BOX-JENKINS TRANSFER FUNCTION FRAMEWORK APPLIED TO SAVING-INVESTMENT NEXUS IN THE SOUTH AFRICAN CONTEXT

Ntebogang Dinah Moroke*

Abstract

This paper studied the relationship between investment and savings in South Africa for the period 1990 quarter 1 to 2014 quarter 3. The unit root test confirmed the non-stationarity of the series prior to first differencing. The correlation coefficient and the model assessing a full capacity mobility hypothesis were significant and passed all the diagnostic examinations. The estimated parameter provided evidence of imperfect capital mobility. ARIMAX (5, 1, 0) out-performed all the five models and was used for pre-whitening process. This model was later used to produce a two year forecasts of investment. The error forecast measure provided enough evidence to conclude that ARIMAX (5, 1, 0) provided valid forecasts. These results are recommended when embarking on future saving-investment plans in South Africa.

Key Words: Box-Jenkins Transfer Function, Full Capacity Mobility Hypothesis Saving, Investment, South African

* North West University, South Africa

1. Introduction

Using a simple linear regresssion for model estimation could give misleading results about the relationship between the output and input variables. Possible problems to this may involve for example, (1) feedback from the output to the input series, (2) ommitted time-lagged input term, (4) an autocorrelated disturbance series and (4) common autocorrelation patterns shared by the variables that could produce spurious correlations.

This paper tries to set out a practical model identification procedure designed to handle the four problems mentioned. The details for model identifcation used in this study differ from those proposed by Box and Jenkins (1976) but the idea is similar. The authors extensively studied the autoregressive moving average (ARMA) models. They effectively put together, in a comprehensive manner relevant information required to understand the use of univariate time series ARMA models. These models are specifically effective when the purpose of the study is forecasting of an output variable on the basis of its past values. Only one variable is needed to achieve the task in this instance.

Key objective of this investigation is to build a model that extends the Box-Jenkins ARMA framework. This model is intended to show the relationship between the output $\{Y_t\}t \in z$ and the input time series $\{X_t\}t \in z$ implementing the Box-Jenkins transfer function framework. The reseacher hopes that any model chosen might be interesting in its own right but the interest will also be in using it to forecast savings from future values of investment. Priestley (1981) vouch for this framwork especially when both time series fall into the class of ARMA type models. The said framework combines characteristics of both the univariate process and those of simple linear regression analysis. Thus, the framework combines time series approach with the causal approach to perform forecasting.

Makridakis, Wheelwright and Hyndman (1998) prescribe the use of this framework as a powerful tool when appropriate conditions for its use exist. The transfer function framework generalise the three phases of Box-Jenkins univariate procedure such as the identification, estimation and cheking, and application. To extend this model, an input variable is introduced as an exogenous variable resulting to a bivariate framework simply known as the ARIMAX. The thrust of this paper is the identification and application with an investment as an exogenous variable to model and forecast savings in South Africa. Both the autocorrelation and cross correlation structures of the input and output variables are exploited in the identification stage.

The saving-investment investigation is relevant for different countries as it may hold the key to the positive correlation between saving and economic growth. Additionally, if capital accumulation is deemed the engine of growth, understanding the interaction between saving investment is crucial to assessing the validity of the traditional recipe that raising saving is the surest way to increase growth.

This is a notion that implicitly require each country's extra saving to be automatically translated into higher domestic investment (Wahid et al., 2008). Conventional thinking holds that saving is an essential element in promoting investment, therefore economic growth. Esso and Keho (2010) emphasise that low levels of domestic saving in some developing countries condemn them to an uncomfortable choice between low investment and growth or excessive reliance upon foreign capital which makes them vulnerable.

The application of Box-Jenkins transfer function framework in this study to saving-investment relationship is advantageous in that trends in these two series will be viewed for better understanding and decision making. The framework will also help in constructing and selecting the model that reasonably describe the nexus between these variables. The best model among competing models will be chosen on the basis of the associated error forecasts. The framework is recommended especially when one or more input variables are incorporated in the analysis and longtime forecasts are to be produced. This duty cannot be performed with the traditional Box-Jenkins univariate framework. Furthermore, the Box-Jenkins transfer function framework allows the researcher to make a reflection about the data based on its past, present and future state (Moroke, 2005).

Based on the findings of this study, better suggestions may be formulated and submitted to policy makers. It is almost impossible for policy makers to come up with feasible plans about saving-investment when proper and reliable information about these variables is not available. When valid information is provided, appropriate strategies may be articulated and devised. Moroke (2014) emphasise on the application of proper modelling techniques so as to achieve improved forecasting accuracy. The application of Box-Jenkins transfer function framework has not been exhausted; as a result this paper explores its effectiveness and validity. This analysis could further contribute to existing body of knowledge and a better understanding of saving-investment association in middle income developing countries with relatively poor social indicators like South Africa may be gained. Policy decisions aimed at promoting economic growth and development in the country may be attained if saving-investment process is understood.

There is no evidence in literature about studies that applied the Box-Jenkins transfer function framework to savings-investment nexus. Nonetheless, the framework has been employed by researchers on numerous fields. Among others, Bambang et al., (2009) used the framework to build transfer function model for rainfall index data in Indonesia. The study compared the forecast accuracy among ARIMA, ASTAR, Single input transfer function, and multiple input transfer function models. Another study by

Khin et al., (2011) proposed transfer function model to predict electricity prices based on both past electricity prices and demand. The rational to build this model was also discussed in this study. Arumugam and Anithakumari (2013) used the Transfer function model to fit the model of natural rubber production where sales were used as an influential variable. Eni et al., (2013) adopted the framework to develop a three input Transfer function to forecast rainfall in Calabar.

2. Data

The study uses quarterly data retrieved from the South African Reserve Bank covering the period 1990 Q1 to 2014 Q3. A total of 98 observations are used for the analysis. The choice of this period will allow us to track trends of saving-investment before, during and after democracy of South Africa. The number of observations used will help in safeguarding the normality assumption. Two variables considered are ratio of gross savings to gross domestic product measured in percentages and gross fixed capital formation in millions. These variables are subjected to log transformations in order to take care of irregularities borne as a result of non-constant time series stochastic terms such as the mean (μ_t) and variances (σ^2).

Sadowski (2010) suggested log transformation as the optimal variance stabilizing factor specifically when standard deviation of the original time series increases in a linear fashion with the series mean. This transformation as highlighted by Montgomery et al. (2008) does well to physically interpret the variance percentage change. This can be shown in a time series with $y_i, \ldots y_t,$ where the interest is more in the percentage change in y_T, say,

$$x_t \frac{100(y_t - y_{t-1})}{y_{t-1}} \qquad (1)$$

The approximate percentage change in y_t can be calculated from the differences of the log-transformed series $x_t \cong 100[log(y_t) - log(y_{t-1})]$. The Statistical Analysis Software (SAS) version 9.3 is used for data analysis. It is reasonable to assume a unidirectional causal relationship between the variables used in transfer function framework, *i.e.*, past values of savings (X_t) influence future values of investment (Y_t), not vice versa. Otherwise, if a bidirectional relationship is evidenced, the use of this framework becomes futile. The study is interested in using the following model to assess a full capacity mobility hypothesis:

$$\frac{Investment}{GDP} = \beta_0 + \beta_1 \left(\frac{Savings}{GDP}\right)_t + \varepsilon_t, \qquad (2)$$

where β_0 and β_1 are the estimated coefficients using the ordinary least squares regression method. In

particular, β_1 is the savings-investment retention coefficient and ε_t is the error term.

3. Methodology

A transfer function model differs from the ARMA model in that the latter is univariate and the former can accommodate multivariate time series. No relationships can be determined with ARMA model as only one variable is involved in the analysis. The transfer function model overrides ARMA models in that both the output and the input variables are used in model building. This study uses a single input (investment) to construct a forecasting model. Assume the two series X_t and Y_t have properly been transformed so that both they are stationary. In a single input, single output linear system, the output series Y_t and the input series X_t are related through a linear filter as;

$$Y_t = v(B)X_t + \eta_t, \qquad (3)$$

where $v(B) = \sum_{j=-\infty}^{\infty} v_j B^j$ is referred to as the transfer function filter according to Box and Jenkins (1976), v_j is an impulse response weights and B is a coefficient of X_t, η_t is the noise series of the system that is independent of the input series X_t. When X_t and η_t are assumed to follow some ARMA models, the transfer function model [3] is known as ARIMAX.

The coefficients in the transfer function model are the impulse response weights. Transfer function models are said to be stable if the sequence of impulse response weights is absolutely summable (Wei, 2006), i.e., $\sum |v| < \infty$. In a stable system, a bounded input always produces a bounded output. A transfer function modelling serves the purpose of identifying and estimating the transfer function $v(B)$ and a noise model η_t based on the available information of the input series, X_t and the output series, Y_t. One of the difficulties associated with this model is that the information on the two series is finite and this causes the transfer function to contain an infinite number of coefficients. Such difficulties can be alleviated by representing the transfer function as;

$$v(B) = \frac{\omega_s(B)B^b}{\delta_r(B)}. \qquad (4)$$

Operators $\omega_s(B) = \omega_0 - \omega_1 B - \cdots - \omega_s B^s$ and $\delta_1(B) = 1 - \delta_1 B - \cdots - \delta_r B^r$ are polynomials in B, where B is a parameter estimating the delay between the variables. The aim of this study is to extend the univariate ARMA process;

$$\emptyset_x(B)X_t = \theta_x(B)\alpha_t, \qquad (5)$$

where α_t follows a white nose process.

3.1 Identification of the model

The first step in the identification of the transfer function model is to pre-whiten the input X_t and the output Y_t series. The pre-whitening filter may be applied to both series. It helps in removing the corrupting influence of the autocorrelation within the input series while maintaining the same functional relationship between the two series (Yaffee and MacGee, 2000). Instead of solving for X_t, the equation is inverted to solve for ε_t. The pre-whitening filter is formulated from the existing ARMA model.

Pre-whitening of the series is preceded by differencing to achieve stationarity as;

$$\omega_t = (1 - B^d)X_t, \qquad (6)$$

and

$$Z_t = (1 - B^d)Y_t \qquad (7)$$

The transfer function analysis is not fussy about the degree of differencing. This need not be the same for the X_t and Y_t. The stationarity of the series is tested with the Augmented Dickey-Fuller (ADF) unit root test estimated from;

$$\Delta Y_t = \alpha_0 + \beta_0 Y_{t-1} + \sum_{i=1}^{k} \beta_i \Delta Y_{t-i} + \varepsilon_t. \qquad (8)$$

The symbol ∇ is the first difference operator; t is the time drift; k represents the number of lags used and ε_t is the error term; $\alpha's$ and $\beta's$ are the model bounds. The ADF test may include a constant and time trend depending on the analysis. Assuming that the series, $\{Y_t t^T - 1\}$ follows the AR (p) process, Hamilton (1994) shows that the rejection or acceptance of the null hypothesis of a unit root is based on running the regression;

$$Z_t = \mu + (\phi_1 - 1)Y_{t-1} + \sum_{j=1}^{p-1} C_j \Delta Z_{t-j} + \varepsilon_t, \qquad (9)$$

where $Z_{t-1} = Y_{t-1} - Y_{t-j-1}$ for $j = 0, 1, 2, \ldots, p - 1$ and ε_t is a white noise process. The ADF test statistic is given as;

$$\tau_{ADF} = \frac{\phi_1 - 1}{se(\hat{\phi}_1)}, \qquad (10)$$

where $t_{\phi_1 - 1}$ is the test statistic of $\phi_1 - 1$, $se(\hat{\phi}_1)$ is the standard error of $\phi_1 - 1$. The null hypothesis of a unit root $H_0: \phi_1 - 1$ is rejected if [10] is less than the appropriate critical value at some level of significance. Alternatively the test statistic rejects the null hypothesis if the corresponding probability value exceeds the level of significance.

Once stationarity conditions have been addressed, a tentative model is identified by examining the autocorrelations calculated from the autocorrelation (ACF) and partial autocorrelation function (PACF). The observed ACFs and PACFs are compared with the theoretical characteristics found in Box and Jenkins (1976) to determine the orders of $ARMA\ (p, q)$.

The filtered series are obtained from differenced series as;

$$\alpha_t = \frac{1 - \emptyset_x(B)}{1 - \theta_x(B)} X_t ,\tag{11}$$

and

$$\beta_t = \frac{1 - \emptyset_x(B)}{1 - \theta_x(B)} Y_t, \tag{12}$$

respectively.

In the second step of this framework, the sample cross correlation function (CCF) $\hat{\rho}_{\alpha\beta}(k)$ between [11] and [12] is calculated to estimate v_k as;

$$\hat{v}_k = \hat{\rho}_{\alpha\beta}(k) \frac{\hat{\sigma}_\beta}{\hat{\sigma}_\alpha}. \tag{13}$$

The significance of the CCF is assessed by its comparison with the associated standard error $(n - k)^{-1/2}$.

3.2 Estimation of the model

The general preliminary transfer function model is of the form;

$$Y_t = \frac{\hat{\omega}_s(B)}{\hat{\delta}_r(B)} B^b X_t + \hat{\eta}_t. \tag{14}$$

However, this section requires an estimation of the noise model as;

$$\hat{\eta}_t = Y_t - \hat{v}(B)X_t = Y_t - \frac{\hat{\omega}_s(B)}{\hat{\delta}_r(B)} B^b X_t. \tag{15}$$

The appropriate model for the noise is identified by observing the associated ACF and PACF. Later the transfer function and the noise models are combined to produce the function;

$$\alpha_t = \frac{\omega_s(B)}{\delta_r(B)} X_{t-b} + \frac{\theta_x(B)}{\emptyset_x(B)} \eta_t. \tag{16}$$

The parameters $\beta' = (\omega', \delta', \theta', \emptyset') = (\omega_0,\ \omega_1, \dots, \omega_s,\ \delta_1, \dots, \delta_r;\ \theta_1, \dots, \theta_q,\ \emptyset_1, \dots,\ \emptyset_p)$ in [16] are estimated from the past data (X_t, Y_t), $t = 1, 2, \dots, n$.

A conditional least squares method is used to estimate the parameters of both the noise and the final transfer function models. The following formulas describe a conditional least squares;

$$\hat{N}_t(k) = E(N_{t+k}|N_t, N_{t-1}, \dots), \tag{17}$$

assuming $N_t = 0$ for $t < 0$,

$$\sigma_t^2(k) = \sigma^2 \sum_{j=0}^{k-1} \psi_j^2, \tag{18}$$

where ψ_j are coefficients of the power series expansion of $\frac{MA}{\Delta x AR}$.

The results of the estimated parameters are compared for optimal goodness-of-fit. Yaffee and MacGee (2000) recommend the ultimate reasonable magnitude of parameter estimates and statistically significant associated $t - \text{ratios} > 1.96$. The authors further recommend that non-significant parameters should be trimmed from the model.

3.3 Diagnostic checking

A battery of diagnostic tests is used to validate model assumptions and to also evaluate its adequacy. These tests involve the residuals of the model. One of the basic assumptions is that the residuals follow a white noise process. Another assumption concerns the residual correlations. To check the two aforementioned residual assumptions, the associated residual estimates are needed. The error term estimates are automatically calculated at the estimation stage along with the conditional least squares estimation for model parameters.

Model adequacy: This assumption concerns the examination of the error terms. For the assumption to hold, the standardised residuals should look random in an ACF. If the error terms are statistically different from zero, the model is not adequate. The adequacy of the transfer function models is dependent on the significant estimated parameters. Yaffee and MacGee (2000) advise that the decay parameters should conform to the bounds of stability for transfer function models. The authors further advise that if the model is one of first-order decay, then parameter estimates should not be too close to the value of 1.00. Another recommended guideline by these authors is to have parameters not too close to 0.96, otherwise the model may be unstable and further differencing may also be needed. A model with insignificant parameters may require pruning of these parameters.

Consequently, the Ljung-Box (1978) test may be used. This test is defined as;

$$Q = n(n + 2) \sum_{t=1}^{k} \frac{\rho_\eta^2(k)}{n - t}. \tag{19}$$

The statistic is asymptotically distributed as χ^2 with m degrees of freedom (Vogelvang, 2005), where

ρ is called the sample ACF, n is the length of time series, k is the highest order for autocorrelation for which to test and ρ_j^2 is the jth autocorrelation. If the observed value exceeds the critical value, the noise model is considered inadequate

Autocorrelation check: The study uses the Durbin-Watson (DW) as a measure of residual autocorrelation. The DW has been successfully used in the past to test for first-order autocorrelation of the regression model residuals. The test compares the residuals for time period t with the residuals from time period $k-1$. The DW test statistic is calculated as;

$$DW = \frac{\sum_{t=1}^{n}(\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^{n} \varepsilon_t^2} \qquad (20)$$

The decision rule for the DW test is as follows;
If $d < dL$ reject $H_0 : \rho = 0$
If $d > dU$ do not reject $H_0 : \rho = 0$
If $dL < d < dU$ test is inconclusive.

3.4 Forecasting with transfer function models

Once the transfer function model has passed all the diagnostic tests discussed, the model can be used for forecasting of the output series by using past values of the input series.

Suppose X_t and Y_t are stationary process in model [5] and [14], where $\omega_s(B)$, $\delta_r(B)$, $\theta(B)$, $\emptyset(B)$ are finite order of polynomial of B; the roots of $\omega_s(B) = 0$, $\delta_r(B) = 0$, $\theta(B) = 0$ and $\emptyset(B) = 0$ are all outside the unit circle, and a_t and α_t are independent zero mean white noise series with variances σ_η^2 and σ_α^2 respectively.

Let

$$u(B) = \frac{\omega_s(B)B^b, \emptyset_p(B)}{\delta_r(B)\theta_q(B)} = u_0 + u_1 B + \cdots \qquad (21)$$

and

$$\psi(B) = \frac{\theta_q(B)}{\emptyset_p(B)} = 1 + \psi_0 + \psi_1 B + \cdots \qquad (22)$$

This paper forecast h leads into the forecast horizon. This is achieved on the basis of a model that includes both a transfer function and a noise component as highlighted by (Box et al., 1994 and Granger, 1999). The transfer function forecasting model proposed by these authors is given as;

$$\begin{aligned}
Y_{t+h} &= \delta_1 Y_{t+h-1} + \cdots \\
&+ \delta_{p+d+r} Y_{t+h-p-d-r} + \omega_0 X_{t+h-1} + \cdots \\
&+ \omega_{p+d+s} X_{t+h-b-p-d-s} + \varepsilon_{t+h} \\
&- \theta_1 \varepsilon_{t+h-1} - \cdots - \theta_{q+r}\varepsilon_{t+h-q-r},
\end{aligned} \qquad (23)$$

where t is the time period, h the lead time period, p the order of auto-regression, d the order of differencing, r the order of decay, b the delay, s the order of regression and q the order of moving average. The forecast error variance and forecast interval limits are defined as;

$$Var(h) = \sigma_\varepsilon^2 \sum_{j=0}^{h-1} \eta_j^2 + \sigma_\varepsilon^2 \sum_{j=0}^{h-1} \psi_j^2, \qquad (24)$$

where
η_j = error of the transfer function, $\widehat{Y}_{t+h} = \pm 1.96[V(h)]^{1/2}$.

To evaluate the predictive power of the model, the study uses error some forecast measures. Different authors prefer different measures. Mean absolute percentage error (MAPE) is preferred as it is not vulnerable to outlier distortion. Mean square forecast error (MSFE) is also regarded as an ideal measure and is found not to be susceptible to distortion due to estimates that approaches zero (Fildes et al., 1998). The two forecast errors are explored for the sake of this study.

4. Empirical analysis

This section provides and discusses the preliminary and primary analyses results.
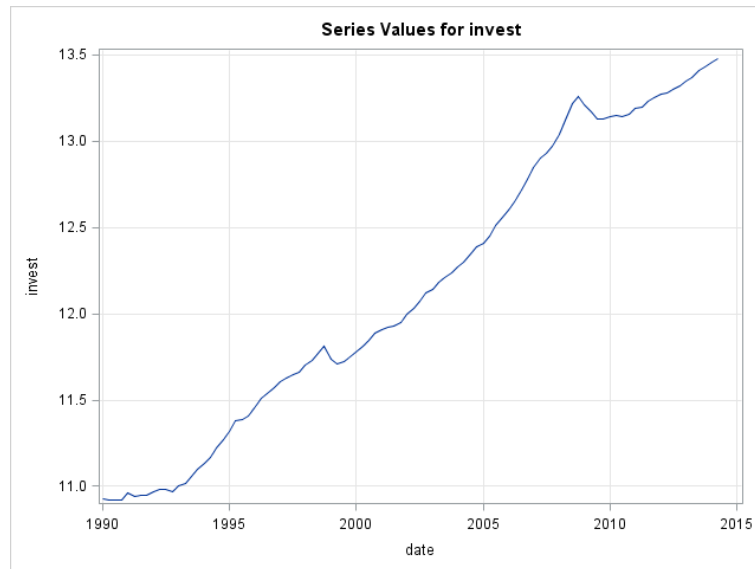
4.1 Preliminary results

Certain underlying assumptions need to be considered when analyzing macroeconomic time series. Stationarity and normality are two main assumptions made about time series data. The secondary task of this section is to investigate whether or not the two time series data under investigation comply with these assumptions. Prior to the analysis, the data were transformed into logarithmic form to iron out irregularities such as noise and heteroscedasticity. This happens so because the scale which the variables are measured drastically is compressed (Gujarati and Porter, 2009).
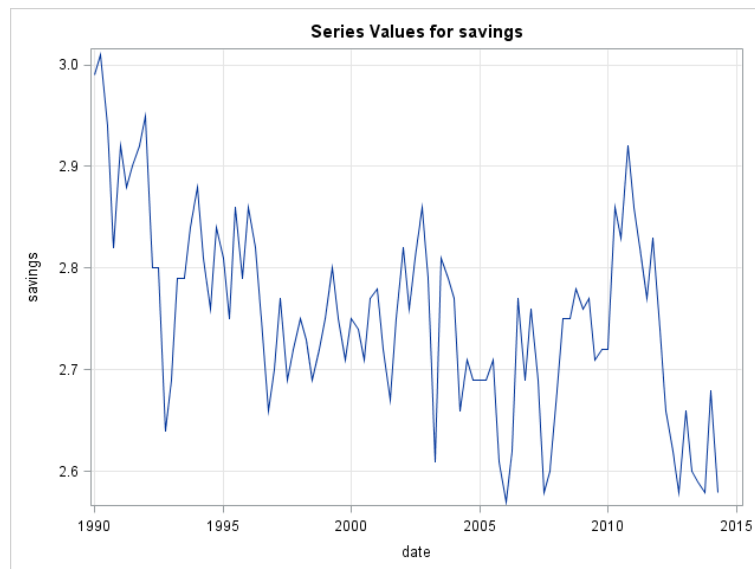
By observation, the descriptive statistics measuring these assumptions confirmed that the two series are approximated by a normal distribution. The observed probabilities for the Jacque-Bera statistics calculated as 0.09 and 0.51 for investment and savings respectively are greater than 0.05 significance level. This is a confirmation that the variables are approximately normal.

Figures 1 and 2 are time series display of the two variables. A visual inspection of these series confirms that they are non-stationary, suggesting that differencing may be required to induce stationarity. The plot suggests that investment in South Africa is explained by an upward moving trend, while savings follows an irregular pattern. This suggests that no co-integration and causal effect could be expected between the variables. The results allow the application of Box-Jenkins transfer function framework for further analyses.
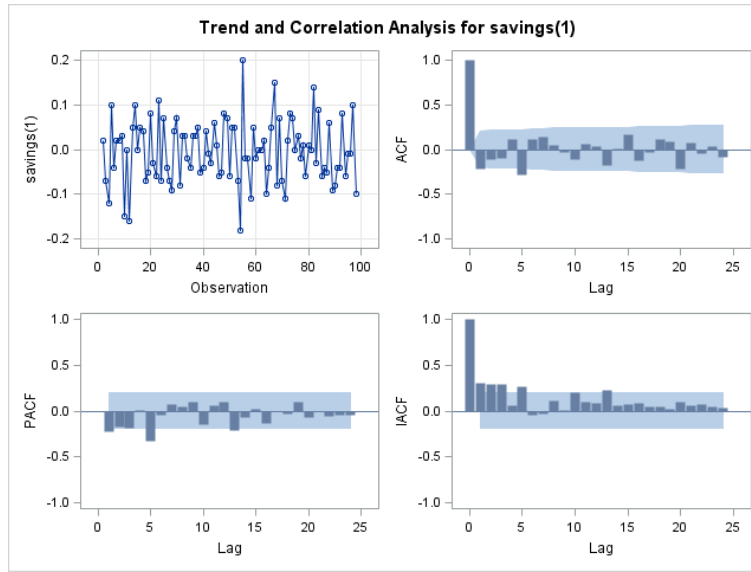
Figure 1. Plot of log (investment)

Figure 2. Plot of log (savings)

Figures 3 and 4 present an overlay of the single-differenced logarithmic data. Also displayed in these figures are the ACFs and the PACFs of these series. A visual examination of this single-differenced logarithmic series suggest that the two series are integrated with order one. However, formal tests of integration are required to support this view. To formally test for stationarity, the ADF unit root test is used with lags up to 5. The results of this test are summarized on the appendix in Tables 1 and 2. The selection of the optimal lag length is done on the basis of the Akaike information criterion (AIC). The remarks underneath are also with alluded to Figures 1 and 2.

Figure 3. Overlay plot of single difference log (savings)

Figure 4. Overlay plot of single difference log (invest)

From Figure 1, log (invest) exhibits an upward trend, suggesting that the ADF test should be of the form;

$$\Delta Y_t = \alpha + \beta_t + (\rho - 1)Y_{t-1} + \sum_{i=1}^{k} \beta_i \Delta Y_{t-i} + \varepsilon_t, \quad (25)$$

The log (savings) exhibits no upward trend and the following ADF model is suggested;

$$\Delta Y_t = \alpha + (\rho - 1)Y_{t-1} + \sum_{i=1}^{k} \beta_i \Delta Y_{t-i} + \varepsilon_t. \quad (26)$$

Based on these assertions, ADF tests conducted for log (savings) and log (invest) confirmed that the series are stationary in the presence of features [25] and [26] after first difference. Figure 3 and 4 reveal the differenced stationary savings and investment plots located on the upper left corners. This allows the analysis to be further conducted with this information in mind.

Upon realizing that the data satisfies the conditions for stationarity, the study continues to assess if there is a relationship between savings and investment using a correlational analysis. However, these analyses are just meant to confirm the nexus of savings and investment level. It does not form part of the transfer function framework and only important selected results are highlighted. The correlation coefficient confirms a negative correlation (-0.507)

between the variables. A full capacity mobility hypothesis was also assessed. Please note that the results of the estimated model are not shown here. The initial estimated model proved to be significant, and a reasonable amount of variation in the two variables is explained. However, the Durbin-Watson statistic suggests the presence of serial correlation. As a result, the model was re-estimated taking into consideration the presence of autocorrelation factor. The re-estimated model also proved to be significant, this time no serial correlation was revealed. The coefficient of savings was even larger and nowhere near zero proving evidence of imperfect capital mobility. This coefficient simply accentuates the presence of the relationship between the variables. Capital mobility as defined by Ogbokor and Musilika (2014) is the ability of private funds to move across

national boundaries in pursuit of higher returns. This mobility is dependent on the absence of currency restriction on the inflows and outflows of capital. Likewise, Chow's test showed no indication of a structural break in the data.

4.2 Transfer function model results

The ACFs and PACFs displayed on Figures 3 and 4 are used to identify a tentative model for the input series. The patterns of ACFs and PACFs of AR *(p)* and MA *(q)* processes look different and may as a result identify different models. Trial and error method is used to identify the models describing X_t. Summary of the results is presented in Table 3.

Table 3. Model identification for Savings

Model	AIC
$ARIMAX\ (5,1,0)$	-252.1
$ARIMAX\ (5,1,5)$	-245.9
$ARIMAX\ (0,1,5)$	-250.2

The AIC is used in selecting the model that best explain the variables. $ARIMAX\ (5,1,0)$ has the least AIC. Box and Jenkins (1976) suggest that upon selection of the appropriate model describing X_t, the same model should be used in pre-whitening the

values of X_t and Y_t. In this instance; pre-whitening of the series is a preliminary step to determining the relationship between the variables. The results of the selected model are summarised in Table 4.

Table 4. Conditional Least Squares estimation of $ARIMAX\ (5,1,0)$

Parameter	Estimate	Standard Error	*t*-Value	Approx Pr > |t|	Lag
AR1,1	-0.27733	0.09943	-2.79	0.0064	1
AR1,2	-0.26086	0.10376	-2.51	0.0137	2
AR1,3	-0.23993	0.10457	-2.29	0.0240	3
AR1,4	-0.05955	0.10400	-0.57	0.5683	4
AR1,5	-0.32670	0.10082	-3.24	0.0017	5

The pre-whitened filter contains autoregressive factors according to the output in Table 4. The mathematical form of the model is presented as;
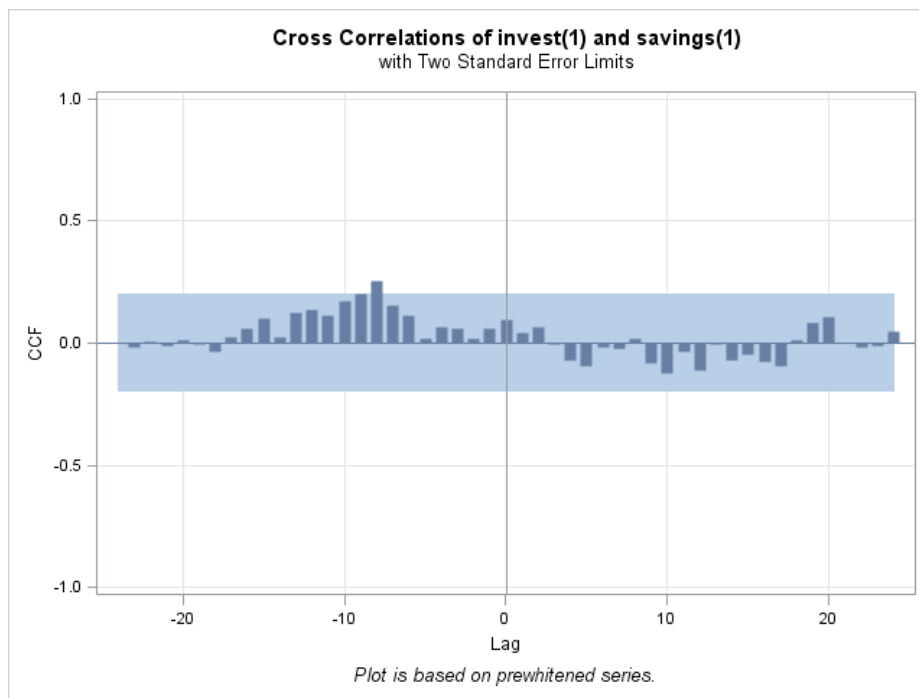
$$1 - 0.27733\ B^{**}(1) - 0.26086\ B^{**}(2) -$$
$$0.23993\ B^{**}(3) - 0.05955\ B^{**}(4) - 0.3267 \qquad (25)$$
$$B^{**}(5).$$

The point estimates of this model are all significant (t-ratios exceed 1.96) except for the parameter estimate of AR 1,4 and all have negative parameters. These results are in accordance with those of the correlational and full capacity mobility hypothesis obtained in the preliminary analyses.

Moreover, the estimates are all less than one and are in accordance with the recommended guidelines by Yaffee and MacGee (2000). Note that a first-order regular differencing was used to obtain stationary savings and investment values. Equation [27] represents both the pre-whitened X_t and Y_t values. None of the parameters is discarded. Further analysis is based on this pre-whitened series.

Next, the CCF between pre-whitened series is calculated to help identify a preliminary transfer function model describing relationship between savings and investment. The results are presented in Figure 5.

Figure 5. Cross correlation check of log differenced investment and savings

The first thing noted is that there are no spikes in the CCF at all lags save for lag eight. This implies that $r_{-1}(\alpha_t, \beta_t)$ is not statistically different from zero i.e. present values of savings are not related to past values of investment. This further confirms that savings is a leading indicator of investment. The spike at lag 8 follows that $r_8(\alpha_t, \beta_t)$ is statistically different from zero, which says that investments in the present quarter are related to savings eight quarters or two years ago. Simply, it can be concluded that it take two years for savings to affect investment in South Africa. The results obtained in this section are a prerequisite for applying the transfer function models.

Presented next in Table 4 is the output for the estimation and diagnostic checking of the tentative model.

Table 5. Cross correlation Check of Residuals with input savings

To Lag	Chi-Square	DF	Pr > ChiSq	Cross correlations					
5	0.55	3	0.9070	-0.017	-0.044	-0.053	-0.013	-0.031	0.029
11	1.28	9	0.9985	-0.047	-0.071	0.031	-0.008	0.024	-0.020
17	3.15	15	0.9995	-0.032	0.035	0.107	0.051	-0.071	-0.049
23	4.72	21	0.9999	0.010	0.063	-0.002	0.013	-0.064	-0.108

Since the observed probabilities of chi-square from column 4 of Table 5 are large, it is concluded that the pre-whitened input series α_t is statistically independent of the error component η_t. This is a necessary condition for the validity of transfer function modelling. To determine a noise model η_t [14], note that the ACF in Appendix B dies down quickly and the PACF which has a spike at lag 1 also dies down quickly. The mathematical function of the noise model in Appendix C according to [15] is represented as;

$$\hat{\eta}_t = -0.13323 - \frac{1 + 0.51323\,B**(1) - 0.21783\,B**(2) - 0.08015\,B**(3) + 0.46459\,B**(4) + 0.72777\,B**(5)}{1 - 0.04787\,B**(1) - 1.01565\,B**(2) - 0.32022\,B**(3) + 0.37145\,B**(4) + 0.16099\,B**(5)}\,B^b X_t. \tag{28}$$

An appropriate estimated final transfer function model employing the non-seasonal autoregressive operator of order 5 becomes;

$$\alpha_t$$
$$= \frac{1 + 0.34897\,B**(1) - 1.45196\,B**(2) - 0.7827\,B**(3) - 0.40376\,B**(4) - 0.4928\,B**(5)}{1 + 1.14276\,B**(1) + 0.26819\,B**(2) + 0.79391\,B**(3) + 1.08853\,B**(4) + 0.20643\,B**(5)} X_{t-b} \quad (29)$$
$$+ \frac{1}{1 - 0.86318\,B**(1) + 0.42825\,B**(2) - 0.36364\,B**(3) + 0.35403\,B**(4) - 0.21753\,B**(5)} \eta_t$$

The parameters of the noise model [28] have insignificant *t*-ratios (see appendix C). However, all the parameters of the autoregressive factors for the final transfer function model are significant with *t*-ratios greater than 1.96 according to Yaffee and MacGee (2000). This is a confirmation that the final transfer model is adequate and may be used for forecasting. The calculated observed standard errors

are (0.039 and 0.0199) and the AICs are (-295.665 and -404.033) for the noise and final transfer function models respectively. These measures further confirm the adequacy of the transfer function model [29]. The residuals of this model [29] are further assessed for randomness and the results are summarised in Table 6.
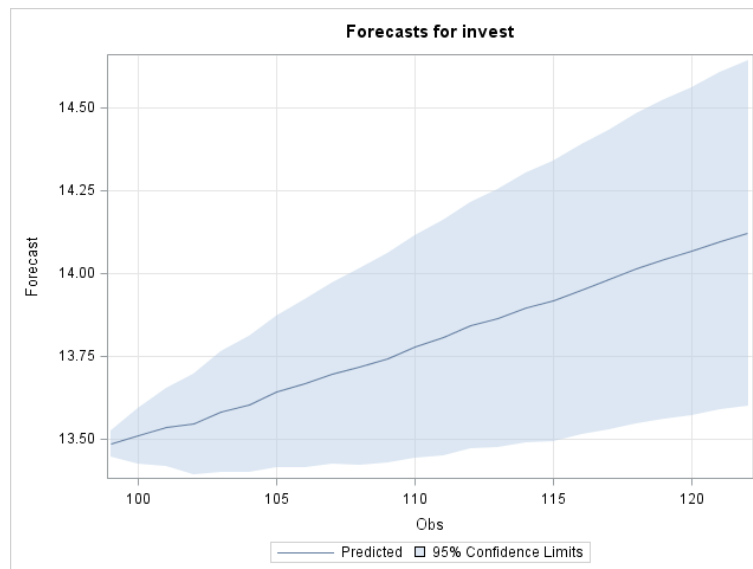
Table 6. Autocorrelation Check of Residuals of Final transfer function model

To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	0.77	1	0.3804	0.023	-0.042	0.028	-0.022	0.066	-0.026
12	10.55	7	0.1595	-0.094	0.102	-0.030	0.130	-0.037	-0.246
18	11.76	13	0.5477	0.013	0.003	-0.031	-0.030	-0.083	-0.048
24	15.80	19	0.6704	-0.015	-0.046	0.052	-0.102	-0.031	0.133

The chi-square test has all insignificant observed probabilities at all lags implying that the residuals are not correlated. Reference can also be made to the autocorrelation coefficients in Table 6. The next and

final step provides the forecasts of investment using the final transfer function model [29]. Numerical results of the forecasts are given as Appendix E and plotted in Figure 6.

Figure 6. A two year forecasts of investment using $ARIMAX\,(5,1,0)$

The forecasts shows that investment in South Africa is expected grow to in the next two years. These forecasts are estimated within confidence bounds of 95%. The estimated MAPE associated with these forecasts is 0.1335. This forecast error is so small and provides a sensible reason to conclude that the fitted model is valid and reliable. The forecasts produced with this model are with no doubt reliable and may be referred to when embarking on new policies.

5. Conclusions

This paper is an attempt to investigate the relationship between saving and investment in South Africa. Data spanning the period 1990 quarter 1 to 2014 quarter 3 was used. As a preliminary step, the series were tested for stationarity using the ADF unit root test. The test confirms both savings and investment are non-stationary in their level but stationary in their first difference. This gives a degree of confidence on the models estimated and incorporates the stylized

facts. The coefficient correlation matrix obtained implies a negative relationship between the two variables. Feldstein and Horioka (1988) argue that the association between domestic savings and investment is perfect in a closed economy but the presence of capital mobility breakdown is considered.

However, by employing Chows' structural break test, the study find that in the estimated model, the fact that the estimated coefficient is nowhere near zero provides evidence of imperfect capital mobility. In other words, this confirms that savings and investment are related. This is a good indication for South African economy implying that investment-promoting strategies are likely to be a success and economically efficient. This however does not mean that the resources should not be focused more on savings-promoting initiatives. It simply means that the current in-house strategies could be augmented to ensure even better performance by this sector. Furthermore, the government could also consider a reduced income tax rates and rethink economic and political steadiness to help stabilize domestic savings. Availability of more employment opportunities to residents of the country could also boost savings behavior.

The primary analysis of data was conducted following the Box-Jenkins transfer function framework. As a first step in building a transfer function model, the order of the operators and the pure delay is identified. As suggested, good preliminary estimates of the impulse response weights should be significant to help in guessing the orders of the model. The final transfer function model outperformed the noise model according to the error forecast measures. At face value, it is tempting to use this model to forecast trends of investment into the future reasonably well. The model is significant according to battery of tests used. Additional analysis shows forecasts of savings from the pre-whitened transfer function model increasing over time, backing the conclusion drawn from the preliminary analysis. Literature supports that low investment levels in South African economy are consistently identified as the principal factor behind the suboptimal growth rates. Despite the increasing recognition of the importance of investment, there is alarming little analytical research available in South Africa on the savings-investment behavior. On this basis, the study recommends an inclusion of more determinants of investment in the model. Econometric techniques that include the estimation of non-linear models or neural networks may also be used. The study can be expanded upon by adding different countries in order to form cross sectional or panel data study. When formulating policies that concern investment–saving, the study recommends the application of multivariate transfer function framework where variables such as growth among others is also factored in the model.

References

1 Arumugam, P. and Anithakumari, V. (2013), "Seasonal Time Series and Transfer Function Modelling for Natural Rubber Forecasting in India", International Journal of Computer Trends and Technology, Vol. 4 No. 5, pp. 1366-1370.

2 Bambang, W.O. and Suhartono, (2009), "Development of Rainfall Forecasting Model in Indonesia by using ASTAR, Transfer Function and ARIMA Methods," European Journal of Scientific Research, Vol. 38 No. 3, pp. 386-395

3 Box, G.E.P. and Jenkins, G.M. (1976), Time series analysis: forecasting and control, Holden-Day, San Fransisco.

4 Box, G.E.P., Jenkins, G.M. and Reinsell G. (1994), Time series analysis: forecasting and control, Holden-Day, San Francisco.

5 Eni, D., Ibiang, E.I., Jeremiah, A. and Ewona, O.A. (2013), "Three -Input Transfer Function Modeling of Rainfall in Calabar, Cross River State of Nigeria", Journal of Emerging Trends in Engineering and Applied Sciences, Vol. 4 No. 4, pp. 630-635.

6 Esso, L.J. and Keho, Y. (2010), "The savings-Investment Relationship: Co-integration and causality Evidence from Uemoa Countries', International Journal of Economics and Finance, Vol. 2 No. 1, pp.174-181.

7 Feldstein M. and Horioka, C. (1980), "Domestic Saving and International Capital Flows," The Economic Journal, Vol. 90, pp. 314-29

8 Fildes, R., Hibon, M., Makridakis, S., and Meade, N. (1998), ''Generalizing about Univariate Forecasting Methods: Further Empirical Evidence,'' The International Journal of Forecasting, Vol. 14 No. 3, pp. 342.

9 Granger, C. W. J. (1999), "Comments on the Evaluation of Econometric Models and of Forecasts, paper presented at Symposium on Forecasting", New York University, New York, NY.

10 Gujarati, D.N. and Porter, D.C. (2009), Basic econometrics, McGraw-Hill Irwin, New York.

11 Hamilton, J. (1994), Time series analysis, Princeton University Press, New Jersey.

12 Khin, A.A., Zainalabidin, M. and Nasir. S. (2011), "Comparative Forecasting Models Accuracy of Short-term Natural Rubber Prices," Trends in Agricultural Economics, Vol. 4 No. 1, pp. 1-17.

13 Ljung, G. and Box, G.E.P. (1978), "On a measure of a lack of fit in time series models", Biometrika, Vol. 65 No. 2, pp. 297-303.

14 Makridakis, S., Wheelwright, S.C. and Hyndman, R.J. (1998), "Forecasting: Methods and Applications", John Wiley and Sons, Inc.

15 Montgomery, D.C, Jennings, C.L. and Kulahci, M. (2008), "Introduction to time series analysis and forecasting", Wiley-Interscience, Hoboken, New Jersey.

16 Moroke, N.D. (2005), "An Application of Box-Jenkins Transfer Function Analysis to Consumption-income Relationship in South Africa", a dissertation submitted to the North West University of South Africa.

17 Moroke, N.D. (2014), "The robustness and accuracy of ARIMA in modelling and forecasting household debts in South Africa, Journal of Economics and Behavioural Studies, Vol. 6 No. 9, pp. 748-759.

18 Ogbokor, C.A. and Musilika, O.A. (2014), "Investigating the Relationship between Aggregate Savings and Investment in Namibia: A Causality Analysis", Research Journal of Finance and Accounting, Vol. 5 No. 6, pp. 85-89.

19 Priestley, M.B. (1981), Spectral analysis and time series, Academic Press, London.

20 Sadowski, E.A. (2010), "A Time Series Analysis: Exploring the Link between Human Activity and Blood Glucose Fluctuation", Theses and Dissertations (Comprehensive), Paper 1021.

21 Vogelvang, B. (2005), Econometrics: Theory and Applications with EViews, Prentice-Hall, New York.

22 Wahid, A, Salahuddin, M. and Norman, A. (2008), "Saving Investment Correlation in south Asia: A Panel Approach", European Journal of Economics, finance and Administrative Science, No. 11.

23 Wei, W.W.S. (2006), Time Series Analysis: Univariate and Multivariate Analysis, Wesley Pearson, Edison.

24 Yaffee, R. and McGee, M. (2000), Introduction to Time Series Analysis and Forecasting with Applications of SAS and SPSS, Academic Press, Orlando, Florida.

Appendix A

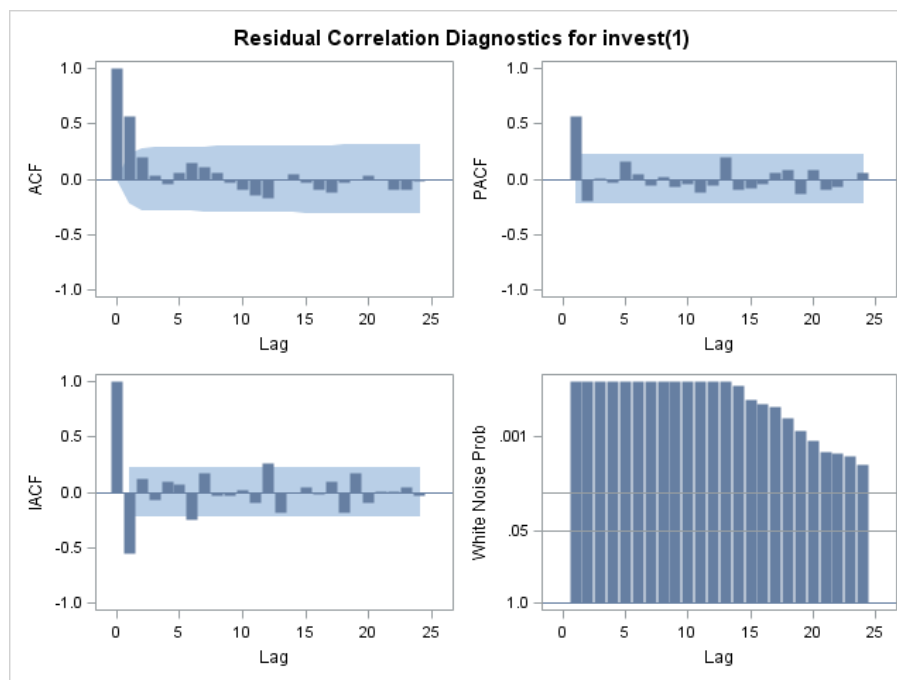**Table 1.** ADF Unit Root Test for log differenced savings

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-117.444	0.0001	-12.10	<.0001		
	1	-163.251	0.0001	-8.92	<.0001		
Single Mean	0	-117.779	0.0001	-12.09	<.0001	73.15	0.0010
	1	-165.718	0.0001	-8.94	<.0001	39.92	0.0010
Trend	0	-117.911	0.0001	-12.04	<.0001	72.49	0.0010
	1	-166.042	0.0001	-8.89	<.0001	39.53	0.0010

Table 2. ADF Unit Root Test for log differenced investment

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-20.1163	0.0012	-3.32	0.0011		
	1	-17.8612	0.0024	-2.94	0.0037		
Single Mean	0	-41.1352	0.0009	-5.14	<.0001	13.20	0.0010
	1	-47.3633	0.0009	-4.86	0.0002	11.81	0.0010
Trend	0	-41.1881	0.0003	-5.10	0.0003	13.05	0.0010
	1	-47.4446	0.0003	-4.82	0.0009	11.68	0.0010

Appendix B

Figure 1. Residual correlation diagnostics

Appendix C

Table 1. Noise model

Parameter	Estimate	Standard Error	t-Value	Approx Pr > |t|	Lag	Variable	Shift
SCALE1	-0.13323	0.07057	-1.89	0.0630	0	savings	5
NUM1,1	-0.51323	1.08218	-0.47	0.6367	1	savings	5
NUM1,2	0.21783	0.92720	0.23	0.8149	2	savings	5
NUM1,3	0.08015	0.88726	0.09	0.9283	3	savings	5
NUM1,4	-0.46459	0.82213	-0.57	0.5737	4	savings	5
NUM1,5	-0.72777	0.81605	-0.89	0.3754	5	savings	5
DEN1,1	0.04787	1.01592	0.05	0.9625	1	savings	5
DEN1,2	1.01565	0.61315	1.66	0.1019	2	savings	5
DEN1,3	0.32022	1.44776	0.22	0.8256	3	savings	5
DEN1,4	-0.37145	0.61688	-0.60	0.5489	4	savings	5
DEN1,5	-0.16099	0.80153	-0.20	0.8414	5	savings	5

Appendix D

Table 2. Final transfer function model

Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag	Variable	Shift
MU	0.02925	0.0061435	4.76	<.0001	0	invest	0
AR1,1	0.86318	0.12216	7.07	<.0001	1	invest	0
AR1,2	-0.42825	0.15995	-2.68	0.0093	2	invest	0
AR1,3	0.36364	0.16311	2.23	0.0291	3	invest	0
AR1,4	-0.35403	0.16048	-2.21	0.0308	4	invest	0
AR1,5	0.21753	0.12384	1.76	0.0836	5	invest	0
SCALE1	-0.03566	0.03054	-1.17	0.2470	0	savings	5
NUM1,1	-0.34897	1.73564	-0.20	0.8413	1	savings	5
NUM1,2	1.45196	2.84491	0.51	0.6115	2	savings	5
NUM1,3	0.78270	2.64965	0.30	0.7686	3	savings	5
NUM1,4	0.40376	2.45091	0.16	0.8696	4	savings	5
NUM1,5	0.49280	1.55397	0.32	0.7521	5	savings	5
DEN1,1	-1.14276	0.69498	-1.64	0.1048	1	savings	5
DEN1,2	-0.26819	0.64534	-0.42	0.6790	2	savings	5
DEN1,3	-0.79391	0.19403	-4.09	0.0001	3	savings	5
DEN1,4	-1.08853	0.55475	-1.96	0.0539	4	savings	5
DEN1,5	-0.20643	0.68825	-0.30	0.7652	5	savings	5

Appendix E

Table 1. Forecasts of investment by transfer function model

Obs	Forecast	Std Error	95% Confidence Limits	
99	13.4845	0.0200	13.4453	13.5237
100	13.5088	0.0422	13.4260	13.5916
101	13.5347	0.0607	13.4158	13.6536
102	13.5447	0.0779	13.3920	13.6974
103	13.5825	0.0926	13.4011	13.7640
104	13.6049	0.1049	13.3993	13.8104
105	13.6413	0.1168	13.4124	13.8703
106	13.6675	0.1286	13.4155	13.9196
107	13.6974	0.1399	13.4232	13.9716
108	13.7177	0.1510	13.4217	14.0136
109	13.7439	0.1615	13.4273	14.0604
110	13.7771	0.1714	13.4411	14.1130
111	13.8057	0.1809	13.4512	14.1603
112	13.8430	0.1900	13.4707	14.2154
113	13.8645	0.1987	13.4751	14.2540
114	13.8960	0.2072	13.4899	14.3020
115	13.9164	0.2153	13.4944	14.3384
116	13.9515	0.2232	13.5141	14.3890
117	13.9807	0.2308	13.5282	14.4331
118	14.0138	0.2382	13.5469	14.4807
119	14.0425	0.2454	13.5615	14.5236
120	14.0664	0.2524	13.5717	14.5612
121	14.0974	0.2592	13.5893	14.6055
122	14.1218	0.2659	13.6007	14.6429