which can be used for unconditional forecasting. The reader is reminded that routine "data cleaning"[2] procedures should be undertaken before applying these methods to any given set of data.

---

1 Steckler (16) p.431, suggests that this may be one of reasons why large scale econometric models have not performed well to date.

2 Data cleaning is discussed in a number of texts on economic statistics e.g., Freund & Williams (7) p.391-393, Neter & Wasserman (14), p.610-611.

A number of time series models are presented in section 3. These range from naive techniques, justified largely in terms of their intuitive appeal, to more sophisticated methodologies based on statistical theory. Their common feature is that forecasts on a given variable are prepared without regard to the behaviour of other variables in the system. Causal models, on the other hand, take into account the relationships between variables in the system. Section 4, discusses a number of causal models which can be used for unconditional forecasting.

Section 5 compares time series and causal models. The more sophisticated techniques, especially those discussed in sections 3.6, 4.1, 4.2 and 4.3, are not explained exhaustively here. The reader should consult the references listed for this purpose.

However sufficient detail is provided to enable the reader to work through sections 8, 9 and 10. These respectively use exponentially weighted moving averages, Box-Jenkins, and a standard econometric model to forecast prices of New Zealand wool on a clean c.i.f., U.K, dry combing basis for the 1974/5 season.

2. TERMINOLOGY

The following terms are used frequently throughout this paper. Other terminology which is specific to a particular method will be defined at the appropriate time.

2.1 Time Series

A time series is a chronologically ordered set of observations on some variable. Observations are collected at discrete and equi-distant intervals of time.

2.2 Forecast Lead Time

The forecast lead time (or forecast horizon) is the number of periods from the time when a forecast is prepared to the time to which the forecast refers. For example, given a series of quarterly data, a forecast prepared sometime during the first quarter of 1974 and referring to the fourth quarter of 1974 has a lead time of three quarters.

2.3 The Secular Trend

The (secular) trend indicates the direction of a time series, in terms of a general upward or downward movement over a long period of time. These movements are usually related to factors such as population growth, technological change, economic growth etc,.

2.4 Cyclical Movements

The business cycle is a historical phenomenon associated with many time series in unregulated economies. It consists of up and down movements which have a periodicity greater than one year around a secular trend. Sometimes trend and cyclical components are not separated but are regarded as a single series.

2.5 Seasonal Variations

These are periodic variations in a time series which tend to repeat themselves every twelve months. The major factors responsible for seasonality are weather and social customs. For example, agricultural production displays marked seasonal fluctuations as do sales of air conditioners and skiis. The influence of social custom is seen in the increase in retail sales which typically occurs prior to Christmas.

2.6 Random or Irregular Fluctuations

These are erratic disturbances (sometimes called noise) which follow no regular pattern and therefore cannot be predicted. They are regarded as the residual once systematic variations have been accounted for.

2.7 Mean Square Error

The mean square error is sometimes used to measure the goodness of fit of a given model to a set of sample data and is defined as:

$$\text{M.S.E.} \quad = \quad \frac{\sum_{t=1}^{n} (z_t - \hat{z}_t)^2}{n}$$

where:

z_t is the actual observation in period t,
$\hat{z}_t$ is the value predicted by the model for period t,
n is the number of observations in the sample.

2.8 The Short, Medium and Long Term

For the purpose of this paper, short term forecasts are those with a lead time of from one to three months, medium term forecasts are from three months to two years and long term forecasts have a forecast horizon greater than two years.

3. TIME SERIES MODELS

Time series analysis regards movements of a given variable z_t to be solely a function of time. Any relationships which exist between z_t and other variables in the system are completely ignored. The aim is to identify recurring patterns in the series, using historical data, and to extrapolate these patterns into the future. It is assumed, therefore, that stable patterns exist and will continue into the future. Since patterns are less likely to be stable in the long run, many of these techniques are best suited to short to medium term forecasting with a lead time of one year or less.

The successful application of any forecasting technique depends upon the availability of accurate data, but for the applied econometrician the acquisition of reliable data can be one of the more irksome tasks. Time series models therefore have an advantage over causal models in that they require only the past history of the variable to be forecast as input. However they differ among themselves in the number of observations necessary for effective application.

3.1 Simple Moving Averages

Simple moving averages (17) can be used to forecast a time series whose basic pattern is horizontal; that is, observations vary only slightly from period to period. It is possible that over a long period of time the mean of the series will gradually change but over shorter periods it is assumed to be approximately constant.

Under these conditions an estimate of the current average level of the series will provide reasonably accurate short term forecasts. A simple approach is to take an average of the most recent observations and to use this as a forecast for the next period.

$$\hat{z}_{t+1} = \frac{1}{n}(z_t + z_{t-1} + \ldots + z_{t-n+1}) \qquad (1)$$

where:

z_t is the actual observation in period t,
$\hat{z}_{t+1}$ is the forecast for period t+1, made at the end of period t, and
n is the number of observations included in the moving average.

When an additional observation becomes available it is included in the moving average and the oldest observation is dropped. To forecast more than one period in advance either

(a) the one-step-ahead forecast is reused on the assumption that the series remains horizontal, or

(b) previous forecasts are substituted in the moving average in place of values of the variable which have not yet been observed.

There are three main criticisms of simple moving averages:

(1) The choice of the number of observations to be included in the moving average is essentially arbitary, although if the researcher has the patience he/she can experiment with a number of values and choose the one yielding the smallest mean square error when forecasting on past data. The more observations included, the greater the smoothing effect and the slower the adjustment to any change in the underlying pattern.

(2) Equal weights are given to all observations used in calculating the average, while prior observations are completely ignored. In many cases a weighting system which places greatest importance on the most recent observations is more logical.

(3) The requirement that the underlying pattern of the data be horizontal severely limits the practical use of simple moving averages. Should there be a definite shift in the level of the series moving averages will be slow to adapt to the change.

Because of these limitations, a simple moving average model is rarely a strong contender in any forecasting situation. However, if short-term forecasts on a large number of items are required at frequent intervals, and if accuracy is not at a premium, then the simplicity of the method and its modest data requirements, (say 5 to 10 observations per variable) may recommend it.

3.2 Double Moving Averages

Double moving averages (17) can be used to project a time series which exhibits a trend as well as random fluctuations. The method is based on the fact that a simple moving average will always lag behind the original series if the data has a trend. Similarly, a moving average of the simple moving average will always lag behind the latter. Furthermore, if plotted on a graph, the double moving average tends to be about the same distance below the simple moving average as the simple moving average is below the original series.

The forecast is prepared by taking the difference between the simple and the double moving average and adding this figure to the simple moving average.[3]

Double moving averages require twice as many observations as simple moving averages. Although double moving averages can handle a trend, the method is subject to the

---

3 Sometimes an adjustment factor is used in an effort to improve forecasts, see Wheelwright and Makridakis, (17), pp 41-43.

same criticisms as simple moving averages; the choice of the number of observations in the averages is essentially arbitary and equal weighting is given to all observations.

3.3 Exponential Smoothing

Exponential smoothing (4), (9), (17), is an advance on simple moving averages in that it uses a weighted average of all past observations as a forecast for the next period. The model is derived from the formula for a weighted average:

$$\frac{\sum_{i=1}^{n} w_i X_i}{\sum_{i=1}^{n} w_i} \qquad (2)$$

where:

X_i is the i-th observation in the sample,
w_i is the weighting given to the i-th observation and
n is the number of observations in the sample.

The weights used in exponential smoothing are the infinite geometric sequence:

$$(1-\alpha),\ \alpha(1-\alpha),\ \alpha^2(1-\alpha),\ \alpha^3(1-\alpha),\ \ldots$$

where α lies between zero and unity. This system of weights is chosen because:

(a) the series of weights converges to unity, and

(b) since $0<\alpha<1$, the weights diminish exponentially thereby attributing greater importance to the more recent observations.

A one-step-ahead forecast, made at the end of period t is given by a weighted average of all past observations:

$$\hat{z}_{t+1} = (1-\alpha)z_t + \alpha(1-\alpha)z_{t-1} + \alpha^2(1-\alpha)z_{t-2} + \ldots \qquad (3)$$

The above form of the model is not suitable for practical purposes since it contains an infinite number of terms. However, the infinite geometric series given by equation (3) converges to give:

$$\hat{z}_{t+1} = (1-\alpha)z_t + \alpha\hat{z}_t$$

or $$\hat{z}_{t+1} = Az_t + (1-A)\hat{z}_t \qquad 0<A<1 \qquad (4)$$

This is the form of simple exponential smoothing used in practice; a one-step-ahead forecast is a weighted average of the most recent observation and the most recent forecast. The same value is used to forecast with lead time greater than one period.

A systematic search procedure is usually used to locate the value of A which minimises the mean square error; A is allowed to assume the values 0.1, 0.2, ... 0.9, the model is used to forecast past values in the series[4] and the mean square error is calculated for each value of A. This will indicate suitable values of A which can be investigated more fully.

By rearranging the model given in equation (4) we obtain:

$$\hat{z}_{t+1} = \hat{z}_t + A(z_t - \hat{z}_t)$$

---

4 $\hat{z}_1$ cannot be calculated in the usual way, so it is set to z_1

which illustrates that a forecast for period t+1 can also be interpreted as the forecast for period t plus a proportion of the error which occurred in forecasting period t. When A is small the forecast for period t+1 approximates the forecast for period t and so a considerable amount of smoothing takes place. On the other hand, if the time series displays a considerable amount of variation then the value of A should be large, so that forecasts will be sensitive to previous errors.

Simple exponential smoothing, like simple moving averages, is only suitable for forecasting a series whose basic pattern is horizontal and when this is so, it usually performs well in a short term forecasting situation. An extension to the model which enables it to be applied to series which exhibit a trend and/or seasonality will be discussed in section 3.4.

Its small data storage requirements make exponential smoothing particularly attractive in a situation when a large number of forecasts are required at frequent intervals. About thirty observations are required to estimate the smoothing constant A, but once the model is set up only two items need to be stored from one period to the next; the latest observation z_t and the most recent forecast $\hat{z}_t$.

3.4 Exponential Smoothing with Trend and Seasonal Components

The movements of many economic variables through time follow trends and/or are marked by seasonal patterns. The extension of the simple exponential smoothing model to incorporate both a trend and seasonal factors was developed by Holt (8) and Winters(18). We shall first consider a model which allows for seasonality but no trend, then extend the model to include both these components.

The Seasonal Model:

Seasonality can be handled using either an additive or a multiplicative seasonal factor. The additive factor is appropriate when the amplitude of seasonal variation is independent of the level of the series, whereas a multiplicative factor should be used when the amplitude of seasonal variation is proportional to the level of the series. Since the latter case is more common, discussion will be limited to the multiplicative case.

The inclusion of a multiplicative seasonal component, F, gives:

$$\tilde{z}_t = Az_t/_{F_{t-L}} + (1-A)\tilde{z}_{t-1} \qquad 0<A<1 \qquad (5)$$

where:

(a) $\tilde{z}_t$ is an estimate of the expected value of the deseasonalised variable in period t, and

(b) $$F_t = Bz_t/\tilde{z}_t + (1-B)F_{t-L} \qquad 0<B<1 \qquad (6)$$

is an estimate of the seasonal component in period t.

(c) L is the periodicity; for example L=12 for monthly data and L=4 for quarterly data.

Forecasts up to one year in advance are prepared by readjusting the deseasonalised value of the variable for the most recent period, t:

$$\hat{z}_{t+1} = \tilde{z}_t F_{t-L+1}$$

$$\hat{z}_{t+2} = \tilde{z}_t F_{t-L+2} \qquad (7)$$

$$\vdots \qquad \vdots$$

$$\hat{z}_{t+L} = \tilde{z}_t F_t$$

To forecast beyond a lead time of L periods (i,e., one year) the same seasonals are reused.[5] This is only recommended if the seasonal pattern is stable.

Successive substitution of $\tilde{z}_{t-1}$, $\tilde{z}_{t-2}$ etc, in equation (5) reveals that the expected value of the deseasonalised variable in period t is a weighted average of seasonally adjusted values over all past periods. The seasonality is removed by dividing each observation by the seasonal factor computed for the same month or quarter of the previous year. $\tilde{z}_t$ is then used in calculating a new seasonal factor for the month or quarter corresponding to period t, which is a weighted average of the current estimate of seasonality, $z_{t/\tilde{z}_t}$ and the previous estimate, F_{t-L}.

<u>The Complete Exponential Model</u>

A time series which displays a trend over a number of years as well as seasonal variation can be represented by a model which includes an additive trend factor:

$$\tilde{z}_t = Az_{t/F_{t-L}} + (1-A)\ (\tilde{z}_{t-1} + R_{t-1}) \quad 0<A<1 \qquad (8)$$

---

[5] This results in a repeat of forecasts every L periods.

where:

$$\text{(a)} \quad F_t = Bz_{t/\tilde{z}_t} + (1-B)F_{t-L} \qquad 0<B<1 \qquad (9)$$

$$\text{and (b)} \quad R_t = C(\tilde{z}_t - \tilde{z}_{t-1}) + (1-C)R_{t-1} \quad 0<C<1 \qquad (10)$$

is an estimate of the trend component for period t. Since $\tilde{z}_t$ is an estimate of the expected deseasonalised value of z in period t, $(\tilde{z}_t - \tilde{z}_{t-1})$ is the latest estimate of the trend. The new trend factor is a weighted average of this latest available estimate and the previous estimate R_{t-1}.

Forecasts up to one year ahead are given by:

$$\begin{aligned} \hat{z}_{t+1} &= (\tilde{z}_t + 1.R_t)F_{t-L+1} \\ \hat{z}_{t+2} &= (\tilde{z}_t + 2.R_t)F_{t-L+2} \qquad (11) \\ \cdot \quad & \quad \cdot \\ \cdot \quad & \quad \cdot \\ \hat{z}_{t+L} &= (\tilde{z}_t + L.R_t)F_t \end{aligned}$$

To forecast with a lead time longer than one year, additional multiples of the trend are added to the deseasonalised value and the same seasonals are reused.

The forecasting procedure for the general exponentially weighted moving average model with trend and seasonals can be summarised as follows:

(i) At the end of period t, $\tilde{z}_t$ is computed from equation (8) using the latest observation in the series z_t along with $\tilde{z}_{t-1}$, R_{t-1} and the appropriate seasonal F_{t-L} found during the previous cycle.

(ii) The seasonal factor is updated by equation (9) to give F_t which replaces F_{t-L}.

(iii) Equation (10) is used to update the trend component, R_t, which replaces R_{t-1}.

(iv) Forecasts are prepared using equation set (11).

An examination of the above steps reveals that there is a need to determine initial values for $\tilde{z}$, the F's and R, as well as the weights A, B and C. The usual procedure is to divide the time series into two parts, the first of which is used to calculate initial values of $\tilde{z}$, R and the F's while the second section is used to derive the optimal set of weights A, B and C.

Let the first section of the series consist of observations $z_1, z_2, \ldots z_n$ and the second section contain $z_{n+1}, z_{n+2}, \ldots z_t$.

Initial $\tilde{z}$: An average of monthly or quarterly data for year 1 is used to initialise $\tilde{z}$, i.e.,

$$\text{initial } \tilde{z} = \bar{z}_{yr\ 1} \qquad (12)$$

Initial Trend: The initial value of R is set at the average trend measured over the first section of the series, i.e.,

$$R_1 = \frac{\bar{z}_{yr\ n/L} - \bar{z}_{yr\ 1}}{n - L} \qquad (13)$$

Initial Seasonals: Preliminary seasonal factors are computed for each period from 1 to n by dividing the observed value for the period by the average value for the corresponding year and adjusting for the trend, i.e.,

$$F_k = \frac{z_k}{\bar{z}_{yr\ i} - \frac{(L+1}{2} - j)R_1} \qquad (14)$$

where:

(a) $k = 1, 2, \ldots n$

(b) $\bar{z}_{yr\ i}$ is the average for the year corresponding to F_k, i.e., $i = 1, 2, \ldots n/L$

(c) j is the position of the period within the year, e.g., with monthly data January (j=1), February (j=2), etc,.

These preliminary seasonal factors are then averaged by month or quarter and normalised so that they add to L.

$$F_j = Av(F_j) \frac{L}{\sum_{j=1}^{L} Av(F_j)} \qquad j = 1, 2, \ldots L \qquad (15)$$

Once $\tilde{z}_1$, R_1 and initial seasonals $F_1, F_2, \ldots F_L$ have been determined, equations (8), (9) and (10) are applied to the first part of the series. After n cycles, starting values for $\tilde{z}$, R and the F's are available.

Determining optimal values for A, B and C: Taking the second section of the series, an iterative search procedure is used to find values of A, B and C which minimise the mean square error when forecasting with a given lead time. For example, when the lead time is 1 period:

$$M.S.E._1 = \sum_{i=n}^{t-1} \frac{(z_{i+1} - \hat{z}_{i+1})^2}{t - n} \qquad (16)$$

when forecasting 2 periods ahead:

$$M.S.E._2 = \sum_{i=n}^{t-2} \frac{(z_{i+2} - \hat{z}_{i+2})^2}{t - n - 1} \tag{17}$$

and so on.

Alternatively, some combination of mean square errors can be minimised, for example:

$$a_1 M.S.E._1 + a_2 M.S.E._2 + \ldots + a_L M.S.E._L \tag{18}$$

The search is more detailed than for a single weight but is carried out in essentially the same manner. A, B and C are successively given the values 0.1, 0.2, ... 0.9 and the combination resulting in the smallest mean square error can be further refined, if necessary.

The full exponentially weighted moving average model is suitable for forecasting a number of series up to one year in advance, and provided seasonal and trend components remain stable over time, longer term forecasts can also be prepared. Usually about 5-6 seasons of data are required to compute a suitable weighting system, but once the model has been estimated only a small amount of information needs to be stored from one cycle to the next.

3.5 Classical Time Series Analysis

The classical treatment of a time series (1), (7), (17) assumes that observations result from the interaction of ~~one~~ or more of three systematic sources of variation: a secular trend (T), a cyclical component (C), and, if data is measured over intervals of less than one year, seasonal fluctuations (S). Sometimes the trend and cyclical movements are regarded as a composite component called the trend-cyclical (TC). In addition to these regular sources of variation there is also an irregular factor (u).

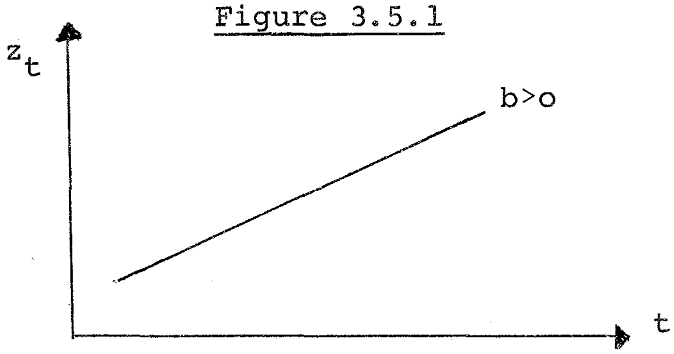
Analysis proceeds by decomposing the series into its separate components, projecting each separately, then synthesising these separate projections into a final forecast.

We shall initially consider the classical approach to a time series which contains only a secular trend plus random fluctuations:

$$z_t = T + u_t \tag{19}$$

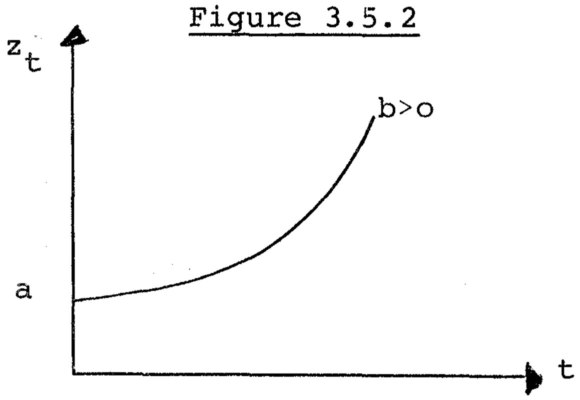
The behaviour of a series of this type can often be described by some mathematical function, fitted to the data using the method of least squares with z_t as the dependent variable and time, t, as the independent variable. (Since the method of least squares appears in almost every elementary econometrics text, it will not be discussed here). Figures 3.5.1., 3.5.2., 3.5.3 and 3.5.4 depict various mathematical functions which are useful in representing secular trends of various kinds and which can be fitted using least squares.

Figure 3.5.1



Linear function $z_t = a + b.t$

Figure 3.5.2



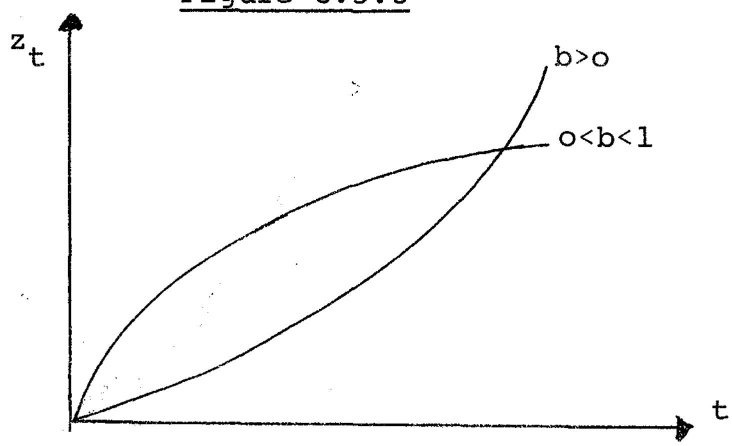
Exponential function $z_t = a.b^t$

This function must be transformed into linear form by taking logarithms of both sides to give:

$$\log(z_t) = \log(a) + \log(b).t \qquad (20)$$

from which the parameters log(a) and log(b) can be estimated using the method of least squares. Estimates of a and b are subsequently obtained by taking antilogs.

Figure 3.5.3



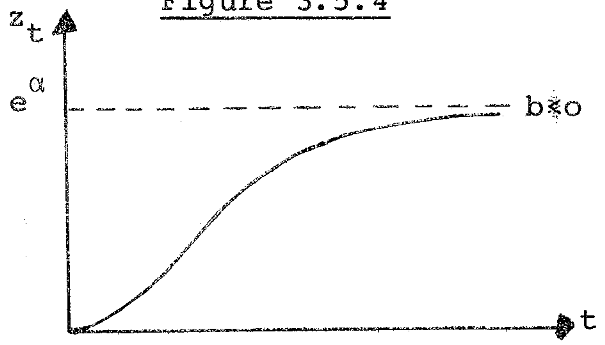
Power Function $z_t = a.t^b$

A logarithmic transformation is also required on this function giving:

$$\log(z_t) = \log(a) + b.\log(t) \qquad (21)$$

Least squares provide estimates of log(a) and b, so a must be estimated by taking antilogs.

Figure 3.5.4



Logistic curve $z_t = \exp(a + b/t)$

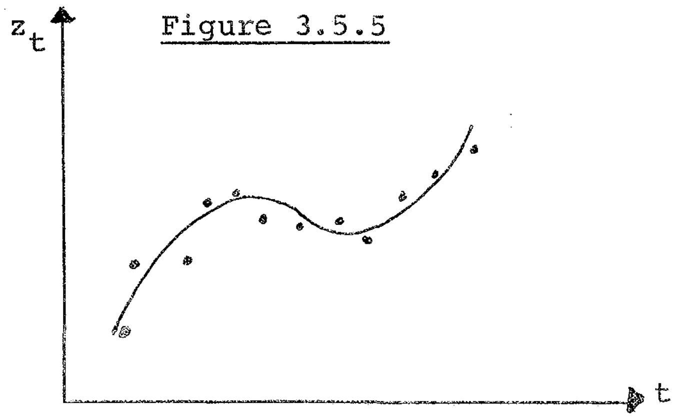
A natural logarithmic transformation is used to transform this equation into a form which linearly relates $\log(z_t)$ to $1/t$:

$$\log(z_t) = a + b.\frac{1}{t} \qquad (22)$$

Should the series display a cyclical component as well as a trend, a polynomial function can be fitted as an estimate of the trend-cyclical. For example, a cubic function could be used to describe the data in figure 3.5.5.

$$z_t = a + b.t + c.t^2 + d.t^3 \qquad (23)$$

Figure 3.5.5



If data are recorded at intervals of less than one year seasonality may be present in addition to trend and cyclical movements. The two most common models used in this case are:

<u>Additive model</u>: $z_t = S + T + C + u_t$ (24)

or

<u>Multiplicative model</u>: $z_t = S \times T \times C + u_t$ (25)

The additive model is appropriate when the amplitudes of seasonal and cyclical movements are independent of the general level of the series, while the multiplicative model is used when their amplitudes are proportional to the general level of the series. As was the case for exponential smoothing, discussion will be confined to the multiplicative model.

First, seasonal variation is removed by taking a centred moving average of twelve months, four quarters etc.,[6] leaving the trend and cyclical components. A measure of seasonal variation is then obtained by dividing the original series by the trend-cyclicals. These seasonal factors are then arranged by month, quarter etc., and a modified mean, calculated for each group, is used as an index of seasonality.

Next, the original series is deseasonalised by dividing each observation by the corresponding seasonal index. The deseasonalised series can be described by a trend-cyclical function as discussed previously without any attempt to separate the cyclicals from the trend component, or alternatively the two can be treated separately. The latter is accomplished by fitting a trend line to the deseasonalised data, calculating the trend factor for each

---

6 A more elaborate way of separating seasonal variation from trend-cyclical movements is by the X-II method developed by the U.S. Bureau of Census (11), (12).

period by substituting the appropriate value of t into the trend equation, then dividing the deseasonalised data by the trend factor, in which case cyclical movements plus noise remain. Cyclical movements can be measured by:

(a) fitting a polynomial to the cyclical irregulars,
(b) taking a weighted average of the noisy cyclicals (usually the weights are based on binomial coefficients e.g., (1,2,1) or (1,4,6, 4,1)) to smooth out any remaining irregular factors, or,
(c) fitting a Fourier series to the data by spectral analysis.

Having isolated the systematic components of the time series each is projected individually, then individual projections are combined to give the final forecast.

If a mathematical function has been used to represent either the trend, the cyclicals or the trend-cyclicals it can easily be extrapolated by substituting the appropriate value of t into the estimated equation. The resulting value(s) are multiplied by the existing seasonal index, to give the final forecast.

Classical time series analysis can be used for short, medium or long term forecasting, although in long term forecasting the seasonal factor is rarely taken into account. Five or six seasons of historical data are usually required to measure the underlying patterns in the data.

The main criticism of the method is the assumption that a series can be neatly decomposed into basic components such as the trend, cyclicals and seasonals. Furthermore the technique does not have a sound statistical basis. Thus, although trend projections obtained using least squares can be evaluated in terms of probability theory, seasonality is measured in a deterministic fashion.

3.6 The Box-Jenkins Method of Forecasting

The Box-Jenkins method of forecasting (2), (9), (13), (17) is an iterative model building procedure which is particularly applicable to time series whose patterns are not easily recognisable. The basic assumption is that an observed time series has been generated by a stochastic process and therefore can be regarded as a sample, drawn from a joint probability distribution.

In addition, the stochastic process which has generated the data is assumed to be stationary; all observations have the same theoretical mean, the same variance and covariances between any pair of observations depend only on the number of periods separating them. Stationarity is a very restrictive assumption, however, many series which are non stationary display stationary behaviour in their first, or higher, order differences. This is known as homogeneous non stationarity. As a result the models which have been developed for stationary series, can be applied equally well to certain non stationary series provided the appropriate degree of differencing is carried out beforehand.

Application of the Box-Jenkins methodology can be broken down into five basic steps:

(a) Postulate a general class of model capable of describing the behaviour of a given series.

(b) Identify a particular model from this general class which can be tentatively entertained.

(c) Estimate the parameters of the hypothesised model.

(d) Carry out statistical tests to determine whether the estimated model is a good fit to the data.

(e) If the model is satisfactory, prepare the required forecasts, otherwise carry out further identification procedures.

The general stochastic process postulated by Box and Jenkins as capable of describing the behaviour of many economic time series is the autoregressive, integrated, moving average (ARIMA) model, which has the form:

$$w_t - \phi_1 w_{t-1} - \ldots - \phi_p w_{t-p} = \delta + u_t - \theta_1 u_{t-1} - \ldots - \theta_q u_{t-q} \tag{26}$$

where $w_t = \nabla^d z_t$

and z_t is the actual observation in period t

∇ is the difference operator, defined by $\nabla z_t = z_t - z_{t-1}$

d is the degree of differencing required to achieve stationarity

δ is a constant

u_t is a stochastic disturbance term which is assumed to be normally distributed with zero mean and constant variance σ_u^2.

(u_t is referred to as white noise)

$\phi_1, \phi_2, \ldots \phi_p$ are autoregressive parameters

$\theta_1, \theta_2, \ldots \theta_q$ are moving average parameters

This general linear stochastic process gives rise to three classes of model according to the values of p and q. When p equals zero the model is a pure moving average process:

$$w_t = \mu + u_t - \theta_1 u_{t-1} - \ldots - \theta_q u_{t-q} \tag{27}$$

When q equals zero the result is a pure autoregressive process:

$$w_t = \phi_1 w_{t-1} + \phi_2 w_{t-2} + \ldots + \phi_p w_{t-p} + \delta + u_t \tag{28}$$

and when p and q are both non zero, the model is the mixed process given in equation (26).

Model identification consists of:

(i) determining the degree of differencing required to produce a stationary series, i.e., determining the value of d.

(ii) determining the number of autoregressive and moving average parameters to be in the model, i.e., determining values for p and q respectively.

In practice the values of these parameters rarely exceed two. These procedures are carried out using the autocorrelation function and the partial autocorrelation function for the given series. An explanation of these two sets of parameters follow.

The definition of autocovariances:

$$\begin{aligned}\gamma_j &= \mathrm{Cov}(z_t\ z_{t+j})\\ &= E(z_t - E(z_t))\ (z_{t+j} - E(z_{t+j}))\\ &= E(z_t - \mu)\ (z_{t+j} - \mu)\ \text{(since all values in the series have the same mean } \mu) \qquad (29)\end{aligned}$$

implies that if a higher than average observation tends to be followed by another higher than average observation j periods later, γ_j will be positive. Similarly, if a higher than average observation tends to be followed by a lower than average observation j periods later γ_j will be negative. As a result the behaviour of autocovariances provides a great deal of information about the general appearance of a given time series. However, autocovariances depend upon the units in which the variable is measured, so it is preferable to work with autocorrelations:

$$\rho_0 = \frac{\gamma_0}{\gamma_0} = 1 \qquad (30)$$

$$\rho_j = \frac{\gamma_j}{\gamma_0} \qquad j \geqq 1$$

The theoretical autocorrelation function is the graph of ρ_j against j.

Theoretical autocorrelations are estimated by sample autocorrelations:

$$c_j = \hat{\gamma}_j = \frac{1}{T} \sum_{t=1}^{T-j} (z_t - \bar{z}) (z_{t+j} - \bar{z}) \qquad (31)$$

where

$$\bar{z} = \frac{1}{T} \sum_{t=1}^{T} z_t \qquad (32)$$

T is the number of observations in the series and

$$r_j = \hat{\rho}_j = c_{j/c_0} \qquad (33)$$

The sample autocorrelation function is the graph of r_j against j.

Autocorrelations for a non stationary time series are large even after many lags, since there is a tendency for the series to remain on one or other side of its mean for long periods of time. In contrast, autocorrelations for a stationary time series die out quickly. The value of d necessary to induce stationarity in a given time series is therefore determined by differencing the series until sample autocorrelations die out quickly.

The theoretical partial autocorrelations are the $\phi_1, \phi_2, \ldots \phi_p$ which relate the theoretical autocorrelations for a pure autoregressive process in the in the Yule-Walker equations:[7]

$$\begin{aligned}
\rho_1 &= \phi_1 + \phi_2\rho_1 + \phi_3\rho_2 + \ldots + \phi_p\rho_{p-1} \\
\rho_2 &= \phi_1\rho_1 + \phi_2 + \phi_3\rho_1 + \ldots + \phi_p\rho_{p-2} \\
&\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\
&\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\
\rho_p &= \phi_1\rho_{p-1} + \phi_2\rho_{p-2} + \ldots \phi_p
\end{aligned} \tag{34}$$

and

$$\rho_j = \phi_1\rho_{j-1} + \phi_2\rho_{j-2} + \ldots + \phi_p\rho_{j-p} \qquad j>p \tag{35}$$

Sample partial autocorrelations $\hat{\phi}_1, \hat{\phi}_2, \ldots \hat{\phi}_p$ are found by replacing theoretical autocorrelations by sample autocorrelations and solving the Yule-Walker equations for successive values of p.

For example, when p=1

$$r_1 = \hat{\phi}_1$$

when p=2, the equations

$$r_1 = \hat{\phi}_1 + \hat{\phi}_2 r_1$$

$$r_2 = \hat{\phi}_1 r_1 + \hat{\phi}_2$$

are solved to find $\hat{\phi}_2$, and so on.

The sample partial autocorrelation function is the graph of $\hat{\phi}_j$ against j.

---

7 The Yule-Walker equations are derived using the algebra of expectations. See Nelson (13) page 46 and Box and Jenkins (2) page 55.

Given a set of values for p, d and q the behaviour of the theoretical autocorrelation function and the theoretical partial autocorrelation function can be determined. Sample autocorrelation and partial autocorrelation functions[8] are compared to their theoretical counterparts and a tentative model is identified. For example, consider the first order moving average process, (ARIMA(0,0,1)):

$$z_t = \mu + u_t - \theta_1 u_{t-1} \tag{36}$$

The mean of the process is:

$$\begin{aligned} E(z_t &= \mu + E(u_t) - \theta_1 E(u_{t-1}) \\ &= \mu \end{aligned}$$

and the variance is

$$\begin{aligned} \gamma_0 &= E(u_t - \theta_1 u_{t-1})^2 \\ &= \sigma_u^2(1 + \theta_1{}^2) \end{aligned}$$

The autocovariance with lag of one period is

$$\begin{aligned} \gamma_1 &= E(u_t - \theta_1 u_{t-1})(u_{t-1} - \theta_1 u_{t-2}) \\ &= -\theta_1 \sigma_u^2 \end{aligned}$$

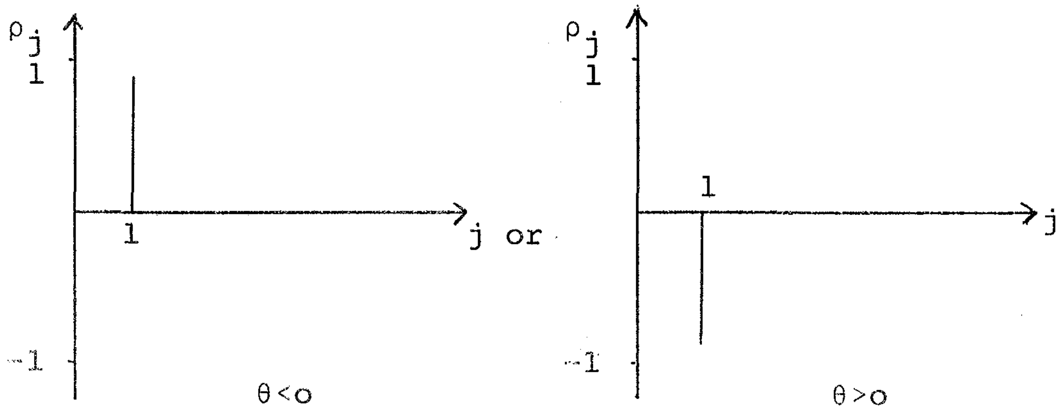
Autocovariances with lags greater than one period are all equal to zero. Therefore, the theoretical autocorrelation function is

$$\begin{aligned} \rho_0 &= 1 \\ \rho_1 &= -\theta_1/(1 + \theta_1{}^2) \\ \rho_j &= 0 \qquad \text{for } j \geq 2 \end{aligned}$$

8 These are calculated using either the original or differenced series according to stationarity requirements.

Diagrammatically these are represented by one or other of the graphs in Figure 3.6.1.

Figure 3.6.1.



If a time series has been generated by a first order moving average process we would expect sample autocorrelations to display behaviour similar to that of the theoretical autocorrelations, that is, cut off abruptly after a lag of one period.

Once a model has been tentatively identified its parameters are estimated using the principle of maximum likelihood. The likelihood function is however, non linear in parameters and so its maximum is located using an iterative procedure, known as nonlinear least squares (10).

Diagnostic checks on the fitted model are carried out in two ways:

(i) By deliberately including more terms than are believed to be appropriate and testing the significance of their coefficients.

(ii) If the correct model has been identified then the residuals around the fitted model should be randomly distributed. These residuals are tested for randomness using the chi-square statistic derived by Box and Pierce (3) and if they are found to be random, the model can be used for forecasting. If on the other hand the residuals are autocorrelated in some way, then the tentative model is rejected. The pattern of autocorrelated residuals can be used as a guide to identify a more suitable model.

The Box-Jenkins methodology can be used to forecast many types of time series including those which display a trend, cyclical and seasonal variation, although it is restricted to series which display either stationary or homogeneous non stationary behaviour.

It is superior to other time series techniques in that it provides a means of determining the optimal model for a given series of observations in terms of forecasts on past data. However, a certain amount of expertise is required to use the Box-Jenkins approach since model identification procedures are rarely clear cut. About 50 observations are needed for model identification and estimation.

The technique is based on statistical theory and therefore the researcher is able to evaluate the goodness of fit of any estimated model and to assess the reliability of forecasts by placing confidence intervals around point predictions.

4. CAUSAL MODELS

The time series models discussed in Section 3 all have one general deficiency; the inability to predict turning points in a series (other than those due to seasonality). The turning points of a series occur at both peaks and troughs.

Turning point predictions require a causal model and preferably one in which future values of the dependent variable are determined by conditions which are known at the time the forecast is prepared. Causal models relate the behaviour of the variable to be forecast to that of other variables in the system. They invariably require more data than time series models, and usually more time and expertise to develop.

4.1 Leading Indicator Models (15)

A leading indicator is a variable whose movements through time are functionally related to movements in the variable to be forecast but a lag of one or more periods takes place between a change in the leading indicator and the corresponding change in the dependent variable. A simple leading indicator model is therefore of the form:

$$z_t = f(x_{t-m}) \qquad (37)$$

where

z is the dependent variable which is to be forecast,
x is the leading indicator, and
m is the time lag.

The relationship between the dependent variable and the leading indicator is usually linear, but may be non linear, and parameters are estimated using the method of least squares.

The greatest problem is to find a suitable leading indicator, that is, one which can be justified on economic a priori grounds. Spurious correlations will not do; they are bound to bring grief at some stage. However, if a suitable leading indicator is available the model usually performs well. The maximum lead time is m periods, since values of the indicator itself will not be known beyond this stage. Theoretically this type of model can be used for long, medium or short term forecasting, but unfortunately in order to establish lagged relationships, it is often necessary to reduce the interval over which data is collected, in which case the forecast horizon is also reduced. Consequently in practice leading indicator models are more often used for short and medium term forecasting.

The model may include more than one leading indicator:

$$z_t = f(x_{t-m}; y_{t-n}) \tag{38}$$

where x and y are both leading indicators for z but not necessarily with the same delay in response.

Parameter estimation is carried out using the method of least squares, which allows both point and interval predictions to be made. Multiple leading indicator models are subject to the same problems as any multiple regression model, particularly multicollinearity since all variables are measured over time.

4.2 Box-Jenkins Transfer Function Models

Box and Jenkins (2) have extended their methodology to allow for the inclusion of explanatory variables. Such models are called transfer function models or Box-Jenkins leading indicator models. We shall consider models which have only one leading indicator, although it is possible to include more than one.

The basic assumption of the transfer function model is that pairs of observations (X_t, Y_t) in two related time series have been generated by a joint stochastic process. The process itself is assumed to be stationary and bivariant normal, which implies that:

(a) both variables have constant means,
(b) both have constant variances,
(c) autocovariances for each series depend only on the number of periods separating observations, and
(d) cross covariances depend only on the number of lags involved.[9]

If stationarity is not apparent both series are differenced until it is achieved.

The model building procedure is similar to that for a univariate stochastic process, i.e.,

(a) Postulate a general class of model,
(b) Identify a tentative model from this general class,
(c) Estimate the parameters of the hypothesised model,
(d) Determine whether the model is an adequate fit to the data,
(e) If statistical tests indicate that the model is satisfactory, it can be used for forecasting. Otherwise an alternative model is investigated.

The general form of a discrete linear transfer function model is:

$$y_t - \delta_1 y_{t-1} - \cdots - \delta_r y_{t-r} = \omega_0 x_{t-b} - \omega_1 x_{t-b-1} - \cdots - \omega_s x_{t-b-s} + N_t \qquad (39)$$

where $y_t = \nabla^d Y_t$ and $x_t = \nabla^d X_t$

---

9 Cross covariances with reversed lags are not necessarily equal.

and Y is the variable to be forecast,

X is the leading indicator,

∇ is the difference operator, defined by $\nabla Y_t = Y_t - Y_{t-1}$

d is the degree of differencing required to achieve stationarity

b is the lagged period between a change in X and a resulting change in Y

N_t is noise which is independent of the variable X and can be represented by an ARIMA(p,d,q) process of the form discussed in Section 3.6:

$$N_t - \phi_1 N_{t-1} - \dots - \phi_p N_{t-p} = \delta + u_t - \theta_1 u_{t-1} - \dots - \theta_q u_{t-q}$$

where u_t is white noise.

Model identification consists of determining values for r, d, and s in the transfer function and p, d and q in the noise function. The cross correlation function is used to identify a transfer function model in the same way as the autocorrelation function is used to identify a univariate stochastic model.

The definition of cross covariance:

$$\gamma_{xy}(k) = E(x_t - \mu_x) \quad (y_{t+k} - \mu_y) \qquad (40)$$

implies that if a higher than average x value is followed by a higher than average y value k periods later, then the cross covariance will be positive. Similarly, if a higher than average x value is followed by a lower than average y value k periods later then the cross covariance will be negative. As a result the behaviour of the cross covariance function provides information about the relationship between the two variables. Again it is necessary to work with cross correlations rather than cross covariances to obtain a measure of association which is independent of the units in which the two variables are measured.

The theoretical cross correlation coefficient is the ratio of the cross covariance to the product of the individual standard deviations.[10]

$$\rho_{xy}(k) = \frac{\gamma_{xy}(k)}{\sigma_x \sigma_y} \qquad \text{where } k = 0, \mp 1, \mp 2, \ldots \tag{41}$$

The theoretical cross correlation function is the graph of $\rho_{xy}(k)$ against k.

Theoretical cross correlations are estimated by sample cross correlations:

$$c_{xy}(k) = \hat{\gamma}_{xy}(k) = \frac{1}{T} \sum_{t=1}^{T-k} (x_t - \bar{x})(y_{t+k} - \bar{y}) \quad k=0,\ 1,2 \ldots \tag{42}$$

where $\bar{x}$ and $\bar{y}$ are the sample means of the two series and T is the total number of observations in each series,

$$r_{xy}(k) = \hat{\rho}_{xy}(k) = \frac{c_{xy}(k)}{s_x\ s_y} \qquad k = 1, 2, \ldots \tag{43}$$

and $\qquad s_x = c_{xx}(0) \ ; \qquad s_Y = c_{yy}(0)$

The sample cross correlation function is the graph of $r_{xy}(k)$ against k, and it is the primary tool used to determine values for r and s.

The parameters of a tentatively entertained model are estimated using the method of maximum likelihood and

10 This is the familiar correlation coefficient encountered in statistics. It is called the "cross" correlation coefficient here to distinguish it from autocorrelation coefficients.

statistical tests are applied to determine the goodness of fit of the estimated model to the sample data. If the model is satisfactory it can be used for forecasting. If diagnostic checks indicate the model is inadequate a new model is identified. About 50 observations are required to identify and estimate the model.

Provided a suitable leading indicator can be found, Box-Jenkins transfer models provide satisfactory forecasts for a number of series and have the advantage of being able to predict turning points.

The inclusion of past and present values of both the dependent variable and the explanatory variable represents the dynamic nature of economic behaviour.

4.3 Econometric Models

Forecasting using econometric models consists of constructing a statistical model of an economic system in an attempt to analyse and measure cause and effect relationships between the variables in the system. The model may be a single equation regression model or a simultaneous equation model.

Econometric models of either the single or simultaneous equation type are often used for forecasting. In general they are more suitable to conditional forecasting situations rather than unconditional forecasting, but under certain conditions they can be used for the latter.

One such situation is when the set of all possible values for each exogenous variable is known with certainty at the time forecasts are prepared. This can occur if exogenous variables are either,

(a) lagged by one or more period, or
(b) under the direct control of the party for whom the forecasts are prepared, or
(c) unlagged but fixed at infrequent intervals by an outside body (such as government) with the result that the current level will not change within the forecast horizon.

If exogenous variables are unknown when forecasting takes place, the researcher may decide to project the values of exogenous variables using one or other of the time series techniques discussed in Section 3. One may be forgiven for asking why the endogenous variable(s) was not extrapolated in the first case. This approach is only justified when exogenous variables follow very stable patterns which can be projected with a high degree of accuracy. Even so, the variance of these projected values should be taken into account when preparing confidence interval forecasts for endogenous variables.

Rather than extrapolate exogenous variables using time series analysis it may be possible to project exogenous variables using a separate leading indicator model for each variable. Again the variance of these projections should be included in confidence interval predictions of endogenous variables.

The unfortunate drawback with either of the last two approaches is that the resulting confidence interval predictions are often so wide as to be of little practical value.

The use of subjective estimates for unknown exogenous variables is not recommended in general.

5. COMPARISON OF FORECASTING TECHNIQUES

Wheelwright and Makridakis (17) and Chambers, Mullick and Smith (5) have compared forecasting techniques using a number of criteria such as:

(a) the lead time to which the method is best suited,
(b) ability to detect turning points,
(c) the statistical basis of the technique,
(d) data requirements for setting up the system and keeping it updated,
(e) time and expertise required to set up the system and keep it updated,

The forecasting techniques presented in Sections 3 and 4, of this paper will now be compared using these criteria.

Forecast Lead time:

Few models are able to predict the long term with any degree of accuracy; the trend and cyclical projections obtained using classical time series analysis are as good as any for this purpose. For medium term forecasting, causal models are more accurate than time series models provided suitable leading indicators can be found. They also provide good short term forecasts. Time series models are best suited to short term forecasting and of them the Box-Jenkins model is the most accurate since it provides an optimal model in terms of forecasts on past data. Exponential smoothing is in fact a special case of the Box-Jenkins general model[11] and so will not forecast as well as Box-Jenkins except under special circumstances.

<u>Ability to predict turning points</u>:

If turning point prediction is the primary concern then a causal model must be used.

---

11 See Box-Jenkins (2) p.106 and Nelson (13) p.61.

Statistical basis of the technique:

In most cases it is important to be able to evaluate a model in terms of its forecasting performance on past data and its reliability in making predictions about the future. These requirements depend upon the model having a sound statistical basis. The Box-Jenkins and causal models fall into this category.

Data Requirements:

Computer facilities have eliminated many of the data storage problems which previously existed. The main difficulty experienced by the econometrician is in obtaining valid and accurate data in the first place. This is another reason for viewing large scale econometric models with scepticism. The most sophisticated system, if based on inaccurate data, can only lead to poor performance. Time series models are attractive in that they require only the past history of the variable to be forecast, causal models require at least one more set of data. If very long series are required one may find that the early part of the series is not relevant because of changes in the system being analysed.

Time and Expertise Required:

Naive time series models require the least effort all round to set up and maintain. Providing a leading indicator is obvious, simple leading indicator models follow close behind. Box-Jenkins models, both time series and transfer function types, require considerable expertise and time to develop but once operational are simple to update. Econometric models also require expertise in the model building stage but can be used for forecasting by the layman.

6. CONCLUSIONS REGARDING TECHNIQUES

The researcher should be aware of the range of quantitative forecasting techniques available to him/her. This paper discusses a number of models which can be used for unconditional forecasting. These range from simple time series projection techniques to explanatory models which require some expertise in econometric theory to apply.

7. THE WOOL PRICE FORECASTING PROBLEM

Agricultural commodity prices tend to be subject to their fair share of fluctuation from season to season. In countries such as New Zealand, which are heavily dependent upon agriculture, accurate forecasts of the prices of major agricultural products can be of considerable benefit to economic planning.

The problem to be considered here is that of providing quarterly forecasts of New Zealand wool prices, up to one year in advance. Three series have been chosen as representative of fine, medium and coarse wools,

(a) Fine wools: price of 58's
(b) Medium wools: price of 48's
(c) Coarse wools: price of 46's

all on a clean, c.i.f., U.K., dry combing basis.[12]

Three of the methodologies previously discussed in this paper have been used in preparing the forecasts, and their applications are represented in the subsequent sections.

An exponentially weighted moving average model with trend and seasonal components of the kind developed by Winters (18) is presented in section 8. Section 9 deals with the time series approach of Box and Jenkins (2). Finally an econometric model for forecasting wool prices is discussed in section 10.[13]

Results from the three methods of forecasting are compared in section 11.

---

12 The data was obtained from various issues of "Wool Intelligence".

13 This model was developed by L.D. Woods, Dept. of Ag. Economics & Marketing, Lincoln College. Some slight modifications have been made to its original form at Mr. Wood's suggestion.

8. FORECASTING WOOL PRICES USING EXPONENTIALLY WEIGHTED MOVING AVERAGES

8.1 Model Specification

An exponentially weighted moving average model including seasonal and trend components of the following form was estimated using quarterly data,

$$\tilde{P}_t = A \frac{P_t}{F_{t-4}} + (1-A) \quad (\tilde{P}_{t-1} + R_{t-1}) \qquad o \leq A \leq 1 \tag{44}$$

where

$$F_t = B \frac{P_t}{\tilde{P}_t} + (1-B) F_{t-4} \qquad o \leq B \leq 1 \tag{45}$$

$$R_t = C(\tilde{P}_t - \tilde{P}_{t-1}) + (1-C) R_{t-1} \qquad o \leq C \leq 1 \tag{46}$$

P_t = actual price in period t

$\tilde{P}_t$ = estimated expected deseasonalised price in period t

F_t = estimate of the seasonal factor for period t

R_t = estimate of the trend component for period t

A, B, C are parameters to be estimated from the data

A forecast made at the end of period t, of price T periods in the future, is given by,

$$\hat{P}_{t+T} = (\tilde{P}_t + T.R_t) \ F_{t-4+T} \qquad T = 1,2,3,4 \tag{47}$$

8.2 Data Collection

Fifty eight observations were used in estimating the model, i.e., fourteen and a half years of quarterly data.

Prices of 58's 48's and 46's on a clean c.i.f., U.K., dry combed basis were used to represent prices of fine medium and coarse wools respectively.[14] These prices were deflated by a weighted consumer price index for the major importers of New Zealand wool.

Forecasts were produced for the deflated price series, then converted back to current prices.

8.3 Model Estimation

The parameters A, B, and C for fine, medium and coarse wools were chosen so as to minimise the standard deviation of forecast errors with a lead time of one quarter:

$$\sigma_1 = \left\{ \frac{\Sigma e_{t,\,1}^{2}}{N - 1} \right\}^{\frac{1}{2}} \tag{48}$$

where $e_{t,\,1} = P_{t+1} - \hat{P}_{t+1}$

N = number of observations used in the estimation process.

The minimum σ_1 and corresponding estimates for A, B and C were located using an exhaustive search on a three dimensional grid, consisting of various possible combinations for A, B and C. The best weights were:

Table 8.1

	A	B	C	σ_1
Fine wool	0.98	0.06	0.08	0.20
Medium wool	0.98	0.16	0.52	0.07
Coarse wool	0.98	0.20	0.48	0.07

14 Source: "Wool Intelligence"

8.4 Model Evaluation

The forecasting ability of the model can be assessed in terms of,

(i) The standard deviations of forecast errors for lead times, 1,2,3 and 4 quarters,

$$\sigma_T = \left\{ \frac{\Sigma e_{t,T}^2}{N - 1} \right\}^{\frac{1}{2}} \qquad T = 1,2,3,4 \qquad (49)$$

(ii) The coefficient of variation for lead times 1,2,3 and 4 quarters. This is the ratio of σ_T to the average of past prices.

$$CV_T = \frac{\sigma_T}{\bar{P}} \qquad T = 1,2,3,4 \qquad (50)$$

For example, in table 8.2 CV_1 = 0.1951 for fine wools means that 68% of forecasts made one period in advance will be within 19.51% of actual prices.

Similarly CV_1 = 0.1023 for medium wools indicates that 68% of all forecasts made one period in advance will lie within 10.23% of the actual price

Table 8.2

	Fine Wools	Medium Wools	Coarse Wools
σ_1	0.20	0.07	0.07
σ_2	0.35	0.13	0.13
σ_3	0.47	0.19	0.19
σ_4	0.57	0.25	0.24
CV_1	0.1951	0.1023	0.1041
CV_2	0.3381	0.2002	0.2025
CV_3	0.4542	0.2933	0.2939
CV_4	0.5507	0.3862	0.3774

8.5 Summary of Results

The following quarterly estimates of wool prices for the 1974-75 season were computed using the exponentially weighted moving average model. Annual estimates are weighted averages of quarterly figures.

Predicted Price, 1974-75 Season
(clean c.i.f. U.K. drycombed
in new pence/kg)

Fine Wools (58)

1974	July - Sept.	208
1974	Oct - Dec.	207
1975	Jan - March.	217
1975	April - June.	224
	Annual Average	214

Medium Xbd (48)

1974	July - Sept.	127
1974	Oct. - Dec.	123
1975	Jan - March.	118
1975	April - June.	112
	Annual Average	119

Coarse Xbd (46)

1974	July - Sept.	127
1974	Oct - Dec.	126
1975	Jan - March.	123
1975	April - June.	117
	Annual Average	123

8.6 Conclusions

The models have predicted a slight fall in the prices of medium and coarse wools but a rise in the price of fine wools.

However an examination of the coefficients of variation given in Table 2.2 indicates that,

(a) as expected in all three models, forecasts become less reliable as the lead time increases, and

(b) the model used for predicting fine wool prices is considerably less reliable than those for medium and coarse wools. Forecasts made four quarters in advance for the latter wool types are almost as reliable as those made only two periods in advance for fine wools.

Consequently the predicted rise in fine wool prices is regarded with suspicion, as is the slight price advantage of coarse wools over medium wools.

The forecasting ability of the three models with a lead time of one quarter can be seen in figures 8.1, 8.2 and 8.3 in the Appendix.

9. FORECASTING WOOL PRICES USING BOX-JENKINS TIME SERIES ANALYSIS

9.1 Model Specification

It is proposed to use the ARIMA model. Recall that its form is:

$$w_t - \phi_1 w_{t-1} - \dots - \phi_p w_{t-p} = \delta + u_t - \theta_1 u_{t-1} - \dots - \theta_q u_{t-q} \qquad (51)$$

where $w_t = \nabla^d P_t$

and
- P_t is the actual price in period t
- ∇ is the difference operator, defined by $\nabla P_t = P_t - P_{t-1}$
- d is the degree of differencing required to achieve stationarity
- δ is a constant
- u_t is a stochastic disturbance term with zero mean and constant variance, σ_u^2
- $\phi_1, \phi_2, \dots \phi_p$ are autoregressive parameters
- $\theta_1, \theta_2, \dots \theta_q$ are moving average parameters

Model identification consists of determining values for p, d and q and is performed using the sample autocorrelation function and the sample partial autocorrelation function. These statistics for the three price series are given in figures 9.1, 9.2 and 9.3 in the Appendix.

In each case, autocorrelation coefficients for the raw data die out after a small number of lags, indicating that the original series is stationary and d is equal to zero.

Furthermore, the initial decay in autocorrelations together with the cut-off in partial autocorrelations after two lags suggest that a second order autoregressive,[15] AR(2), scheme is appropriate for each price series, i.e.,

$$P_t = \phi_1 P_{t-1} + \phi_2 P_{t-2} + \delta + u_t \qquad (52)$$

9.2 Data Collection

At least fifty observations are required for model identification and estimation so the same quarterly time series were used as for the exponentially weighted moving average model. (i.e., prices of 58's, 48's and 46's on a clean, c.i.f., U.K., dry combed basis). These were deflated by an aggregate consumer price index for the principal importers of New Zealand wool.

9.3 Model Estimation

Initial estimates of the two autoregressive parameters ϕ_1 and ϕ_2 were obtained by solving the Yule-Walker (2) equations:-

$$\begin{aligned} r_1 &= \hat{\phi}_1 + \hat{\phi}_2 r_1 \\ r_2 &= \hat{\phi}_1 r_1 + \hat{\phi}_2 \end{aligned} \qquad (53)$$

where r_1 and r_2 are sample autocorrelation coefficients with lags of 1 and 2 periods respectively. The resulting values were then used to initiate an iterative procedure for obtaining final estimates of parameters using the Marquardt algorithm for nonlinear least squares (10). Results are given in table 9.1.

---

15 The following models were also tentatively entertained but were found to be inferior to the AR(2) scheme:- ARIMA(0,1,1); ARIMA(1,1,1); ARIMA(1,0,1).

Table 9.1

Fine wools: 58's

$$P_t = \underset{(0.122)}{1.310}\, P_{t-1} - \underset{(0.126)}{0.432}\, P_{t-2} + 0.135 + u_t$$

$$\alpha_u^{\,2} = 0.020 \qquad \chi^2 = 4.182 \qquad \text{(27 d.f.)}$$

Medium wools: 48's

$$P_t = \underset{(0.113)}{1.505}\, P_{t-1} - \underset{(0.114)}{0.565}\, P_{t-2} + 0.046 + u_t$$

$$\sigma_u^{\,2} = 0.003 \qquad \chi^2 = 24.531 \qquad \text{(27 d.f.)}$$

Coarse wools: 46's

$$P_t = \underset{(0.117)}{1.450}\, P_{t-1} - \underset{(0.119)}{0.504}\, P_{t-2} + 0.042 + u_t$$

$$\sigma_u^{\,2} = 0.003 \qquad \chi^2 = 18.164 \qquad \text{(27 d.f.)}$$

9.4 Model Validation

(i) Tests of significance on the parameters of the models given in table 9.1 show that all are significantly different from zero at the 1% level.

(ii) The adequacy of the fitted models can further be assessed by an examination of the residuals corresponding to the least squares estimates. The chi-square statistic suggested by Box and Pierce[16] was used to test for serial correlation in the residuals,

$$\chi^2 = N \sum_{j=1}^{K} r_j^{\,2} \tag{54}$$

16 G.E.P. Box & D.A.Pierce, "Distribution of Residual Autocorrelations in Autoregressive, Moving Average, Time Series Models", Journ. Amer. Stat. Assocn., Vol 64, 1970. (2) (A satisfactory model will produce residuals which have the properties of random numbers).

where r_j are the residual autocorrelations and N is the number of observations in the series. K was set to 30. A non significant chi-square statistic in each case supported the AR(2) specification of the series.

9.5 Predictions

The three price series were projected from one to four periods ahead by using conditional expectations,

$$\hat{P}_t(\ell) = E(P_{t+\ell} \mid P_{t+\ell-1}, P_{t+\ell-2}, \ldots) \tag{55}$$

where ℓ is the forecast lead time

Lead time = 1
$$\hat{P}_t(1) = \hat{\phi}_1 P_t + \hat{\phi}_2 P_{t-1} + \hat{\delta}$$

Lead time = 2
$$\hat{P}_t(2) = \hat{\phi}_1 \hat{P}_t(1) + \hat{\phi}_2 P_t + \hat{\delta}$$

Lead time = 3
$$\hat{P}_t(3) = \hat{\phi}_1 \hat{P}_t(2) + \hat{\phi}_2 \hat{P}_t(1) + \hat{\delta}$$

Lead time = 4
$$\hat{P}_t(4) = \hat{\phi}_1 \hat{P}_t(3) + \hat{\phi}_2 \hat{P}_t(2) + \hat{\delta}$$

9.6 Summary of Results

The following quarterly estimates of wool prices for the 1974-75 season were computed using the Box-Jenkins method of forecasting. Annual estimates are weighted averages of quarterly figures.

Predicted Prices, 1974-75 Season
(clean c.i.f. U.K., dry combed
in new pence/kg)

Fine wools (58)

1974	July - Sept.	166
1974	Oct. - Dec.	155
1975	Jan. - March.	148
1975	April - June.	145
	Annual Average	151

Medium Xbd (48)

1974	July - Sept.	126
1974	Oct. - Dec.	121
1975	Jan. - March.	117
1975	April - June.	114
	Annual Average	118

Coarse Xbd (46)

1974	July - Sept.	126
1974	Oct. - Dec.	122
1975	Jan. - March.	119
1975	April - June.	116
	Annual Average	120

9.7 Conclusions

The Box-Jenkins method indicates that wool prices will fall during the coming season. The predictive ability of the autoregressive models are depicted in figures 9.4, 9.5 and 9.6 in the Appendix. Statistical tests indicate that each model is a good representation of the series it describes.

10. FORECASTING WOOL PRICES USING AN ECONOMETRIC MODEL

10.1 Model Specification

The model is a multiple regression equation of the form:

$$P_t = f(Q_t; Y_t; S_t; u_t) \tag{56}$$

where

t = time period of 12 months from July to June
P = average price of wool
Q = supply of wool
Y = income
S = supply of substitutes for wool
u = stochastic disturbance term

10.2 The Variables

The Market

The market for New Zealand wool was assumed to consist of the following countries; France, Germany, Italy, Belgium, Netherlands, U.K., U.S.A and Japan.

Wool Prices

As before, prices of 58's 48's and 46's each of a clean, c.i.f., U.K., dry combing basis were used to represent prices of fine, medium and coarse wools respectively. These series were chosen because:

(a) the use of individual prices rather than a group average avoids aggregation problems.

(b) these were the only prices which enabled a sufficient number of annual observations to be obtained for estimation procedures.

Quarterly prices were initially collected and converted from sterling to $U.S., using quarterly exchange rates, to allow compatibility with the income variable.

These prices were then deflated by a quarterly consumer price index, (base: 1970), for countries comprising the market for New Zealand wool. Finally, four quarter weighted averages were taken to give average annual prices.

Wool Supply

Wool supply was taken as production plus stocks for the following countries; New Zealand, Australia, Argentine, South Africa, Uruguay, U.K and U.S.A. The variable was expressed in kilograms, clean, per capita by dividing total supply by total population of "the market".

Income

GNP for each country in "the market" was converted to \$U.S. deflated to 1970 values and divided by the market population to give real income per capita.

Supply of Substitutes

Annual per capita supply of polyester and acrylic fibres per annum was used to represent the supply of substitutes.

10.3 Model Estimation

The model was estimated for each price series by applying ordinary least squares to annual data from 1960-61 to 1973-74 inclusive. The results were:

Fine wool (58's)

$$\log P_t = -0.05 \underset{(1.36)}{-2.37} \log Q_t + \underset{(2.81)}{2.88} \log Y_t \underset{(0.48)}{-0.67} \log S_t \qquad (57)$$

$$F = 13.62^{***} \qquad R^2 = 0.80 \qquad d = 2.14$$

Medium wool (48's)

$$\log P_t = 1.29 \underset{(1.11)}{-3.05} \log Q_t + \underset{(2.29)}{1.69} \log Y_t \underset{(0.39)}{-0.58} \log S_t \qquad (58)$$

$F = 27.75^{***}$ $\qquad R^2 = 0.89$ $\qquad d = 2.16$

Coarse wool (46's)

$$\log P_t = 1.56 \underset{(1.10)}{-3.29} \log Q_t + \underset{(2.26)}{1.56} \log Y_t \underset{(0.39)}{-0.57} \log S_t \qquad (59)$$

$F = 30.71^{***}$ $\qquad R^2 = 0.90$ $\qquad d = 2.10$

10.4 Model Validation

All coefficients have correct signs and approximately correct magnitudes according to economic a priori criteria.

The Durbin-Watson statistic indicates that autocorrelation is absent from all models, so statistical criteria can be reliably used to assess the goodness-of-fit of each equation to past data.

In each case the F statistic, obtained by performing an analysis of variance on the model, is significant at the 0.1% level. This indicates that there is a strong double-logarithmic relationship between price and the explanatory variables. High R^2 values support this conclusion. The relatively large standard error for the income coefficient and the supply of substitutes coefficient in each model is explained by the presence of strong multicollinearity between income and the supply of substitutes. However, this is not regarded as important since the models are to be used for prediction only and the multicollinearity can reasonably be assumed to continue in the future.

The ability of the three equations to forecast past data is seen in figures 10.1, 10.2 and 10.3 in the Appendix.

To determine the reliability of each model in making future predictions it is necessary to investigate the structural stability through time. To do this each model was re-estimated using only 13 observations, 1960/61 to 1972/73, and the change in coefficients was tested for significance using an F statistic. The results were as follows:

Fine Wools (58)

n	d.f.	Constant	Log Q	Log Y	Log S	Σe^2	F
14	10	-0.05	-2.37	2.88	-0.67	0.29407	3.26
13	9	-3.95	-1.88	6.59	-1.26	0.21586	

Medium Wools (48)

n	d.f.	Constant	Log Q	Log Y	Log S	Σe^2	F
14	10	1.29	-3.05	1.69	-0.58	0.19464	1.64
13	9	-1.13	-2.75	3.99	-0.95	0.16457	

Coarse Wools (46)

n	d.f.	Constant	Log Q	Log Y	Log S	Σe^2	F
14	10	1.56	-3.29	1.56	-0.57	0.19043	1.80
13	9	-0.93	-2.98	3.92	-0.95	0.15869	

In each case the F statistic is given by:

$$F = \frac{\left(\sum_{i=1}^{14} e_{1i}^{2} - \sum_{i=1}^{13} e_{2i}^{2}\right)/1}{\sum_{i=1}^{13} e_{2}^{2}/9} \quad \text{on 1, 9 d.f.} \qquad (60)$$

and is non-significant indicating that in aggregate parameter estimates do not change significantly as the sample size is increased from 13 to 14 observations. The models can therefore be used for forecasting.

10.5 Prediction

To use the model for prediction it is necessary to estimate the values of exogeneous variables in 1974-75. These were derived as follows:

Wool Supply

1973/74 production figures for the major producing countries were obtained from the October issue of "Wool Intelligence", and updated according to later issues.

1974/75 production was assumed to equal 1973/74 production

1973/74 beginning stocks were obtained from the December 1973 version of "Wool Intelligence".

1973/74 end stocks (=1974/75 beginning stocks) and 1974/75 end stocks were assumed to equal carry overs to 1973/74. "World" population is assumed to rise by 1%, therefore wool supply per capita is assumed to fall by 1%.

Supply of Substitutes

1973 production was assumed to be 20% higher than 1972 figures. 1974 production was assumed to be 9% higher than 1973. Combining the latter with a population increase of 1% implies an increase in synthetic production per capita of 8%.

Income

GNP figures for 1973 and 1974 were based on predictions ex O.E.C.D. Economic Outlook, December 1973, and National Institute Economic Review, November 1973. Predictions for consumer price indices for various countries were obtained from the same source. GNP per capita (in current $U.S.) is assumed to rise by 5% while GNP per capita in real terms (i.e., 1970 $U.S.) is assumed to fall by 1.67%.

Exchange rates for the 3rd and 4th quarters of 1974 were assumed to remain at their 2nd quarter level.

Mid-year population figures for 1973 and 1974 are linear extrapolations of the past series.

10.6 Summary of Results

The following annual estimates of wool prices for the 1974-75 season were obtained from the econometric model.

Predicted Price, 1974-75 Season
(clean c.i.f., U.K., dry combed
in new pence/kg)

Fine wools (58)	230
Medium Xbd (48)	146
Coarse Xbd (46)	145

Quarterly measures of seasonality were derived using the "Census Method II" (11) and used to obtain the following predictions.

Predicted Price, 1974-75 Season
(clean c.i.f., U.K., dry combed
in new pence/kg)

Fine wools (58)

1974	July - Sept.	203
1974	Oct. - Dec.	257
1975	Jan - March.	263
1975	April - June.	196

Medium Xbd (48)

1974	July - Sept.	129
1974	Oct. - Dec.	163
1975	Jan. - March.	167
1975	April - June.	125

Coarse Xbd (46)

1974	July - Sept.	128
1974	Oct. - Dec.	162
1975	Jan. - March.	166
1975	April - June.	124

10.7 Conclusions

The econometric model is predicting wool prices for the 1974-75 season which are higher than current prices. However, this model overpredicted prices for 1973-74 so the estimates for 1974-75 are regarded with suspicion. Estimates for 1974-75 are, however, lower than estimates for 1973-74 obtained from the 13 observation version of the model. This would tend to support a fall in prices.

11. COMPARISON OF THE THREE APPROACHES

The exponentially weighted moving average model with seasonal and trend components omitted is given by,

$$\hat{P}_t = A\ P_t + (1-A)\ \hat{P}_{t-1} \tag{61}$$

where

P_t = actual price in period t
$\hat{P}_t$ = expected price in period t
A = smoothing constant

It can be shown[17] that the above model is equivalent to the Box-Jenkins ARIMA(0,1,1) process,

$$P_t = P_{t-1} + u_t - \theta u_{t-1}$$

where

P_t = actual price in period t
u_t = stochastic disturbance term in period t
θ = moving average parameter which is equal to the smoothing constant A in the exponentially weighted moving average model.

However, the identification procedures of the Box-Jenkins approach suggested an ARIMA(0,1,1). As a further check the latter process was fitted to each of the price series and in each case was shown by statistical criteria to be an inferior fit to the data.

Finally, the mean square error was calculated using the last 30 observations[18] of each series and the forecasts for lead times of one to four quarters. These are shown in table 11.1.

---

17 C.R. Nelson (13), p.61-62.

18 The entire series could not be used since the exponentially weighted moving average model uses the first 28 observations for initial estimation of seasonal and trend factors and only predicts the last 30 observations.

Table 11.1

Lead Time	Exponential Smoothing			Box-Jenkins		
	58's	48's	46's	58's	48's	46's
1 quarter	0.040	0.004	0.004	0.032	0.003	0.003
2 quarters	0.123	0.017	0.017	0.086	0.012	0.011
3 quarters	0.221	0.037	0.034	0.140	0.020	0.020
4 quarters	0.325	0.064	0.055	0.186	0.033	0.033

The smaller mean square errors for the Box-Jenkins process again suggest its superiority over the exponential smoothing model.

The econometric model is a more sophisticated approach than the time series analysis of Box and Jenkins in the sense that it takes into account the influences of other variables on wool prices. However, in this application the econometric model has two draw-backs,

(i) the model relates wool prices in period t to levels of explanatory variables in the same time period. As a result it is necessary to predict the value of explanatory variables for the coming season before wool prices can be predicted. Errors made in forecasting values of exogeneous variables will be reflected in wool price predictions.

(ii) the strong multicollinearity between income and supply of substitutes results in predictions which have large variances associated with them.

The Box Jenkins model is therefore considered to be superior to the econometric model for the preparation of final forecasts.

11.1 Conclusions

Final forecasts are those computed from the Box-Jenkins model. Indications are that wool prices will fall during the 1974/75 season.

Figure 8.1

Exponentially Weighted Moving Averages

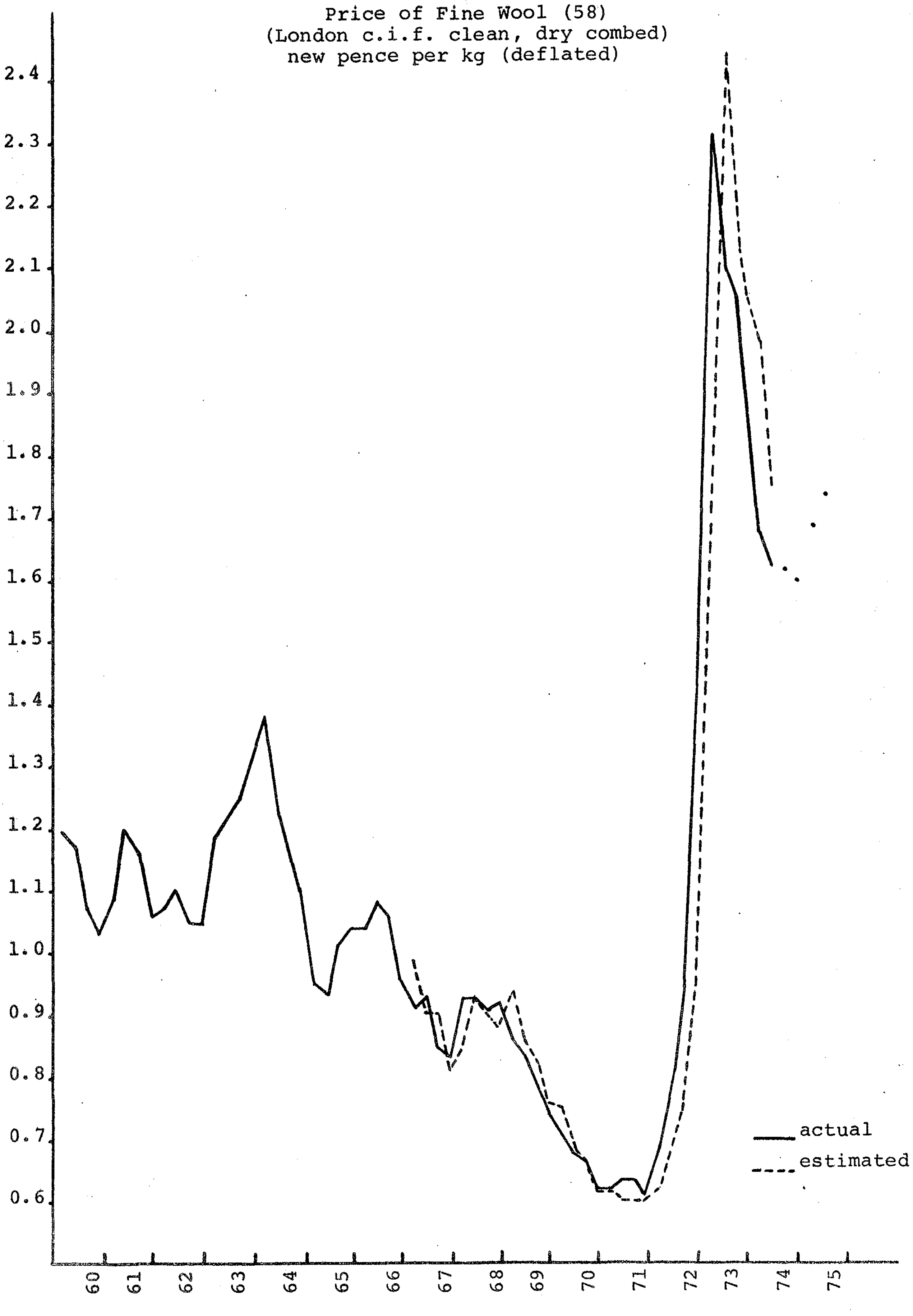


Figure 8.2

Exponentially Weighted Moving Averages

Price of Medium Wool (48)

(London c.i.f. clean, dry combed)
new pence per kg (deflated)

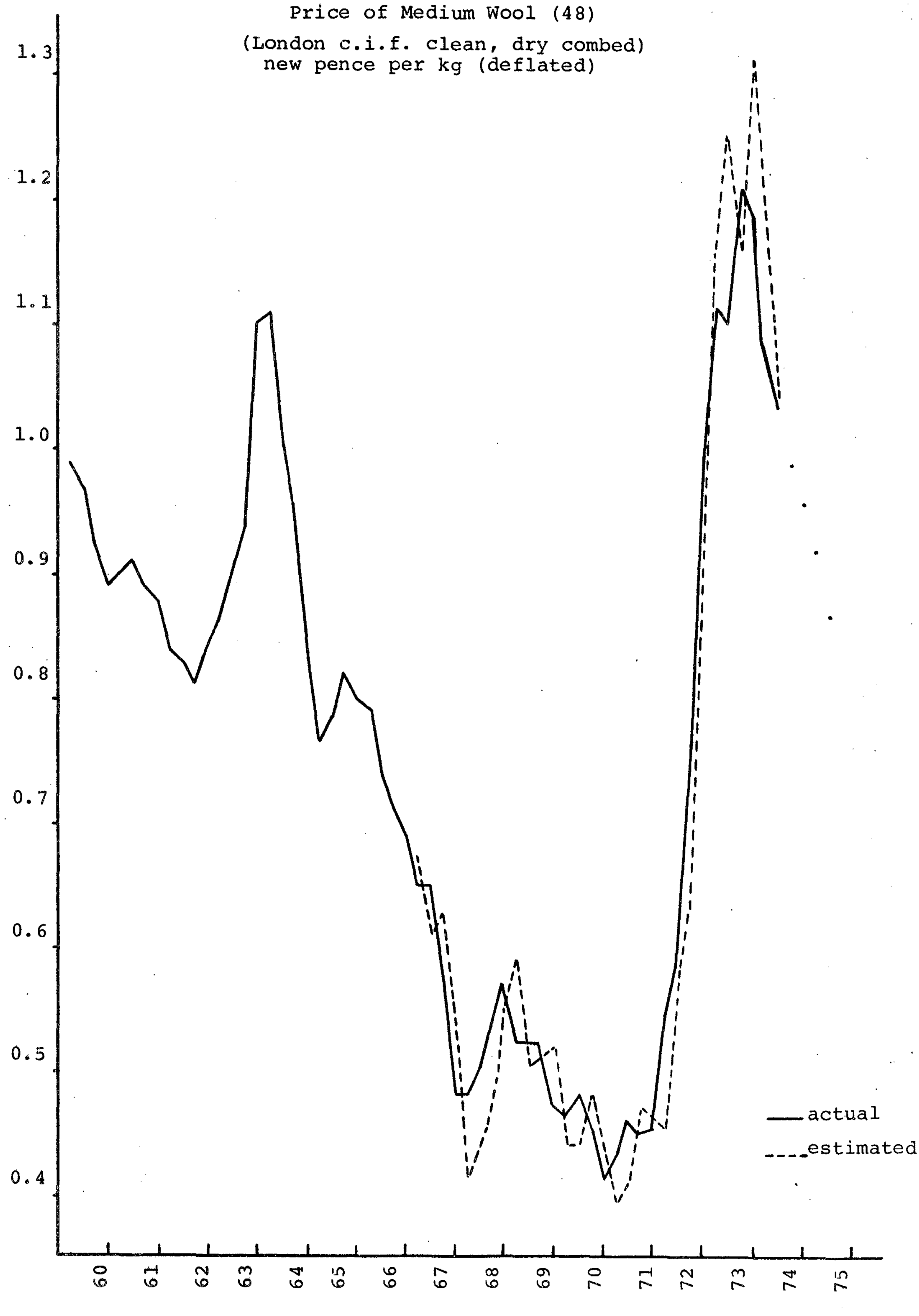


Figure 8.3

Exponentially Weighted Moving Averages

Price of Coarse Wool (46)

(London c.i.f. clean, dry combed)
new pence per kg (deflated)

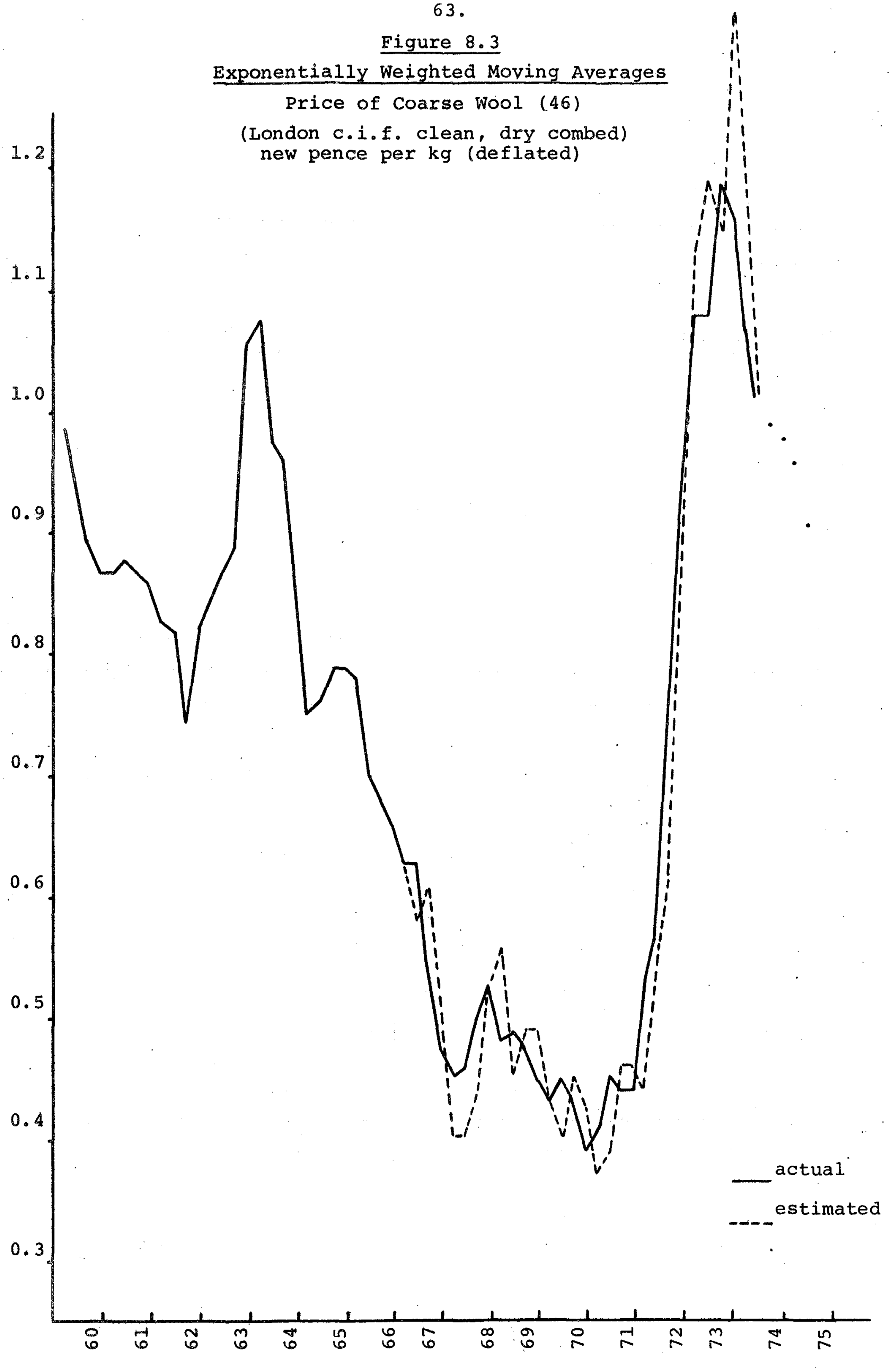


Figure 9.1

Fine Wool (58)

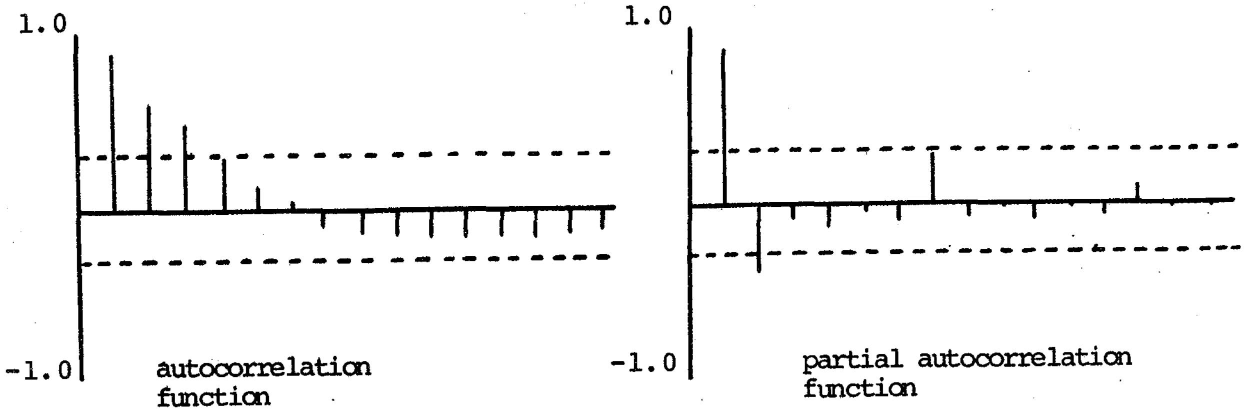


Figure 9.2

Medium Wool (48)

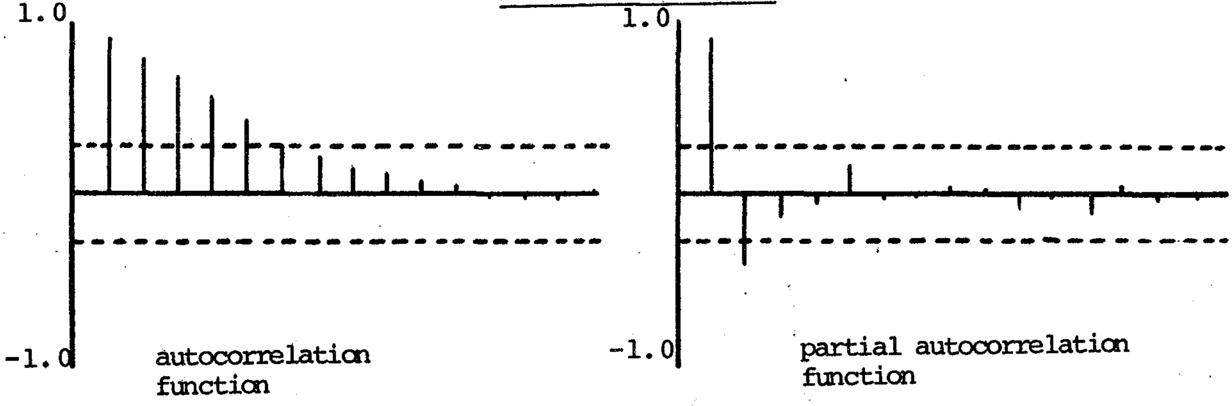
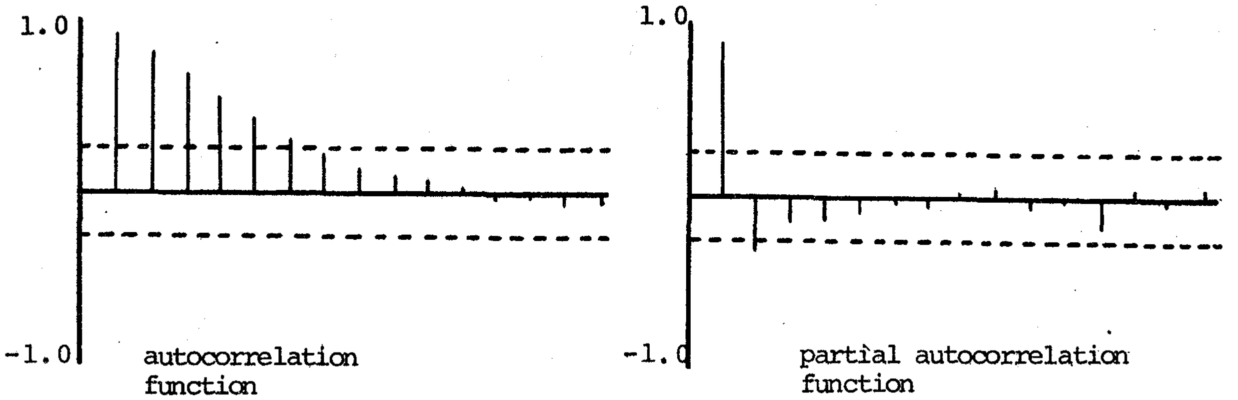


Figure 9.3

Coarse Wool (46)



----- indicates plus or minus one standard deviation

Figure 9.4

Box-Jenkins Model

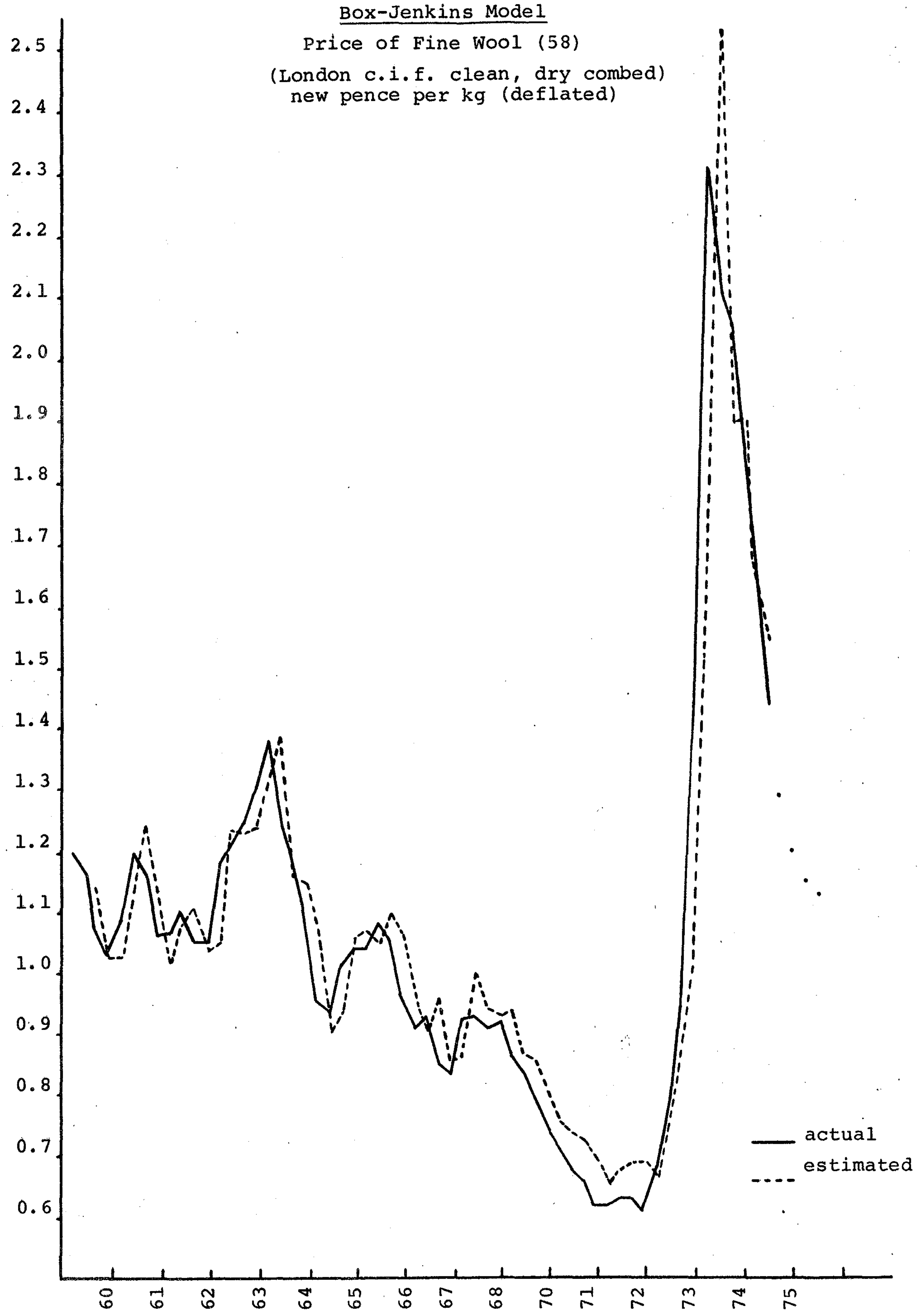


Figure 9.5

Box-Jenkins Model

Price of Medium Wool (48)

(London c.i.f. clean, dry combed)
new pence per kg (deflated)

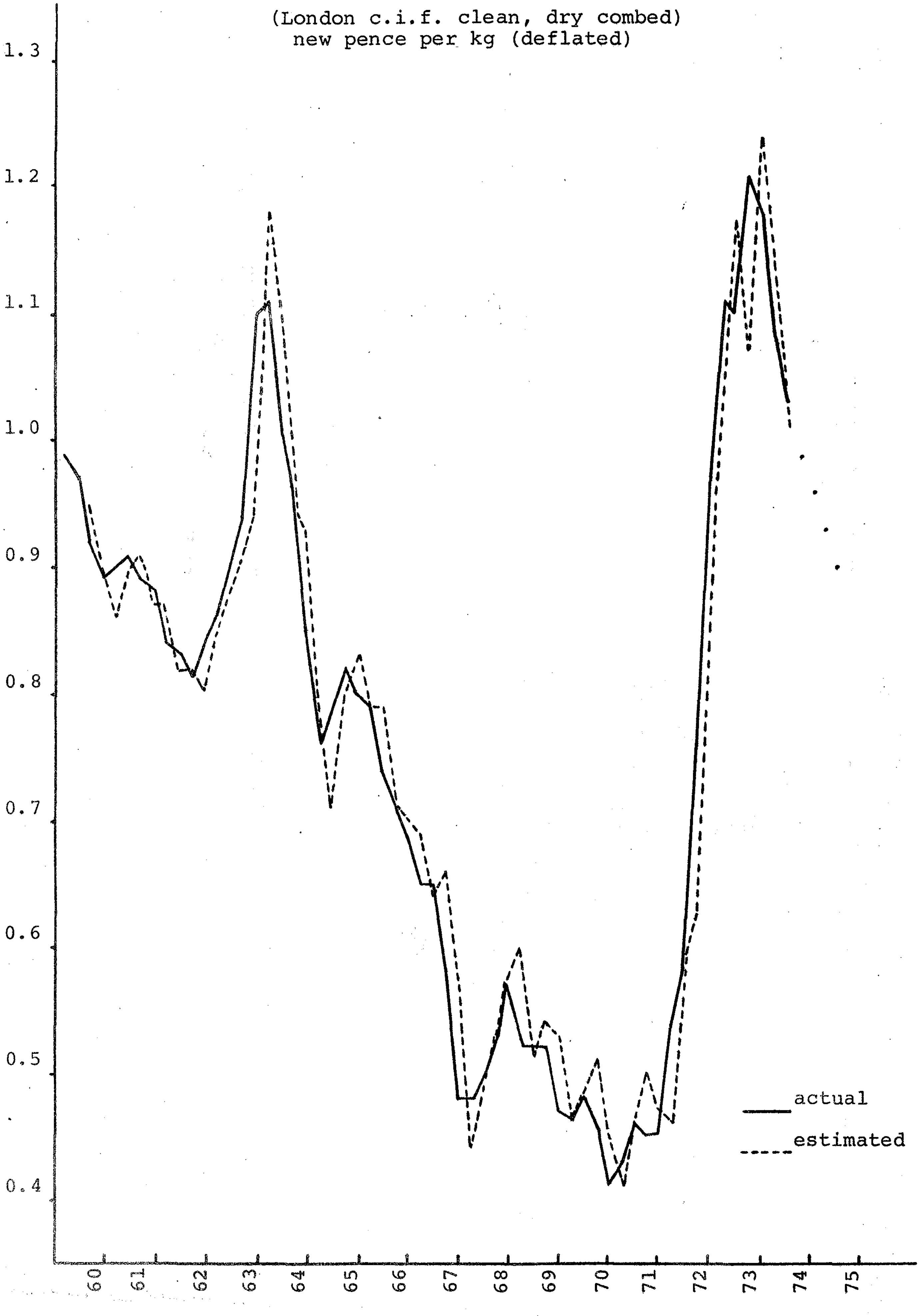


Figure 9.6

Box-Jenkins Model

Price of Coarse Wool (46)

(London c.i.f. clean, dry combed)
new pence per kg (deflated)

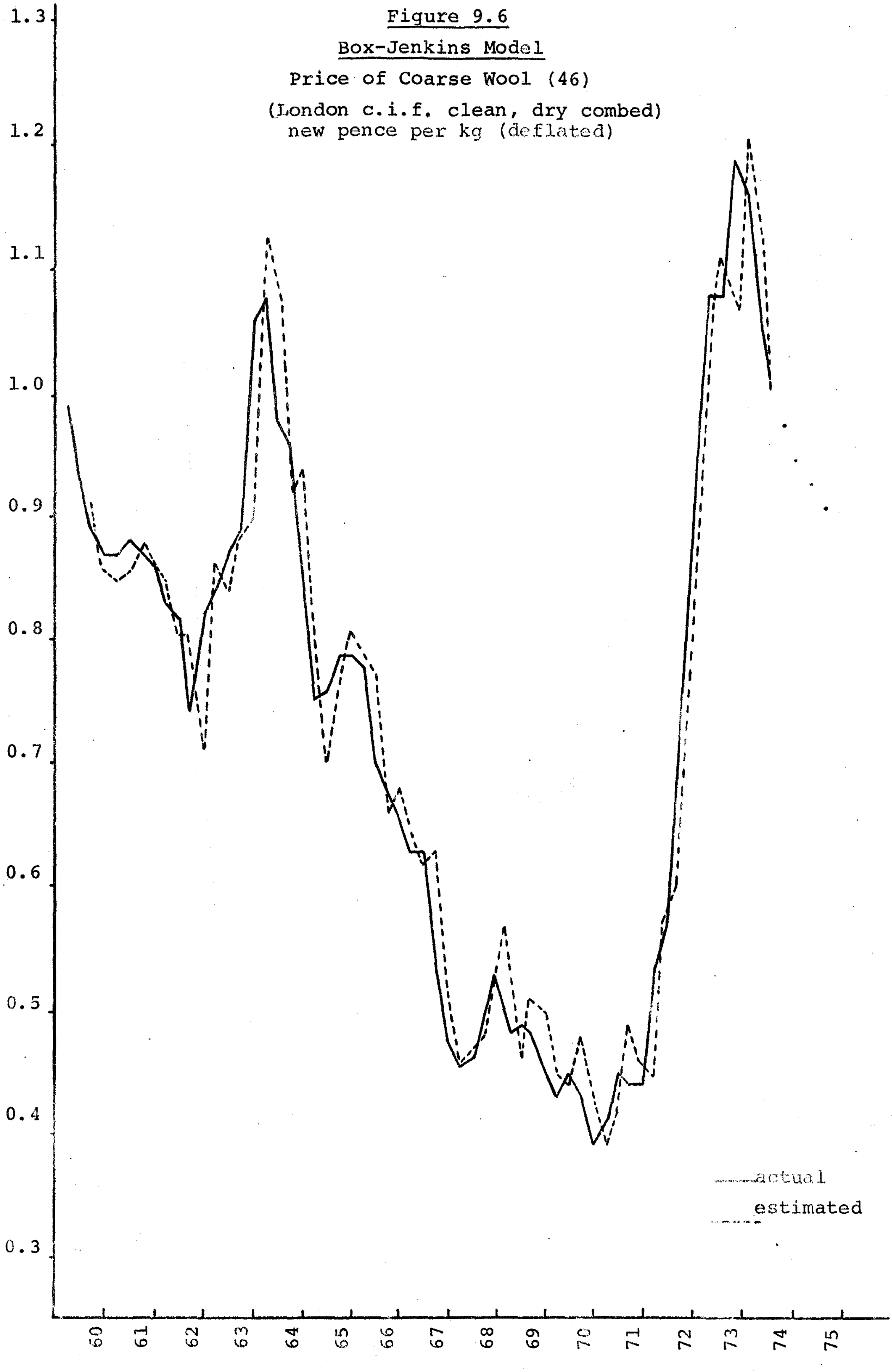


Figure 10.1

Econometric Model

Price of Fine Wool (58)

(London c.i.f. clean, dry combed)
$U.S. per kg (deflated)

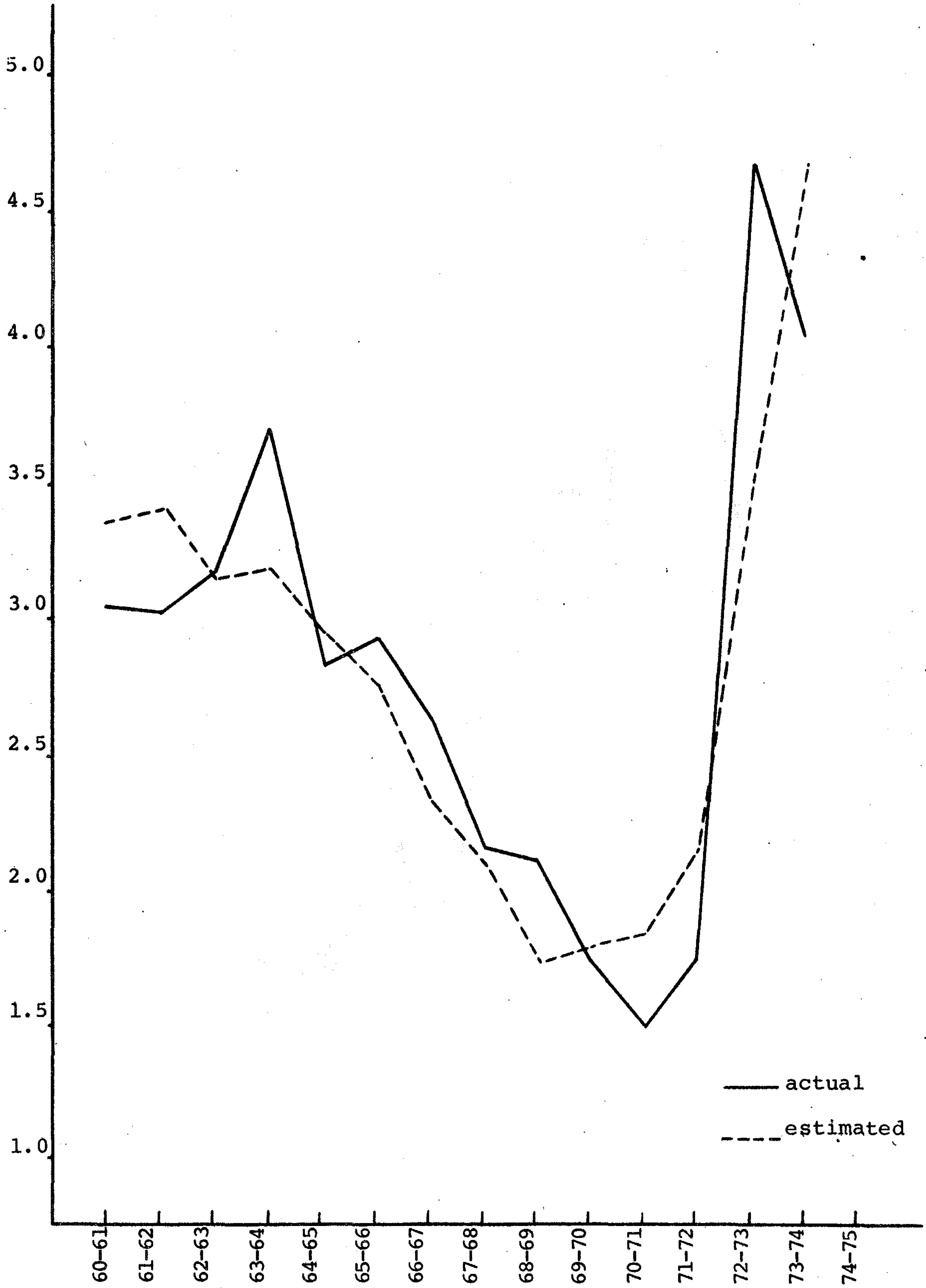


Figure 10.2

Econometric Model

Price of Medium Wool (48)

(London c.i.f. clean, dry combed)
$U.S. per kg (deflated)

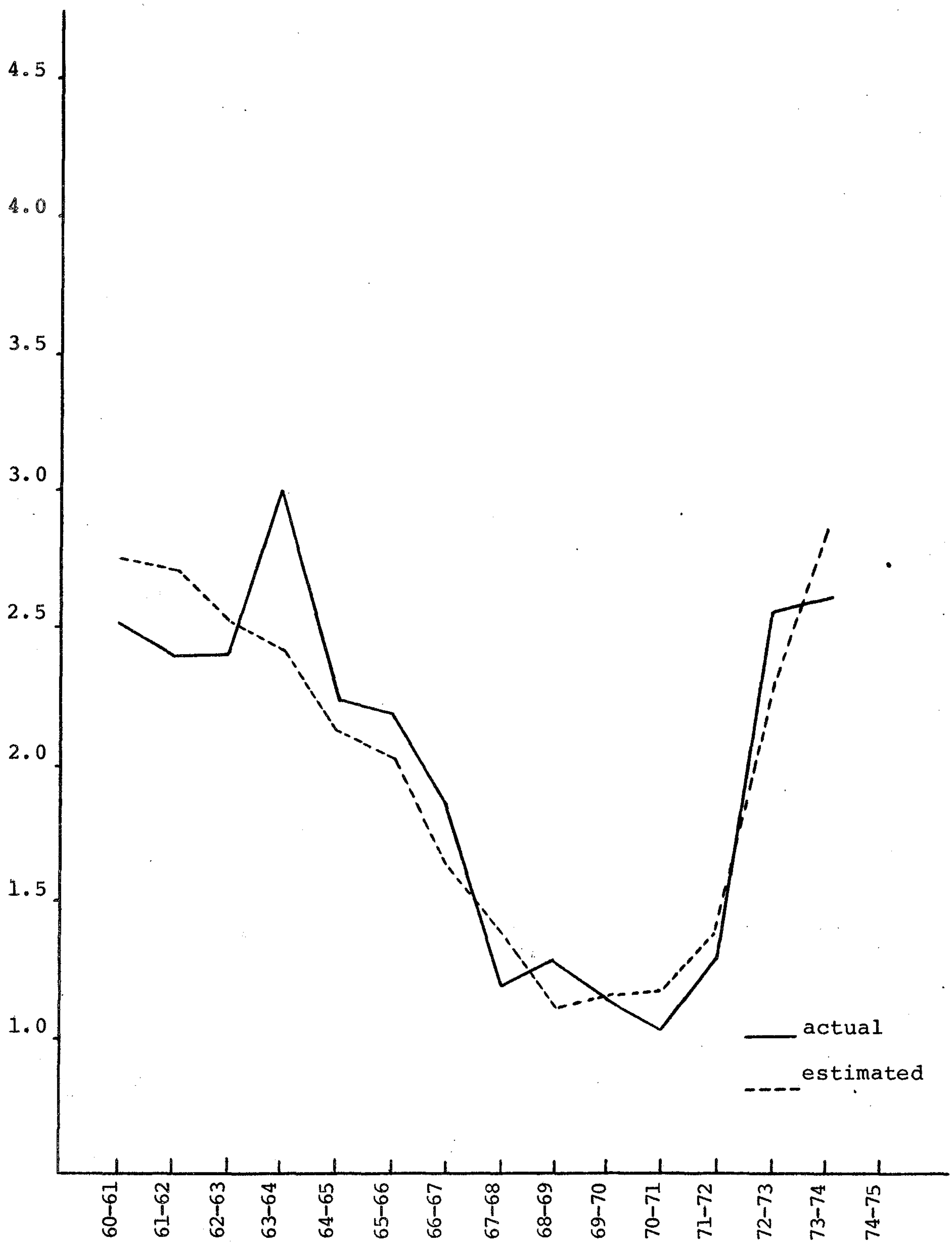
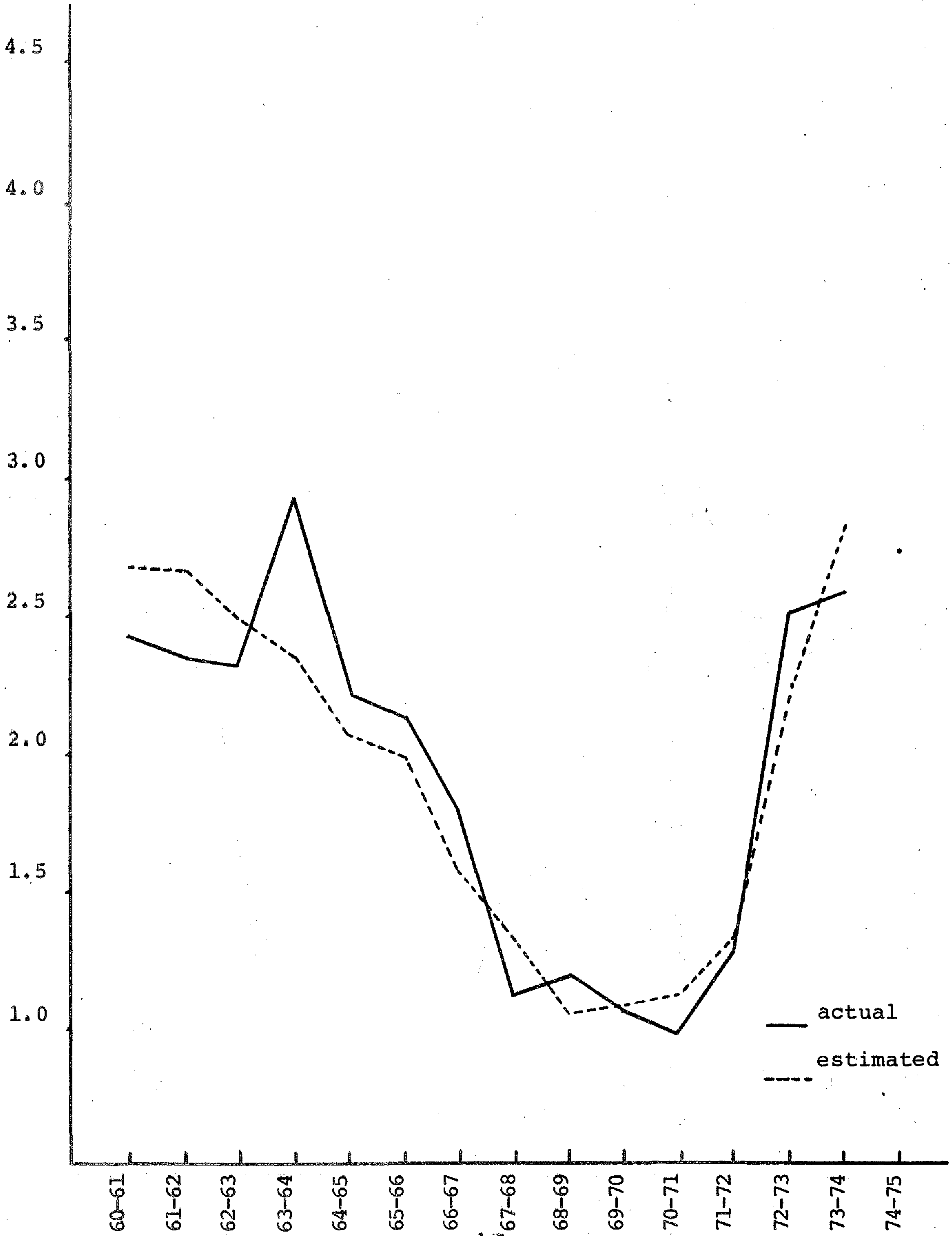


Figure 10.3

Econometric Model

Price of Coarse Wool (46)

(London c.i.f. clean, dry combed)
$U.S. per kg (deflated)



R E F E R E N C E S

(1) BOOT J.C.G. & COX E.B. 1970. Statistical Analysis for Managerial Decisions. McGraw-Hill.

(2) BOX G.E.P, & JENKINS G.M. 1973. Time Series Analysis, Forecasting and Control. Holden-Day.

(3) BOX G.E.P. & PIERCE D.A. 1970. Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. Journ. Amer. Stat. Assoc. 64.

(4) BROWN R.G. 1963. Forecasting and Prediction of Discrete Time Series. Prentice-Hall.

(5) CHAMBERS J.C. MULLICK S.K. & SMITH D.D. 1971. How to Choose the Right Forecasting Technique. Harvard Business Review: July-August,

(6) CRANE D.B. & CROTTY J.R. April, 1967. A Two-Stage Forecasting Model: Exponential Smoothing and Multiple Regression. Management Science 13 No.8.

(7) FREUND & WILLIAMS. 1969. Modern Business Statistics revised by Benjamin Perles & Charles Sullivan. Prentice-Hall.

(8) HOLT C.C. MODIGLIANI F., MUTH J.F., & SIMON H.A., 1960. Planning Production, Inventories and Work Force. Prentice-Hall.

(9) KENDALL M.G. 1973. Time Series. Charles Griffin & Co. Ltd.

(10) MARQUARDT D.W. 1963. An Algorithm for Least Squares Estimation of Non-Linear Parameters. Journ. Soc. Ind. Appl. Math. 11.

(11) MCLAUGHLIN R.L. Time Series Forecasting, Marketing Research Technique Series, No.6, American Marketing Association.

(12) MCLAUGHLIN R.L. & BOYLE J.J. Short-Term Forecasting. Marketing Research Technique Series, No.13, American Marketing Association.

(13) NELSON C.R. 1973. Applied Time Series for Managerial Forecasting. Holden-Day.

(14) NETER J. & WASSERMAN W. 1966. Fundamental Statistics for Business and Economics. Allyn and Bacon Inc.,

(15) SPENCER, CLARK & HOGUET. 1961. Business and Economic Forecasting. Richard D. Irwin, Inc,.

(16) STECKLER H.O. July-Oct 1968. Forecasting with Econometric Models: An Evaluation, Econometrica, 36.

(17) WHEELWRIGHT S.C. & MAKRIDAKIS S. 1973. Forecasting Methods for Management. Wiley.

(18) WINTERS P.S. 1960. Forecasting Sales by Exponentially Weighted Moving Averages. Management Science 6 No.3.

RECENT PUBLICATIONS

RESEARCH REPORTS

48. *Proceedings of an N.Z. Seminar on Project Evaluation in Agriculture and Related Fields,* R. C. Jensen (ed.), 1968.
49. *Inter-Industry Structure of the New Zealand Economy, 1961-5,* B. J. Ross and B. P. Philpott, 1968.
50. *Fresh Vegetable Retailing in New Zealand,* G. W. Kitson, 1968
51. *Livestock Targets in North Canterbury Hill Country:* The *Impact of Changing Prices,* J. L. Morris, H. J. Plunkett and R. W. M. Johnson, 1968.
52. *Sectoral Capital Formation in New Zealand, 1958-65,* T. W. Francis, 1968.
53. *Processing Peas: A Survey of Growers' Returns, 1967-8,* B. N. Hamilton and R. W. M. Johnson, 1968.
54. *Fertiliser Use in Southland,* R. W. M. Johnson, 1969.
55. *The Structure of Wool and Wool Textile Production, Trade and Consumption, 1948-68,* B. P. Philpott, G. A. Fletcher and W. G. Scott, 1969.
56. *Tower Silo Farming in New Zealand—Part I: A Review,* D. McClatchy, 1969.
57. *Supply and Demand Projections of the United Kingdom Meat Market in 1975,* D. R. Edwards and B. P. Philpott, 1969.
58. *Tower Silo Farming in New Zealand—Part II: Economic Possibilities,* D. McClatchy, 1969.
59. *Productivity and Income of New Zealand Agriculture, 1921-67,* D. D. Hussey and B. P. Philpott.
60. *Current Trends in New Zealand Beef Production and Disposal,* D. McClatchy.
61. *Land Development by the State: An Economic Analysis of the Hindon Block, Otago,* E. D. Parkes.
62. *An Economic Analysis of Soil Conservation and Land Retirement on South Island High Country,* R. W. M. Johnson, 1970.
63. *A Regional Analysis of Future Sheep Production in New Zealand,* R. W. M. Johnson, 1970.
64. *An Economic Assessment of the Middle Class and Upper Middle Class Market in Malaya as a Potential Outlet for New Zealand Meat and Dairy Products,* K. Y. Ho, 1970.
65. *Capital Formation in New Zealand Agriculture, 1947-67,* R. W. M. Johnson, 1970.
66. *Distribution Costs and Efficiency for Fresh Fruit and Vegetables,* G. W. Kitson, 1971.
67. *The Optimisation of a Sixteen Sector Model of the New Zealand Economy,* T. R. O'Malley, 1973
68. *An Analysis of Lands and Survey Development Projects, 1945-69,* H. J. Plunket, 1972.
69. *Quantitative Techniques for Forecasting: A Review with Applications to New Zealand Wool Prices for 1974-5,* Joan Rodgers, 1974.

MARKET RESEARCH REPORTS

1. *Processing Plant Location Studies: I: Theory and a Simple Application to N.Z. Wool Selling Centres,* W. O. McCarthy, J. L. Rodgers and C. R. Higham, 1972.
2. *Processing Plant Location Studies: II: Policy Alternatives for N.Z. Wool Selling Centres,* C. R. Higham, J. L. Rodgers and W. O. McCarthy, 1972.
3. *Doing Business in Japan,* W. O. McCarthy (ed.), 1972.
4. *The Japanese Distribution System and Implications for New Zealand Traders,* G. W. Kitson, 1973.
5. *Prospects and Strategies in Promoting Tourism Between Japan and New Zealand,* G. W. Kitson, 1973.
6. *Market Assessment,* W. O. McCarthy (ed.), 1973.
7. *Optimum Site, Number and Location of Freezing Works in the South Island, New Zealand — A Spatial Analysis,* R. J. Brodie and W. O. McCarthy, 1974.

DISCUSSION PAPERS

3. *Economic Evaluation of Water Resources Development,* R. C. Jensen, A.N.Z.A.A.S., Christchurch, 1968.
4. *An Illustrative Example of Evaluation Procedures,* A. C. Norton and R. C. Jensen, N.Z. Assn. of Soil Conservators, May 1968.
5. *The Shape of the New Zealand Economy in 1980,* B. P. Philpott and B. J. Ross, N.Z. Assn. of Economists, August 1968.
6. *Economic Problems of New Zealand Agriculture,* R. W. M. Johnson, A.N.Z.A.A.S., Christchurch, 1968.
7. *Recent Trends in the Argentine Beef Cattle Situation,* R. W. M. Johnson, November, 1968.
8. *Price Formation in the Raw Wool Market.* C. J. McKenzie, B. P. Philpott and M. J. Woods, N.Z. Assn. of Economists, February 1969.
9. *Agricultural Production Functions,* A. C. Lewis, N.Z. Assn. of Economists, February 1969.
10. *Regional Economic Development in the Context of the Changing New Zealand Economy,* B. P. Philpott, Nelson Development Seminar, April 1969.
11. *Quarterly Estimates of New Zealand Meat Price, Consumption and Allied Data, 1946-65,* C. A. Yandle.
12. *Indicative Economic Planning with a Sixteen Sector Projection Model of the New Zealand Economy,* B. J. Ross and B. P. Philpott, A.N.Z.A.A.S., Adelaide, 1969.
13. *Recent Developments in the Meat Industry with particular reference to Otago and Southland,* R. G. Pilling, Otago Branch, N.Z. Econ. Soc., October 1969.
14. *The Future Profitability of Beef Production in New Zealand,* R. W. M. Johnson, N.Z. Inst. Agr. Sc., August 1970.
15. *Demand Prospects for Beef,* B. P. Philpott, N.Z. Inst. Agr. Sc., August 1970.
16. *The Structure of Wool and Wool Textile Production, Trade and Consumption, 1958-69,* B. P. Philpott and W. G. Scott, June 1970.
17. *Trends in the Terms of Exchange and Productivity in the New Zealand Dairy Industry,* R. W. M. Johnson, June 1970.
18. *Marketing Margins for New Zealand Lamb and for all Lamb and Mutton in the United Kingdom,* A. C. Lewis and S. M. C. Murray, July 1970.
19. *A Pilot Optimisation Model for the 1972-3 N.D.C. Plan,* B. P. Philpott and T. R. O'Malley, August 1970.
20. *Recent Trends in Capital Formation in New Zealand Agriculture, 1964-9,* R. W. M. Johnson and S. M. Hadfield, 1971.
21. *Productivity and Income of New Zealand Agriculture, 1921-67,* (Supplement to Research Report No. 59), S. M. Hadfield, 1971.
22. *Some Aspects of the Economics of Nitrogen Storage Farming in New Zealand, B. P. Philpott,* I. D. Greig and A. Wright, 1972.
23. *Economic Aspects of Stone Fruit Marketing in New Zealand,* A. W. Smith 1972.
24. *New Zealand, The Ten, and Future Market Strategies,* C.C.C. Bulletin, No. 559, W. O. McCarthy 1972.
25. *The Wool Acquisition Controversy,* C.C.C. Bulletin, No. 577, W. O. McCarthy 1974.
26. *Productivity,* C.C.C. Bulletin No. 579, B. J. Ross, 1974.
27. *Investment on the Rural Scene,* paper presented to N.Z. Inst. of Valuers Seminar, B. J. Ross, 1974.

Additional copies of Research Reports, apart from complimentary copies are available at $1.00 each; except Research Report No. 48 and Market Research Report No. 3, which are $NZ3.00. Remittance should accompany order. Discussion Papers are unbound and are available free on request.