International Journal of Economics and Financial Issues

ISSN: 2146-4138

available at http: www.econjournals.com

International Journal of Economics and Financial Issues, 2016, 6(2), 728-735.

Effectiveness of a Cluster of Determinants to Increase Economic Growth Rate: A Combined Statistical Criteria Approach

Chee Yin Yip[1], Hock Eam Lim[2], Hooi Hooi Lean[3]*

[1]Department of Economics, Faculty of Business and Finance, Universiti Tunku Abdul Rahman, Kampar, Perak, Malaysia, [2]College of Arts and Sciences (Economics), Universiti Utara Malaysia, Malaysia, [3]Economics Program, School of Social Sciences, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia. *Email: hooilean@usm.my

ABSTRACT

This paper attempts to estimate the effectiveness of a cluster of determinants to increase gross domestic product (GDP) growth rate by using a combined statistical criteria approach. First, combining three ranking measures i.e., partial regression coefficients, adjusted R^2 and Bayesian information criterion (BIC) into one single ranking procedure for finding and ranking the impact of each determinant - Y-procedure. Second, ranking the effectiveness of a cluster of determinants, each of which has been Y-procedure ranked using F-statistics, adjusted R^2 and BIC in increasing GDP growth rate - Y-average. The results show that sets of top five or more variables should be considered as one entity with respect to increasing GDP growth rate, and the degree of effectiveness increases if their Y-average of relative measures increases. On application of this Y-procedure and Y-average to Australian GDP growth rate, it is found that investment, current account balance, gross foreign liability, export and import have the highest impact and thus, these five variables should be given priority when constructing the relevant economic policies and allocation of funds towards increasing GDP growth rate specifically for the case of Australia.

Keywords: Prioritize, Allocation, Ranking Measures, Y-procedure, Y-average
JEL Classifications: C13, C18

1. INTRODUCTION

This paper attempts to estimate the effectiveness of combinations of determinants in increasing gross domestic product (GDP) growth rate. GDP growth rate is a common indicator for the performance of the economy, whether it is expanding or contracting as compared with previous period. GDP growth rate is widely used as a measure of the national economic activities and therefore the economic wellbeing of the nation and its people in general. It is mainly because of this reason that every country attempts to design and implement its economic policy and allocation of funds by considering the would-be high impact determinants which can increase GDP growth rate. Theoretically, the fundamental determinants of a country's economic growth could be identified through the variables included in calculating a country's GDP by expenditure approach: Consumptions, investment, government expenditure, import, and export. Numerous studies have been carried out to find the long-run growth path. The choice of focus is normally based on economic theories and conventional wisdom as well as statistical criteria. Studies based on economic theory, for example, the growth model as proposed by Solow (1956) and Swan (1956) which uses a well behaved neoclassical production function, a single homogenous good, exogenous labor augmenting technical process, full employment and exogenous labor force growth have been assumed for economic growth. Other examples are Mankiw et al. (1992) and Pack (1994). However, recent growth theorists allege that the standard neoclassical model fails to explain the observed difference in per capita income across countries. As a result, an endogenous growth model based partly on conventional wisdom as well as economic theory and which assumes constant and increasing returns to capital is developed. Examples of endogenous growth model are Gregorio (1991) who found that the productivity growth, macroeconomic stability and investment (physical and human resource) had played an important role in determining the economic growth in Latin American countries. Next, due to globalization, economic growth of a country could be affected by external shocks (Easterly et al., 1993). Furthermore, Ristanovic (2010) suggests that among the determinants of economic growth are exports, imports, inflation rate, direct foreign investments, real interest rate, real exchange rate, consumptions and investments. On the other hand, conventional wisdom suggests that each determinant produces different impact and that there are interrelations among the determinants, for example export and import are directly related to inflation rate. All these three variables are GDP determinants and there is at least one interrelation among them. All these economic wisdom suggest that the effect of GDP determinants would better be quantified and considered in the form of cluster or combination. Examples of literature in support of this view are Petri (1997), and Thomas and Wang (1996). The different implications of exogenous and endogenous growth models have spurred empirical studies in recent years. Ironically, none of these studies focus on any statistical combined effect of determinants on GDP growth rate. Conventional wisdom tells us that these determinants from the real sector of the economy are the driver of economic growth in the long-run (Romer, 1992) and the economic growth of a country though determined by random and non-random factors could be controlled or tuned properly to some extend by dealing with the appropriate determinants. This paper works on how to deal with a cluster of appropriate determinants using statistical theory as well as economic theory and conventional common sense as backing, for the purpose to induce higher GDP growth rate.

Literature in the field of economic growth abounds with theoretical and empirical analyses of determinants of economic growth. By conventional wisdom, a country's socioeconomic characteristics, political stability and appropriate macroeconomic policies, are the significant determinants of economic growth. Nevertheless the process of economic growth is complex, determined and impacted by various factors which are interrelated. Positive interaction between socio-economic factors which fosters economic growth is brought about by appropriate economic policy and the allocation of resources. Therefore the formulation and implementation of economic policies and allocation of funds towards achieving higher GDP growth rate is among the top priorities of many nations. Towards achieving this goal, we suggest a way to formulate and implement economic policies and allocation of funds by quantifying and ranking the relative impact of combination of GDP determinants. It is well recognized that determinants of GDP growth rate experience causality impact among each other and that each determinant Granger causes GDP growth rate in one way or another (Vojinovic, 2008). This further supports our conception that it would be more appropriate to study quantitatively the effect of different combination of determinants on GDP growth rate. Furthermore, due to constraints of resources, it serves policy makers well if we can quantify and rank the relative importance of each combination of GDP determinants so that prioritization of resources can be done according to the relative impact each combination of determinants on GDP growth rate. The above argument is supported by literature reviews that economic growth can be enhanced by the combined effect of various variables Petri (1997). However, since we study the effects of different combinations of GDP determinants, it would be sensible to quantify and rank all the individual determinants and only then, we construct combinations of these determinants. Thus, our central hypothesis is that GDP growth rate can be greater enhanced by focusing on a selected group of determinants that have the best positive interaction impact on economic growth while the selection of the determinants is achievable if the relative impact of each determinant or combined determinants can be statistically quantified. The following is a brief literature reviews on statistical quantification of a set of determinants. Hereafter, determinants and independent variables are used interchangeably.

The usefulness of a set of high impact independent variables in almost every field of study is well recognized. The relative impact or importance of independent variables has been effectively applied in many aspects. On a broad front, Kruskal and Majors (1989) found that statistical significance as a measure of relative importance has been applied indiscriminately. In the economic and financial front, Kacperczyk et al. (2005) showed that on average, funds which place more emphasis on specific industries which they have access to their information, perform better than normal diversification of portfolios. However, this placement of funds is on a non-random basis. Doukas et al. (2006) suggested that future returns for a stock will be higher if there is a greater disagreement among investors about the stock's value. However this divergence of analysts' opinion has not been quantified for its relative importance. Ait-Sahalia and Brandt (2001) proposed that the dependence of the optimal portfolio weights on the predictive variables be determined directly as opposed to normal practice. His approach supports the study of relative impact. However, he did not indulge in averaging process. This brief literature review suggests stochastic dependency and more than one measure of relative impact of the independent variables should be considered for any relative impact study of single or combinations of GDP determinants. The central difficulty is that very few of these existing studies address the problem of stochastic dependency among the independent variables. Nevertheless, to-date studies that employed more than one concept of relative importance on the same set of data or introduced new measure are still lacking. This paper proposes a procedure and a ranking measure, the former we name as the Y-procedure and the latter is Y-average. We use Y-procedure to quantify and rank each individual independent variable (determinant) by addressing the issues of stochastic inter-dependency and subsequently using Y-average to rank clusters of Y-procedure ranked determinants. Y-average will be defined in Section 3. To deal with the issue of stochastic interdependency, we propose to use partial regression coefficients, adjusted R^2 and Bayesian information criterion (BIC) for the Y-procedure. For Y-average, we take the simple average of F-statistics, adjusted R^2, standard error of regression and BIC. For the second issue, we use an approach which is very similar to model averaging procedure by Hansen (2007). With that, we define Y-procedure relative impact, S as a combined relative (ranking) measure in the numeric format of the explanatory power of the independent variables. Then, we rank S by an ascending sequence of positive integers g starting from 1 to m (denote the number of exogenous variables).

The rest of this paper is organized as follows: Section 2 introduces briefly the selected literature reviews that form the basis for the conceptualization of the Y-procedure and Y-average. Section 3 describes the steps to use Y-procedure and Y-average for the analysis of the empirical results and the validation of the Y-procedure. Section 4 describes the empirical analysis and finally section 5 covers the conclusion of this study and suggestions for further research.

2. CONCEPTUALIZATION OF THE Y-PROCEDURE AND Y-AVERAGE

Both Y-procedure and Y-average mechanism are motivated by the unsettled issues concerning model selection and model averaging (combination of models) procedure.

Model selection procedure has a long history and it is an excellent tool used to choose the best model out of a number of models for the purpose of estimation and/or prediction. However recent studies showed that its inference may be biased, and this is mainly due to the existence of uncertainty in model selection process. Results from studies by Potscher (1991), Leeb and Potscher (2003; 2005; 2006) and Hansen (2007) validated the existence of this model selection uncertainty. This finding should have motivated further research to determine the real nature of this uncertainty. Unfortunately, until then, the real nature of model selection uncertainty has not been understood yet (Yuan and Yang, 2005). Instead, alternative method like model averaging has been proposed to take over the role of model selection. Since then, model averaging has become increasingly popular because it can reduce estimation variance while controlling omitted variables (Hansen, 2007). However, the issue of choosing the weights for model averaging has not been settled yet. Thus, model selection and model averaging each has its own weakness. Which one should be used and under what conditions is an open ended question. We are not clear whether model averaging can be better than model selection in some or all aspects. Moreover, we are not sure when model averaging is preferable to model selection. Under this uncertainty of these two well-known procedures, it is best for us to use the universal concept of taking the best features out of everything and in this case out of these two procedures. Model selection is a procedure whereby we use statistical criteria like BIC criterion to select the best model. However, we have other criteria like adjusted R^2, and coefficients of regression are mainly used to estimate the impact of independent variables on dependent variable. For Y-procedure, we combine BIC, adjusted R^2 and coefficients of regression to rank each determinant while Y-average we combine F-statistics, BIC, adjusted R^2 and standard error of regression to rank the clusters of Y-procedure ranked determinants. Coefficient of regression is used only in Y-procedure because it estimates the impact of each determinant on GDP growth rate. We omit other poor performance statistical criteria like R^2 because including them in can greatly affect the final inference for simple averaging procedure. Each of these statistical criteria has some uncertainty which influences the final inference. The Y-procedure and Y-average mechanism works on reducing these uncertainties by using simple averaging of these three statistical criteria and at the same time reducing estimation variance through using combination procedure similar to that of model averaging technique.

Our model of interest is the first order autoregressive model (AR(1)) with a single exogenous independent variable, X (determinant). ARX(1) is selected based on the following research findings: There are at least three articles in support of our decision not to include any regressor in our AR model. Banerjee and Marcellino (2006) compares the forecasting accuracy of models using leading indicators and simple AR model for forecasting GDP growth. Their results indicate that pure AR model has a better forecasting ability. Ang et al. (2007) investigates whether macroeconomic variables, asset markets, or surveys best forecast US inflation. They found that survey tend to yield improved forecasts for most macroeconomic variables. Lastly, Granger and Newbold (1986) who found that forecast from simple models only marginally less accurate than models built by using complex technique. They suggested that only when the benefits of the complex techniques outweigh the additional costs of using them, should they be the preferred choice. We include only one exogenous variable X into the AR(1) model just to ensure that the potential endogeneity problem has been eliminated thoroughly and that the coefficient of this exogenous variable fully estimate its potential impact on the dependent variable. Thus, the best model also implies that the independent variable (determinant) has the highest impact on the dependent variable (GDP). The Y-procedure combines three common model selection criteria (BIC, adjusted R^2 - $\bar{R}_j^2$, and partial regression coefficients - β_j) into one single relative measure by simple averaging technique similar to that of selecting the best model and thereby the best exogenous variable. The reasons for this way of combining are as follows: The third criterion (partial regression coefficients) is the single crucial measure of the impact created by the exogenous variable (determinant) on GDP growth rate. This measure is most efficient if the best model has been identified by the first two model selection criteria. The argument is that partial coefficient is not the optimum impact exerted by the corresponding determinant if the model is misspecified and not the best. Thus the first two criteria must go hand in hand with the third criterion. Therefore we combine the three criteria to become a single relative measure of the impact created on GDP growth rate. Y-average use simple averaging technique to combine F statistics, BIC, adjusted R^2 and standard error of regression into a single measure to estimate the impact of each cluster of determinants. We will explain the intrinsic uncertainty associated with each of these criteria in the next section.

3. THE Y-PROCEDURE AND THE Y-AVERAGE

3.1. The Model and the Variables

We use the first order AR model with exogenous variables (ARX) as given in Equation (1):

$$y_t = c + \rho y_{t-1} + \beta_j x_{jt} + u_t, \quad j = 1,\ldots,m;\ t = 1,\ldots,n \tag{1}$$

$$Cov(x_{jt}, u_t) = 0$$

Where, x_{jt}, y_t, m and n denote the exogenous independent variables, the dependent variable, number of exogenous variables and number of observations while $\{c, \rho, \beta_j\}$ is a set of parameters. u_t is the error term with constant variance. The second panel of Equation (1) is the exogenous condition. The reasons for selecting this ARX model are: AR(1) is recognized as a very efficient forecasting model (Banerjee and Marcellino, 2006; Ang et al., 2007), but its error term u_t which is assumed to be a homoskedastic may not be able to capture the full endogeneity which is caused by using one period lag past historical data (first lag of y_t) for predicting the present value of y_t. x_{jt} is an exogenous variable specially included to capture all the remaining endogeneity. It is recognized that the more endogeneity an exogenous variable x_{jt} can capture, the better is the model and the more impact x_{jt} on the variation of y_t. The coefficient β_j of this exogenous variable (determinant) is the proxy measure of the impact which x_{jt} exert on the dependent variable y_t. The selection of this exogenous variable is based on literature reviews on determinants of economic growth and data availability. However, for this study, we selected 15 exogenous variables based on literature reviews. Then we discard those insignificant exogenous variables after running a regression with GDP as dependent variable and all the 15 exogenous variables as independent variables. After running this filtering process, we are left with 10 exogenous variables. They are: Consumer price index (CPI) of alcohol and tobacco, CPI of South Korea[1], current account balance, disposable income on consumption, export, gross foreign liability, import, investment, net income (net investment income from balance of payment) and retail trade spending.

3.2. Estimating ARX Model

We estimate Equation (1) by maximum likelihood estimation method for each of the exogenous variable, x_{jt} but the set of y_t is kept constant. We record down the partial regression coefficients (β_j), adjusted R^2 ($\bar{R}_j^2$) and Bayesian information criterion value (BIC_j). Thus for each x_{jt}, we have one set of values of parameter measure $\{\beta_j, \bar{R}_j^2, BIC_j\}$. β, $\bar{R}^2$ and BIC are three separate measures for the relative impact of x_{jt}. By sorting $\{\beta_j, x_{jt}\}$ in descending order, we would obtain one relative impact of the exogenous independent variables (determinants). The same process can be done for $\{\bar{R}_j^2, x_{jt}\}$ but in the reverse order for $\{BIC_j, x_{jt}\}$.

3.3. Partial Regression Coefficient, β

Regression coefficient is used to measure the amount of change in the dependent variable due to one unit change in independent variable. Assuming that AR(1) is the best approximated data generating process, we use only one single exogenous variable in the ARX model at any one time, the regression coefficient can also be used as a measure of how fit the model is and indirectly this measure the impact of the single variable exerted on the dependent variable. As such, it can be used as a model selection criterion for a single independent variable model. However, regression coefficient has one substantial uncertainty. It is often considered as deterministic in regression but in practice, it is more of a random than deterministic vector. This introduces uncertainty in using regression coefficient as a model selection criterion, and thereby as a measure of the relative impact of the respective independent variable.

Let β_j^* and s_j^* be the coefficient and standard deviation of regression. β_j^* can be used as a ranking measure because change in x will result of a change in y_t. However, this coefficient has a shortcoming as pointed out by Darlington (1990) who observed some inconsistency in the definition of standardized coefficients. We overcome this problem by using Bring's (1994) definition of a consistent partial standard deviation s_j. The partial standardized regression coefficient is given by:

$$\beta_j = \beta_j^* \frac{s_j^*}{\sqrt{\mathrm{VIF}}} \tag{2}$$

Where, VIF is the variance inflation factor. Partial standardized regression coefficient β_j is assumed to be random and in absolute value.

3.4. Adjusted R^2, $\bar{R}^2$

Adjusted R^2 is used to measure the explanatory power of the independent variable with the formula:

$$\bar{R}^2 = 1 - \frac{(n-1)SSE}{(n-k-1)SST} \tag{3}$$

Where, SSE, SST, n and k denote respectively sum square error, sum square total, sample size and the number of regressors. SSE is directly connected to error term of regression which consists basically of missing variable, error in variable and simultaneity, Thus, SSE introduces appreciable amount of uncertainty to adjusted R^2. This uncertainty is very difficult to measure if it can be measure at all.

3.5. BIC

Using Equation (2.56) of Franses and Dijk (2000) with p = 1 and $k = 2$, we obtain Equation (4):

$$\mathrm{BIC}(2) = \ln n^2 \hat{\sigma}^{2n} \tag{4}$$

Since n is essential fixed, BIC depends only on the standard error $\hat{\sigma}$ which is random in nature and quite a direct measure for the quality of the model. It is obvious that the smaller the $\hat{\sigma}$ value the better is the model and so is the explanatory power of the particular exogenous variable x_{jt}.

BIC works on trade-off process between variance and the number of parameters. However, there may be different best trade-off to suit the different set of data and different practical situations. Thus, this introduces uncertainty to the model selection process.

1 South Korean CPI is chosen because it is a good exogenous variable for Australian GDP.

Table 1: Overall ranking

Variables (j=1 to 10)	Individual ranking (F_w)			Relative impact	
	β	$\bar{R}^2$	BIC	Weighted average (S)	Y procedure ranking (g)
CPI (alcohol and tobacco)	2	10	10	8.4081	10
CPI (South Korea)	3	8	9	7.4320	8
Current account balance	5	3	2	2.9711	2
Disposable income	10	7	6	7.1700	7
Export	1	1	8	3.9883	4
Gross foreign liability	6	4	3	3.9711	3
Import	7	5	4	4.9711	5
Investment	4	2	1	1.9711	1
Net income	8	6	5	5.9711	6
Retail trade spending	9	9	7	8.1462	9

CPI: Consumer price index, BIC: Bayesian information criterion

3.6. Y-procedure

For a particular x_{jt}, we have three separate measures for its relative impact $\{\beta_j, \bar{R}_j^2, BIC_j\}$. Each measure is to be ranked and denoted by f_w where $w = 1$ to q, such that q denotes the number of measures. β_j and $\bar{R}_j^2$ are directly proportional to the explanatory power of x_{jt} while BIC_j is inversely proportional. The Y-procedure combines all these three measures into a single ranking procedure.

3.6.1. Steps for the Y-procedure

1 Let $f_w = 1$ to m. β_j in $\{\beta_j, x_{jt}, j\}$ is arranged in descending order of magnitude together with the corresponding x_{jt} and j. Let the resulting series be $\{\beta_f, x_{at}, f_1\}$. The same procedure is repeated for adjusted R^2 and BIC. The results are $\{\bar{R}_f^2, x_{bt}, f_2\}$ and $\{BIC_f, x_{ct}, f_3\}$ where x_{at}, x_{bt}, and x_{ct} are x_{jt} in different arrangement. f_1, f_2, and f_3 are ranking number starting from 1 to m (individual ranking).

2 For $\{x_{at}, f_1\}$, $\{x_{bt}, f_2\}$ and $\{x_{ct}, f_3\}$, x_{at}, x_{bt}, and x_{ct} are arranged back to original position of x_{jt} by sorting a, b and c in increasing order from 1 to m together with the corresponding f_1, f_2 and f_3. We obtain $\{x_{jt}, f_{11}, f_{22}, f_{33}\}$ where f_{11}, f_{22} and f_{33} are taken from f_1, f_2, f_3 and are not arranged in sequential order.

3 We divide the y_t series into z portions. For each portion of y_t, we regress it on β_j, $\bar{R}_j^2$ and BIC_j. We denote the coefficient of regression of β_j, $\bar{R}_j^2$ and BIC_j by b_j, r_j and p_j. We take the average of all the b_j, r_j and p_j and denote as $\bar{b}$, $\bar{r}$ and $\bar{p}$. These $\bar{b}$, $\bar{r}$ and $\bar{p}$ are the weights for combining f_{11}, f_{22} and f_{33} into a single ranking number as follows:

$$S_j = \bar{b}f_{11} + \bar{r}f_{22} + \bar{p}f_{33} \qquad (5)$$

Where, S_j corresponds directly to j and x_{jt}. We rank the S_j by the Y-procedure ranking[2].

3.7. Y-average

3.7.1. Steps for the Y-average

1. We let X_{it} be an exogenous variable with Y-procedure ranked i. We construct 9 combinations of 2 exogenous variables each as shown below:

   $(X_{it}, X_{(i+1)t})$ for i = 1 to 9

   We run regression for each of the 9 combinations and for each regression,

$$GDP_t = \beta_j X_{it} + \beta_{(i+1)} X_{(i+1)} + u_t \qquad (6)$$

   We record F-statistics, BIC, adjusted R^2 and standard error of regression. Then we use equation (7) to compute the Y-average.

   Y-average = [F-statistic +1/BIC + $\bar{R}_j^2$ +1/standard error]/4 (7)

2. We repeat the computation in item 1 using combinations of 3, 4, 5, 6 and 7 exogenous variables each. Equation (8) shows all the regressions required. We analyze the Y-average values for each clusters of Y-procedure ranked determinants.

$$\begin{bmatrix} y_t(1) \\ ; \\ y_t(39) \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{9} & \sum_{i=1}^{9} & 0 & 0 & 0 & 0 & 0 \\ \sum_{i=1}^{8} & \sum_{i=1}^{8} & \sum_{i=1}^{8} & 0 & 0 & 0 & 0 \\ \sum_{i=1}^{7} & \sum_{i=1}^{7} & \sum_{i=1}^{7} & \sum_{i=1}^{7} & 0 & 0 & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ \sum_{i=1}^{4} & \sum_{i=1}^{4} & \sum_{i=1}^{4} & \sum_{i=1}^{4} & \sum_{i=1}^{4} & \sum_{i=1}^{4} & \sum_{i=1}^{4} \end{bmatrix} \begin{bmatrix} X_{it} \\ X_{(i+1)t} \\ X_{(i+2)it} \\ X_{(i+3)it} \\ X_{(i+4)it} \\ X_{(i+5)it} \\ X_{(i+6)it} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \\ . \\ . \\ . \\ . \\ u_{39t} \end{bmatrix} \qquad (8)$$

4. EMPIRICAL ANALYSIS

We apply the Y-procedure ranking for each of the 10 exogenous variables (determinants) of the GDP growth rate of Australia[3]. Our exogenous variables are CPI of alcohol and tobacco, CPI

2 See definition of g in page 2.

3 All data are obtained from Reserve Bank of Australia (September 1959-September 2011).

of South Korea[4], current account balance, disposable income on consumption, export, gross foreign liability, import, investment, net income (net investment income from balance of payment) and retail trade spending. GDP and the determinants are found to be I(1) based on the unit root tests (DF-FGS and Kwiatkowski–Phillips–Schmidt–Shin) and are co-integrated (Johansen co-integration test). Thus, no differencing is needed for the variables. The independent variables are statistically exogenous to GDP (Davidson and MacKinnon augmented regression test). Table 1 shows the ranking results with weights[5] for β, $\bar{R}^2$, BIC and also the Y-procedure ranking. We use the Y-procedure ranking number to denote the determinants. Thus, 1 denotes top Y- procedure ranked determinant, 2 second ranked, 3 third ranked and so on. The Y-procedure ranking number denotes the relative impact each individual independent variable has on the dependent variable.

As we have come to the conclusion that impact of independent variables (determinants) should be considered in groups (clusters or combinations) in the introductory section, we now proceed to construct clusters of these determinants and run the regressions as shown in equation (8) and execute each steps as in Section (3). Then, we rank the impact of each cluster of independent variables (determinants) using Y-average measure. Empirical analysis of the results reveals that that for different combinations of clusters of 2 exogenous variables, Y-average rank has no clear relationship with Y-procedure ranking. Similar trend is found for combinations of 3 or 4 exogenous variables. However, for combinations of 5, 6 or 7 variables, the Y-average measure is related directly to the Y-procedure ranking of the individual independent variables.

Overall, it is found that the Y-procedure is not that accurate for single determinant sets of two, three and four independent variables. However, when a set of five or more top ranked variables are considered, the resulting accuracy is good and consistent. Results of the validation are shown in Table 2. Beside Y-average validates the Y-procedure, it also suggests that its top ranked cluster of 5 determinants which has the highest impact on GDP growth rate for the case of Australia are investment, current account balance, gross foreign liabilities, export and import.

To put our results in a clearer footing, we define (1,2,3,4,5,6,7,8,9,10) as the set of Y-procedure ranking number denotation of each individual variable. We further define:

Definition 1: 5a = (1,2,3,4,5), 5b = (2,3,4,5,6)
5c = (3,4,5,6,7), 5d = (4,5,6,7,8)
5e = (5,6,7,8,9), 5f = (6,7,8,9,10)

Table 3 provides details of Definition 1. By analyzing results in Table 2, it is found that Y-average measure ranks 5a as the top cluster of 5 determinants, 5b second top, followed by 5c, 5d, 5e, and 5f.

With that, we plot Figure 1a and b which show that once a group of determinants with respect to Australian GDP growth rate is ranked by the Y-average, sets of five top ranked variables 5a always have

Table 2: Y-average ranked clusters of determinants and validation Y-procedure

Variable by ranking	Five variables						Six variables					Seven variables			
	1	2	3	4	5	6	1	2	3	4	5	1	2	3	4
Investment	x						x					x			
Current account balance	x	x					x	x				x	x		
Gross foreign liability	x	x	x				x	x	x			x	x	x	
Export	x	x	x	x			x	x	x	x		x	x	x	x
Import	x	x	x	x	x		x	x	x	x	x	x	x	x	x
Net income		x	x	x	x	x	x	x	x	x	x	x	x	x	x
Disp. Income			x	x	x	x		x	x	x	x	x	x	x	x
CPI (South Korea)				x	x	x			x	x	x		x	x	x
Retail trade spending					x	x				x	x			x	x
CPI (alcohol and tobacco)						x					x				x
F-statistic	117,560	115,631	115,122	87,066	81,406	64,731	100,401	98,633	75,334	75,696	55,243	87,410	65,896	66,665	53,397
BIC	3606.5	3609.9	3610.8	3260.2	3272.7	2963.5	3611.5	3615.2	3262.6	3261.7	2968.3	3616.8	3266.9	3264.7	2955.7
$\bar{R}_j^2$	0.99971	0.9997	0.9997	0.99965	0.99962	0.99957	0.99971	0.99970	0.99965	0.99965	0.99957	0.99971	0.99965	0.99965	0.99961
Standard error	1365.6	1377.0	1380.0	1430.6	1479.5	1503.5	1368.1	1380.3	1423.9	1420.5	1506.8	1371.6	1424.1	1415.9	1433.7
Y-average	29,390	28,908	28,780	21,766	20,351	16,183	25,100	24,658	18,833	18,919	13,811	21,852	16,474	16,666	13,349
Y-average ranking	1	2	3	4	5	6	1	2	4	3	5	1	3	2	4

(Y-average) = F-statistics+1/BIC + $\bar{R}_j^2$ + 1/standard error. Y-average ranking is the best if the Y-average is largest. BIC: Bayesian information criterion

4 South Korean CPI is chosen because it is a good exogenous variable for Australian GDP.

5 See Equation (5).

Figure 1: (a) Validation statistics I, (b) validation statistics II

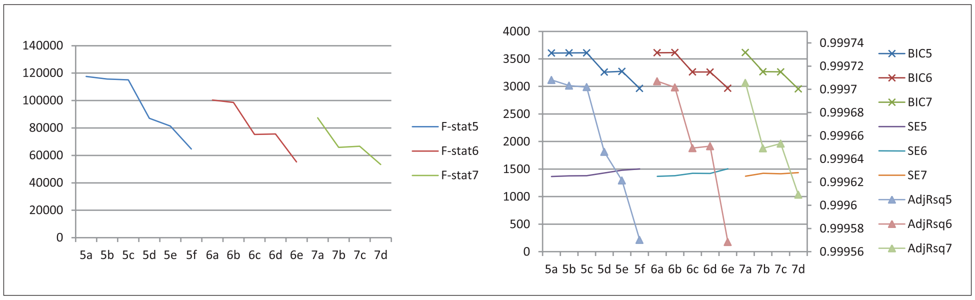

Table 3: Variables included in combined set of 5

Set of 5	Variables included
5a	Investment (building and structure) (1); current account balance (2); gross foreign liability (3); export (4); import (5)
5b	Current account balance (2); gross foreign liability (3); export (4); import (5); net income (6)
5c	Gross foreign liability (3); export (4); import (5); net income (6); disposable income on consumption (7)
5d	Export (4); import (5); net income (6); disposable income on consumption (7); CPI (South Korea) (8)
5e	Import (5); net income (6); disposable income on consumption (7); CPI (South Korea) (8); retail trade spending (9)
5f	Net income (6); disposable income on consumption (7); CPI (South Korea) (8); Retail trade spending (9); CPI (alcoholic and tobacco) (10)

Values in parenthesis represent the Y-procedure ranking [Table 1]. 5a to 5f refer to different combinations of five top ranked variables (see definition 2 and Table 3); 6a to 6d and 7a to 7d refer to combinations of six and seven top ranked variables

higher impact on the dependent variable than the set of the second five top ranked variables 5b and this trend is also true of 5c, 5d, 5e and 5f. We summarize our final result as follows:

Let I = Impact of set of variables

$$I(5a) > I(5b) > I(5c) > I(5d) > I(5e) > I(5f)$$

Thus for Australian GDP growth rate, it is best to focus on investment, current account balance, gross foreign liabilities, export and import which constitute the top five independent variables that have the highest impact on its GDP growth rate.

The above result is found to be true also for sets of six or seven variables.

5. CONCLUSION

The results show that for real data, it is difficult to assess the impact of individual independent variable (determinant) accurately. However, the proposed Y-procedure reveals that if the top ranked variables are considered in sets of five, six and seven, the set with the highest Y-average is the best set of high impact independent variables and this result is consistent for sets of six or seven variables. This result is in line with empirical economic models which normally have more than five independent variables. For our empirical analysis on Australian GDP growth rate, it is found that economic policy and allocation of funds should be given top priority for the top five independent variables that is investment, current account balance, gross foreign liabilities, export and import.

REFERENCES

Ait-Sahalia, Y., Brandt, M. (2001), Variable selection for portfolio choice. Journal of Finance, 56, 1297-1351.

Ang, A., Bekaert, G., Wei, M. (2007), Do macro variables, asset markets, or surveys forecast inflation better? Journal of Monetary Economics, 54(4), 1163-1212.

Banerjee, A., Marcellino, M. (2006), Are there any reliable leading indicators for US inflation and GDP growth? International Journal of Forecasting, 22(1), 137-151.

Bring, J. (1994), How to standardize regression coefficients. American Statistician, 48(3), 209-213.

Darlington, R.B. (1990), Regression and Linear Models. New York: McGraw-Hill.

Doukas, J.A., Kim, C.F., Pantzalis, C. (2006), Divergence of opinion and equity return. Journal of Financial and Quantitative Analysis, 41(3), 573-606.

Easterly, W., Kremer, M., Pritchett, L., Summers, L.H. (1993), Good policy or good luck? Country of growth performance and temporary shocks. Journal of Monetary Economics, 32, 459-483.

Franses, P.H., Dijk, D.V. (2000), Nonlinear Time Series Models in Empirical Finance. Cambridge: Cambridge University Press.

Granger, C.W.J., Newbold, P. (1986), Forecasting Economic Time Series. 2nd ed. San Diego: Academic Press.

Gregorio, J.D. (1991), Economic growth in Latin America. Journal of Development Economics, 39, 59-84.

Hansen, B.E. (2007), Least squares model averaging. Econometrica, 75, 1175-1189.

Kacperczyk, M., Sialm, C., Zheng, L. (2005), On the industry concentration of actively managed equity mutual funds. Journal of Finance, 60(4), 1983-2011.

Kruskal, W., Majors, R. (1989), Concepts of relative importance in recent scientific literature. American Statistician, 43(1), 2-6.

Leeb, H., Potscher, B.M. (2003), The finite sample distribution of post model selection estimators and uniform versus non-uniform approximations. Econometric Theory, 19, 100-142.

Leeb, H., Potscher, B.M. (2005), Model selection and inference: Facts and fiction. Econometric Theory, 21, 21-59.

Leeb, H., Potscher, B.M. (2006), Can one estimates the conditional distribution of post model selection estimators. The Annals of Statistics, 34, 2554-2591.

Mankiw, N.G., Romer, D., Weil, D.N. (1992), A contribution to the empirics of economic growth. The Quarterly Journal of Economics, 107(2), 407-437.

Pack, H. (1994), Endogenous growth theory: Intellectual appeal and empirical shortcomings. Journal of Economic Perspectives, 8(1), 55-72.

Petri, P.A. (1997), Common Foundations of East Asian Success. Lessons from East Asia – Ann Arbor: University of Michigan Press. p541-567.

Potscher, B.M. (1991), Effects of model selection on inference. Econometric Theory, 7, 163-185.

Ristanovic, V. (2010), Macroeconomic determinants of economic growth and world economic-financial crisis. Economics and Organization, 7(1), 17-33.

Romer, P.M. (1992), The origins of endogenous growth. Journal of Economic Perspectives, 8(1), 3-22.

Swan, T.W. (1956), Economic growth and capital accumulation. Economic Record, 32(2), 334-361.

Solow, R.M. (1956), A contribution to the theory of economic growth. The Quarterly Journal of Economics, 70(1), 65-94.

Thomas, V., Wang, Y. (1996), Distortions, interventions and productivity growth: Is East Asia different. Economic Development and Cultural Change, 44(2), 265-288.

Vojinovic, B. (2008), What is the Causality Between Economic Growth and its Major Determinants in the Case of Transition Economies. The 10th Bi-Annual EACES Conference of Pattern of Transition and New Agenda for Comparative Economics, Higher School Economics, Moscow: August 29-30, 2008.

Yuan, Z., Yang, Y. (2005), Combining linear regression models: When and how? Journal of the American Statistical Association, 100(472), 1202-1214.