Int. Stat. Rev., Vol. 40, No. 2, 1972, pp. 123–137/*Longman Group Ltd/Printed in Great Britain*

Box-Jenkins Methods: An Alternative to Econometric Models [1]

Thomas H. Naylor and Terry G. Seaks
Duke University

and

D. W. Wichern
University of Wisconsin

Introduction

As we look back over the 1960's, it would not be inappropriate to label the 1960's "The Age of the Large-Scale Econometric Model". With the advent of the Brookings [12], FRB-MIT [10], OBE [16] and Wharton [14] econometric models of the economy of the United States, the name of the game was "How many equations are there in your model?" Some initial success in forecasting the behaviour of the U.S. economy with their big models generated a spirit of optimism on the part of econometricians. For example, Klein and Evans [14] had this to say about their model, "The past record in forecasting provides sound judgment for the view that we have a model which is thoroughly in tune with the current short-run state of the U.S. economy. This analysis gives us confidence in applying the system to a wide variety of problems in the analysis of economic theory and policy."

But as the 1960s drew to a close, some econometricians were beginning to express some doubts about the predictive capabilities of the large-scale models. For example, Cooper and Jorgenson concluded that,

> First, no single quarterly econometric model currently available is overwhelmingly superior to all other quarterly models in predicting the components of the national income and product accounts. Second, the econometric models are not, in general, superior to purely mechanical methods of forecasting. However, there are modules of the econometric models which are definitely superior to purely mechanical models. Third, the econometric models are, in general, structurally unstable [9].

Steckler ([28], p. 463) found that, "the results suggest that econometric models have not been entirely successful in forecasting economic activity". The Cooper-Jorgenson study examined the predictive performance of seven different models while Steckler considered only six. Both studies included earlier versions of the OBE [16] and Wharton [14] models, but neither study treated the Brookings [12] or the FRB-MIT [10] models.

But when 1969 rolled around, what had been a decade of extreme optimism on the part of econometric model builders was soon to end in the form of a complete disaster. The point was made quite succinctly in an article in *Business Week* entitled "Bad Year for Econometrics", which concluded that "Poor forecasts have brought once high-flying econometricians

---

Social System Simulation Program, Duke University, Working Paper No. 57.
[1] This research was supported by National Science Foundation Grant GS-2981. We are grateful to Fanny Grandis, our Research Assistant, for her valuable assistance in preparing this manuscript.

I

back to earth" [7]. In reflecting on the results of 1969 forecasts by the large-scale econometric models, Liebling and Russel put it this way:

If the past year's experience is symbolic or portentous, we had better look critically at econometric model forecasts.

But an even stronger implication emerges for consideration. During periods of unusual events, such as developed over 1968–69 when inflationary expectations were apparently important in many decision responses, all the large-scale econometric models at all applicable or useful tend to be in error, because the equations are historical in nature. Perhaps a fundamental difficulty is that the equations do not easily capture the spirit of inflation. At the very least, some effort must be made with respect to determining whether, *inherently*, econometric models are useful during periods of rapidly changing circumstances ([21], p. 6).

That large-scale econometric models may give rise to very bad forecasts is now a well-established fact. What is less clear is the exact cause of the poor predictive performance of these models. However, one can be reasonably certain that both improper specification and errors of estimation lie at the heart of the problem. There appears to be sufficient evidence to cast serious doubts on the strategy of merely adding more equations to an already complex model. Alternatively, there is no reason to believe that we can get much more mileage out of our present simultaneous equation estimating techniques, e.g. two-stage least squares, limited information maximum likelihood methods, full-information maximum likelihood methods, three-stage least squares, etc. To be sure, these estimation techniques represent improvements over ordinary least squares, but when they are applied to models which have been improperly specified, the results may not really be worth the added computational effort.

What emerges is the conclusion that econometricians who are interested in forecasting should begin to look around for alternative specifications for their forecasting models and for estimation techniques that are appropriate for these specifications. In this paper we shall describe a relatively new statistical approach to the parametric modelling of discrete time series developed by Box and Jenkins [2, 3] which has enjoyed considerable success as a forecasting tool [2, 3, 29, 30]. In spite of the promise which this approach seems to offer as a forecasting technique, econometricians have shown only limited interest in it.

We shall, therefore, attempt to summarize Box-Jenkins methods for econometricians in such a manner that they can be easily compared with more conventional econometric methods. To facilitate this comparison we shall apply Box-Jenkins methods to four economic times series for the U.S. economy over the period 1963–67. We shall then compare our forecasts using Box-Jenkins methods with the forecasts obtained by the Wharton [14] model over this same period.

Since optimal forecasts of future values of a time series are determined strictly by the stochastic model that describes the series, most of the forecasting effort must be directed toward obtaining a suitable time series model. Box and Jenkins suggest the use of an iterative four-stage procedure of: (1) specification, (2) identification, (3) estimation, and (4) diagnostic checking to determine an appropriate model.

Model Specification

In this section we describe two alternative specifications for Box-Jenkins models. However, for purposes of comparison with econometric models we shall review the specification of standard econometric models.

Econometric Models. With econometric models we assume that the economic system in question can be described by a set of simultaneous equations of the following form,

$$Ax_t + By_t + \sum_{j=1}^{p} B_j y_{t-j} + Cz_t + D = u_t \tag{1}$$

where

x_t = an $m \times 1$ vector of exogenous variables
y_t = an $n \times 1$ vector of endogenous variables
y_{t-j} = an $n \times 1$ vector of lagged endogenous variables when $j = 1, \ldots, p$
z_t = a $q \times 1$ vector of policy instruments
u_t = an $n \times 1$ vector of stochastic disturbance terms
A, B, C, D = coefficient matrices whose parameters have been estimated by standard econometric techniques.

(Of course, the model may also be non-linear.)

It should be noted that with econometric models the endogenous or output variables are assumed to be part of a system of *interdependent* equations. The output variables of the Box-Jenkins models described below are treated as though they were *independent*.

Box-Jenkins Models. A class of linear models useful in representing many types of stationary and non-stationary time series are the autoregressive integrated moving average (ARIMA) models. These models are extremely flexible, contain few parameters and can be easily modified to handle seasonality.

Suppose that observations $\{y_t\}$, $t = 0, \pm 1, \pm 2, \ldots$, constituting a time series become available at equally spaced time periods. Let $\{a_t\}$, $t = 0, \pm 1, \pm 2, \ldots$, be independent random deviates with mean zero and variance σ^2. Then there are two different approaches to representing this series. First, we may use the autoregressive model which makes a current deviation from the mean linearly dependent on previous deviations. If $\dot{y}_t = y_t - E(y_t)$ for all t, then we may write a simple autoregressive model as

$$\dot{y}_t = \phi_1 \dot{y}_{t-1} + \phi_2 \dot{y}_{t-2} + \ldots + \phi_p \dot{y}_{t-p} + a_t \tag{2}$$

where the ϕ_i represent the coefficients of the population model.

An alternative to this approach is to use the moving average model where $\dot{y}_t$ is made linearly dependent on the previous a_t. The moving average model then takes the general form

$$\dot{y}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q}. \tag{3}$$

As Box and Jenkins point out ([2], p. 93], the behaviour of one type of model cannot be represented by the other type unless an infinite number of terms are included on the right-hand side. However, the two models can be easily combined to yield our autoregressive integrated moving average model that will have a manageable number of terms. This ARIMA, or "Box-Jenkins" model has the general form

$$\dot{y}_t - \phi_1 \dot{y}_{t-1} - \ldots - \phi_p \dot{y}_{t-p} = a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q}. \tag{4}$$

A more compact representation is obtained by the use of the finite difference calculus. Let the difference operator ∇ be defined by

$$\nabla y_t = (1-B)\, y_t = y_t - By_t = y_t - y_{t-1} \tag{5}$$

where B is the backward shift operator and in general $B^p y_t = y_{t-p}$. The notation ∇^d indicates the differencing operation is employed d times. Then Box-Jenkins models may be written in the compact form

$$\phi_p(B)\, \dot{w}_t = \theta_q(B)\, a_t \tag{6}$$

where

$$\dot{w}_t = w_t - E(w_t);\; w_t = \nabla^d y_t$$
$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$$
$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q.$$

Equation (6) is referred to as an ARIMA (p, d, q) model where the numbers p and q refer to the orders of the polynomials $\phi(B)$ and $\theta(B)$ and d indicates the degree of differencing.

In order to satisfy invertibility and stationarity requirements (see [3]), the roots of $\phi(B)$ and $\theta(B)$ must lie outside the unit circle in the complex plane. In many practical situations p and q are small, and if $d>0$, the mean of w_t is often not appreciably different from zero so that we can set $\dot{w}_t = w_t$.

It may happen that either p or q is zero so that (6) reduces to a purely moving average (MA) model of order q, that is, a MA (d, q) model or a purely autoregressive (AR) model of order p, that is, an AR (p, d) model. To see that (6) could indeed have been written as an infinite moving average or infinite autoregression, let $\Phi(B) = (B)\nabla^d$; then with $E(w_t) = 0$:[1]

$$\begin{aligned} y_t &= \Phi^{-1}(B)\,\theta(B)\,a_t = \Psi(B)\,a_t \\ &= (1+\Psi_1 B+\Psi_2 B^2+\ldots)\,a_t \end{aligned} \tag{7}$$

and

$$\begin{aligned} a_t &= \Phi(B)\,\theta^{-1}(B)\,y_t = \Pi(B)\,y_t \\ &= (1-\Pi_1 B-\Pi_2 B^2-\ldots)\,y_t. \end{aligned} \tag{8}$$

Note that (8) is of the form $y_t = f(y_{t-1}, y_{t-2}, \ldots)+a_t$ so that y_t depends linearly on previous values $y_{t-1}, y_{t-2}, \ldots$ with the Π's as coefficients.

Equation (6) provides a rich class of time series models. Methods for selecting a tentative model from within this class (i.e. the identification problem) as well as procedures for estimating the model parameters (i.e. the estimation problem) and testing for adequacy of fit (i.e. the diagnostic checking problem) will be discussed later in this paper.

The ARIMA models can be extended to include processes in which recurrent seasonal patterns with a known period s occur. Parsimony can frequently be achieved by using multiplicative models of the form

$$\phi_p(B)\,\Phi_P(B^s)\,\nabla^d\nabla_s D y_t = \theta_q(B)\,H_Q(B^s)\,a_t \tag{9}$$

where $\Phi_P(B^s)$ and $H_Q(B^s)$ are polynomials in B^s of degree P and Q respectively and $\nabla_s D$ indicates the difference operator

$$\nabla_s y_t = (1-B^s)\,y_t = y_t - y_{t-s} \tag{10}$$

is applied D times. The polynomials $\phi_p(B)$ and $\theta_q(B)$ and the difference operator ∇^d have been defined in equation (6). The reader is referred to [3, 4] for an example and a more complete discussion of incorporating seasonals.

Selecting a model from the general class of ARIMA models (6) is the objective of the identification stage of the investigation. The principal tools used in identification are the autocorrelation function and the partial autocorrelation function. Since these functions exist only for stationary series, it is necessary to manipulate the original time series until it can be assumed to be stationary. This can usually be accomplished by differencing the series an appropriate number of times. Generally, if large values of ρ_k, the autocorrelation function, persist, non-stationarity is suspected. In practice, failure of an estimate of ρ_k to die out suggests that it may be profitable to treat the underlying stochastic process as stationary in ∇y_t or some higher difference.

If the autocorrelation function for various stationary processes is known, an appropriate model can be tentatively selected by differencing the original series until stationary and then comparing the observed behaviour of the estimated (sample) autocorrelation function with the theoretical values. The autocorrelation function will, in general, depend on the parameter vectors $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$. In fact, initial estimates of these quantities can be obtained by equating the appropriate second moments.

---

[1] If $E(w_t)$ is non-zero, (7) will contain a polynomial in t of degree d. Also, if $d = 0$, the y_t appearing in (7) may represent the deviation of each observation from its mean.

For example, the ARIMA (0, 1, 1) process has the autocorrelation function

$$\rho_k = \frac{-\theta}{1+\theta^2} \qquad k=1$$
$$\rho_k = 0 \qquad k>1 \tag{11}$$

where $-1<\theta<1$. If r_1 is an estimate of ρ_1, then an initial estimate $\hat{\theta}$ of θ is obtained by solving the equation

$$r_1 = \frac{-\hat{\theta}}{1+\hat{\theta}^2} \tag{12}$$

and choosing the root with modules less than one. Although estimates obtained in this way are not efficient, they are useful as starting values for procedures which yield efficient estimates.

The estimator of ρ_k used in this paper is the " minimum mean square error " estimator advocated by Jenkins and Watts [26]. The sample autocorrelation is defined as

$$r_k = \frac{\frac{1}{n}\sum_{t=1}^{n-k}(y_t-\bar{y})(y_{t+k}-\bar{y})}{\frac{1}{n}\sum_{t=1}^{n}(y_t-\bar{y})^2} \qquad k=0, 1, 2, \ldots \tag{13}$$

where the numerator and denominator are the minimum mean square error estimators of the theoretical autocorrelations of lag k and 0 respectively and $\bar{y} = \frac{1}{n}(y_1+\ldots+y_n)$ is the sample mean.

Under the assumption that the sample autocorrelation function mimics the behaviour of the autocorrelation function, failure of the sample statistic to die out quickly is taken as an indication of possible non-stationarity. If this is so, the sample autocorrelation function of the differenced series is computed. This procedure is continued until the estimated autocorrelations of $w_t = \nabla^d y_t$ are effectively zero after a moderate number of lags. At this point the sample autocorrelation hopefully has a pattern of behaviour indicated by the population autocorrelation of a member of the class of ARIMA models. This model is then fitted to the appropriately differenced series.

As an example, consider the sample autocorrelation given in Table I. The sample autocorrelation of the original series ($d = 0$) decays rather slowly indicating possible non-stationarity. However, the sample autocorrelations of the differenced series ($d = 1$) are effectively zero when judged by their approximate standard error limits of $\pm 1/\sqrt{n}$. For this example, $1/\sqrt{n} \approx 0{\cdot}10$. This implies that the first differences may be assumed to be independent random variables so that a tentative model for the stock price series is the ARIMA (0, 1, 0) model

$$\nabla y_t = a_t. \tag{14}$$

This model is also recognized as the classical random walk.

Table I. *The sample autocorrelation function for a stock price series*

Difference d	Lags k						
0	1–6	0·92	0·85	0·78	0·70	0·65	0·57
	7–12	0·51	0·45	0·39	0·34	0·28	0·25
	13–18	0·22	0·18	0·14	0·10	0·08	0·07
	19–24	0·06	0·05	0·04	0·01	−0·01	−0·04
1	1–6	−0·04	0·03	−0·03	−0·17	0·13	−0·07
	7–12	0·06	0·03	−0·13	0·11	−0·15	0·01
	13–18	−0·00	−0·04	0·04	−0·13	−0·03	−0·05
	19–24	−0·06	0·09	0·09	−0·11	0·08	−0·12

Below we reproduce a table (Table II) developed by Box and Jenkins ([2], p. 97) which lists the tentative models that can be associated with a particular pattern of behaviour by ρ_k. For the reader who is specifying a first model, their work should be most helpful. Before discussing estimation of a specified Box-Jenkins model, we review the comparable material of conventional econometrics to help highlight the differences.

Table II. *Behaviour of theoretical autocorrelation function of* d*th difference of series for various simple* (p, d, q) *models* [2]

Order	$(1, d, 0)$	$(0, d, 1)$
Behaviour of ρ_k	$\rho_k = \phi^k$ decays exponentially	only ρ_1 non-zero
Preliminary estimates from	$\phi_1 = \rho_1$	$\rho_1 = \dfrac{-\theta_1}{1+\theta_1^2}$
Admissible region	$-1<\phi_1<1$	$-1<\theta_1<1$
Order	$(2, d, 0)$	$(0, d, 2)$
Behaviour of ρ_k	mixture of exponentials or damped sine wave	only ρ_1 and ρ_2 non-zero
Preliminary estimates from	$\phi_1 = \dfrac{\rho_1(1-\rho_2)}{1-\rho_1^2}$ $\phi_2 = \dfrac{\rho_2-\rho_1^2}{1-\rho_1^2}$	$\rho_1 = \dfrac{-\theta_1(1-\theta_2)}{1+\theta_1^2+\theta_2^2}$ $\rho_2 = \dfrac{-\theta_2}{1+\theta_1^2+\theta_2^2}$
Admissible region	$-1<\phi_2<1;\ \phi_2+\phi_1<1;\ \phi_2-\phi_1<1$	$-1<\theta_2<1;\ \theta_2+\theta_1<1;\ \theta_2-\theta_1<1$
Order	$(1, d, 1)$	
Behaviour of ρ_k	decays exponentially after first lag, $\rho_k = \phi\rho_{k-1}$ $(k \geqq 2)$	
Preliminary estimates from	$\rho_1 = \dfrac{(1-\theta_1\phi_1)(\phi_1-\theta_1)}{1+\theta_1^2-2\phi_1\theta_1}$	$\rho_2 = \rho_1\phi_1$
Admissible region	$-1<\phi_1<1$	$-1<\theta_1<1$

Estimation

Econometric Models. In building econometric models such as (1), the most frequently used estimating techniques usually employ a criterion based on minimum mean squared error. In the regression model we employ linear estimators, β_i, such that for

$$y_t = \beta_0+\beta_1x_{t1}+\beta_2x_{t2}+\ldots+\beta_kx_{tk}+u_t \tag{15}$$

we have a minimum for Σu_t^2. Because this is a single equation technique (as opposed to a simultaneous equation estimation technique), it is frequently referred to as Ordinary Least Squares (OLS). If the population model of the x's and y's satisfies the well-known assumptions of homoskedastic and serially independent error terms with zero mean, then the linear estimators β_i that minimize Σu_t^2 will be unbiased and minimum variance estimators. Great generality is afforded by the fact that the linearity of the model exists only in the parameters, and the exogenous variables may in fact be manipulated in any non-linear way to yield the x's for which coefficients are estimated.

For estimating a simultaneous equation system like (1), OLS is inappropriate because of the interdependency of the regressors x and the error term u. The presence of this correlation will yield biased, inconsistent estimators if OLS is applied to systems estimations. To remedy this, econometricians have devised many ingenious techniques to consistently estimate the model coefficients. Probably the most widely used has been the method of Two Stage Least Squares (2SLS). The reader should consult any standard econometrics text for a thorough discussion of the method and its properties, but for our comparative purposes here a few points need only to be noted. First, 2SLS yields consistent coefficient estimates. That is, as sample size increases infinitely, the bias and variance of the estimators go to zero. Additionally, the method is fairly simple computationally. In its first stage, each endogenous variable is regressed with OLS on all exogenous variables. Then these estimated endogenous variables are employed in a second stage of OLS where they are used along with the exogenous variables as the regressors. At the end of the process S2LS yields consistent estimates and has gotten around the surplus of exogenous variables that accompany the usual overidentification.[1]

Earlier reference was made to the Wharton Model. This is a large, 76 equation simultaneous model developed by the University of Pennsylvania [14]. Using quarterly time series data from 1948 to 1964, the coefficients of the model were estimated by 2SLS. For purposes of comparison we here present a few typical equations from the Wharton Model. Consumption of durables in billions of 1958 dollars was given by the function

$$C = -11{\cdot}52 + 0{\cdot}1570Y - 0{\cdot}0574K_{-1} \qquad (16)$$

where Y is disposable personal income in billions of 1958 dollars and K_{-1} represents the stock of durables, lagged one quarter, in billions of 1958 dollars. A typical equation from the production sector that serves to illustrate non-linearity of the variables is

$$\log_e X = 0{\cdot}645 + 0{\cdot}7547 \log_e (N.H) + 0{\cdot}2402 \log_e (K.CP) + 0{\cdot}881\ PROD \qquad (17)$$

where X is gross output from the manufacturing sector in billions of 1958 dollars, N is millions of manufacturing employees, H is an index of hours worked in manufacturing, K is the stock of manufacturing investment in billions of 1958 dollars, CP is an index of capacity utilization, and $PROD$ is a productivity trend. The variables C and X enter into the right-hand sides of several other equations of the model, so that it is indeed a simultaneous model.

The variables that are selected for the right-hand sides above are chosen with an eye on several diagnostic checks that are discussed in more detail below in a comparison with the Box-Jenkins models. Without going into this further here, it suffices to say that the variables are selected so that they best explain the left-hand variable. Certain *a priori* relations are usually known from economic theory, and the specific formulation is selected with an eye on the diagnostic checks discussed below.

Box-Jenkins Models. Here again the least squares criterion plays a vital role, but our model is no longer linear. It is further different from an OLS model in that we do not estimate coefficients of exogenous or predetermined variables, but rather seek coefficients vectors ϕ and θ for our model (6) such that when these coefficients are employed in the model selected in the identification procedure we have a "best" model in the least squares sense.

Remembering our requirement that the roots of $\theta(B)$ and $\phi(B)$ lie outside the unit circle, we may write equation (6) as

$$\begin{aligned} a_t &= \theta_q^{-1}(B)\, \phi_p(B)\, \nabla^d y_t \\ &= f(\boldsymbol{\beta}, y) \end{aligned} \qquad (18)$$

where $\boldsymbol{\beta} = (\boldsymbol{\phi}, \boldsymbol{\theta})$ is the $(p+q)\times 1$ vector used to designate all the unknown parameters. Best estimates of $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$ in the least squares sense are those which minimize the sum of squared "errors" $\sum_{t=1}^{n} a_t^2$. If the distribution of the a_t's is normal, the least squares estimates are to a close approximation, the maximum likelihood estimates. The likelihood function actually involves a factor other than $\sum_{t=1}^{n} a_t^2$ which is a function of the parameters but its influence is small.

Very often a_t is non-linear in the parameters $\boldsymbol{\beta}$. That is, $\partial a_t/\partial \beta_i$ is not independent of all of the remaining parameters $\beta_1, \ldots, \beta_{i-1}, \beta_{i+1}, \ldots, \beta_{p+q}$. If this is the case, estimates of the

[1] A thorough discussion of 2SLS is available in any econometrics textbook. See, for example, A. S. Goldberger. *Econometric Theory*. New York: John Wiley and Sons, Inc., 1964.

parameters can be found using non-linear least squares (e.g. Hartley [17] and Marquardt [22]). An excellent introductory account of non-linear estimation motives and methods can be found in Draper and Smith [11].

Non-linear estimation procedures have been used with considerable success in the time series modelling (see, for example, Box, Jenkins and Wichern [5]). The approach can be outlined as follows.

Expanding $\boldsymbol{a}_t = f(\boldsymbol{\beta}, \boldsymbol{z})$ in a Taylor series about the point $\boldsymbol{\beta} = \boldsymbol{\beta}^0$ and ignoring terms of order higher than the first gives, with $a_t^0 = f(\boldsymbol{\beta}^0, z)$

$$a_t = a_t^0 + \sum_{i=1}^{p+q} (\beta_i - \beta_i^0)(\partial a_t/\partial \beta_i)\Big|_{\boldsymbol{\beta} = \boldsymbol{\beta}^0}$$

or

$$a_t^0 = -\sum_{i=1}^{p+q} (\beta_i - \beta_i^0)(\partial a_t/\partial \beta_i)\Big|_{\boldsymbol{\beta} = \boldsymbol{\beta}^0} + a_t. \tag{19}$$

The relationship (19) is now linear in the parameters $\gamma_i = \beta_i - \beta_i^0$ and has a form which is easily adapted to existing non-linear estimation computer programs.

For ARIMA models the derivatives, $\partial a_t/\partial \beta_i$, can be calculated explicitly; however, numerical estimates of the derivatives are generally more convenient. A provision for calculating derivatives is present in most non-linear estimation programs and these programs usually include a built-in capability to cope, at least partially, with convergence difficulties.

Two procedures which are often used in determining the parameter values that yield the minimum residual sum of squares are "steepest descent" and the Gauss linearization (19). Marquardt [22] has found that there are distinct advantages in arranging that the initial iterations be based principally on steepest descent and be constrained so that changes which are made in the parameters are not too large. His method then allows for a larger and larger component of the Gauss iteration to be amalgamated as the calculation proceeds. The computer program we used to obtain the parameter estimates given later is essentially based on Marquardt's method. The reader is referred to Box and Jenkins ([3], p. 504) for a discussion of Marquardt's algorithm as it relates to the problem at hand and to [22] for the original algorithm. The iterative non-linear estimation procedure then requires only a capability for calculating a_t^0 and initial estimates of the parameter values.

To illustrate the Box-Jenkins Model more clearly and also indicate its potential predictive power as an alternative to econometric models, ARIMA models were fitted to four of the equations from the Wharton Model. Utilizing the iterative fitting techniques and a computer program based on Marquardt's method, the following four models were obtained to characterize the behaviour of investment in fixed plant and equipment, GNP, GNP price deflator, and the unemployment rate as given in [14]. They were fitted to the actual data for the period 1948 through 1964.[1]

Investment: $(1-0{\cdot}4997B)(1-B)\,y_t = a_t$ (20)

Price deflator: $(1-0{\cdot}4576B)(1-B)\,y_t = 0{\cdot}002473 + a_t$ (21)

Unemployment rate: $(1-1{\cdot}564B+0{\cdot}7190B^2)(y_t-4{\cdot}969) = a_t$ (22)

GNP: $(1-0{\cdot}4249B)(1-B)\,y_t = 3{\cdot}156 + (1-0{\cdot}3819B^4)\,a_t.$ (23)

Though these probably look strange indeed to an econometrician used to simultaneous

[1] These variables were selected for fitting with ARIMA models because they are typical key variables that economists often wish to explain and variables that gave economists trouble in 1969. There are no strictly comparable equations in the Wharton Model since these variables are defined in identities. However, the earlier examples of Wharton equations would be strictly comparable to the type of econometric equation that would describe these four variables were they not part of identities.

difference equations, the forecasting abilities of these equations to be discussed shortly will indicate their latent power. First we compare diagnostic checking procedures of conventional and Box-Jenkins models.

Diagnostic Checking

Econometric Models. With econometric models there are several well-known diagnostic checks that are usually used to evaluate equations. These include the t and F statistics, the R^2, and the Durbin-Watson test. The t statistic measures the probability that a single regression coefficient could be at its observed level simply by chance. Hence the statement of 0·01 significance for a coefficient means that such a value would occur because of sampling variability in only 1 out of 100 cases. The F statistic is similar to the t, but it measures the significance of all the coefficients in the equation simultaneously. The R^2 statistic (or $\bar{R}^2$ if adjusted for degrees of freedom) indicates the explanatory power of the right-hand variables in a regression. It ranges from 0 to 1. An R^2 of 0·90 indicates that the right-hand variables explain 90 per cent of the variability of the left-hand, or dependent variable. The Durbin-Watson statistic is not a test of the specific model, but is rather a test of whether OLS and its assumptions are validly applied to a particular body of data. This statistic measures the presence of serial correlation in the residual errors. If it is present to too great a degree, the application of OLS is not the best strategy, and a scheme such as generalized least squares is called for.

To illustrate these statistics, we present again the two Wharton equations along with the $\bar{R}^2$, the Durbin-Watson, and beneath each coefficient the t statistic:

$$\begin{array}{lll} C = -11{\cdot}52 + 0{\cdot}1570Y - 0{\cdot}0574K_{-1} & & \\ \qquad\qquad\quad (5{\cdot}74) \qquad (2{\cdot}28) & & \\ \qquad\qquad\quad \bar{R}^2 \qquad\quad \text{D.W.} & & \\ \qquad\qquad\quad 0{\cdot}965 \qquad 1{\cdot}29 & & \end{array} \tag{24}$$

$$\begin{array}{l} \log_e X = 0{\cdot}645 + 0{\cdot}7547 \log_e (N.H) + 0{\cdot}2402 \log_e (K.CP) + 0{\cdot}881 PROD \\ \qquad\qquad (8{\cdot}48) \qquad\qquad\qquad (4{\cdot}18) \qquad\qquad\qquad (26{\cdot}2) \\ \qquad\qquad \bar{R}^2 \qquad\quad \text{D.W.} \\ \qquad\qquad 0{\cdot}984 \qquad 0{\cdot}82 \end{array} \tag{25}$$

Then by knowing the number of observations (68 in this case), we could look up probability levels for the coefficients. In fact, all are significant at 0·01 level, except for K_{-1} which is only significant at the 0·05 level. Looking at the $\bar{R}^2$, we see the equations have excellent explanatory power. A look up of the Durbin-Watson would show that we cannot reject the hypothesis of absence of serial correlation in this case.

Box-Jenkins Models. Diagnostic checking is concerned both with testing the adequacy of the ARIMA model and indicating what might be wrong if inadequacies exist.

One approach is to deliberately overparameterize the model in the direction of feared inadequacy and test to see if a significant reduction in the residual sum of squares has been achieved. Overparameterization can be an effective technique if used with caution. If the parameters are being estimated by a non-linear least-squares procedure, convergence difficulties can result because of a "trade-off" between the autoregressive and moving average polynomials. For example, the ARIMA (1, 1, 0) model

$$(1-\phi_1 B)\,\nabla y_t = a_t \tag{26}$$

is identical to the ARIMA (2, 1, 1) model

$$(1-\phi_1' B - \phi_2' B^2)\,\nabla y_t = (1-\theta_1 B)\,a_t \tag{27}$$

with

$$(1-\phi_1' B - \phi_2' B^2) = (1-\theta_1 B)(1-\phi_1 B).$$

Thus if the model (26) were adequate, overparameterizing in the form of (27) will lead to ill-conditioning and convergence difficulties.

Other tests of model adequacy are based upon detecting departures from randomness among the residuals. This is sensible since the general ARIMA model (6) implies that if the process were fitted using the true parameter values the resulting errors would be random deviates. Consequently, on the assumption of model adequacy the a_t's calculated with the least-squares estimates of the parameters $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$ should have properties very similar to independent random deviates. Any other behaviour would indicate model inadequacy.

Let us denote by $\hat{a}_t$, $t = 1, \ldots, n$, the values of a_t^0 at convergence. A simple test of model adequacy is then to examine the sample autocorrelation function of the residuals $\hat{a}_t$. If the sample autocorrelations are effectively zero when compared with standard error limits, the hypothesis of the model adequacy remains tenable.

Recently Box and Pierce [6] have been able to derive the variance covariance matrix for all of the sample autocorrelations of $\hat{a}_t$. Their work indicates the standard errors of these estimated autocorrelations depend heavily on the parameters in the entertained model for small lags and can be substantially less than those associated with larger lags.

The Box-Pierce results confirm and extend the results obtained by Durbin [13] for a first-order autoregressive process. The former authors were able to show that a portmanteau test for detecting lack of fit which takes account of these effects consists in referring

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{m} r_k^2 \tag{28}$$

to a chi-square distribution with the number of degrees of freedom equal to $m-p-q$.

A test used to detect departure from randomness of a periodic nature is the periodogram test proposed by Jenkins and Watts [19]. The necessity for such a test can arise, for example, if after fitting a model containing seasonal variation, the inadequacy of the model is reflected by periodic effects in the residuals. Periodicity is most easily detected in the frequency domain. The Jenkins-Watts test makes use of the linearity property of the integrated spectrum of a white noise process. A more detailed discussion of the periodogram test can be found in [19] and is mentioned here only for completeness.

Simulation

Econometric Models. The ultimate test of an econometric model must finally be how closely it simulates the behaviour of the system it is supposed to describe. This verification can be accomplished in two really equivalent ways. First the model may be solved for the successive time periods over which it was estimated and then the actual values of endogenous variables compared with the values predicted by the model. Goodness of fit and effects of changes in exogenous policy variables are then typically explored through the *F*-test, multiple comparisons, multiple ranking procedures, and more recently spectral analysis [25]. A second way to test the validity of the model is to forecast future behaviour and see how well the predictions work in practice. While this obviously opens many more useful and fascinating possibilities than the first approach, it has also proved to be a great stumbling-block for large econometric models as discussed at the outset of the paper.

Even apart from disappointing forecast results, simulation still may suffer from the problem of perverse time paths, inadequate estimating techniques, and unstable coefficients. Each will be reviewed briefly.

When a model has been carefully grounded in economic theory and validly estimated, it may yet yield results in complete conflict with economic reality. Time paths may exhibit explosive behaviour and logically positive variables may become negative. No answer exists to these problems so far. Work by Howrey and Kelejian is helping to develop methods whereby we can spot likely trouble areas analytically before simulating the model [18], but the mathematics of analyzing a simultaneous non-linear system for solution properties is very difficult and much work remains to be done in this area.

Though many significant advances have occurred in estimating techniques, a problem remains in the dichotomy between the static and dynamic properties of such well known estimating techniques as OLS, 2SLS, LISE, 3SLS, and FIML. That is, though these techniques may guarantee various desirable properties of estimators such as consistency and efficiency, they in no way promise us a model that will give a valid closed-loop simulation. Until there is developed some method that judges goodness of fit in terms of dynamic simulation properties instead of just static criterion such as F and t-tests, we are likely to have models that " blow up" or "run into the ground".

Unstable coefficients are yet another unresolved simulation problem. While coefficients are clearly random variables in the statistical sense, little has been done in the way of exploring the result of treating coefficients in a simulation run as random variables. Some work with the Klein-Goldberger Model indicates that shocking the coefficients during simulation may make the model difficult to solve and also cause the model to yield quite different results than from a deterministic simulation run. This short review should indicate some of the many unresolved problems that face simulators.

In light of these problems, the Box-Jenkins forecasts which we now explore offer some very attractive features.

Box-Jenkins. If estimation and diagnostic checking stages have given us the desired ARIMA model, we may use it for forecasting the time path of the variable just as we may simulate an econometric model to generate time paths. We implicitly assume that the effect of estimation errors in the parameters can be ignored. (For a Bayesian approach to the effects of estimation errors on forecasts, see [32].)

Suppose at time n we wish to forecast an observation l time periods ahead. Following Box and Jenkins we call this an origin n forecast for lead time l. Whittle [31] has shown that at time n the minimum mean square error forecast of the future observation y_{n+l} is given by

$$\hat{y}_n(l) = \underset{n}{E}(y_{n+l}) \tag{29}$$

where $\underset{n}{E}(l)$ denotes the expected value at time n and is conditional on all the information up to and including this point. Using (8) and recalling that $E(a_t) = 0$, the one-step ahead forecast at time $t-1$ can be written

$$\hat{y}_{t-1}(1) = \underset{t-1}{E}(y_t) = \Pi_1 y_{t-1} + \Pi_2 y_{t-2} + \dots \tag{30}$$

so that

$$a_t = y_t - \hat{y}_{t-1}(1).$$

That is, the "shock" a_t which enters the system at time t can be interpreted as the one-step ahead forecast error.

Given any ARIMA model the forecast $\underset{n}{E}(y_{n+l})$ can be easily computed since

$$\begin{aligned} \underset{n}{E}(y_{n+j}) &= \hat{y}_n(j) \qquad j > 0 \\ &= y_{n+j} \qquad j \leqq 0 \end{aligned} \tag{31}$$

and

$$\begin{aligned} \underset{n}{E}(a_{n+j}) &= 0 \qquad j > 0 \\ &= a_{n+j} \qquad j \leqq 0 \end{aligned} \tag{32}$$

with

$$a_{n+j} = y_{n+j} - \hat{y}_{n+j-1}(1).$$

As an example, the model

$$(1-\phi B)(1-B)\, y_t = a_t - \theta a_{t-1} \tag{33}$$

can be written

$$y_t = (1+\phi)\, y_{t-1} - \phi y_{t-2} + a_t - \theta a_{t-1}. \tag{34}$$

Replacing t with $n+l$ and taking expectations at time n gives

$$\hat{y}_n(1) = (1+\phi)\, y_n - \phi y_{n-1} - \theta a_n$$

$$= (1+\phi)\, y_n - \phi y_{n-1} - \theta\{y_n - \hat{y}_{n-1}(1)\} \tag{35}$$

$$\hat{y}_n(2) = (1+\phi)\, \hat{y}_n(1) - \phi y_n \tag{36}$$

$$\hat{y}_n(l) = (1+\phi)\, \hat{y}_n(l-1) - \phi \hat{y}_n(l-2) \qquad l \geqq 3. \tag{37}$$

Forecasts for all non-seasonal ARIMA processes *eventually* satisfy the homogeneous difference equation

$$\phi(B)\, \nabla^d \hat{y}_n(l) = 0. \tag{38}$$

That is, the nature of the forecast function is ultimately determined by the generalized autoregressive operator $\Phi(B) = \phi(B)\, \nabla^d$. An analogous result holds for seasonal models.

The recursive procedure just outlined provides an easy method for obtaining forecasts and is ideally suited for an electronic computer. To extend the procedure to a general ARIMA model, note that the generalized autoregressive operator $\Phi(B) = \phi(B)\, \nabla^d$ is a polynomial in B of degree $p+d$ so that

$$y_t - \Phi_1 y_{t-1} - \ldots - \Phi_{p+d}\, y_{t-p-d} = a_t - \theta_1 a_{t-1} - \ldots - \theta_q a_{t-q}. \tag{39}$$

Setting $t = n+l$ and taking expectations at time n gives the lead-l forecast

$$\hat{y}_n(l) = \underset{n}{E}(y_{n+l}) = \Phi_1 \underset{n}{E}(y_{n+l-1}) + \ldots + \Phi_{p+d} \underset{n}{E}(y_{n+l-p-d})$$

$$+ \underset{n}{E}(a_{n+l}) - \theta_1 \underset{n}{E}(a_{n+l-1}) - \ldots - \theta_q \underset{n}{E}(a_{n+l-q}). \tag{40}$$

The expectations on the right-hand side are evaluated using (31) and (32).

It is easily shown that the forecast error, $e_n(l)$, is given by

$$e_n(l) = y_{n+l} - \hat{y}_n(l) = (1 + \Psi_1 B + \ldots + \Psi_{l-1} B^{l-1})\, a_{n+l} \tag{41}$$

with variance

$$\operatorname{var}[e_n(l)] = (1 + \Psi_1^2 + \ldots + \Psi_{l-1}^2)\, \sigma^2. \tag{42}$$

Under the assumption of normality, conditional probability limits can be obtained by referring $e_n(l)$ to a $N[0, (1+\Psi_1^2+\ldots+\Psi_{l-1}^2)\,\sigma^2]$ distribution. For example, 50 per cent probability limits for z_{n+1} are

$$\hat{y}_n(l) \pm U_{0\cdot 25}\, [(1+\Psi_1^2+\ldots+\Psi_{l-1}^2)\, \sigma^2]^{\frac{1}{2}}. \tag{43}$$

The meaning of these limits is that *given the series up to time n*, the probability distribution associated with the future observation y_{n+l} is the normal distribution and with 50 per cent probability the value of y_{n+l} when it becomes available will lie between these limits. The notation $U_{\frac{1}{2}\varepsilon}$ is used to designate the deviate which is exceeded by a proportion $\frac{1}{2}\varepsilon$ of the unit normal distribution.

Empirical Results

To indicate how Box-Jenkins models may be used in practice, we have computed one-step-ahead forecasts for the models in equations (20) to (23) and have compared this data with the simulation results given in the Wharton Model for the same variables. Our results appear to reflect very favourably on the Box-Jenkins methods in the four stochastic processes under investigation. The comparison is based on quarterly observations from 1963.2 to 1967.4.[1] The Wharton simulation results are given in ([14], pp. 166–168), and the Box-Jenkins forecasts were generated by the models (20) to (23). Computing the absolute average error for each series, we found these results

	Wharton Average absolute error	Box-Jenkins Average absolute error
I_p (investment in billions)	1·09	0·59
P (*GNP* price deflator in percentages)	0·22	0·11
Un (unemployment in percentages)	0·186	0·109
GNP (in billions)	2·51	2·01

The interpretation is straightforward. In forecasting investment over the period, the Wharton Model error was on average 1·09 billion dollars, while the Box-Jenkins predition was off by 0·59 billion dollars on the average. As the data shows, the Box-Jenkins results were significantly better in all cases, and except for *GNP* they provide better forecasts by a factor of almost two to one.

While far from "proof" of the superiority of Box-Jenkins methods, the above forecast errors raise some significant questions.[2] Certainly for the data in question the ARIMA forecast would be preferred to the Wharton estimate. Even beyond the question of accuracy, there is the consideration of difficulty of computation. If the Box-Jenkins models did exactly the same as the Klein-Evans predictions, one might easily choose the former because of the less work involved. While a large econometric model requires dozens of regressions and then a complicated simulation run, with Box-Jenkins models the amount of work required is less by an order of magnitude. As described here, a model is fitted directly to the variable of interest and forecasts are then obtained from the model and a white noise series.

Summary

The relatively new Box-Jenkins methods have been described and compared to the standard econometric models. Based on their simplicity and possibly improved forecasts, the methods seem to offer an attractive alternative to conventional econometric methods. Certainly they point to an area where econometricians could fruitfully direct more research.[3]

One possible point of such continuing research that has gone unmentioned so far concerns the use of exogenous or policy variables. Often the prediction of a series may be improved by the use of a leading exogenous variable (or policy variable) that anticipates the movements of the variable of interest. So far there is very little literature in this area, although understanding is coming mainly from the work of Box and Jenkins [3]. When studying a variable y and a leading variable x we may fit ARIMA models to both variables and combine the models by a transfer function to yield what Box and Jenkins call a dynamic model. If our model of x has been validly derived as discussed in [3], we obtain a model that more accurately describes y than the simple application of an ARIMA model to y alone. A major field of new

---

1 Owing to a misprint in the book [14], two quarters were omitted. Also, U and P were not added to the model until 1964.2, so we had fourteen observations on U and P and seventeen on Ip and *GNP*. The Box-Jenkins errors are based on all nineteen quarters.

2 We are grateful to an anonymous referee for drawing our attention to an article by Box and Newbold (to appear in the *Journal of the Royal Statistical Society*) and the Ph.D. thesis of D. J. Reid which have attempted to compare forecasts by Box-Jenkins and other methods. They should be welcomed by econometricians interested in forecasting.

3 One should not infer from these statements that Box and Jenkins themselves have advocated the replacement of econometric models with univariate and dynamic time series modelling techniques for the purpose of forecasting.

investigation in econometrics could centre around combining ARIMA and dynamic models to represent a simultaneous system similar to that handled by traditional econometric methods. Judging from recent failures of large econometric models and the successes described above with fitting Box-Jenkins models to four Wharton variables, investigations along such lines might prove to be fruitful for econometricians.

Of course, the forecasting and computational advantages of Box-Jenkins methods have to be weighed against their inherent shortcomings: (1) they are void of explanatory power; (2) they are not based on economic theory; and (3) they are essentially sophisticated smoothing techniques and not economic models.

In summary, if one is primarily interested in forecasting, the Box-Jenkins methods may have considerable appeal. But there is some risk with Box-Jenkins methods. If they yield poor forecasts, we may be at a complete loss to explain "Why?", since they have no underpinnings in economic theory. Furthermore, if our goal is to "explain" the behaviour of an economic system and not merely to grind out forecasts, then Box-Jenkins methods may be totally unacceptable. Since they are void of economic theory, they cannot be used to test hypotheses and establish confidence intervals for complex economic phenomena.

References

[1] Barnard, G. A., Jenkins, G. M., Winston, C. B. (1962). Likelihood inference and time series. *Journal of the Royal Statistical Society: A*, **CXXV**, 321–372.

[2] Box, G. E. P., Jenkins, G. M. (1968). Some recent advances in forecasting and control. *Journal of the Royal Statistical Association: C*, **XVII**, 91–109.

[3] Box, G. E. P., Jenkins, G. M. (1970). *Time Series Analysis, Forecasting and Control.* San Francisco, Holden-Day.

[4] Box, G. E. P., Jenkins, G. M., Bacon, D. W. (1967). Models for forecasting seasonal and non-seasonal time series. *Advanced Seminar on Spectral Analysis of Time Series*, B. Harris (editor). New York, John Wiley and Sons.

[5] Box, G. E. P., Jenkins, G. M., Wichern, D. W. (1967). Least squares analysis with a dynamic model. Technical Report No. 105, Department of Statistics, University, Madison.

[6] Box, G. E. P., Pierce, D. A. (1968). Distribution of residual autocorrelations in integrated autoregressive moving average time series models. Technical Report No. 154, Department of Statistics, University of Wisconsin, Madison.

[7] *Business Week* (1970). Bad year for econometrics.

[8] Chambers, J. M. (1970). Computers in statistical research: simulation and computer-aided mathematics. *Technometrics*, **XIII**, 1–15.

[9] Cooper, Ronald L., Jorgenson, Dale W. (1969). The predictive performance of quarterly econometric models of the U.S. Conference on Econometric Models of Cyclical Behavior, Harvard University, November 14–15.

[10] De Leeuw, Frank, Gramlich, Edward (1968). The Federal Reserve-MIT Econometric Model. *Federal Reserve Bulletin* (January), 1–40.

[11] Draper, N. R., Smith, H. (1966). *Applied Regression Analysis.* New York, John Wiley and Sons.

[12] Duesenberry, James S. *et al.* (1969). *The Brookings Model: Some Further Results.* Chicago, Rand McNally.

[13] Durbin, J. (1970). Testing for serial correlation in least-squares regression when some of the regressors are lagged dependent variables. *Econometrica*, **XXXVIII** (May), 410–421.

[14] Evans, Michael K., Klein, Lawrence R. (1968). *The Wharton Econometric Forecasting Model* (Second Edition). Philadelphia, Wharton School, University of Pennsylvania.

[15] Fromm, Gary, Taubmann, Paul (1968). *Policy Simulations with an Econometric Model.* Washington, D.C., The Brookings Institution.

[16] Green, George (1969). Short and long term simulations with the OBE econometric model. Conference on Econometric Models of Cyclical Behavior, Harvard University, November 14–15.

[17] Hartley, H. O. (1961). The modified Gauss-Newton method for the fitting of non-linear regression functions by least-squares. *Technometrics*, **III**, 269–280.

[18] Howrey, Philip, Kelejian, H. H. (1969). Computer simulation versus analytical solutions. *The Design of Computer Simulation Experiments*, Thomas H. Naylor (editor). Durham, N.C., Duke University Press.

[19] Jenkins, G. M., Watts, D. G. (1968). *Spectral Analysis and Its Applications.* San Francisco, Holden-Day.

[20] Klein, Lawrence R., Evans, Michael K. (1969). *Econometric Gaming.* New York, Macmillan Company.

[21] Liebling, H. I., Russel, J. M. (1969). Forecasting business investment by anticipation surveys and econometric models in 1968–1969. *Proceedings of the American Statistical Association* (August).

[22] Marquardt, D. W. (1963). An algorithm for least squares estimation of non-linear parameters. *S.I.A.M. Journal*, **II**, 431–441.
[23] Naylor, Thomas H. (editor). (1969). *The Design of Computer Simulation Experiments*. Durham, N.C., Duke University Press.
[24] Naylor, Thomas H. (1971). *Computer Simulation Experiments with Models of Economic Systems*. New York, John Wiley and Sons.
[25] Naylor, Thomas H., Wertz, Kenneth, Wonnacott, Thomas (1969). Spectral analysis of data generated by simulation experiments with econometric models. *Econometrica*, **XXXVII** (April), 333–352.
[26] Pierce, D. A. (1968). Distribution of residual correlations in dynamic/stochastic time series models. Technical Report No. 173, Department of Statistics, University of Wisconsin, Madison.
[27] Stekler, H. O. (1966). Forecasting with an econometric model: comment. *American Economic Review*, **LVI** (December), 1241–1248.
[28] Stekler, H. O. (1968). Forecasting with econometric models: an evaluation. *Econometrica*, **XXXVI** (July-October), 437–463.
[29] Stralkowski, C. M., Quisel, K. (1969). An example in which a leading indicator is utilized to improve the forecast. Conference on Times Series Models for Marketing Forecasts, University of Wisconsin, May.
[30] Tiao, G. C., Thompson, H. E. (1969). Analysis of telephone data: a case study of forecasting seasonal time series, Conference on Time Series Models for Marketing Forecasts, University of Wisconsin, May.
[31] Whittle, P. (1952). Tests of fit in time series. *Biometrika*, **XXXIX**, 309–318.
[32] Wichern, D. W. (1969). *Studies in the Identification and Forecasting of Non-Stationary Time Series*. Ph.D. Thesis, Department of Statistics, University of Wisconsin, Madison, Wisconsin.

Résumé

Dans cet article, nous présentons et nous discutons quelques développements assez récents dans le domaine de l'analyse des séries chronologiques non stationnaires, développements qui semblent offrir des alternatives intéressantes aux méthodes économétriques classiques. Nos recherches sont basées sur les travaux récents de G. E. P. Box et G. M. Jenkins. Nous présentons leur modèle fondamental, et nous discutons les propriétés et l'estimation du modèle. Celui-ci est comparé aux techniques économétriques classiques sous forme d'identification, d'estimation et de validation. Nous signalons les similitudes et les différences des deux techniques.

Pour faciliter la comparaison et pour illustrer les possibilités des méthodes de Box et Jenkins, nous avons discuté quatre variables du modèle économétrique de Wharton (investissements de capitaux dans l'équipement, PNB, ajustement de PNB pour changements de prix, chômage) en ce qui concerne leurs équations économétriques, leurs prévisions, et leurs contre-parties de Box-Jenkins. Nous avons ajusté quatre modèles de Box et Jenkins à ces variables sur la période de 1948 à 1964. Ces modèles ont ensuite été utilisés pour projeter les variables sur cette période de temps et les résultats de cette projection ont été comparés aux données réelles et aux résultats obtenus par la simulation du modèle de Wharton. Nous avons trouvé que les modèles de Box et Jenkins donnent de meilleures prévisions (sous forme d'erreur moyenne) dans le rapport d'à peu près deux à un.

Ces modèles semblent très prometteurs pour l'économétricien qui s'intéresse surtout aux prévisions ou qui cherche des alternatives à la méthodologie économétrique classique.