EXPLAINING AND FORECASTING INVESTMENT EXPENDITURE IN CANADA: COMBINED STRUCTURAL AND TIME SERIES APPROACHES, 1961-2000

ISLAM, Asadul*

Abstract: This paper examines the structural flexible accelerator model of investment with time series model. The Box-Jenkins methodology of ARIMA specification is used for the estimated residuals of the multivariate Flexible Accelerator Model. I then re-estimate the time series model and structural model simultaneously to model the Canadian investment over the period and see how the model forecasts and fits well. The results indicate that the combined structural and ARIMA modeling gives better fit to the actual investment forecast than structural or ARIMA modeling by itself.

JEL classification: C22, E22

Keywords: Investment, Accelerator Model, AR, MA, ARIMA

1. Introduction

Expansion of investment has been considered as one of the main catalysts for the long term economic growth and employment. However investment expenditures are typically volatile and therefore its movement has important consequences for productive capacity, employment demand, personal income and balance of payment. So it is critical that the trends and causes of variations in aggregate investment are well understood. However there is yet any convincing model to explain and forecast changes in aggregate investment expenditure to the desired degree of precision. Tevlin and Whelan (2000), for example, show that the existing time series model cannot explain the behavior of investment of U.S economy. In this paper I show that the traditional structural model is not able to capture the behavior of investment. This is also demonstrated in many of the recent empirical literature. The aim of this paper is therefore to contribute to the literature of investment by combining the structural model of investment with the time series model. The combined model turns out to forecast well and can explain the investment behavior of the economy better than simply applying the structural or time series model alone.

In the paper such as those written by Chenery (1952), Koyck (1954), Lucas (1967), Jorgenson (1971), Lovell (1971), Epstein and Denny (1983) and many others used the flexible accelerator model[1] or gradual adjustment hypothesis of investment to link between a country or a firm's "desired capital stock" or that level which would be predicted on the basis of economic theory and observed market conditions and observed series of actual capital stock. Some authors, e.g., Hay(1970) argues that it is inappropriate to assume that firms plan capital so as to adjust inventories by some fraction of the gap between their current magnitude and their equilibrium level. While I do not contend with

* Asadul Islam, Department of Economics, Monash University, Clayton Vic 3800 Australia, e-mail: asadul.islam@buseco.monash.edu.au

[1] The term flexible accelerator model is used because the model is a generalized form of the older *accelerator* model of investment, in which investment is proportional to change in the level of GDP.

the suggestions of how the firms can close the gap between current and desired magnitude, instead in this paper, I examine the combined flexible accelerator model and time series model to account for investment of the Canadian economy.

The rest of the paper is organized as follows. Section 2 presents the basic model. Section 3 describes the data and presents summary statistics. In section 4 I estimate the flexible accelerator model and specify the ARIMA model. This section also presents the estimation results that combine the time series model with flexible accelerator model of investment. Section 5 goes with the model's ability to forecast, and section 6 concludes.

2. The Basic Model and Estimation Strategy

2.1. The Basic Model: One of the earliest empirical models of aggregate investment behaviour is the accelerator model based on the assumption of a fixed capital/output ratio. A more general version of the original accelerator model is called the flexible accelerator model (FAM) and was put forward by Koyck (1954). The basic notion behind the flexible accelerator model is that the larger the gap between the existing capital stock and the desired capital stock, the more rapid a firm's rate of investment. In its simplest version the FAM encompasses the hypothesis that investment I_t is proportional to the difference between desired capital, K_t*, and previous capital, $K_{t-1.}$ The speed of adjustment is constant proportionality, η. That is-

$$K_t - K_{t-1} = \eta[K_t^* - K_{t-1}] + u_t \qquad (1)$$

where $0<\eta\leq1$ is the adjustment parameter and u_t is a random error term.

According to equation (1), firms plan production so as to adjust capital by some fraction of the gap between their current magnitude and their equilibrium. Rewriting equation (1) as -

$$K_t = K_{t-1} + \eta[K_t^* - K_{t-1}] + u_t \qquad [1(a)]$$

Equation 1(a) implies that in order to increase the capital stock from K_{t-1} to the level of K_t, a country or firm has to achieve an amount of investment $I_t = K_t - K_{t-1}$. Equation 1(a) can be written in terms of net investment as –

$$I_t = \eta[K_t^* - K_{t-1}] + u_t \qquad (2)$$

The desired or target capital K_t* is strictly proportional to output Q_t, $K_t^* = \varphi Q_t$ (Jorgenson 1971). Substituting $K_t^*=\varphi Q_t$ in equation (2) we get –

$$I_t = \eta\varphi Q_t - \eta K_{t-1} + u_t$$

Which we can write as –

$$I_t = \beta_0 + \beta_1 Q_t + \beta_2 K_{t-1} + u_t \qquad (3)$$

Where $I_t = K_t - K_{t-1} = \Delta K_t$, β_0= intercept, $\beta_1=\eta\varphi$ and $\beta_2= -\eta$ [$\therefore \varphi = -\beta_1/\beta_2$= desired capital output ratio, and $\beta_{0,}$ $\beta_1\geq0$, $1\leq\beta_2<0$]. Equation (3) can be written more compactly as-

$$Y_t = X_t'\beta + u_t \qquad (4)$$

where $Y_t = [I_t]$, T×1 column vector of observations on the dependent variable, $X_t' = [1\ Q_t\ K_{t-1}]$, T×3 matrix giving T observations on variable Q_t and K_{t-1}, β =3×1 column vector of unknown parameters and u_t = T×1 column vector of stochastic error term

2.2. The Estimation Strategy. Equation (4) represents our structural model. The model has an additive error term, u_t, that accounts for unexplained variance in I_t; that is, it accounts for that part of the variance of I_t that is not explained by Q_t and K_{t-1}. I first estimate equation (4) and then construct an integrated autoregressive-moving average (ARIMA) model for the estimated residual series u_t of the regression (4). Then I substitute the ARIMA model for the implicit error term in the original regression equation and combine the two, and re-estimate all the parameters simultaneously. The combined regression time-series model is termed as Multivariate ARIMA or MARIMA model[2]. The MARIMA can be written as-

$$Y_t = X_t'\beta + \theta^{-1}(B)\ \phi(B)\ \varepsilon_t \qquad (5)$$

where, $\theta(B) = (1- \theta_1 B - \theta_2 B^2 - \ldots\ldots\ldots \theta_p B^p)$ $\quad \phi(B) = (1- \phi_1 B - \phi_2 B^2 - \ldots\ldots\ldots\ \phi_q B^q)$, B is the backward shift lag operator (for example, $Bu_t = u_{t-1}$, $B^2 u_t = u_{t-2}$), p and q are the order of AR and MA and ε_t is normally distributed error term which may have different variance from u_t..

I use Box and Jenkins (1976) "identification" techniques to specify the order of ARIMA (*p*, *d*, *q*). Accordingly, in the first stage, I determine the correct order of differencing, *d*, and then the order of ARMA model (*p* and *q*). In the second stage, I estimate the parameters of β_0, β_1, β_2 of the structural FAM and the parameters $\theta_1, \theta_2, \ldots \theta_p, \phi_1, \phi_2, \ldots, \phi_q$ of the time series model simultaneously. The final step is to conduct some diagnostic checks and other model specification tests to examine if the model "makes sense" or "holds together.

3. Data and Descriptive Statistics

The data are from statistics Canada, CANSIM II quarterly time series from 1961 to 2000. Gross Domestic Product (GDP), Q_t, and the business and government fixed capital formation, K_t, are at 1992 constant prices. The net investment, I_t, series is derived from the capital formation as indicated by the structural flexible accelerator model.

Table-1 summarizes the mean and standard deviation (S.D) of the key variables by different time periods. The averages of investment, capital formation and GDP for the entire time period are \$287m, \$135132m and \$23315m with standard deviation of \$2569m, \$48767m and \$11099m respectively.

[2] The model is also called ARIMAX model to distinguish from ARIMA.

Figure 1: Net Investment in Canada from 1961:2-2000:4 (in million CD$)

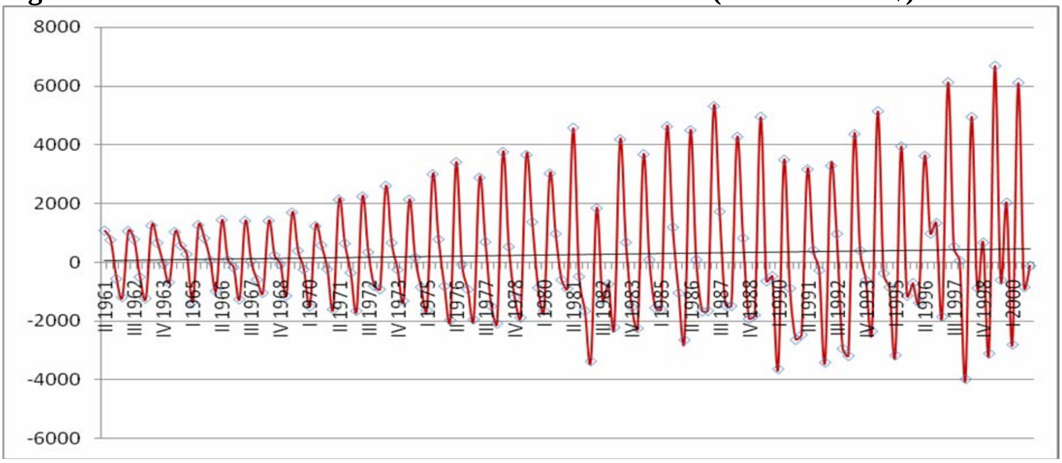
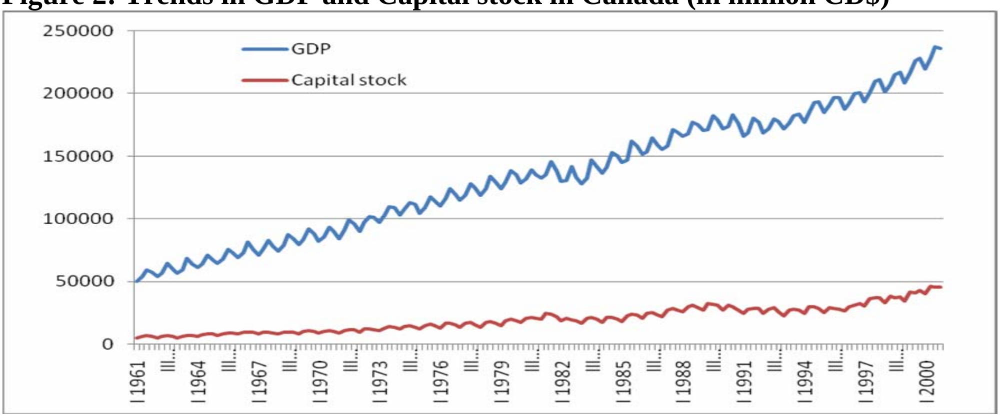


Table 1: Descriptive Statistics for Key Variables (in million CD$)

Time period	GDP	S.D (GDP)	Investment (I)	S.D (I)	Capital Stock (K)	S.D (K)
1961-1970	72495.82	11300.71	176.36	1300.61	10592.23	1962.95
1971-1980	114707.45	14390.64	261.53	2196.21	17536.90	3244.56
1981-1990	155196.68	16757.47	211.63	3006.46	27640.33	4978.39
1991-2000	196561.90	19704.60	493.98	3343.22	37171.68	6917.33

Figure 1 displays the net investment behaviour in Canada over the different quarters from 1961 to 2000. As shown in Figure the net investment is more volatile and showing oscillating pattern.

Figure 2: Trends in GDP and Capital stock in Canada (in million CD$)



The time series of GDP and capital stocks are illustrated in Figure 2 and their autocorrelation functions are plotted in the appendix. Both the GDP and capital formation series show upward trend; but time series of capital shows more or less volatile. Visual inspection indicates that the average value of *I(t)* is very different from the increase of the stock of capital. It might be surprising to see that the stock of capital increases during the decades 1981-90 and 1991-2000 are very alike, while the average values of it are very different. However this is true since the capital stock in 1981-90 were more volatile and there was a big reduction in capital stock in the early and late 1980. Since investment is a flow measure, the difference between *K(t)* and *K(t-1)*, the reduction in capital stock in 1981-90 did not affect much to its average value while that reduction affected mean value of investment. On the other hand, there has been a secular increase in capital stock from the mid-1990s and that contributed to increase in the average value of investment.

An examination of their autocorrelation function (see Appendix 1) shows that, for GDP and capital formation series, the autocorrelation functions decline as the number of lags becomes large, but only very slowly. The autocorrelation function for the investment series spike after each four lags.

4. Estimation and Identification of the Model

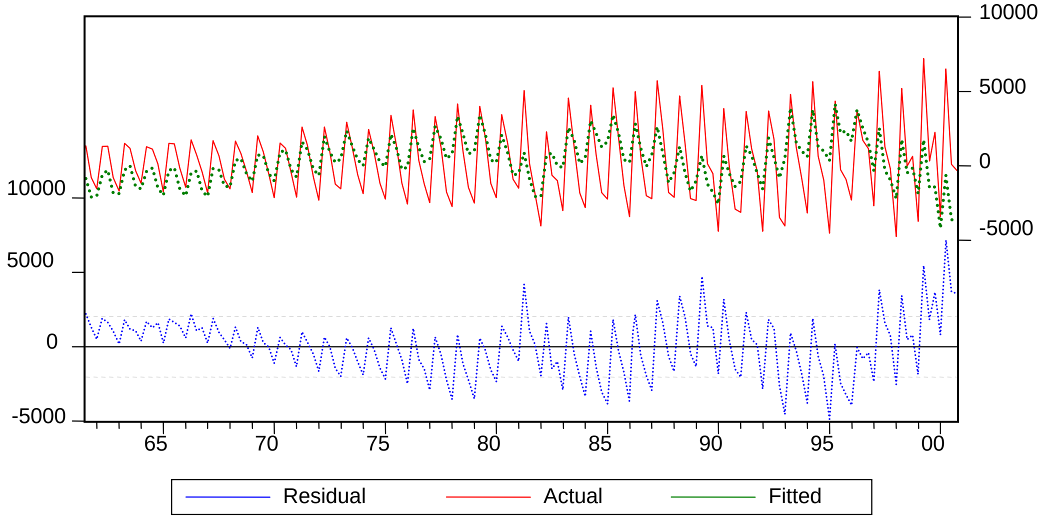
4.1. Estimation of the Structural form. I first estimate the flexible accelerator model of investment using ordinary least squares (OLS) and obtain the following result (standard errors are in parenthesis)-

$$\underset{}{I_t} = \underset{(661.8)}{-4649.21} + \underset{(0.015)}{0.148Q_t} - \underset{(0.068)}{0.655K_{t-1}}; \qquad R^2=0.378 \quad DW=1.58 \quad F=47.08 \qquad (6)$$

The R^2 value indicates that a considerable variation of I_t can not be explained by Q_t and K_{t-1} i.e. the FAM given in equation (1) is not enough to account for investment. The Durbin-Watson statistic greater than R^2 indicates that the estimated regression doesn't suffer from spurious regression (Granger and Newbold 1974). The actual and fitted investment series and residuals are plotted in Figure 3.

It can be seen from Figure 3 that there is a large variation between actual and estimated value of investment. The residuals appear to have a high degree of positive autocorrelation, which is consistent with the Durbin-Watson statistic. The residual plot also shows that there is residual seasonality and there may be cycle in the residual, an impression confirmed by residual correlogram. The residual sample autocorrelation function has large spikes, far exceeding the Bartletts bands, at the seasonal displacement, 4,8,12 and so on. The Augmented Engle-Granger (AEG) test (for three lags as determined by Akaike Information Criteria (AIC) and Schwartz information criteria (SIC) on the differenced series, e_t, rejects the hypothesis of unit root. The AEG test statistic is –7.74, which is significant at the 1% level.

Figure 3: Actual, fitted and Residual from the Regression



4.2. Multivariate ARIMA Specification

I first determine the order of AR and MA process using the estimated residuals from the structural model. Then I determine the order of ARMA process (see Appendix 2 for details).[3] After determining possible orders of ARMA, I specify an ARIMA for the estimated residuals and then estimate the combined regression-time-series model. This model is likely to provide better forecast than the regression equation alone or time series model alone since it includes structural (economic) explanation of that part of the variance of Y_t that can be explained structurally, and a time-series explanation of that part of the variance of Y_t that cannot be explained structurally. Equation (5) is referred to as a *transfer function model* or alternatively, a multivariate autoregressive-moving average (MARMA) model.

Now I combine the structural part and time series part of the model and reestimate all the parameter simultaneously. That is, I estimate the parameters of the following model:

$I_t = \beta_0 + \beta_1 Q_t + \beta_2 K_{t-1} + \theta^{-1}(B)\, \phi(B)\, \varepsilon_t$

I specify an ARIMA (4, 0, 3) in Appendix 2, which is equivalent to writing as-

$$\theta(B)e_t = (1-\theta_1 B-\theta_2 B^2-\theta_3 B^3-\theta_4 B^4)\, e_t \qquad (7)$$

So $\phi(B) = (1- \phi_1 B- \phi_2 B^2 - \theta_3 B^3)$

As shown in Appendix 2, in terms of our original residual series u_t, there is a fourth differencing in the residual series of u_t, so equation (7) is written as -

[3] Available at the journal web site: http://www.usc.es/econometaeid.htm

$\theta(B)e_t = (1-\theta_1 B-\theta_2 B^2-\theta_3 B^3-\theta_4 B^4)(1-B^4)\,u_t$

The error process is equivalent to-

$\psi(B)\,u_t = (1-\psi_1 B-\psi_2 B^2-\ldots-\psi_8 B^8)u_t$

So, in this case I have ARIMA(8,0,3) model of u_t which is equivalent to saying ARIMA(4,0,3) of e_t. I now estimate the MARIMA or ARIMAX model.

Table 2 shows that MARIMA (8,0,5) has a chi-square value of 37.99, and it is statistically significant at the 95 percent confidence level and so it does not pass the *diagnostic test*. MARIMA (8,0,3) has the chi-square of 28.79 and so it is insignificant with (4,25) degrees of freedom and, therefore, we can accept the hypothesis that residuals are white noise. We see that the MARIMA (8,0,3) is more promising because it has the lowest chi-square statistic. Furthermore it has the smaller AIC and SIC value compared to ARIMA (8,0,5). So we can choose MARIMA (8,0,3) by all the available criteria.

Table 2: Test statistic value for MARIMA

	(8,0,3)	(8,0,5)
Chi-square	28.79	37.99
Akaike info criterion	16.03	16.07
Schwarz criterion	16.32	16.42

The results of the Multivariate ARIMA (8, 0, 3) estimation are as follows:

$$I_t = \underset{(677.42)}{-5631.08} - \underset{(.070)}{0.7167}{*}K_{t-1} + \underset{(0.0159)}{0.16587}Q_t + \{(1-0.5285B-1225B^2-0.5685B^3)/(1+0.0805B$$
$$+\ 0.4572B^2-0.1332B^3+0.6664B^4-0.0889B^5-0.4401B^6+0.1155B^7+0.2836B^8)\}\ \varepsilon_t$$

$R^2=0.935$ DW=2.018 F=149.86

Note that R^2 is very higher, and the DW is very close to 2. The sample autocorrelation for the residuals of this equation are all very small, so that the residuals appear to be white noise (see Appendix 1, Figure A4). The histogram of residual looks symmetric as confirmed by the skewness near zero. The residual kurtosis is higher than 3, but does not reject the hypothesis of normal distribution at the 5% or 10% level.

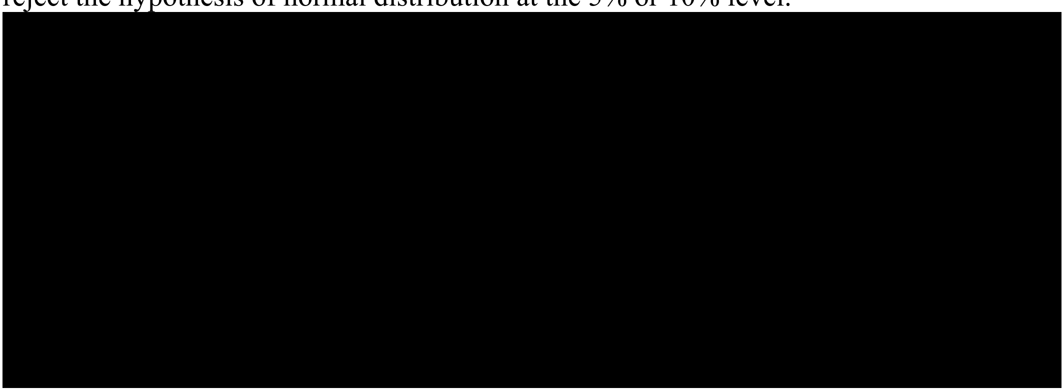


Figure 4 and Figure 5 show the residual sample autocorrelation and partial autocorrelation, which displays no patterns, and mostly inside the Bartletts bands (not shown in figure).

Specification Test:

I carry out the specification test through most appropriate and widely used two test statistic, Wald Test and Lagrange Multiplier test. The Wald test rejects the hypothesis that all the AR or MA or ARMA terms are zero. The values of the test statistic which are distributed as chi-square with 1 degree of freedom are as below:

Table-3: Value of Wald test statistic

	Chi-square	P-value
AR	2015.86	0.0000
MA	2290.43	0.0000
ARMA	62.849	0.0000

To see if there is any possibility that the errors exhibit autocorrelation, I conduct LM test. The LM test rejects the hypothesis of serial correlation.

5. Forecasting:

I now investigate whether the model forecast well. For this I reestimate the equation by changing the sample to 1961:2-1996:4 and forecast for the period from 1997:1 to 2000:4. I choose here two types of forecast; dynamic forecasts, which calculates *multi-step* forecasts starting from the first period in the forecast sample. That is forecast for the 1997:1 uses the forecasted value for 1996:4, and so on for the later periods. After that I make the static forecasts, which calculate a sequence of *one-step-ahead* forecasts, using actual, rather than forecasted values foe lagged dependent variable. That is, static forecast uses the actual value of investment in earlier period, rather than the earlier forecasts.

[redacted]

In dynamic forecasting, I start at the beginning of the forecast sample, and compute a complete set of *n*-period ahead forecasts for each period in the forecast interval. Thus, I start at period t and forecast dynamically to $t+n$, and compute a one-step ahead forecast for $t+1$, a two-step ahead forecast for $t+2$, and so forth, up to an *n*-step ahead forecast for $t+n$. It may be useful to note that as with n-step ahead forecasting, I simply initialize a Kalman filter at time $t+1$ and run the filter forward additional periods using no additional signal information. For dynamic forecasting, however, only one *n*-step ahead forecast is required to compute all of the forecast values since the information set is not updated from the beginning of the forecast period.

[redacted]

The static forecasts are the forecast that one would have made in real life if one uses this equation every quarter since 1996:4. Not surprisingly, the static forecasts are a lot more accurate than are the dynamic forecasts. The forecasted actual series for investment are shown in Figure 6 and Figure 7 Our model has generated forecast that are more accurate in case of static forecast. It picks up the broad seasonal cycle. This model may be acceptable as a forecasting tool.

6. Conclusion

In this paper I model the Canadian investment over a period of more than 40 years and found that the popularly known flexible accelerator model cannot predict the investment behavior of the Canadian economy. So, I model the investment through multivariavte approach of autoregressive and moving average model of time series econometrics. The results of the multivariate model shows that one can better count the investment by combining structural flexible accelerator model with that of time series ARIMA model and the model forcasts very well. So, the empirical evidence suggests that a model builder should go for the ARIMA specification of the investment when the structural model cannot properly account the time series behavior of the data.

References

Box,G.E.P., and Genkins(1970), " Time Series Analysis: Forecasting and Control" San Francisco: Holden-Day

Chenery, H. B (1952), "Overcapacity and Acceleration Principle" *Econometrica*, Vol.20 No.1, p.1-28

Cramer,R.H and R.B. Miller(1976)," Dynamic Modeling of Multivariate Time Series in Bank analysis", *Journal of Money, Credit and Banking*, Vol.8, No.1 February p.85-96

Darling and M. Lovell (1971), "Inventories, Production Smoothing, and Flexible Accelerator", *Quarterly Journal of Economics*, Vol.85, No.2, May

Engle,R.f and C.W.J. Granger(1987), "Cointegration and Error correction: Representation, Estimation and Testing," *Econometrica*, Vol.55, p.251-276

Epstein, L.G and M. Denny, (1983), "The Multivariate Flexible Accelerator Model: Its Empirical Restriction and Application to U.S Manufacturing", *Econometrica*, Vol.51, No.3, May, p.647-674

Fisher, M.E and J. Seater (1993), "Long-run Neutrality and Superneutrality in an ARIMA framework", *The American Economic Review*, Vol.83, No.3 p. 402-415

Granger,C.W.J and P. Newbold(1974), " Spurious Regression in Econometrics," *Journal of Econometrics*, Vol.2,p.111-120

Hay, G.A (1970), "Adjustment Cost and Flexible Accelerator", *Quarterly Journal of Economics*, Vol. 84, No.1, February, p-140-143

Joregensen,D.W and J. Stephenson(1969), " Issues in the Development of Neoclassical Theory of Investment Behavior" *Review of Economics and Statistics*, Vol.51,p.346-353

Joregensen, D. W (1971), "Econometric Studies of Investment Behavior: A survey", *Journal of Economic Literature*, Vol.9, No.4, December, P.1111-1147

Koyck, L.M (1952), "Distributed Lag and Investment Analysis" North Holland, Amsterdam

Ljung,G and G. Box(1979), " On a Measure Of Lack Of Fit in Time Series Models," *Biometrika*, Vol. 66, p.265-270

Lucas, R.E (1967), "Optimal Investment Policy and The Flexible Accelerator", *International Economic Review*, Vol.8, No.1, February.

Pindyck, R.S and D.L Rubinfeld (1991), "Econometric Models and Economic forecasts", McGraw-Hill

Stockton, D.J. and J.E. Glassman(1987), "An Evaluation of the Forecast Performance of Alternative Models of Inflation" *The Review of Economics and Statistics*, Vol. 69, No. 1. February, p. 108-117

Tevlin, S and K. Whelan (2000), "Explaining Investment Boom in the 1990s", Federal reserve Bank, New York, Working Paper.

Appendix on line at the journal web site

Journal published by the EAAEDS: http://www.usc.es/economet/eaa.htm

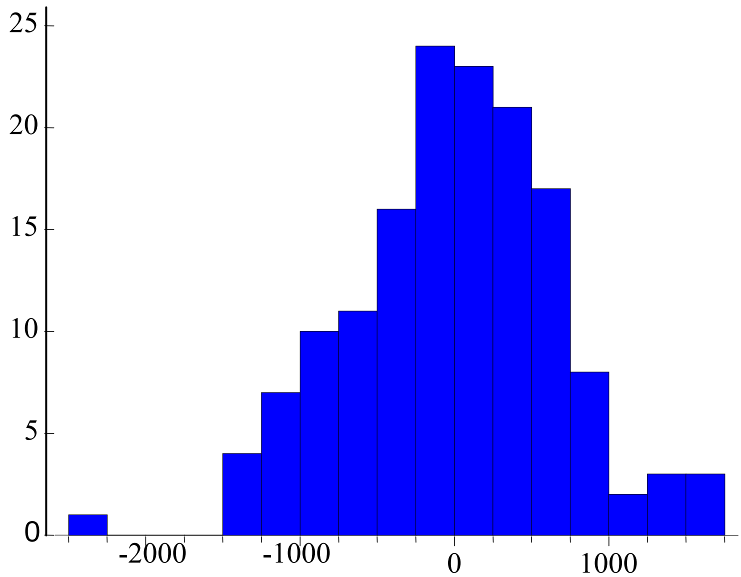
Appendix 1

[redacted]

[redacted]

[redacted]

Figure- A4: Histogram for Residual of ARMAX model



Appendix 2

This appendix discusses the stationary property of the estimated residuals from the flexible accelerator model. It then specifies an ARIMA model for the residual series.

Stationarity of the Residuals

Here I first test unit root to determine the stationarity property of the estimated residual. The appropriate test statistic is the Augmented Engle-Granger (AGE)[4] test. The value of the AGE statistic is 0.579. The Engel-Granger 1%, 5% and 10% critical values of the τ statistic (t-statistic) are, respectively, -2.59, -1.94, and –1.62.
In the above instance, the AEG test is in the above case is carried out by the following procedure:

$\Delta u_t = \mu + \gamma u_{t-1} + \delta_1 \Delta u_{t-1} + \delta_2 \Delta u_{t-2} + \delta_3 \Delta u_{t-3} + \delta_4 t + \xi_t$ ....(7)
Or in general-
$\Delta u_t = \mu + \gamma u_{t-1} + \sum^{k}_{j=1} \delta_j \Delta u_{t-j} + \xi_t$
where, k is the number of lag
μ, γ and δ are parameters,

[4] Since the estimated residual is based on the estimated parameters, the Augmented Dickey-Fuller (ADF) test critical significance values are not quite appropriate. Engle and Granger (1987) have calculated these values and ADF test in our context is known as AEG test.

ξ_t is assumed to be white noise and $\gamma = \rho\text{-}1$.

The null and alternative hypothesis are- H_o: $\gamma = 0$, H_1: $\gamma < 0$. The order of the lag of AEG test has been chosen by the AIC and SIC criteria, i.e. I choose the minimum AIC and SIC value for the AEG test (in all cases).

From Figure A5 it is obvious that there are seasonal cycles in the residuals for every four-lag. To make the residual series stationary, we take a four-quarter difference and obtain the following-

$$e_t = u_t - u_{t-4} = (1-B^4)u_t$$

where B is the backward shift lag operator. As can be seen in figure-4, the sample autocorrelation function for this 4-quarter differenced series doesn't exhibit seasonality. The autocorrelation function for e_t falls down to zero after four lags and it also shows that the e_t is stationary. The AEG test (for three lags as determined by AIC and SIC) on the differenced series e_t rejects the hypothesis of unit root. The AGE test statistic is –7.74, which is significant at the 1% level.

Determining the order of ARMA

At first I try to determine the order of AR(p) process. The partial auto correlation function jumps down to zero after three lags indicate the order of the autocorrelation may be three or four. AIC suggest the order of AR to be five while SIC suggest the corresponding AR order should be four (Table A1). At this stage I retain the both possibilities and try to estimate MA(q) models. The sample autocorrelation function drops down to zero after four lags is an indicative of the order of MA. Here AIC suggest that the order of MA should be five, in contrast, SIC indicate that the order should be three. So, I keep both possibilities to investigate the order of ARMA.

Table A1: AIC and SIC value for AR and MA

	AR		MA	
Order	AIC	SIC	AIC	SIC
5	16.119	16.241	16.031	16.130
4	16.137	16.237	16.067	16.146
3	16.365	16.445	16.059	16.118
2	16.368	16.428	16.301	16.341
1	16.376	16.415	16.507	16.527

I now turn to the ARIMA specification of the differenced residual series e_t. I employ portmanteau test (Q-test), in addition to AIC and SIC, to diagnose the inadequacy of fitted ARIMA model. The portmanteau statistic that is frequently used to test H_0 is- $Q = T \sum_{i=1}^{K} \hat{h}_i^2$, where $\hat{h}_i$ is an estimate of the i^{th} order autocorrelation of the residuals from ARIMA or Multivariate ARIMA and K is constant typically chosen to be 36. Box and Pierce (1970) demonstrated that Q is asymptotically distributed as chi-square with *K-p-q* degrees of freedom[5]. The result is based on the two important properties of $\hat{h}_i$: (1) the distribution of vector $(\hat{h}_1 \ldots \hat{h}_k)$ is asymptotically nonsingular normal $(0, T^{-1}\omega)$ and (2) $(\hat{h}_1 \ldots \hat{h}_k) = (h_1 \ldots h_k)\omega$, where h_i is the i^{th} order autocorrelation of true disturbances and the $(k \times k)$ covariance matrix ω is idempotent of rank (K-p-q)[6]. However, Ljung and Box (1978) propose the modified Q statistic, $Q = T(T+2)\sum_{i=1}^{K}(T-i)^{-1} \hat{h}_i^2$, which has been shown to have better finite sample properties than the Box-Pierce test in the context of ARIMA model.

The previous specification of AR and MA terms reflect that the e_t can be modeled as ARIMA(4,0,3) or ARIMA(4,0,5) or ARIMA(5,0,3) or ARIMA(5,0,5). We begin with ARIMA (5, 0, 3) model. The result is-

ARIMA(5,0,3):

$$(1-.892B+.404B^2+.0594B^3-.297B^4-.139B^5)\, e_t=(1-.606B-.318B^2-.465B^3)\, \varepsilon_t \quad \ldots\ldots(8)$$

AIC=16.07 *SIC=16.23* $\chi^2(8,36)= 30.72$

With 28 degrees of freedom, the chi-square value is below the critical 95 percent level. Thus one can conclude that residuals from ARIMA (5, 0, 3) model are not autocorrelated. But one needs to consider other ARIMA order term and choose on the basis of lowest value of the statistic e.g., AIC, SIC and χ^2.

As a next step I increase the number of MA terms and estimate ARIMA (5, 0, 5) model. The result is-
ARIMA (5, 0, 5):

$$(1+.729B-.360B^2-.195B^3-.148B^4-.0193B^5)e_t=(1-.258B-.695B^2-.588B^3-.001B^4-.660B^5)\varepsilon_t \quad \ldots\ldots\ldots\ (9)$$

---

[5] If there is *original data series* then the Q statistic is chi-square with *K* df. Since the residuals from ARIMA model are estimated, Q will be chi-square with *K-p-q* df. For the original data series, p=q=0.

6 These properties, which hold true for AR and ARMA series, do not necessarily hold true for the dynamic linear models. For instance, Dezhbakhsh (1990) argue that portmanteau (Box-Pierce or Ljung- Box) tests are inappropriately applied to linear models with lagged dependent variable and exogenous regressors.

AIC=16.02 SIC=16.22 $\chi^2(10,36)= 24.996$

The chi-square statistic has dropped, and it is also insignificant, leading to accept this specification too.

I now try specifications that are lower order, beginning with ARIMA (4, 0, 3):
ARIMA(4,0,3):

$$(1+.129B+.397B^2-.08B^3-.338B^4)\, e_t=(1-.387B-.201 B^2-.463B^3)\, \varepsilon_t \tag{9}$$

AIC= 16.05 *SIC=16.19* $\chi^2(7,36)= 29.726$

The chi-square statistic has increased, however, there are now 29 degrees of freedom and this value is insignificant at the 95 percent level. Furthermore, the value of SIC decreases compared to earlier two specifications. So, we can also select ARIMA (4, 0, 3).

In the next step, I increase the number of moving average terms, and estimate an ARIMA (4,0,5) model, and the results are shown below:
ARIMA (4,0,5):

$$(1+.791B-.351B^2-.252B^3-.118B^4)\, e_t=(1-.320B-.657 B^2-.594B^3-.048B^4-..640B^5)\, \varepsilon_t \tag{10}$$

AIC= 15.99 SIC=16.17 $\chi^2(9, 36) = 22.89$

Here the magnitudes of the three statistics have reduced, but the degree of freedom here is lower compared to ARIMA (4, 0, 3). The chi-square statistic is 22.89, which, with 27 degrees of freedom, is insignificant.

All of these chi-square statistics are insignificant even at the 90 percent level, allowing me in each case to accept the hypothesis that the residuals are white noise. It is, however, clear a low-order ARIMA model can describe the differenced residual of the model. Of these four ARIMA (p.d.q), ARIMA (4,0,3) and ARIMA (4,0,5) models look promising. But ARIMA(4,0,3) has the highest degrees of freedom and this should also be taken into account.