ESA/STAT/AC. 223/S3.7

Third International Seminar on Early Warning and Business Cycle Indicators

17 – 19 November 2010
Moscow, Russian Federation

Construction of a Composite Leading Indicator for India

T.Rajeswari
Central Statistical Office India

Construction of a Composite Leading Indicator for India

T.Rajeswari

Affiliation : Central Statistical Office, Ministry of Statistics and Programme Implementation, National Accounts Division , New Delhi, India

Abstract

The Central Statistical office of India brings out quarterly estimates of GDP with a time-lag of 2-3 months and annual advance estimates with a time-lag of 10 months. The Central Statistical Office neither forecasts GDP growth rates nor have a composite leading indicator. In India at present there is no institutional mechanism for business cycle analysis/compilation of composite leading indicators as in the case of countries like UK, Australia, Canada, USA etc. Due to the changing structure of the Indian economy and its increased openness and market orientation, the need for a leading indicator is more relevant than ever before for alternate policy measures and also for analysing how policy actions are transmitted to activity levels. In the light of the above, this paper attempts to construct a composite leading indicator (CLI) for tracking the future path of GDP growth in India. The leading indicator can be broadly defined as a variable with meaningful economic linkage to a reference series whose turning points precede the turning points of the reference series. The reference series chosen in this paper is the real Quarterly Gross Domestic Product (QGDP). But, in view of the predominant role of weather and lower role of market forces in determining the agricultural output, non-agricultural GDP (in real terms) at quarterly frequency has been considered as the main reference series. The crux of the Leading Indicator approach to predict economic activity lies in the selection of leading indicators, whose movements precede the changes of the reference variable. Therefore, to cover a wide range of independent factors that determine economic activity, indicators from various broad sectors like the real sector, financial sector and external sector are considered. The indicators used for the study have been collected from the Central Statistical Office (CSO) and the Reserve Bank of India(RBI) database. Based on the cross-correlations analysis, potential indicators are selected with the appropriate lead period, for the reference series. Indicators have been chosen based on cross correlation in growth rates with the reference series and also cross correlation of cyclical components. The cyclical components of various time series are estimated by the Hodrick-Prescott filter. Regression based composite Index, with regression parameters as the weights for the composite index and ‘Principal Component’ Analysis (PCA) based Composite Index have been constructed. The regression based leading index has been constructed based on simple regression of the growth rates of reference series on other leading indicators(expressed in growth rates) that represent state of the economy. For deriving the principal components, the indicator series are suitably transformed. After the construction of the composite index , the performance of its out-of-sample forecasts is evaluated using distance measures like Root Mean Square

Error (RMSE) and Mean Absolute Percentage Errors (MAE). Performance mainly refers to the closeness of the predicted with the actuals. The results indicate that both the procedures seem to provide indices that reflect the reference series fairly well.

JEL classification codes: C13, C19**,** E32, E37

Keywords: Hodrick- Prescott filter, estimation, leading indicators, forecasting , statistical method

Author's E-Mail Address : **traji1962@yahoo.com**
r.thondiyil@nic.in

Construction of a Composite Leading Indicator for India[1]

T.Rajeswari
Central Statistical Office
Ministry of Statistics and Programme Implementation
National Accounts Division
New Delhi
India

I. Introduction:

A timely understanding of the direction of economic activity is essential for macroeconomic policy formulation. An unexpected weakening of the growth momentum would point to the need for appropriate change in policy, so as to avoid an overtly contractionary stance. On the flip side, an upside in economic performance would clearly warrant change of emphasis in the policy stimulus package. And in this context, if the growth estimates are only available after a considerable lag, it would surely delay the appropriate policy response. Often the growth estimates are available with considerable time lag. The Central Statistical office of India brings out quarterly estimates of GDP with a timelag of 2-3 months and annual advance estimates with a time-lag of 10 months. It is therefore essential to zero in on leading indicators, to respond to policy makers' needs for a reliable indication of economic activity in advance of release of statistical data. Such indicators are regularly used to closely track growth in the advanced economies.

The objective of a national statistical system is to provide relevant, comprehensive, accurate and objective statistical information which are invaluable for monitoring the country's economic and social conditions, planning and evaluation of government and private-sector programs and investments. To this extent it is very important to have a leading indicator. In large economies like India accurate macro economic policy formulation is crucial for sustained growth. In this regard an early understanding of business cycle is essential . Leading indicator approach to understanding the business cycle requires an information base as it involves combining several statistical series. The NSO's can make available the necessary information required, from the exisiting statistical infrastructure or if necessary by upgrading them. My view is that if there has to be an official CLI for use of the government then NSO's could step in. Since in many economies Central Banks also play a very crucial role, a coordinated effort in this direction would pave the way for an official CLI. Then again the crucial issue here is of educating

[1] The views expressed in this Paper are those of the authors' alone and not of the institution to which they belong.

the public on the CLI – its compilation methodology, data sets used etc and statistical advocacy. In India sentiment indicators like Business Confidence index is compiled and released by organisations like Confederation of Indian Industries(CII) and NCAER . My personal view is that NSO's role in compilation of leading indicators which requires statistical data (which are available with the NSO's through regular annual/quarterly surveys) is thus crucial.

In this backdrop, this paper presents an estimation of composite leading indicators(CLI) for India. Two different methodologies have been considered for constructing the composite index. One is the regression based composite Index, with regression parameters as the weights for the composite index and where the regression analysis has been done with respect to the original growth series. The other method is based on Principal Components Analysis. After the construction of the composite index the performance of its out-of-sample forecasts is evaluated. To test the forecasting accuracy of the composite indicator, the usual distance measures like Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) have been used.

The scheme of the paper is as follows. Section II presents some highlights of Indian experience in construction of composite leading indicators. Section III discusses the approach to construct a composite leading indicator. In Section IV selection of leading indicators is discussed. Section V discusses the compilation of composite leading indicators and the efficiency of proposed CLI in terms of out of sample error in forecasting. Section VI discusses the availability and limitations of relevant economic indicators in the Indian context and further efforts to be made to aid the development of reliable leading indicators for Indian economy.

II. Highlights of Indian experience

Attempts to understand the features of business cycles and forecast their movements have been carried out in India by several researchers. As the objective of the present study is to develop a composite leading indicator(CLI), literature review is restricted to select studies on CLI. Dua and Banerji (2001), identified leading indicators and constructed a CLI index designed to anticipate business cycle and growth rate cycle upturns and downturns. Chitre (2001) studied the business cycles in India for the period 1951-1982 and, inter-alia, presented a list of leading, coincident and lagging indicators (at peaks and at troughs) and the turning points. Mall (1999) studied the cyclical behaviour of output variables such as real GDP, non-agricultural GDP, GDP from manufacturing, trade, Index of Industrial production(IIP), index of sales of private corporate sector, etc. and has concluded that non-agricultural GDP can be taken as a reference series for tracking business cycles in India. Using spectral analysis method, he constructed a composite index of leading indicators to forecast cyclical movements in IIP from manufacturing sector.

A Working Group of the Reserve Bank of India (2002) on economic Indicators examined the information base for the analysis of business cycles and explored the leading indicators approach for study of business cycles and forecasting. The group suggested that the quarterly time series of non-agricultural GDP can be considered as the main reference series. This was in view of the fact that performance of agricultural sector is dependent on weather, rainfall, etc. and the relatively low inter-play of market forces in determining their levels. Also due to a shift in the composition of Indian GDP with declining share of agriculture industry, the report recommends non-agricultural GDP as a reference series. The Report presents a composite index constructed based on principal component analysis by considering the IIP as the reference series. Six series viz., Narrow money (M1), Non-food credit, Whole Sale Price Index(WPI) raw materials, production of coal and aluminium, and rail good traffic originated have been identified as leading indicators.

OECD has developed a CLI for India with reference to the monthly industrial production as the reference indicator for the growth cycle analysis. From a set of 30 economic indicators considered initially, eight economic indicators, *viz.*, Business Confidence Index[3], imports, Money Supply, Exchange rate (Indian Rupee per US Dollar), Deposit interest rate, stock prices (Bombay Stock Exchange SENSEX based on 30 scrips), production of basic goods and production of intermediate goods were identified as the leading indicators for constructing the CLI. The performance of the CLI is evaluated over the period since 1995 for which consistent industrial production data are available and it was found to have a median lead of four months at all turning points over the two cycles registered in industrial production since 1995.

III . Approach to construct CLI – A Brief Description

As the focus of leading indicator approach is to forecast the fluctuations in business activity, the starting point of the analysis would be the selection of a reference series(the economic variable whose cyclical movements are intended to be studied). If one is interested to concentrate on a single series, then the Gross Domestic Product (GDP) may be taken as the reference series, as it covers almost all the economic activities in the economy. Considering the volatility of agriculture and its allied sectors, non-agricultural GDP (in real terms) has been taken as the main reference series as recommended by Working Group of Reserve Bank of India. In view of the importance of short-term economic forecasting for policymakers, it is considered appropriate to consider compilation of leading indicators with periodicity of less than one year. In this paper quarterly data been considered for studying the lead-lag relationship.

Next step is to identify some basic series by optimizing some criterion that measures the strength in leading relationship between basic series and the reference series. Leading indicators are the time series whose cyclical component precedes those in the reference series systematically and hence they are expected to provide useful information regarding the future movement in the reference series. The choice of economic indicators is a very critical issue as the performance of ultimate forecasts depends heavily on the quality of individual indicator. Indicators from various broad sectors (say, real sector, financial sector, government sector, external sector etc.) were considered so as to cover a wide range of independent factors that determine economic activity. Information content of these indicators could then summarized into a composite index.

In the literature, there exist three different approaches to analyse the business cycles, *viz.*, Classical Business Cycles, Growth Cycles and Growth Rate Cycles. A major limitation of the classical business cycle approach, is that it fails to capture the real facet of economic activity when the economy goes through frequent alternating periods of accelerated and decelerated growth. A growth cycle tracks the upswings and downswings through deviations of the actual growth rate of the economy from its long-run trend rate of growth. One basic problem associated with the analysis of growth cycles is the determination of the trend component from the time series. Also, different de-trending methods may generate different growth cycle chronologies. Therefore, growth rate cycles have been analysed in this study. By calculating growth rates, the problems involved in removing trend do not arise, as generally growth rates are stationary. The present work has been confined to the annual (point-to-point) growth rates.

To get an accurate signal, out of the several variables possessing information about the future movement of economic growth, it is advisable to rely on a reasonably diversified group of leading indicators with demonstrated predictive potential. Thus, an index composed of several of these leading series, selected from a variety of economic processes, may provide a better indication of future activity than any one particular series. The selected series are then combined as a weighted index ($W_1X_1+ W_2X_2+ W_3X_3+\ldots+ W_nX_n$) to generate the Composite Leading Indicator (CLI).

One approach to obtain weights is by using the principal component analysis (PCA) for summarising the information content of various leading indicators. The PCA is, however, a purely statistical procedure that yields one or more linear combinations of the series that explain major parts of variations. An alternative to this method is regression based composite Index, with regression parameters as the weights for the composite index and where the regression analysis is been done with respect to the original growth series. It was found that the application of seasonal adjustment filters to the time series generated different growth rates. Also, the finding of Ghysels and Perron (1993) shows that consistency of the OLS estimates is not preserved with lagged dependent variables, when all variables are seasonally adjusted . Hence in this paper, unadjusted data series is used for OLS regression.

After the construction of the composite index next task is to evaluate the performance of its out-of-sample forecasts of the reference series. Performance mainly refers to the closeness of the predicted with the actuals. To test the forecasting accuracy of the composite indicator, we can use the usual distance measures like Root Mean Square Errors (RMSE), Mean Absolute Error (MAE), Mean Square Error (MSE) and Mean Absolute Percentage Errors (MAPE) can be used. However, for our analysis we confine to RMSE and MAE as performance criteria.

IV Selection of Leading Indicators

A leading indicator, is said to be leading by k time periods, when information of the series composing it upto time t–k is required to forecast economic activity at time t. As different variables affect economic activity with different lags, selection of variables that would form the indicators is crucially dependent on it. Based on the criteria of availability of quarterly and a consistent time series data for at least fifteen years, a wide variety of indicators covering all the sectors were considered for the present study. The list of indicators considered in the study has been provided in **Appendix – I**. The data used for the study have been collected from the Central Statistical Office (CSO) and the Reserve Bank of India(RBI) database. Initially 33 indicators were chosen covering the five sectors, *viz.*, monetary, banking, financial market, real sector, and external sector. The sample covers the period – Q1 1994-95 to Q4 2009-10 . The entire data set, ranging from Q1 1994-95 to Q4 2009-10, has been divided into two subsets, viz., estimation set and validation set. The estimation set has been taken from Q1 1994-95 to Q4 2007-08 and validation set has been considered from Q1 2008-09 to Q4 2009-10.

A preliminary exercise in constructing CLIs generally involves exploring the relationship between the cyclical components of the reference series with a large number of other series perceived to be possible leading indicators. TABLE 1 presents the cross-correlation in growth rates of leading indicators with the reference series and cross correlation of cyclical components estimated by the Hodrick-Prescott filter with the smoothening parameter as 1600 (as the time series under consideration are quarterly). Only variables which showed significant correlation with indicator variables are presented here.

TABLE 1

Cross-Correlation in growth rates of leading indicators with the reference series

Variable	Growth rates	HP-Filter
IIPBG	0.63	0.43
IIPCG	0.70	0.67
IIPCONG	0.51	0.33
IIPGEN	0.76	0.57
IMP	0.56	0.39
BSE SENSEX	0.63	0.61
BC	0.69	0.67
DEP	0.56	0.51
CP	0.38	0.48
MO	0.56	0.63
M1	0.72	0.73
WPI ELEC	-0.61	-0.62
WPIMANU	0.24	0.30

Cross Correlation, which essentially is simple product moment correlation between the indicator and the target series for various leads, projects the quality of the indicators as potential leading Indicators. **Appendix - 2** presents cross correlation for all 13 indicators presented in TABLE 1. The lead periods of different indicators considered are given in TABLE 2 below. A lead period of at least six months (2 quarters) have been considered for the construction of composite index.

TABLE 2
Lead periods for the leading indicators

Leading indicator series	Correlation coefficient	Lead associated with maximum correlation in magnitude(in quarters)@
IIPCG	0.70	3 (0.62)
IIPBG	0.63	2 (0.57)
IIPCONG	0.51	2 (0.43)
IIPGEN	0.76	2 (0.62)
IMP	0.56	3 (0.35)
BSE	0.63	2 (0.55)
BC	0.69	3 (0.66)
DEP	0.56	2 (0.44)
CP	0.38	2 (0.41)
M0	0.56	2 (0.46)
M1	0.72	2 (0.67)
WPI ELEC	-0.61	3 (-0.53)
WPIMANU*	0.24	2 (0.08)

* When Augmented Dickey Fuller test (ADF) unit root test was performed on the residuals estimated from cointegrating regression (QGDP on WPIMANU), the variable WPI MANU was found to be cointegrated with QGDP.

@ Figures in brackets indicate the correlation coefficient associated with the lead.

V .Compilation of Composite Leading Index[2]

Through principal component analysis one essentially tries to explain major part of the information content of the multivariate (multiple series) data without following a model set a priori with smallest number of uncorrelated components. The objective is to explain the variance of the observed data through a few linear combinations of the original data which is a weighted sum of all indicator series. If there are K series, K principal components can be obtained satisfying two conditions ,viz., (i) they are uncorrelated to each other (orthogonal); (ii) the first principal component accounts for the maximum possible proportion of the variance of data set , the second principal component accounts for the maximum of the remaining variance, and so on until the last of the principal

[2] All calculations for this paper were done using STATA Software.

components absorbs all the remaining variance not accounted for by the preceding components. Principal components based on indicators Xl(t)'s, l=1,2,… ,k can be expressed as Pj(t) = wj1 X1(t) + wj2 X2(t) + … .. + wjp Xp(t); j =1,2, … .k, where Pj(t) is the j-th principal component and wjl,s l=1,2,… ,k are the coefficients, known as factor loadings, of l-th indicator in j-th principal component. In practice first (or at least first few) principal component(s) normally captures sufficient information to represent the multivariate data. Let k leading indicators are selected to form the composite index and lj, j=1,2,… ,k; be the lead period of jth indicator/series Xj(t), which is suitably transformed. Most of the basic series show some trend in their mean process and therefore, they are passed through a transformation. In this paper, the time series are transformed by considering the growth rate cycle. All transformed series are normalized to give them a common span, such that they are expressed in comparable scale. The normalization of any series X(t) is done by using the formula : X*(t)=[X(t)-Min X(t)]/[Max X(t)-Min X(t)]. All calculations and model estimation is carried out using the normalized series. Thus to assess the prospect of the target series (Non-Agri QGDP) at time point t, one has to combine Xj(t-lj), j=1,2, … ,k values. Therefore, principal components (PCs) may be derived based on past information on Xj(t-lj)'s. Composite index is then constructed by regressing target series on a few PCs. The PCs have been chosen based on the out-of-sample forecast performance of the PCs. While dealing with a number of correlated variables the problem of multicollinearity arises in regression analysis. The advantage of PCA method is that as PCs are uncorrelated to each other, while using them as regressors in regression analysis, one will not face such a problem. Secondly, by dropping few PCs , we are also eliminating a part of the noise components from the data which may yield more reliable estimates. Based on the out of sample forecasts performance first 5 PCs were chosen and the results are given in the table below. The PC based CLIs have been constructed based on three different sets of indicators.

Model 1 : Indicators used are IIP CG, IIPCONG, IIPGen, BC, IMP, Mo, M1 and WPI MANU, CP

ΔCLI(t)=5.31+3.45*ΔtIIPCG(t-3)+0.74*ΔtIIPCONG(t-2)+2.68*ΔtIIPGEN(t-2) +0.30*ΔtBC(t-3)+1.02*ΔtIMP(t-3)+0.12*ΔtMo(t-2)+1.90*ΔtM1(t-2)- 1.34*ΔtWPIMANU(t-2) - 1.05ΔtCP(t-2)

Model II: indicators used are IIPBG, IIP CG, IIPGen, IMP, Mo, WPI ELEC and DEP
ΔCLI(t)=5.54+2.08*ΔtIIPBG(t-2)+2.23*ΔtIIPCG(t-3)+2.76*ΔtIIPGEN(t-2) +0.88*ΔtIMP(t-3)+0.74*ΔtMo(t-2)- 2.21*ΔtWPIELEC(t-3) - 0.90ΔtDEP(t-2)

Model III: Indicators used are IIP CG, IIPGen, IMP, BC, CP , M1 ,WPIMANU
ΔCLI(t)=6.29+1.76*ΔtIIPCG(t-3)+ 2.95*ΔtIIPGEN(t-2) +1.79*ΔtBC(t-3)+1.13*ΔtIMP(t-3)+ 2.55*ΔtM1(t-2)-0.35*ΔtWPIMANU(t-2) -1.39ΔtCP(t-2)

Model IV: Indicators used are IIP BG, IIPCG, IIPGen, IMP, M1 ,WPIELEC

ΔCLI(t)=3.89+2.78*ΔtIIPBG(t-2)+2.02*ΔtIIPCG(t-3)+ 2.52*ΔtIIPGEN(t-2) +1.62*ΔtIMP(t-3)+ 1.41*ΔtM1(t-2)-1.0*ΔtWPIELEC(t-3)

In all the above cases tX(t) refers to transformation of reference series. TABLE 3 below gives the Actuals and forecasts of Non-Agri QGDP using models discussed above. Forecast performances are given in TABLE 4 below. Results are also depicted in the Figure 1.

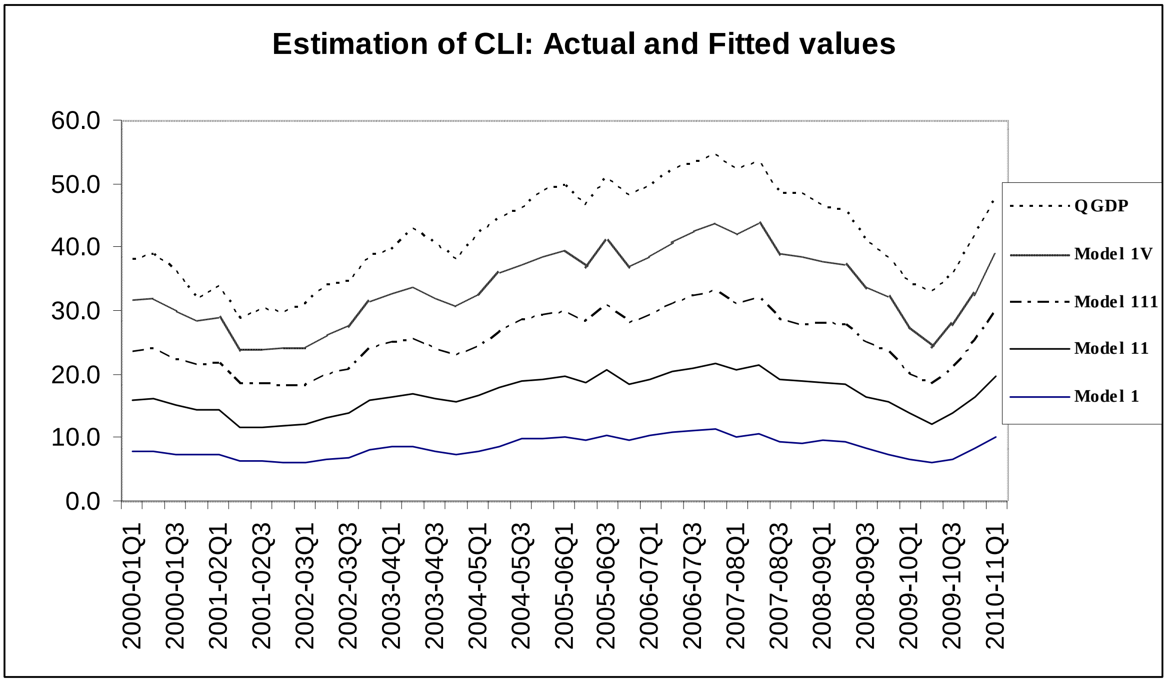
TABLE 3
Actuals Vs Forecasts

	FORECASTS				
	Model 1	**Model 11**	**Model 111**	**Model IV**	**Actuals**
2008-09Q1	9.5	9.1	9.3	9.5	8.8
2008-09Q2	9.3	9.0	9.3	9.3	8.6
2008-09Q3	8.3	8.2	8.6	8.4	7.7
2008-09Q4	7.3	8.3	7.8	8.7	6.4
2009-10Q1	6.5	7.3	6.1	7.4	6.9
2009-10Q2	6.1	6.1	6.2	5.7	9.0
2009-10Q3	6.6	7.3	6.9	7.0	7.8
2009-10Q4	8.3	8.1	8.7	7.4	9.2
2010-11Q1	10.0	9.6	10.2	9.0	9.1

TABLE 4
Forecast Performance

Forecast Performance Criteria	**Model 1**	**Model 11**	**Model 111**	**Model IV**
RMSE	1.24	1.27	1.25	1.55
MAE	1.01	0.98	1.03	1.10

FIGURE 1



From Figure 1 it may be seen that Model IV closely mimics the reference series. Also the signs of coefficients of indicator series in this model are consistent with economic theory.

For the application of a regression model, it is important that the indicators possess stationary property. Thus, all the indicators (annual point-to-point growth rates) are first examined for stationarity. The conventional Augmented Dickey Fuller test is applied to test for stationarity. Some indicator series were not found to be stationary. In such cases, cointegration tests were applied to avoid spurious regressions and also to check if there exists long term or equilibrium relationship between the indicator variable and reference series. A simple method was applied for testing cointegration. Augmented Dickey Fuller (ADF) unit root test was performed on the residuals estimated from cointegrating regressions of QGDP with each of these indicator variables. The annual growth rate of the QGDP was found to be stationary. The findings related to the stationarity property of the indicators are provided in TABLE 5

TABLE 5

Time Series properties of the variables: Augmented Dickey-Fuller Unit Root Test

Variable	DF statistics
IIPCG	-4.130***
IIPBG#	-3.057**
IIPCONG#	-3.505**
IIPGEN	-4.278***
IMP	-3.013**
BC	-3.403**
CP	-2.731*
M1	-4.464***

WPI ELEC #	-3. 643***
WPIMANU	-2.783*

indicator by itself was stationary.

***, ** and * indicates significance at 1%, 5% and 10% levels of significance respectively.

The OLS regressions were performed on the variables which were found to be stationary or cointegrated with the reference series. Two models are presented here based on the efficiency of performance of the leading indicators. Variables considered in Model 1 is same as in Model IV of PCA based CLI.

Model 1 : Indicators used are IIP BG, IIPCG, IIPGen, IMP, M1 ,WPIELEC (model IV of CLI based on PCA)

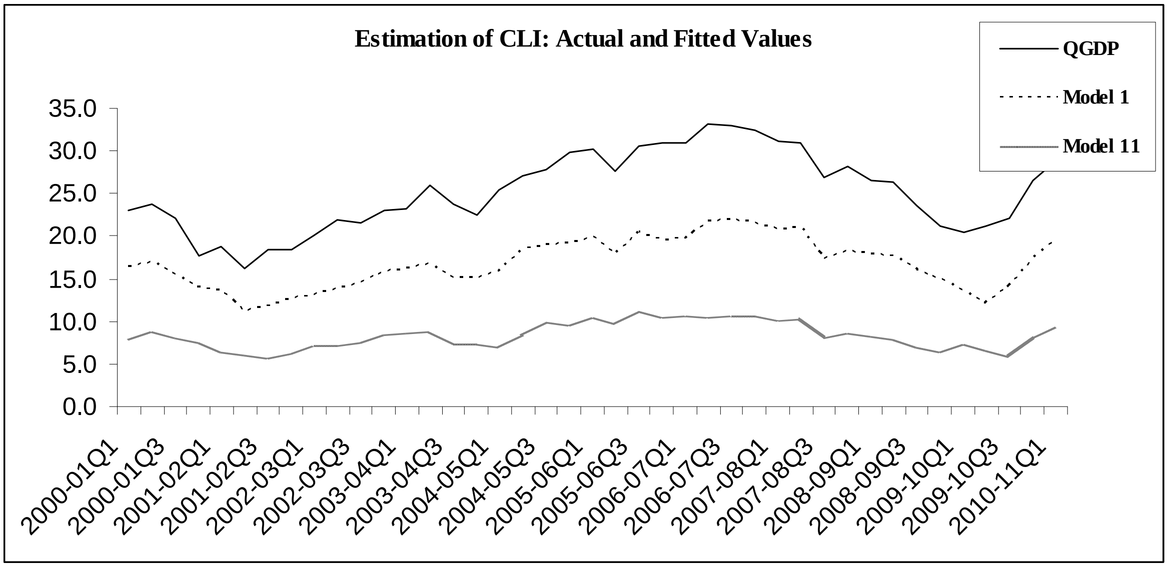
ΔCLI(t)=3.69 +0.16*ΔIIPBG(t-2)+0.03*ΔIIPCG(t-3)+ 0.096*ΔIIPGEN(t-2) -0.02*ΔIMP(t-3)+ 0.23*ΔM1(t-2)-0.04*ΔWPIELEC(t-3)

Model II: CLI based on OLS Regression. Indicators used are , IIP CG, IIPGen, IMP,BC CP , M1 ,WPI MANU

ΔCLI(t)=5.49+0.03*ΔIIPCG(t-3)+ 0.26*ΔIIPGEN(t-2) +0.01*ΔBC(t-3)+0.05*ΔIMP(t-3)+ 0.12*ΔM1(t-2)-0.37*ΔWPIMANU(t-2) – 0.05ΔtCP(t-2)

Results are also depicted in the Figure 11. It may be seen that Model I closely mimics the reference series as in the case of PCA based CLI. Regression results are given in **Appendix 3.**

FIGURE 11



Forecast performances of the two regression models are given in below. .

TABLE 6
Actuals Vs Forecasts

	FORECASTS		
	Model 1	**Model 11**	**Actuals**
2008-09Q1	9.5	8.3	8.8
2008-09Q2	9.7	8.0	8.6
2008-09Q3	8.9	7.0	7.7
2008-09Q4	8.3	6.5	6.4
2009-10Q1	6.2	7.4	6.9
2009-10Q2	5.6	6.6	9.0
2009-10Q3	8.3	6.0	7.8
2009-10Q4	9.2	8.0	9.2
2010-11Q1	10.4	9.4	9.1

TABLE 7
Forecast Performance

Forecast Performance Criteria	**Model 1**	**Model 11**
RMSE	1.53	1.15
MAE	1.21	0.90

The results indicate that RMSE and MAE is quite low in the case of PCA based CLI as well as regression based CLI. In all the models estimated, it was found that forecasts for 2009-10 Q2 exhibited a large variation from actuals. This is due to the turnaround in the growth momentum in the Indian Economy in Q2 2009-10 estimates. This was on account of the continued fiscal expansion, and in particular with the release of 60 per cent of the Sixth Pay Commission arrears in September 2009, Community, Social and Personal services recorded a significant pick-up in growth. Understandably, the regular leading indicators could not capture this growth. Apart from this, the leading indicator provided results with low RMSE. Both the methods adopted have yielded CLI which is quite efficient in generating out of sample forecasts. The plots include predicted values from 2000-01 Q1 to 2010-11 Q1. It is observed that predicted values estimated by all the models move in the same direction as the reference series.

VI Conclusions

It could be concluded that the constructed composite indicator can be used to generate forecasts of QGDP 2-quarters in advance(out of sample forecasts). For the purpose of assessing the accuracy in forecasts, 2 quarters ahead forecasts of QGDP for time points in validation set (Q1 2008-09 to Q4 2009-10) data in the estimation set was used and distance measure like Root Mean square error and Mean absolute error was estimated. The methods adopted in this paper can provide a composite leading indicator that is quite efficient with low RMSE and MAE.

A major problem in the Indian context is non-availability of time series data on many conventional leading indicators. Some of the conventional variables are not presently being compiled at reasonable levels of aggregation, such as the labour working hours, vendor deliveries, new orders or order book, overtime hours. Housing Starts (building permits for new private housing units) is another strong traditional leading indicator used in developed economies where there is no organised database available in the Indian context. But some efforts in this direction has been initiated by National Buildings Organisation (Ministry of Housing and Urban Poverty Alleviation) in collaboration with RBI. Employment and wages are extremely important variables that characterise business cycles in developed countries. In the Indian context, there is no regular flow of data on employment / unemployment at aggregate economy level at quarterly or monthly frequency that can be used for business cycle analysis. Also there is no consolidated information base on average weekly manufacturing hours – considered as a sound leading indicator in many developed countries. Another major impediment for this type of analysis is the non-availability of time series information on features like consumer expectations and business tendency. This kind of data gaps needs to be addressed.

The out of sample forecasts indicate that CLI performs quite well in capturing future path and turning points of QGDP growth rate as the Root Mean Square Error in 2 quarters ahead forecast is within reasonable limits. However further improvement may be achieved by examining more leading indicators. It has also been observed that Indian economy is continually evolving and far too complex to be summarized in a single reference series. As such there is a need to identify other series also for determining reference cycle turning points. There is a need for theoretical research for a better understanding of the complexities in Indian context.

Appendix 1

Initially selected 33 Indicators based on Economic Considerations

Variable Name	
Index of Industrial Production(IIP) Basic Goods	IIPBG
IIP Capital Goods	IIPCG
IIP consumer Goods	IIPCONG
IIP Gen	IIPGEN
Imports	IMP
BSE SENSEX	SEN
Bank credit	BC
Aggregate Deposit	DEP
Currency with public	CP
Broad Money	M3
Narrow Money	M1
Reserve Money	Mo
Wholesale Price Index (All Commodities)	WPIALL
WPI Manufactured Products	WPIMANU
WPI Industrial Raw Material	WPIRM
WPI Electricity	WPIELEC
IIP Electricity	IIPELEC
Non-food credit	NFC
Production of Cement	CEMPRD
Consumer Price index(Industrial Worker)	CPIIW
Call money rate	CMR
Exchange rate (Dollar Vs. Rs.)	EXR
Exports	EXP
Inventory Manufacturing	INVMANU
Gold Price	GP
WPI Food Article	WPIFD
Fiscal deficit	FDEF
Investment in government securities	INVG
Investment in other approved securities	INVA
Investments by FIIS in the Indian capital market	FIIINV
Cash with Banks	CB
Yield of SGL transactions in treasury bills for residual maturities	Yield 19-51
Private Final Consumption Expenditure	PFCE

Appendix -2

Cross-Correlation in Growth Rates with Non-AgriQGDP-lags

	Lead(-)/ Lag (+) in months																
LAG	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8
IIPCG	0.22	0.32	0.39	0.49	0.60	0.62	0.59	0.68	0.70	0.74	0.74	0.60	0.48	0.36	0.27	0.23	0.21
IIPBG	0.17	0.14	0.12	0.26	0.32	0.41	0.57	0.60	0.63	0.59	0.44	0.40	0.41	0.30	0.34	0.24	0.13
IIPCONG	0.45	0.55	0.50	0.34	0.39	0.38	0.43	0.48	0.51	0.36	0.20	0.16	0.17	0.07	0.04	0.01	-0.06
IIPGEN	0.23	0.29	0.30	0.32	0.43	0.54	0.62	0.71	0.76	0.65	0.53	0.48	0.44	0.34	0.26	0.16	0.06
IMP	0.34	0.35	0.43	0.42	0.39	0.35	0.32	0.40	0.56	0.55	0.44	0.34	0.22	0.22	0.34	0.28	0.22
BSE	0.21	0.23	0.20	0.22	0.34	0.45	0.55	0.64	0.63	0.53	0.41	0.29	0.25	0.28	0.24	0.15	0.02
BC	0.63	0.67	0.74	0.75	0.86	0.91	0.90	0.81	0.69	0.51	0.37	0.38	0.35	0.34	0.27	0.09	-0.06
DEP	-0.06	0.00	0.05	0.10	0.21	0.32	0.44	0.53	0.56	0.55	0.61	0.66	0.65	0.68	0.65	0.65	0.59
CP	0.00	0.05	0.10	0.19	0.33	0.39	0.42	0.45	0.38	0.36	0.37	0.38	0.41	0.50	0.45	0.45	0.43
M0	0.04	0.13	0.16	0.23	0.30	0.33	0.46	0.50	0.56	0.58	0.60	0.63	0.70	0.70	0.64	0.56	0.48
M1	0.27	0.35	0.43	0.47	0.57	0.60	0.67	0.72	0.72	0.68	0.64	0.56	0.54	0.50	0.43	0.36	0.25
WPI ELEC	-0.25	-0.33	-0.46	-0.48	-0.51	-0.53	-0.52	-0.57	-0.61	-0.58	-0.56	-0.50	-0.39	-0.39	-0.41	-0.35	-0.31
WPIMANU	0.18	0.25	0.30	0.30	0.26	0.14	0.08	0.13	0.24	0.37	0.46	0.47	0.44	0.34	0.28	-0.59	-0.71

Appendix -3

Regression Results
MODEL 1

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
					R-squared	0.5878
					Adj R-squared	0.5328
					Root MSE	1.2886
					DW	1.35
IIPBG	.1556331	.1344553	1.16	0.253	-0.1151737	0.42644
IIPCG	.0269402	.0397555	0.68	0.501	-0.0531315	0.1070118
IIPGEN	.0958349	.1515459	0.63	0.530	-0.2093943	0.4010641
IMP	-.0156588	.0125961	-1.24	0.220	-0.0410286	0.009711
M1	.2315478	.0866264	2.67	0.010	0.0570734	0.4060222
WPIELEC	-.0423996	.0317918	-1.33	0.189	-0.1064316	0.0216324
CONS	3.694272	1.067414	3.46	0.001	1.544391	5.844154

MODEL 11

Variable	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
				R-squared		0.6281
				Adj R-squared		0.5662
				Root MSE		1.2565
				DW		1.71
IIPCG	0.032651	.0309298	1.06	0.297	-.0297678	0.09507
IIPGEN	0.259354	.0998396	2.60	0.013	.0578693	0.460838
IMP	0.050801	.0154208	3.29	0.002	.0196801	0.081921
BANKCREDIT	0.011053	0.045767	0.24	0.81	-.0813089	0.103414
CP	-0.05475	.0286367	-1.91	0.063	-.1125388	0.003044
M1	0.123165	.0566201	2.18	0.035	.0089007	0.237429
WPIMANU	-.3685978	0.1484975	-2.48	0.017	-0.0689178	-0.06892
CONS	5.495374	.1385471	10.20	0.000	4.408542	6.582206

[In this model all the variables except bank credit are significant. But when this variable was dropped the DW statistic deteriorated. Together these variables explain about 60 percent variations in the growth rate of reference series.]

References

Chitre, V. S. (2001)," Indicators of Business Recessions and Revivals in India :1951-82", Indian Economic Review , Vol. XXXVI, no. 1 ,2001.

Canova, Fabio (1998): "Detrending and Business Cycle Facts", *Journal of Monetary Economics.*

Centre for Monitoring Indian Economy, *Monthly Economic Review*, Various Issues.

Central Statistical Office, India (2010): National Accounts Statistics, 2010

Dua, P. and A. Banerji (1991) , " An Index of Coincident Economic Indicators for the Indian Economy", Journal of Quantitative Economics, Vol. 15, No. 2

____(2001) ,"An Indicator Approach to Business and Growth Rate Cycles : The case of India ", Indian Economic Review, Vol. 36, No. 1 .

Ghysels, E. and P.Perron (1993): Effects of Seasonal Adjustment Filters on tests for a Unit Root", *Journal of Econometrics*, Vol.55.

IMF Working Paper ,2003, "Estimating Indexes of Coincident and Leading Indicators: An Application to Jordan": *Joannes Mongardini and Tahsin Saadi-Sedik,* Available via the Internet: www.**imf.org**/external/pubs/ft/wp/2003/wp03170.pdf

Mall O.P. (1999),"Composite Index of Leading Indicators for Business Cycles in India ", RBI Occasional Papers , Vol 20, no. 3.

Organization for Economic Cooperation and Development, 1998, "Composite Leading Indicators: a Tool for Short-Term Analysis. Available via the Internet: http://www.oecd.org/pdf/M00003000/M00003226.pdf

OECD (2005), Composite Leading Indicators for Major OECD Non-Member Countries Selection of Leading Indicators and Construction of Composite Indicators: Brazil, China, India, Indonesia, Russian Federation and South Africa, OECD, Paris.

Reserve Bank of India (2002), Report of the Working Group on Economic Indicators