

W P S (DEPR) : 3 / 2011

RBI WORKING PAPER SERIES

Real Time Business Conditions Index: A Statistically Optimal Framework for India

Dipankar Biswas
Nivedita Banerjee
Abhiman Das



DEPARTMENT OF ECONOMIC AND POLICY RESEARCH

APRIL 2011

Real Time Business Conditions Index: A Statistically Optimal Framework for India

Dipankar Biswas*
([E-mail](#))

Nivedita Banerjee
([E-mail](#))

and

Abhiman Das
([E-mail](#))

Abstract

The state of the real economy evolves in a continuous fashion. Economic agents and policy makers, while making decisions in real time, require accurate and timely understanding of the state of real activity. In the light of the changing nature of the economy where increasingly more and more activities are being channelised through both organised and unorganized business sectors, the measurement of business condition on real time basis is of paramount importance. In this context, in order to achieve an accurate and timely estimate of the state of real activity in a systematic, replicable and statistically optimal manner, this paper proposes a framework to construct a real-time business conditions index for India. The study is primarily motivated by the seminal work of Aruoba, Diebold and Scotti (2009), for the high frequency business conditions assessment for the U.S. economy. Based on various economic indicators measured at different frequencies, this paper develops a real-time business conditions index for India following a dynamic factor model framework for extracting signals from continuously evolving states. A Kalman filter routine is used for signal extraction from state-space representation as well as evaluation of likelihood function. Empirical results show that this coincident indicator tracks the overall economic activity reasonably well.

JEL Classification: C61, E32, E37

Key Words: Business cycle; Dynamic factor model; Turning points; State-space model; Expansion

* The authors are working in the Department of Statistics and Information Management, Reserve Bank of India, Mumbai. Views expressed in the paper are those of the authors and not of the Reserve Bank of India.

Real Time Business Conditions Index: A Statistically Optimal Framework for India

1. Introduction

The state of the real economy of a country evolves in a continuous fashion. Economic agents and policy makers, making decisions, in real time, require accurate and timely estimates of the state of real activity. In the light of the changing nature of the economy, where more and more activities are being channelised through both organized and unorganized business sectors, the assessment of business condition on real time basis is of paramount importance, particularly for central banks. From mid-1980s until 1998, the Reserve Bank of India (RBI) used a monetary-targeting framework. In the year 1998, the RBI's Working Group on Money Supply, in its report, pointed out that monetary policy exclusively based on money demand could lack precision and hence, it was necessary to monitor a set of additional indicators for monetary policy formulation. Accordingly, the RBI adopted a multiple indicator approach from 1998 wherein, besides monetary aggregates, information pertaining to currency, credit, fiscal position, merchandise trade, capital flows, inflation rate, exchange rate, refinancing and transactions in foreign exchange etc., were juxtaposed with data on output and the real sector activity for drawing policy perspectives. The widening range of variables monitored and studying their dynamic interactions are now possible partly because of the development of more sophisticated econometric models. In this context, in 2002, the RBI's Working Group of Economic Indicators provided importance to deal with the business cycle analysis and to construct a composite index of leading indicators of Indian economy. In 2007, the RBI's Working Group of Leading Indicators for Indian Economy, in its report, recommended two series, viz., monthly Index of Industrial Production (IIP) and quarterly Non-Agricultural GDP, as the reference frame of business cycle in India. The Group also constructed Composite Index of Leading Indicators (CILI) for each of these two reference series following international best practices. As proposed by the Group, the outlook for business cycle movement for 2-3 quarters ahead is regularly examined internally in RBI and serves as an important input to the monetary policy making.

It has been, however, observed that the proposal of the Working Group of Leading Indicators to provide an outlook for business condition of the Indian economy is not sufficient on real time basis due to the following reasons. Firstly, most frequent data used for developing leading indicators is observed on monthly basis. For real time measurement, moving beyond the monthly frequency is a basic pre-requisite. Some important indicators (e.g., asset prices, yield curve term premium) are observed at daily frequency which potentially contains important information on the overall economic activity. Secondly, the report did not take into account the assumption of continuously evolving state of the economy, which is essential to real time measurement. Lastly, the provisional and partially revised data

used for the leading index also affects the performance to predict future movements in aggregate economic activity in the real-time framework (Diebold and Rudebusch, 1991).

Against this backdrop, we propose a framework motivated by the earlier work of Aruoba, Diebold and Scotti (2009), for the high frequency business conditions assessment for India in a systematic, replicable and statistically optimal manner. Giving the latest information of various macroeconomic indicators of different frequencies, our objective is to assess the current state of economic activity based on a real-time index and to update our assessment as more information flows in. Our assessment is as on today, and not beyond it. In that sense, the index is coincident (not leading) to the business condition.

The paper is organised as follows: Section 2 reviews the literature on the real-time data analysis. Section 3 describes the empirical analysis concerned with the development of real-time business conditions index for Indian economy. The description of software used for empirical analysis is mentioned in Section 4. Finally, Section 5 summarises the results, with a few concluding remarks.

2. Literature Review

In empirical econometrics, the use of real-time data is not a recent area of study. A long literature can be mentioned in this regard. Early studies of real-time data focused on the sensitivity of certain statistics to data vintage. Gartaganis and Goldberger (1955) did the first work on real-time data analysis. They mainly confined themselves to the properties of statistical discrepancy between Gross National Product (GNP) and gross national income in United States, after data were revised in 1954. Howrey (1978) focused on the use of preliminary data in econometric forecasting and indicated clearly that the intelligent use of preliminary data would be expected to result in a meaningful reduction in prediction error variances. Diebold and Rudebusch (1991) examined the ability of composite index of leading economic indicators to predict future movements in aggregate economic activities based on real-time analysis. They used the provisional and partially revised data for the leading index that were actually available historically, along with recursive out-of-sample forecasts. They found substantial deterioration of forecasting performance in the real-time framework. Orphanides and Simon van Norden (2002) examined the reliability of several detrending methods for estimating the output-gap in real time. They focused on the extent to which output-gap estimates were updated over time as more information arrived and data were revised. They suggested that, great caution would be required for measuring output-gap on real-time basis.

Later research posed the problem more formally as a signal-extraction problem. Evans (2005) focused on estimating high-frequency GDP, equated business conditions with GDP growth and used state-

space methods to estimate daily GDP growth using data on preliminary, advanced, and final releases of GDP and other macroeconomic variables. Anderson and Gascon (2009) used a state-space model to estimate the “true” unobserved measure of total output in the U.S. economy. The analysis used the entire history (i.e., all vintages) of selected real-time data series to compute revisions and corresponding statistics for those series. The revision statistics, along with the most recent data vintage, were used in a state-space model to extract filtered estimates of the “true” series.

This study is primarily motivated by an empirical study of Aruoba, Diebold and Scotti (2009) on the U.S. economy. They constructed a framework for measuring economic activity at high frequency, potentially in real time. They used a variety of stock and flow data observed at mixed frequencies and performed a prototype empirical application for illustrating the gains achieved by moving beyond the customary monthly data frequency. The four key ingredients of their work are as follows:

- 1. Treatment of business conditions as an unobserved variables, related to the observed indicators.** Latency of business conditions is consistent with economic theory (e.g., Lucas 1977), which emphasizes that the business cycle is not about any single variable, but the dynamics and interactions (or comovements) of many variables.
- 2. Explicit incorporation of business conditions indicators measured at different frequencies.** Important business conditions indicators arrive at a variety of frequencies, including quarterly (e.g., GDP), monthly (e.g., industrial production), weekly (e.g., employment), and continuously (e.g., asset prices), and the incorporation of all of them provides continuously updated measurements.
- 3. Explicit incorporation of indicators measured at high frequencies.** As the goal is to track the high frequency evolution of real activity, it is important to incorporate (or at least not exclude from the outset) the high frequency information flow associated with high frequency indicators.
- 4. Extraction and forecasting of latent business conditions using linear yet statistically optimal procedures, which involve no approximation.** The appeal of exact as opposed to approximate procedures is obvious, but achieving exact optimality is not trivial, due to complications arising from temporal aggregation of stocks versus flows in systems with mixed-frequency data.

They proposed a dynamic factor model that permitted exactly optimal extraction of the latent state of macroeconomic activity being illustrated by a four-variable empirical application with a daily frequency, and in a parallel calibrated simulation (detailed theory mentioned in the technical appendix). The following four indicators with varying frequencies were chosen as business conditions indicators:

1. Yield curve term premium, defined as the difference between 10-years and 3-months U.S. Treasury yield, at daily frequency.
2. Initial claims for unemployment insurance, a weekly flow variable.
3. Employees on non-agricultural payrolls, a monthly stock variable.

4. Real GDP, a quarterly flow variable.

The real activity indicator thus obtained from the empirical analysis threw new lights on the area of business cycle measurement and simultaneously, outperformed the so-called National Bureau of Economic Research (NBER) chronology in some economic as well as statistical sense. First, although the real activity indicator broadly cohered with the NBER chronology, it had a propensity to indicate earlier turning points, especially peaks. Second, the indicator was available at high frequency and hence, a useful “nowcast”, whereas the NBER chronology was available only monthly and with very long lags. Third, it was evident that, incorporation of weekly data in real activity indicator was very helpful for providing real time information, as compared to NBER chronology. However, incorporation of daily data did not improve the performance of the indicator; still a daily state-space setup was needed to accommodate the variation in weeks per month and weeks per quarter. Fourth, based on a simulation calibrated to the empirical results, it was observed that, incorporating high frequency data improved the accuracy of the extracted factor. Lastly, the real time performance (preferably, daily) of the business conditions would be assessed at any point of time by re-estimating the system based on latest-vintage data.

Presently, six macroeconomic indicators are used to construct the Aruoba-Diebold-Scotti Business Conditions Index (ADS Index). These are weekly initial jobless claims, quarterly real GDP, monthly payroll employment, monthly industrial production, monthly real personal income less transfers, and monthly real manufacturing and trade sales. All these are important and widely monitored. The ADS Index is updated weekly, following the release of that week’s new and/or revised component indicator data.

3. Development of Real Time Business Conditions Index for India

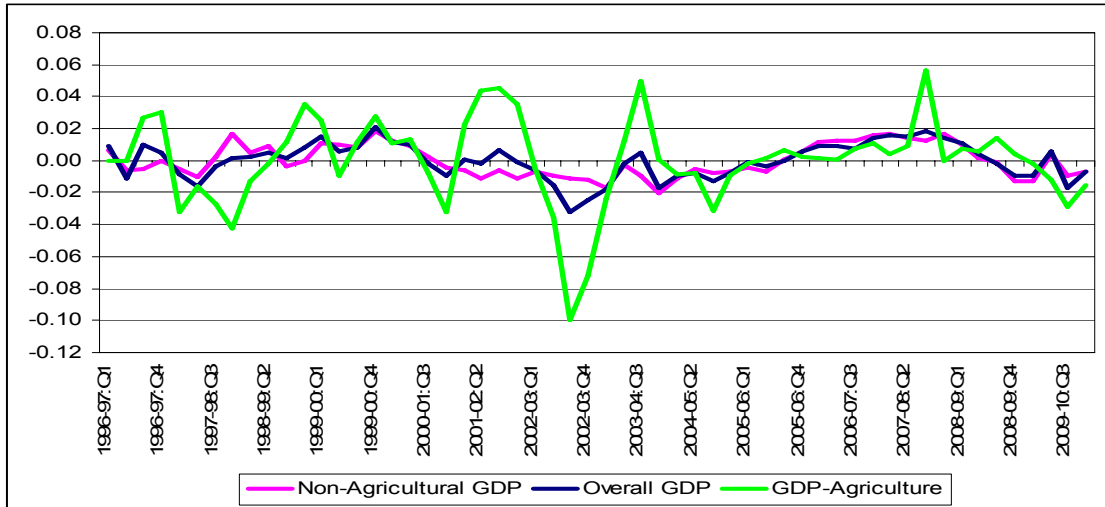
It has been pointed out in Section 2 that, the business conditions of an economy is latent, and it is related to some observed indicators. In order to develop a real time business conditions indicator in Indian context, the prime objective is to select those observed indicators from the existing information base, on the same line as was mentioned in the study of Aruoba, Diebold and Scotti (2009). The information base includes national income aggregates, index of industrial production, capital markets, monetary and banking statistics, price statistics, fiscal statistics, trade data, etc. Some important series along with their source, frequency, availability, and economic as well as statistical justification for inclusion in the selected list are discussed in Table 1.

Table 1: List of Selected Indicators

Sl. No.	Name of the Indicator	Unit of Measurement	Frequency	Source
1	Real Non-Agricultural GDP (NAGDP)	Rs. Crore	Quarterly	Central Statistical Organisation (CSO)
2	Index of Industrial Production (IIP)	---	Monthly	
3	Production of commercial motor vehicles	Thousand Numbers	Monthly	
4	Cargo Handled at Major Ports	Million Tonne	Monthly	
5	Revenue on Railways Freight Traffic	Million Tonne	Monthly	
6	Number of applicants on the live registers of employment exchange	Thousand Numbers	Monthly	
7	Narrow Money (M_1)	Rs. Crore	Fortnightly	Reserve Bank of India (RBI)
8	Money Supply (M_3)	Rs. Crore	Fortnightly	Reserve Bank of India (RBI)
9	WPI Primary Articles	---	Weekly	Ministry of Commerce and Industries
10	Foreign Exchange Reserve	Rs. Crore	Weekly	Reserve Bank of India (RBI)
11	Yield Curve Term Premium (difference between 10-years Govt. Bond and 91 days Treasury yields)	---	Daily	Reserve Bank of India (RBI)
12	BSE Sensex	---	Daily	Bombay Stock Exchange (BSE)

a. Real Gross Domestic Product - Ideally, the Real Gross Domestic Product (GDP) represents almost all aspects of the economic activities. In the literature of business cycle, the cyclical fluctuation of Real GDP is a well-accepted reference frame of the business conditions of the economy, as it includes all the three sectors viz., primary, secondary and tertiary. In India, the Central Statistical Organisation (CSO) of Ministry of Statistics and Programme Implementation releases the quarterly figures of GDP with a two months time lag. In the report of RBI's Working Group of Leading Indicators for Indian Economy, the quarterly Non-Agricultural GDP (NAGDP) was taken as a reference series of business cycle due to the dependence of agricultural sector on monsoon performance. In the context of the Indian economy, the agriculture and allied sector contributes almost 20 per cent to the total GDP. However, the agriculture sector depends heavily on rainfall. The high volatility in agricultural sector may be observed from the movement of cyclical component of GDP Agriculture with standard deviation 0.26 (Chart 1). On the other hand, the standard deviation of each of the Overall GDP and Non-Agricultural GDP cycles is 0.13. Moreover, Chart 1 depicts similar movements of Overall GDP and Non-Agricultural GDP cycles (with correlation 0.87). Based on these observations, Non-Agricultural GDP is considered more preferable than overall GDP.

Chart 1: Cyclical Components of GDP Agriculture, Non-Agricultural GDP and Overall GDP



b. Index of Industrial Production – In the ADS index, out of six macroeconomic series, one is monthly industrial production and accordingly, in this study, monthly Index of industrial Production (IIP) has also been considered. The index is regularly published by CSO with two months lag. It may also be mentioned that, OECD uses monthly Industrial Production as the reference series for business cycle analysis. Chart 2 presents the coincidental movements of Non-Agricultural GDP and IIP cycles with correlation 0.81, and as the Non-Agricultural GDP has nearly 83 per cent share in overall GDP, it is well justified that, IIP also reflects the business conditions of the economy with frequency higher than quarterly intervals (Table 2).

Chart 2: Cyclical Components of Non-Agricultural GDP and Index of Industrial Production (IIP)

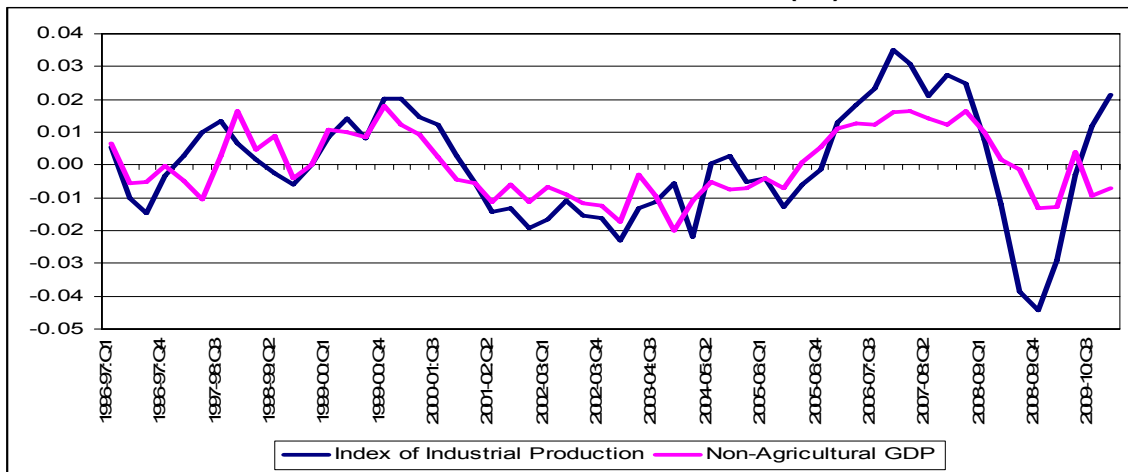


Table 2: Cross-correlation between Non-Agricultural GDP and IIP

i	lag	lead	i	lag	lead
0	0.4321	0.4321	6	0.1256	-0.0078
1	0.3473	0.3130	7	0.1796	0.0219
2	0.2624	0.1940	8	0.2336	0.0472
3	0.1776	0.0749	9	0.2564	0.0348
4	0.0928	-0.0445	10	0.1797	-0.0139
5	0.0716	-0.0374	11	0.1031	-0.0627

c. Indicators observing industrial activities – There are some indicators that show the performance of industrial activities. Transportation of goods by road is a good indicator of industrial production process. Thus ‘Production of commercial motor vehicles’ is taken as an indicator. Moreover, increased levels of production, consumption and trade also get reflected, particularly in a large country such as India, in increased transportation of goods. Thus, the two series, viz., ‘Cargo Handled at Major Ports’ and ‘Revenue on Railways Freight Traffic’ are also considered in this study. The data of these series are regularly released in the monthly capsule report of CSO with two months lag. The movements of cyclical components of the three industrial activities indicators vis-à-vis IIP are presented in Charts 3 to 5. All of them show coincidental movement with IIP cycle.

Chart 3: Cyclical Components of Index of Industrial Production (IIP) and Cargo Handled at Major Ports

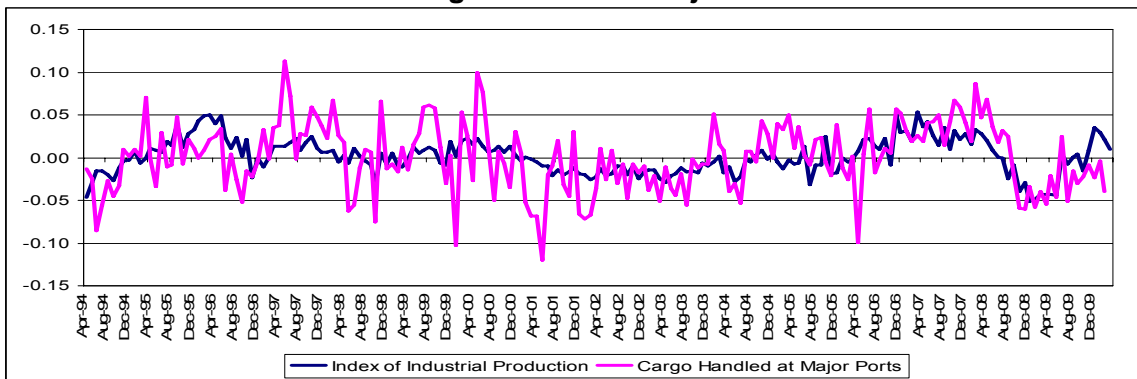


Chart 4: Cyclical Components of Index of Industrial Production (IIP) and Production of Commercial Motor Vehicles

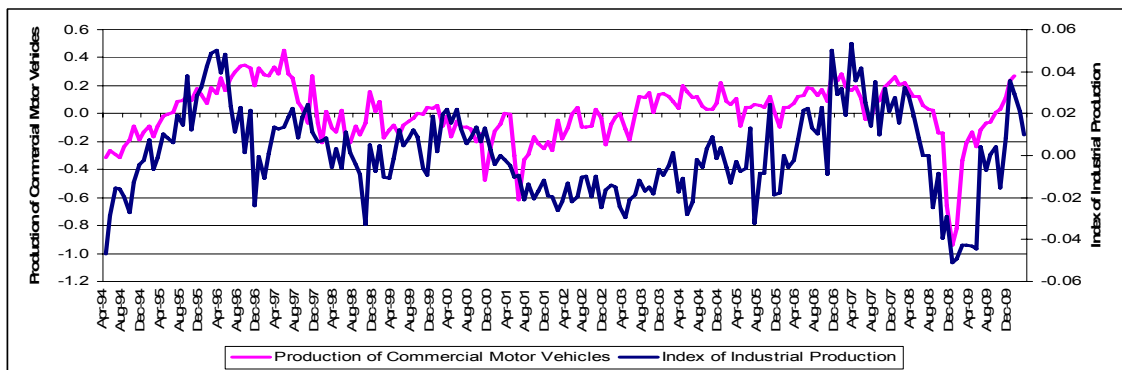
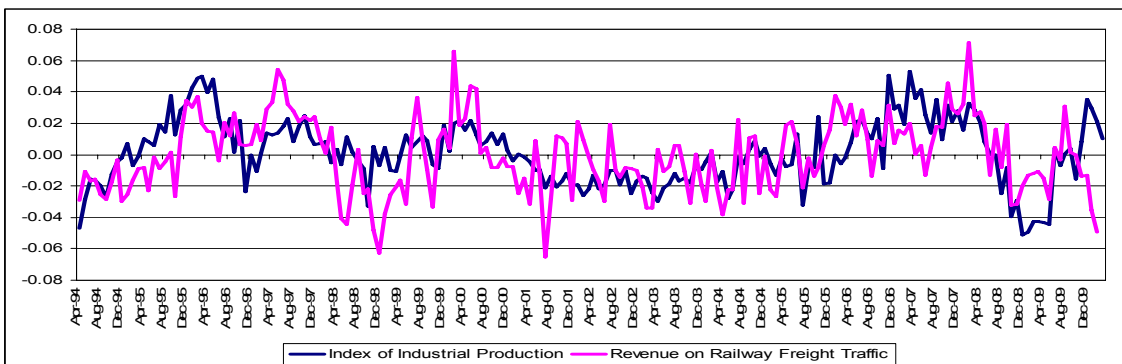


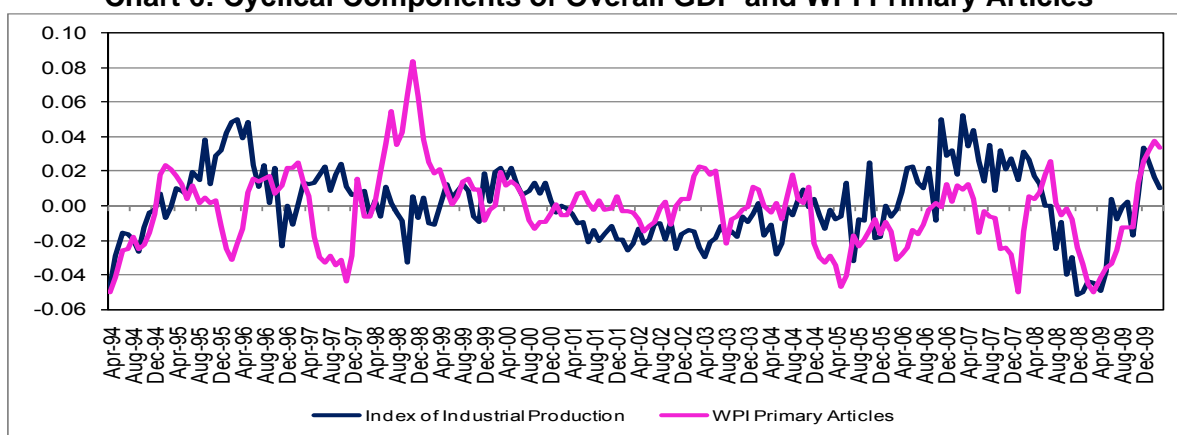
Chart 5: Cyclical Components of Index of Industrial Production (IIP) and Revenue on Railways Freight Traffic



d. **WPI Primary Articles** – In the ADS Index, the price factor was not considered directly, but, the daily yield curve term premium (i.e., difference between 10-years and 3-months U.S. Treasury yield) had taken into account of inflation expectation to some extent. In our study, we consider WPI Primary Articles, compiled by Ministry of Commerce and Industries, due to the following reasons.

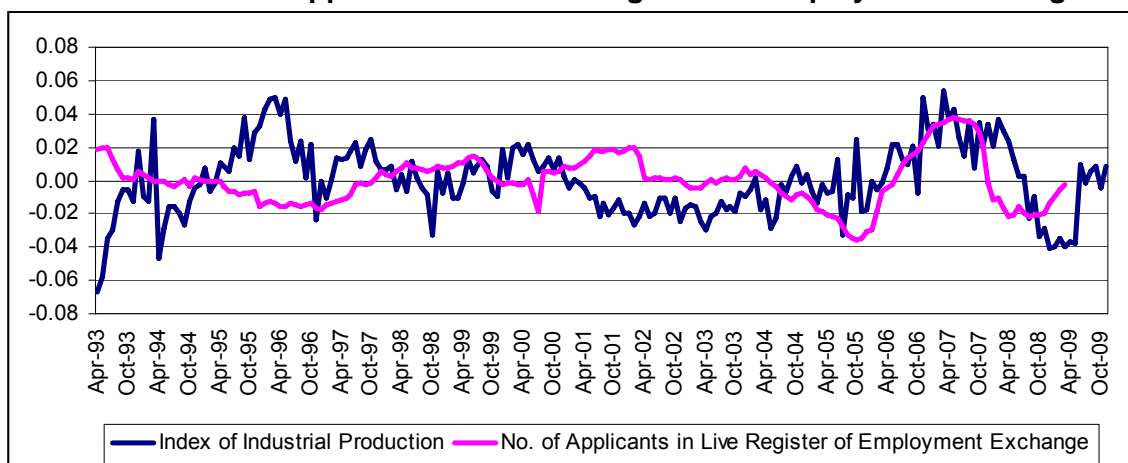
- (i) WPI Primary Articles data are available at weekly frequency. Although WPI Manufactured Products is more related to industrial production than WPI Primary Articles, it is presently compiled on monthly basis. Moreover, although WPI of ‘Fuel and Power’ is available on weekly basis, the prices of some products included in the ‘Fuel and Power’ group are administered by the government.
- (ii) The cyclical movement of WPI Primary Articles is somehow coincident to IIP cycle, although divergence is clearly observed in most cases (Chart 6).

Chart 6: Cyclical Components of Overall GDP and WPI Primary Articles



e. **Employment and Unemployment indicator** – The comprehensive employment series, which is available on a regular basis, is the series giving estimated average daily employment in factories; but it is available on an annual basis. Data on unemployment rate are not compiled on a regular basis. The series, ‘Number of applicants on the live registers of employment exchange’, released in Monthly Abstract of Statistics by CSO, gives some indication of the number of unemployed in the cities, but it suffers from the well-known limitations such as the changing (increasing) number of unemployment exchanges over the years, possibilities of incomplete as well as multiple registrations, registration by those currently employed because they are looking for better jobs or through failure to cancel registration, etc. Despite these limitations, Chitre (2001) found this series as a useful coincident indicator of the industrial production and, therefore, is included in the present study for examination (Chart 7). Another limitation of the series is its long lag period for releasing the data. For example, the Monthly Abstract of Statistics, September-October 2009 publication released the data upto March 2009.

Chart 7: Cyclical Components of Index of Industrial Production (IIP) and Number of applicants on the live registers of employment exchange



- f. **Yield Curve Term Premium** – In the study of Aruoba, Diebold and Scotti (2009), the only series with daily frequency was yield curve term premium, defined as the difference between 10-years and 3-months U.S. Treasury yield. On the same line, in this study, the difference between yields of 10-years government securities and 91-days Treasury Bill is included in the list. These data are available on real-time basis.
- g. **BSE Sensex** – Another daily indicator, as mentioned by Aruoba, Diebold and Scotti (2009), is asset price. As a proxy of asset price, daily Sensex data of Bombay Stock Exchange (BSE) has been examined. It was observed that, in connection with the economic activity of the country, Sensex seems to be too much noisy as well as volatile. Also, in case of checking the cross correlation of BSE Sensex with NAGDP, IIP and weekly Foreign Exchange Reserve, we have seen that the correlation is not encouraging (Table 3).

Table 3: Cross Correlation between cyclical components - BSE-Sensex with NAGDP, IIP and Foreign Exchange Reserve

NAGDP Vs BSE-Sensex			IIP Vs BSE-Sensex			Reserve Vs BSE-Sensex		
i	lag	lead	i	lag	lead	i	lag	lead
0	0.2936	0.2936	0	0.1667	0.1667	0	0.0222	0.0222
1	0.2910	0.2964	1	0.1684	0.1655	1	0.0187	0.0263
2	0.2885	0.2991	2	0.1703	0.1642	2	0.0153	0.0304
3	0.2859	0.3019	3	0.1719	0.1630	3	0.0115	0.0347
4	0.2833	0.3045	4	0.1734	0.1616	4	0.0078	0.0391
5	0.2807	0.3069	5	0.1751	0.1600	5	0.0040	0.0437
6	0.2783	0.3093	6	0.1767	0.1583	6	0.0002	0.0484
7	0.2760	0.3115	7	0.1781	0.1566	7	-0.0035	0.0532
8	0.2737	0.3136	8	0.1799	0.1546	8	-0.0073	0.0577
9	0.2714	0.3156	9	0.1821	0.1528	9	-0.0113	0.0619
10	0.2691	0.3175	10	0.1840	0.1508	10	-0.0154	0.0659

- h. **Foreign Exchange Reserve** - As selected indicators are expected to be highly correlated with each other, we have examined the cross-correlation among different series at different frequencies. It was found that, the Foreign Exchange Reserve, which was considered as a stock variable available at weekly frequency, has a poor correlation with the economic activity, i.e., NAGDP as well as with IIP, and hence is not fit for the requirement (Table 4).

Table 4: Cross Correlation between cyclical components - Foreign Exchange Reserve with NAGDP and IIP

NAGDP Vs Reserve			IIP Vs Reserve		
i	lag	Lead	i	lag	lead
0	-0.0528	-0.0528	0	0.0360	0.0360
1	-0.0514	-0.0545	1	0.0329	0.0393
2	-0.0499	-0.0562	2	0.0295	0.0426
3	-0.0484	-0.0578	3	0.0261	0.0459
4	-0.0470	-0.0596	4	0.0223	0.0489
5	-0.0456	-0.0613	5	0.0183	0.0518
6	-0.0443	-0.0631	6	0.0138	0.0548
7	-0.0431	-0.0648	7	0.0096	0.0579
8	-0.0419	-0.0666	8	0.0053	0.0611
9	-0.0408	-0.0685	9	0.0012	0.0635
10	-0.0398	-0.0704	10	-0.0029	0.0657

i. **Money Supply and Narrow Money** – Besides weekly variables, Narrow Money (M_1) and Money Supply (M_3) have been considered as fortnightly stock variable for testing coincidence with IIP and NAGDP cyclical components. The cross-correlations of M_1 and M_3 with NAGDP and IIP, separately, were derived. It was examined that, Narrow Money had a better relationship with IIP and GDP than Money Supply had, and hence, selected (Table 5 and Charts 8 and 9).

Table 5: Cross Correlation between cyclical components - M_1 and M_3 with NAGDP and IIP

M ₁ Vs NAGDP			M ₁ Vs IIP			M ₃ Vs NAGDP			M ₃ Vs IIP		
i	lag	lead	i	lag	lead	i	lag	lead	i	lag	lead
0	0.6763	0.6763	0	0.5586	0.5586	0	0.1287	0.1287	0	-0.2187	-0.2187
1	0.6464	0.7230	1	0.5972	0.5383	1	0.2700	-0.0008	1	-0.2026	-0.2112
2	0.5806	0.6534	2	0.5527	0.4547	2	0.4322	-0.1132	2	-0.2235	-0.2069
3	0.4357	0.5182	3	0.5087	0.4392	3	0.5224	-0.2116	3	-0.2362	-0.1797
4	0.4030	0.4323	4	0.5059	0.3858	4	0.5782	-0.3264	4	-0.2228	-0.1499
5	0.2699	0.2348	5	0.4454	0.2912	5	0.5886	-0.4753	5	-0.2328	-0.1532
6	0.1092	0.1010	6	0.3878	0.2854	6	0.5751	-0.5606	6	-0.2285	-0.0966
7	0.0243	-0.0267	7	0.3465	0.2390	7	0.5986	-0.6325	7	-0.2549	-0.0589
8	-0.1208	-0.1072	8	0.3334	0.2225	8	0.5342	-0.6227	8	-0.2298	-0.0127
9	-0.2517	-0.1798	9	0.2679	0.2092	9	0.5034	-0.6113	9	-0.2625	0.0602
10	-0.3874	-0.2564	10	0.2404	0.1820	10	0.4240	-0.5991	10	-0.2833	0.1063

Chart 8: Cyclical Components of Index of Industrial Production (IIP) and Narrow Money (M_1)

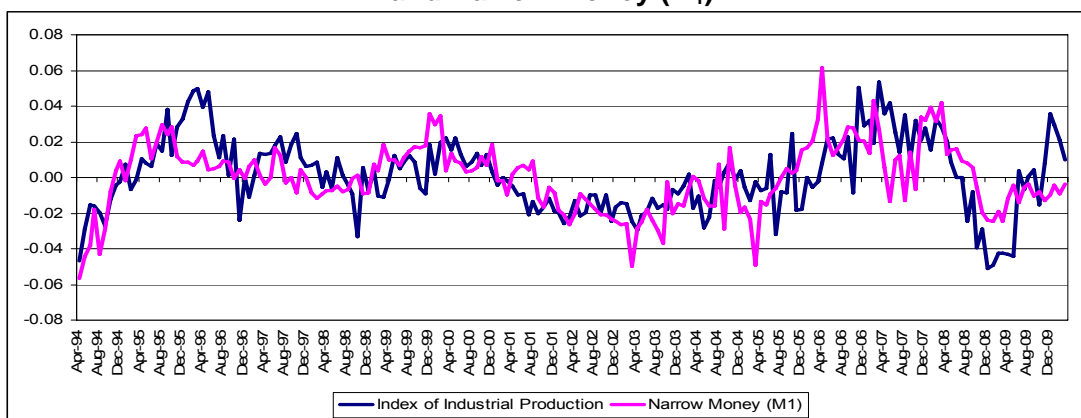
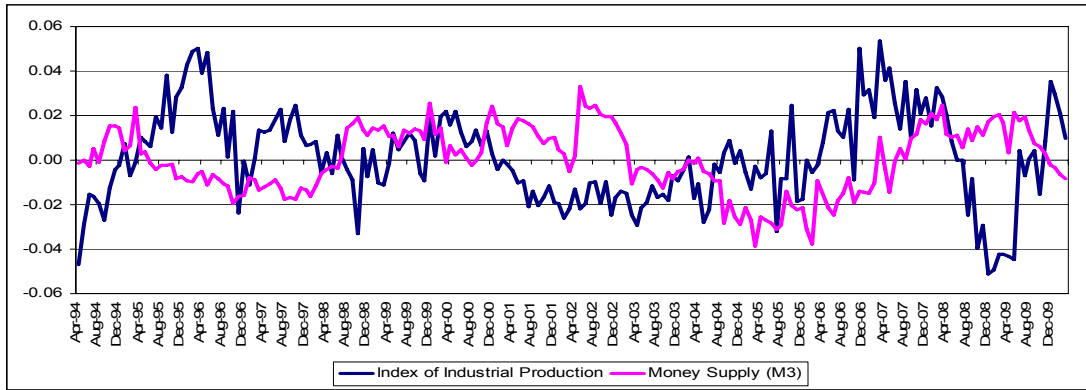


Chart 9: Cyclical Components of Index of Industrial Production (IIP) and Money Supply (M₃)



Based on the above-mentioned indicators as well as the methodology (described in Technical Appendix) the four indicators were finally selected (Table 6).

Table 6: Finally Selected Indicators

Serial no.	Variable Name	Frequency of Availability	Variable Type
1	Yield Curve Term Premium	Daily	Stock
2	M ₁	Fortnightly	Stock
3	Index of Industrial Production	Monthly	Flow
4	NAGDP	Quarterly	Flow

Taking the above-mentioned variables and using the stated methodology, we tried to implement a program in RATS software to construct a real-time business conditions index in Indian context, so that, an accurate and timely estimates of the state of real activity can be achievable. Instead of weekly variable (as used in ADS index), we have considered here a fortnightly available stock variable, i.e. M₁.

4. Software

First, we have extracted the cyclical component from the four selected series as we try to find an indicator for business cycle. We did the extraction based on the following procedure:

- I. Seasonal adjustment of the series using X-12 ARIMA methodology.
- II. HP Filter to remove the smoothed trend from the seasonally adjusted series and finally extracting the cyclical component.

Then the series having cyclical component only has been used in the program as an input. After that, we have defined the matrices of the state-space model following the theory stated in the paper by Aruoba, Diebold and Scotti (2009). In their paper, they considered a weekly flow variable, but, in our case, we have considered a fortnightly stock variable. Again, instead of monthly stock variable, we have considered a flow variable. The matrix coefficients have been changed accordingly. We have used Dynamic Linear Model for estimating the coefficients. We obtain our start-up values in two steps, as follows. In the first step, we use only daily and stock variables, which drastically reduce the dimension of the state vector, resulting in very fast estimation. This yields preliminary estimates of all measurement equation parameters for the daily and stock variables, and all transition equation parameters, as well as a preliminary extraction of the factor, \hat{X}_t (via a pass of the Kalman smoother).

In the second step, we use the results of the first step to obtain start-up values for the remaining parameters, that is, those in the flow variable measurement equations. We simply regress the flow variables on the smoothed state extracted in the first step and take the coefficients as our start-up values. With the model cast in state-space form, and for given parameters, we use the Kalman filter and smoother to obtain optimal extractions of the latent state of real activity.

5. Empirical Results

Based on the four finally selected indicators, viz., daily Yield curve term premium, fortnightly M_1 , monthly Index of Industrial Production (IIP), and quarterly NAGDP, the Real-time Business Conditions Index (RTBCI) was constructed (Chart 10). The latest available information of the selected indicators were used for construction (Table 7).

Table 7: Data Period of Selected Indicators

Serial no.	Variable Name	Frequency of Availability	Variable Type	Period
1	Yield Curve Term Premium	Daily	Stock	1-Dec-2003 to 10-Dec-2010
2	M_1	Fortnightly	Stock	Fortnight ended 12-Dec-2003 to Fortnight ended 19-Nov-2010
3	Index of Industrial Production	Monthly	Flow	Dec-2003 to October 2010
4	Non-Agricultural GDP	Quarterly	Flow	Q4:2003-04 to Q2:2010-11

Chart 10: Real-Time Business Conditions Index (RTBCI) for India

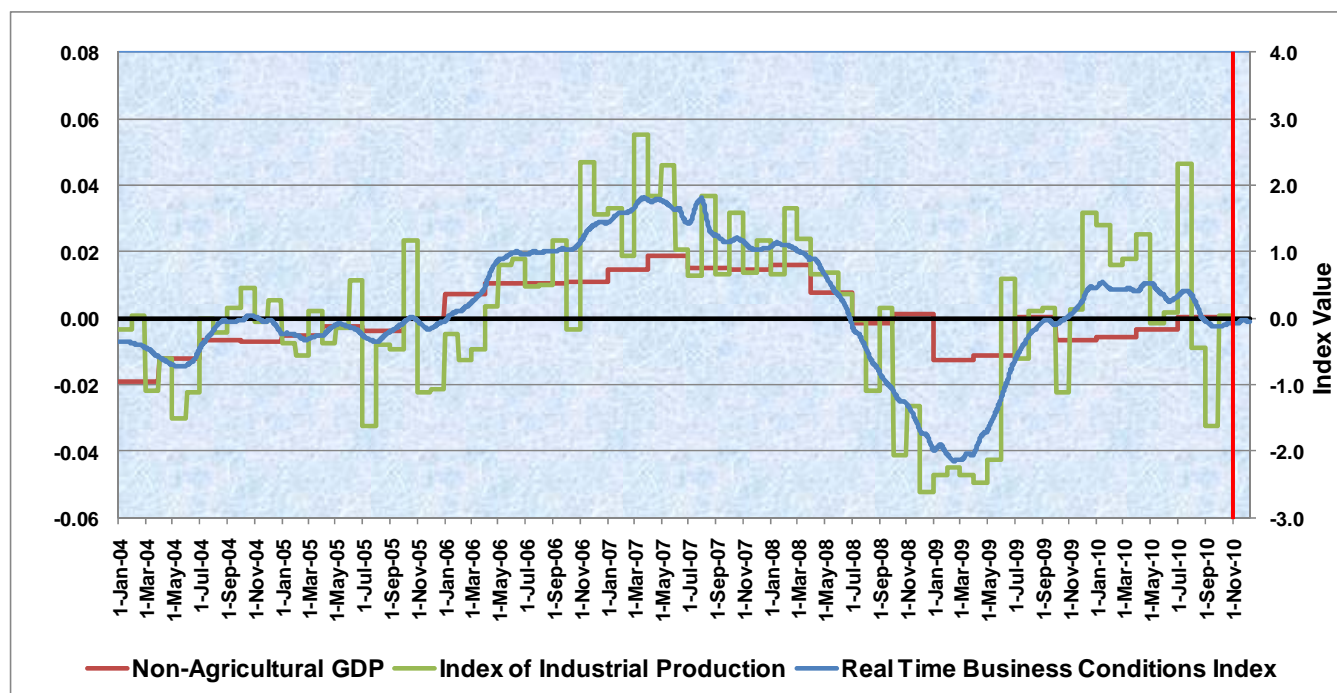


Chart 10 displays the cyclical movements of NAGDP and Index of Industrial Production (IIP), along with the daily movement of RTBCI. It is observed that, there is a coincidental movement among these three series. The movement of RTBCI beyond the vertical redline indicates the state of the economy which is

otherwise not observed from the macro aggregates. The momentum as per the RTBCI, as on December 10, 2010, indicates somewhat low acceleration of economic activity.

In order to see the importance of daily data for real-time measurement, a separate analysis was done to construct an index using the three selected indicators except daily “Yield curve term premium” data. The reason behind this analysis is that, if the fortnightly frequency is considered as the highest available frequency, and if under this assumption, the RTBCI is constructed, then, whether this index performs better than the earlier index, constructed using daily data. Chart 11 displays the cyclical movements of NAGDP and Index of Industrial Production (IIP), along with the daily movement of RTBCI-without daily data. Although this chart indicates co-movement between these three series, the next step is to search for any lag difference between the movements of RTBCI and RTBCI-without daily data. Chart 12 displays the movements of RTBCI and RTBCI-without daily data. In this chart, it is observed that, the turning points of RTBCI, including daily data, is earlier than those of RTBCI-without daily data. The lag difference varies approximately from 1 to 6 weeks. This implies that, if the turning point occurs in the business cycle movement, then RTBCI will capture the turning point earlier than RTBCI-without daily data. For example, the turning point occurred in January 2009 in IIP cycle was captured by RTBCI (value corresponding to January 3, 2009 was the lowest) four weeks earlier than that captured by RTBCI-without daily data (value corresponding to January 31, 2009 was the lowest). This justifies the importance of usage of daily data while constructing the Index.

Chart 11: Movements of NAGDP Cycle, IIP Cycle, and RTBCI –Excluding Daily Data

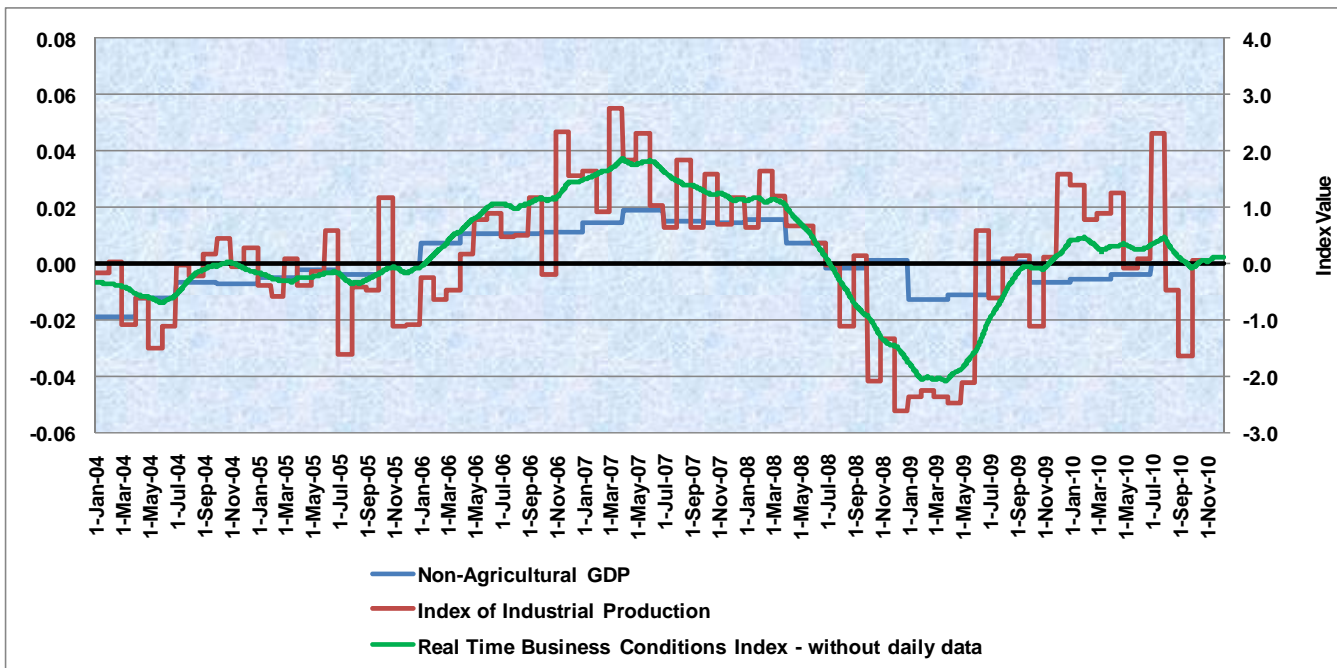
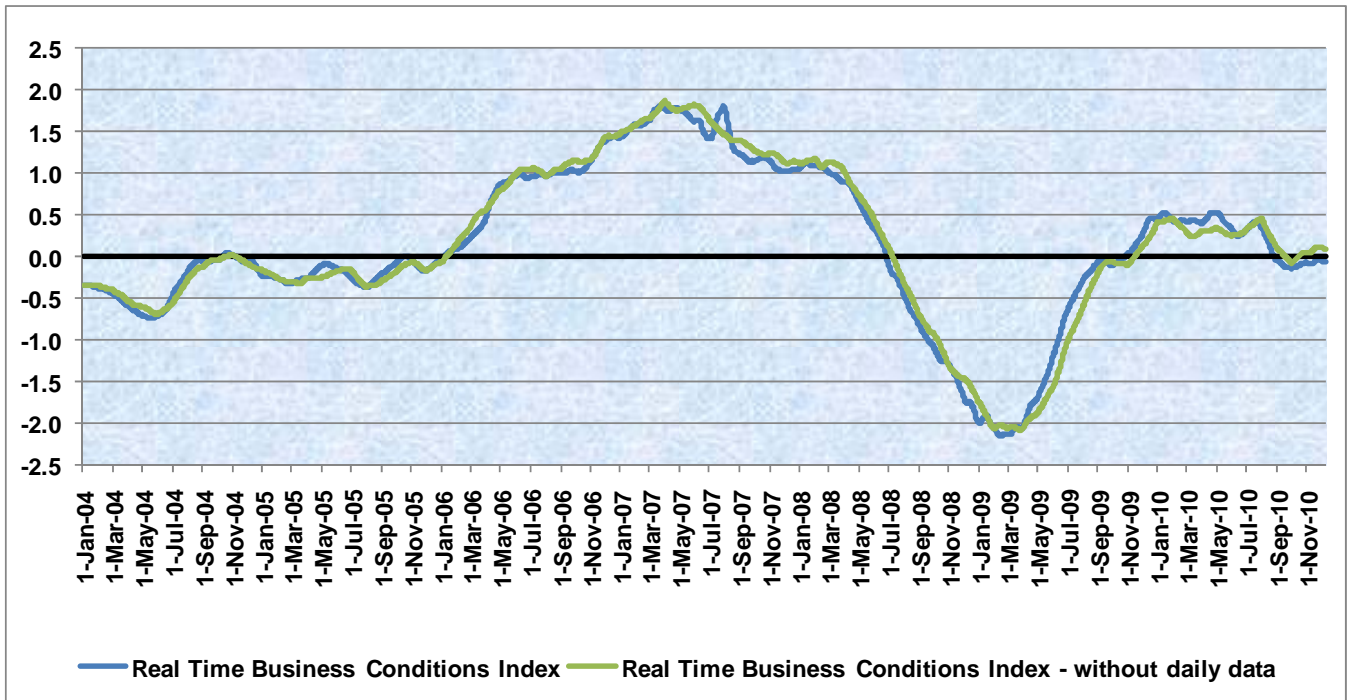


Chart 12: Movements of RTBCI – Including and Excluding Daily Data



6. Concluding Remarks

Real time decision making requires accurate and timely understanding of the state of real activity. In the light of the changing nature of the economy where increasingly more and more activities are being channelised through both organised and unorganized business sectors, the measurement of business condition on real time basis is extremely difficult. In this context, in order to achieve an accurate and timely estimate of the state of real activity in a systematic, replicable and statistically optimal manner, this paper proposes a framework to construct a real-time business conditions index for India. Based on various economic indicators measured at different frequencies, this paper develops a real-time business conditions index for India following a dynamic factor model framework for extracting signals from continuously evolving states. A Kalman filter routine is used for signal extraction from state-space representation as well as evaluation of likelihood function. Empirical results show that this coincident indicator tracks the overall economic activity reasonably well.

Measurement of Real-time Business Condition

a. Modeling framework

We propose a dynamic factor model at daily frequency. The model is very simple at daily frequency, but of course the daily data are generally not observed, so most of the data are missing. Hence we propose explicitly to treat missing data and temporal aggregation, and to obtain the measurement equations for observed stock and flow variables. Following that, we can enrich the model by allowing for lagged state variables in the measurement equations, and by allowing for trend, both of which are important when fitting the model to macroeconomic and financial indicators.

b. Dynamic Factor Model at Daily Frequency

We assume that the state of the economy evolves at a very high frequency; without loss of generality, we can take it as “daily.” In our subsequent empirical work, we want to use a daily base observational frequency, but much higher (intraday) frequencies could be used if desired. Economic and financial variables, although evolving daily, are of course not generally observed daily. For example, an end-of-year wealth variable is observed each December 31, and is unobserved every other day of the year. Let \mathbf{x}_t denote underlying business conditions at day t , which evolve daily with $AR(p)$ dynamics,

$$\mathbf{x}_t = \rho_1 \mathbf{x}_{t-1} + \rho_2 \mathbf{x}_{t-2} + \dots + \rho_p \mathbf{x}_{t-p} + \mathbf{e}_t \quad \dots\dots(1)$$

where \mathbf{e}_t is a white noise innovation with unit variance. We are interested in tracking and forecasting real activity, so we use a single-factor model; that is, \mathbf{x}_t is a scalar, as, for example, in Stock and Watson (1989). Additional factors could be introduced to track, for example, wage/price developments. Let \mathbf{y}_t^i denote the i th daily economic or financial variable at day t , which depends linearly on \mathbf{x}_t and possibly also on various exogenous variables and/or lags of \mathbf{y}_t^i :

$$\mathbf{y}_t^i = c_i + \beta_i \mathbf{x}_t + \delta_{i1} \mathbf{w}_t^1 + \dots + \delta_{ik} \mathbf{w}_t^k + \gamma_{i1} \mathbf{y}_{t-D_i}^i + \dots + \gamma_{in} \mathbf{y}_{t-nD_i}^i + \mathbf{u}_t^i \quad \dots\dots(2)$$

where the \mathbf{w}_t are exogenous variables and the \mathbf{u}_t^i are contemporaneously and serially uncorrelated innovations. It should be mentioned that, we introduce lags of the dependent variable \mathbf{y}_t^i in multiples of D_i , where $D_i > 1$ is a number linked to the frequency of the observed \mathbf{y}_t^i (We will discuss D_i in detail in the next subsection.). Modeling persistence only at the daily frequency would be inadequate, as it would decay too quickly.

c. Missing Data, Stocks vs. Flows, and Temporal Aggregation

We have mentioned above that y_t^i denotes the i th variable on a daily time scale. But most variables, although evolving daily, are not actually observed daily. Hence let \tilde{y}_t^i denote the same variable observed at a lower frequency (call it the “tilde frequency”). The relationship between \tilde{y}_t^i and y_t^i depends crucially on whether y_t^i is a stock or flow variable. If y_t^i is a stock variable measured at a non-daily tilde frequency, then the appropriate treatment is straightforward, because stock variables are simply point-in-time snapshots. At any time t , either y_t^i is observed, in which case $\tilde{y}_t^i = y_t^i$, or it is not, in which case $\tilde{y}_t^i = NA$, where NA denotes missing data (“not available”). Hence we have the stock variable measurement equation:

$$\tilde{y}_t^i = \begin{cases} y_t^i = c_i + \beta_i x_t + \delta_{i1} w_t^1 + \dots + \delta_{ik} w_t^k + \gamma_{i1} y_{t-D_i}^i + \dots + \gamma_{in} y_{t-nD_i}^i + u_t^i & \text{if } y_t^i \text{ is observed} \\ NA & \text{otherwise} \end{cases} \dots\dots\dots(3)$$

Now consider flow variables. Flow variables observed at non-daily tilde frequencies are intra period sums of the corresponding daily values,

$$\tilde{y}_t^i = \begin{cases} \sum_{j=0}^{D_i-1} y_{t-j}^i & \text{if } y_t^i \text{ is observed} \\ NA & \text{otherwise} \end{cases} \dots\dots\dots(4)$$

where D_i is the number of days per observational period (e.g., $D_i = 7$ if y_t^i is measured weekly).

Combining this fact with Equation (2), we arrive at the flow variable measurement equation

$$\tilde{y}_t^i = \begin{cases} \sum_{j=0}^{D_i-1} c_i + \beta_i \sum_{j=0}^{D_i-1} x_{t-j}^i + \delta_{i1} \sum_{j=0}^{D_i-1} w_{t-j}^1 + \dots + \delta_{ik} \sum_{j=0}^{D_i-1} w_{t-j}^k + \gamma_{i1} \sum_{j=0}^{D_i-1} y_{t-D_i-j}^i + \dots + \gamma_{in} \sum_{j=0}^{D_i-1} y_{t-nD_i-j}^i + u_t^{*i} & \text{if } y_t^i \text{ is observed} \\ NA & \text{otherwise} \end{cases} \dots\dots\dots(5)$$

where $\sum_{j=0}^{D_i-1} y_{t-D_i-j}^i$ is by definition the observed flow variable one period ago ($\tilde{y}_{t-D_i}^i$), and u_t^{*i} is the sum of the u_t^i over the tilde period.

Though D_i is a time variable, we treat it as fixed. But, in our subsequent empirical implementation, we allow for time-varying D_i . Again, although u_t^{*i} follows a moving average process of order D_i-1 at daily frequency, it nevertheless remains white noise when observed at the tilde frequency, due to the $(D_i - 1)$ -dependence of an $MA(D_i - 1)$ process. Hence we appropriately treat u_t^{*i} as white noise in what follows, where $\text{Var}(u_t^{*i}) = D_i \cdot \text{var}(u_t^i)$.

d. Trend

The exogenous variables w_t are the key to handling trend. In particular, in the important special case where the w_t are simply deterministic polynomial trend terms [$w^1 t-j = t-j$, $w^2 t-j = (t-j)^2$ and so on] we have that

$$\begin{aligned} & \sum_{j=0}^{D_i-1} [c_i + \delta_{i1}(t-j) + \dots + \delta_{ik}(t-j)^k] \\ & \equiv c_i^* + \delta_{i1}^* t + \dots + \delta_{ik}^* t^k. \end{aligned} \tag{6}$$

Therefore, we have the stock variable equation,

$$\tilde{y}_t^i = \begin{cases} c_i^* + \beta_i x_t^i + \delta_{i1}^* t + \dots + \delta_{ik}^* t^k + v_{i1} \tilde{y}_{t-D_i}^i + \dots + v_{in} \tilde{y}_{t-nD_i}^i + u_t^{*i} & \text{if } y_t^i \text{ is observed} \\ NA & \text{otherwise} \end{cases} \tag{7}$$

And the flow variable equation,

$$\tilde{y}_t^i = \begin{cases} c_i^* + \beta_i \sum_{j=0}^{D_i-1} x_{t-j}^i + \delta_{i1}^* t + \dots + \delta_{ik}^* t^k + \gamma_{i1} \tilde{y}_{t-D_i}^i + \dots + \gamma_{in} \tilde{y}_{t-nD_i}^i + u_t^{*i} & \text{if } y_t^i \text{ is observed} \\ NA & \text{otherwise} \end{cases} \tag{8}$$

This completes the specification of our model, which has a natural state-space form.

e. State- Space Representation, Signal Extraction and Estimation

Here we will discuss our model from a state-space perspective, including filtering and estimation. We will avoid dwelling on standard issues, focusing instead on the nuances specific to our framework, including missing data due to mixed-frequency modeling, high-dimensional state vectors due to the presence of flow variables, and time-varying system matrices due to varying lengths of months.

State-Space Representation

Our model is trivially cast in state-space form as

$$\begin{aligned}
 y_t &= Z_t \alpha_t + \Gamma_t w_t + \varepsilon_t, \\
 \alpha_{t+1} &= T \alpha_t + R \eta_t, \\
 \varepsilon_t &\sim (0, H_t), \\
 \eta_t &\sim (0, Q), \\
 t &= 1, \dots, T.
 \end{aligned}
 \tag{9}$$

where \mathbf{y}_t is an $N \times 1$ vector of observed variables (subject, of course, to missing observations), α_t is an $m \times 1$ vector of state variables, \mathbf{w}_t is a $e \times 1$ vector of predetermined variables containing a constant term (unity), k trend terms and $N \times n$ lagged dependent variables (n for each of the N elements of the \mathbf{y}_t vector), ε_t and η_t are vectors of measurement and transition shocks containing the u_t^i and e_t , and T denotes the last time-series observation. In general, the observed vector \mathbf{y}_t will have a very large number of missing values, reflecting not only missing daily data due to holidays, but also, and much more importantly, the fact that most variables are observed much less often than daily. Interestingly, the missing data *per se* does not pose severe challenges: \mathbf{y}_t is simply littered with a large number of NA values, and the corresponding system matrices are very sparse, but the Kalman filter remains valid (appropriately modified, as we will discuss below), and numerical implementations may indeed be tuned to exploit the sparseness, as we do in our implementation. In contrast, the presence of flow variables produces more significant complications, and it is hard to imagine a serious business conditions indicator system without flow variables, given that real output is itself a flow variable. Flow variables produce intrinsically high-dimensional state vectors. In particular, as shown in Equation (3), the flow variable measurement equation contains x_t and $\max\{D\} - 1$ lags of x_t , producing a state vector of dimension $\max\{\max\{D\}, \rho\}$, in contrast to the ρ -dimensional state associated with a system involving only stock variables. In realistic systems with data frequencies ranging from, say, daily to quarterly, $\max\{D\} \approx 90$. There is a final nuance associated with our state-space system: several of the system matrices are time varying. In particular, although \mathbf{T} , \mathbf{R} , and \mathbf{Q} are constant, \mathbf{Z}_t , Γ_t , and \mathbf{H}_t are not, because of the variation in the number of days across quarters and months (i.e., the variation in D_t across t). Nevertheless the Kalman filter remains valid.

Signal Extraction

With the model cast in state-space form, and for given parameters, we use the Kalman filter and smoother to obtain optimal extractions of the latent state of real activity. As is standard for classical estimation, we initialize the Kalman filter using the unconditional mean and covariance matrix of the state vector. We use the contemporaneous Kalman filter; see Durbin and Koopman (2001) for details.

Let $Y_t \equiv \{\mathbf{y}_1, \dots, \mathbf{y}_T\}$, $\mathbf{a}_{t|t} \equiv E(\boldsymbol{\alpha}_t | Y_t)$, $\mathbf{P}_{t|t} = \text{var}(\boldsymbol{\alpha}_t | Y_t)$, $\mathbf{a}_t \equiv E(\boldsymbol{\alpha}_t | Y_{t-1})$, and $\mathbf{P}_t = \text{var}(\boldsymbol{\alpha}_t | Y_{t-1})$. Then the Kalman filter updating and prediction equations are

$$\begin{aligned}
 \mathbf{a}_{t/t} &= \mathbf{a}_t + \mathbf{P}_t \mathbf{Z}_t' \mathbf{F}_t^{-1} \mathbf{v}_t, \\
 \mathbf{P}_{t/t} &= \mathbf{P}_t - \mathbf{P}_t \mathbf{Z}_t' \mathbf{F}_t^{-1} \mathbf{Z}_t \mathbf{P}_t', \\
 \mathbf{a}_{t+1} &= \mathbf{T} \mathbf{a}_{t/t}, \\
 \mathbf{P}_{t+1} &= \mathbf{T} \mathbf{P}_{t/t} \mathbf{T}' + \mathbf{R} \mathbf{Q} \mathbf{R}', \\
 &\text{where} \\
 \mathbf{v}_t &= y_t - \mathbf{Z}_t \mathbf{a}_t - \Gamma_t w_t, \\
 \mathbf{F}_t &= \mathbf{Z}_t \mathbf{P}_t \mathbf{Z}_t' + \mathbf{H}_t,
 \end{aligned}
 \tag{10}$$

Crucially for us, the Kalman filter remains valid with missing data. If all elements of y_t are missing, we skip updating and the recursion becomes

$$\begin{aligned}
 \mathbf{a}_{t+1} &= \mathbf{T} \mathbf{a}_t, \\
 \mathbf{P}_{t+1} &= \mathbf{T} \mathbf{P}_t \mathbf{T}' + \mathbf{R} \mathbf{Q} \mathbf{R}'
 \end{aligned}
 \tag{11}$$

If some but not all elements of y_t are missing, we replace the measurement equation with

$$\begin{aligned}
 y_t^* &= \mathbf{Z}_t^* \boldsymbol{\alpha}_t + \Gamma_t^* w_t + \varepsilon_t^*, \\
 \varepsilon_t^* &\sim N(0, \mathbf{H}_t^*),
 \end{aligned}$$

where y_t^* is of dimension $N^* < N$, containing the elements of the \mathbf{y}_t vector that are observed. The key insight is that y_t^* and \mathbf{y}_t are linked by the transformation $y_t^* = \mathbf{W}_t \mathbf{y}_t$, where \mathbf{W}_t is a matrix whose N^* rows are the rows of \mathbf{I}_N corresponding to the observed elements of \mathbf{y}_t . Similarly, $\mathbf{Z}_t^* = \mathbf{W}_t \mathbf{Z}_t$, $\Gamma_t^* = \mathbf{W}_t \Gamma_t$, $\varepsilon_t^* = \mathbf{W}_t \varepsilon_t$, and $\mathbf{H}_t^* = \mathbf{W}_t \mathbf{H}_t \mathbf{W}_t'$. The Kalman filter works exactly as described above, replacing \mathbf{y}_t , \mathbf{Z}_t , and \mathbf{H}_t with y_t^* , \mathbf{Z}_t^* , and \mathbf{H}_t^* . Similarly, after transformation the Kalman smoother remains valid with missing data.

Estimation

Thus far we have assumed known system parameters, whereas they are of course unknown in practice. As is well known, however, the Kalman filter supplies all of the ingredients needed for evaluating the Gaussian pseudo log-likelihood function via the prediction error decomposition,

$$\log L = -\frac{1}{2} \sum_{t=1}^T \{N \log 2\pi + (\log |\mathbf{F}_t| + \mathbf{v}_t' \mathbf{F}_t^{-1} \mathbf{v}_t)\} \tag{12}$$

In calculating the log likelihood, if all elements of \mathbf{y}_t are missing, the contribution of period t to the likelihood is zero. When some elements of \mathbf{y}_t are observed, the contribution of period t is

$[N^* \log 2\pi + (\log |F_t^*| + v_t^{*'} F_t^{*-1} v_t^*)]$ where N^* is the number of observed variables, and we obtain F_t^* and v_t^* by filtering the transformed y_t^* system.

f. A Prototype Empirical Application

We now present a simple application involving the daily term premium, fortnightly Narrow Money, monthly IIP, and quarterly NAGDP. We describe in turn the data, the specific variant of the model that we implement subtleties of our estimation procedure, and our empirical results.

Business Conditions Indicators

Our analysis covers the period from December 1, 2003 through July 2, 2010, which is 2406 observations of daily data. (We use a seven-day week.) We use four indicators. Moving from highest frequency to lowest frequency, the first indicator is the yield curve term premium, defined as the difference between 10-year Government securities and 3-month Indian Treasury yields. We measure the term premium daily; hence there are no aggregation issues. We treat holidays and weekends as missing. The second indicator is Narrow Money, a fortnightly stock variable covering the seven-day period from Sunday to Saturday. We set the end-of-fortnight value to the end-of-fortnight daily value, and we treat other days as missing. The third indicator is IIP, a monthly flow variable. We set the end-of-month value to the sum of the daily values within that month, and we treat other days as missing. The fourth and final indicator is NAGDP, a quarterly flow variable. We set the end-of-quarter value to the sum of daily values within the quarter, and we treat other days as missing. Basically, we want the variables chosen to illustrate the flexibility of our framework. Hence we choose four variables measured at four different frequencies ranging from very high (daily) to very low (quarterly), and representing both stocks (term premium, narrow money) and flows (IIP, NAGDP). Since we are seeking for a 'Business Cycle' indicator, we have extracted the cyclical component from the data (except the daily variable). This cyclical series are taken as our input.

Model Implementation

In the development thus far we have allowed for general polynomial trend and general $AR(p)$ dynamics. In the prototype model that we now take to the data, we make two simplifying assumptions that reduce the number of parameters to be estimated by numerical likelihood optimization. First, we de-trend prior to fitting the model rather than estimating trend parameters simultaneously with the others, and second, we use simple first-order dynamics throughout. In future work we look forward to incorporating more flexible dynamics but, as we show below, the framework appears quite encouraging even with simple $AR(1)$ dynamics.

Hence, latent business conditions \mathbf{X}_t follow zero-mean AR (1) process, as do the observed variables at their observational frequencies. For fortnightly narrow money, monthly IIP, and quarterly NAGDP, this simply means that the lagged values of these variables are elements of the \mathbf{W}_t vector. We denote these by \tilde{y}_{t-f}^2 , \tilde{y}_{t-M}^3 , and \tilde{y}_{t-q}^4 , where f denotes the number of days in a fortnight, M denotes the number of days in a month and q denotes the number of days in a quarter. Once again, the notation in the paper assumes M and q are constant over time, but in the implementation we adjust them according to the number of days in the relevant month or quarter. For the term premium, on the other hand, we model the autocorrelation structure using an AR (1) process for the measurement equation innovation, u_t^1 instead of adding a lag of the term premium in \mathbf{W}_t .

The equations that define the model are

$$\begin{aligned} \underbrace{\begin{bmatrix} \tilde{y}_t^1 \\ \tilde{y}_t^2 \\ \tilde{y}_t^3 \\ \tilde{y}_t^4 \end{bmatrix}}_{\mathbf{y}_t} &= \underbrace{\begin{bmatrix} \beta_1 & \beta_2 & \beta_3 & \beta_4 \\ 0 & \beta_2 & 0 & \beta_4 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \beta_2 & 0 & \beta_4 \\ 0 & 0 & 0 & \beta_4 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \beta_4 \text{ or } 0 \\ 0 & 0 & 0 & \beta_4 \text{ or } 0 \\ 0 & 0 & 0 & \beta_4 \text{ or } 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}}_{\mathbf{Z}_t} \underbrace{\begin{bmatrix} x_t \\ x_{t-1} \\ \vdots \\ x_{t-\bar{q}-1} \\ x_{t-\bar{q}} \\ u_t^1 \\ \alpha_t \end{bmatrix}}_{\alpha_t} \\ &+ \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ \gamma_2 & 0 & 0 \\ 0 & \gamma_3 & 0 \\ 0 & 0 & \gamma_4 \end{bmatrix}}_{\mathbf{\Gamma}_t} \underbrace{\begin{bmatrix} \tilde{y}_{t-f}^2 \\ \tilde{y}_{t-M}^3 \\ \tilde{y}_{t-q}^4 \end{bmatrix}}_{\mathbf{w}_t} + \underbrace{\begin{bmatrix} 0 \\ u_t^{*2} \\ u_t^3 \\ u_t^{*4} \end{bmatrix}}_{\mathbf{\varepsilon}_t}, \\ \underbrace{\begin{bmatrix} x_{t+1} \\ x_t \\ \vdots \\ x_{t-\bar{q}} \\ x_{t-\bar{q}+1} \\ u_{t+1}^1 \\ \alpha_{t+1} \end{bmatrix}}_{\alpha_{t+1}} &= \underbrace{\begin{bmatrix} \rho & 0 & \dots & 0 & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & \dots & 1 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & \gamma_1 \end{bmatrix}}_{\mathbf{T}} \underbrace{\begin{bmatrix} x_t \\ x_{t-1} \\ \vdots \\ x_{t-\bar{q}-1} \\ x_{t-\bar{q}} \\ u_t^1 \\ \alpha_t \end{bmatrix}}_{\alpha_t} + \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}}_{\mathbf{R}} \underbrace{\begin{bmatrix} \varepsilon_t \\ \zeta_t \end{bmatrix}}_{\mathbf{v}_t} \\ \begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} &\sim \mathcal{N} \left(\begin{bmatrix} \mathbf{0}_{4 \times 1} \\ \mathbf{0}_{2 \times 1} \end{bmatrix}, \begin{bmatrix} \mathbf{H}_t & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} \end{bmatrix} \right), \\ \mathbf{H}_t &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \sigma_{2t}^{*2} & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_{4t}^{*2} \end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 0 & \sigma_1^2 \end{bmatrix}, \end{aligned}$$

where the notation corresponds to the system discussed in the state-space representation with $N = 4$, $k = 3$, $m = 93$, $p = 1$, and $r = 2$. We use the current factor and 91 lags in our state vector because the

maximum possible number of days in a quarter is 92, which we denote by \bar{q} . (If there are q days in a quarter, then on the last day of the quarter we need the current value and $q - 1$ lags.) Also, in every quarter we adjust the number of nonzero elements in the fourth row of the \mathbf{Z}_t matrix to reflect the number of days in that quarter.

References

Anderson, Richard G. and Gascon Charles S., (2009), "Estimating U.S. Output Growth with Vintage Data in a State-Space Framework", *Federal Reserve Bank of St. Louis Review*, July/August, 91(4), pp. 349-69.

Aruoba, S. B. (2008), "Data Revisions Are Not Well-Behaved," *Journal of Money, Credit and Banking*, 40, 319–340.

Aruoba, S. B., Diebold Francis X. and Scotti Chiara, (2009), "Real-Time Measurement of Business Conditions", *Journal of Business and Economic Statistics*, October, Vol. 27, No. 4, pp. 417-427.

Burns, A. F., and Mitchell, W. C., (1946), *Measuring Business Cycles*, New York: NBER.

Diebold, Francis X. and Rudebusch, Glenn D., (1991), "Forecasting Output with the Composite Leading Index: A Real-Time Analysis", *Journal of the American Statistical Association*, September, 86(415), pp. 603-10.

-----, (1996), "Measuring Business Cycles: A Modern Perspective," *The Review of Economics and Statistics*, 78, 67–77.

Diebold, F. X., Rudebusch, G. D., and Aruoba, S. B., (2006), "The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach," *Journal of Econometrics*, 131, 309–338.

Durbin, J., and Koopman, S. J., (2001), *Time Series Analysis by State-Space Methods*, Oxford: Oxford University Press.

Evans, M. D. D., (2005), "Where Are We Now?: Real Time Estimates of the Macro Economy," *The International Journal of Central Banking*, September, 127–175.

Gartaganis, Arthur J. and Goldberger, Arthur S., (1955), "A Note on the Statistical Discrepancy in the National Accounts", *Econometrica*, April, Vol. 23, No. 2, pp. 166-173.

Hamilton, J. D., (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57, 357–384.

Harvey, A.C., (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.

Howrey, E. P., (1978), "The Use of Preliminary Data in Econometric Forecasting", *The Review of Economics and Statistics*, April, Vol. 60, No. 2, pp.193-200.

Howrey, E. P., (1984), "Data Revision, Reconstruction, and Prediction: An Application to Inventory Investment," *Review of Economics and Statistics*, 66, 386–393.

Liu, H., and Hall, S. G., (2001), "Creating High-Frequency National Accounts with State-Space Modelling: A Monte Carlo Experiment," *Journal of Forecasting*, 20, 441–449.

Mariano, R. S., and Murasawa, Y., (2003), "A New Coincident Index of Business Cycles Based on Monthly and Quarterly Series," *Journal of Applied Econometrics*, 18, 427–443.

Orphanides, A. and Norden, Simon van, (2002), "The Unreliability of Output-Gap Estimates in Real Time", *The Review of Economics and Statistics*, November, Vol. 84, No. 4, pp. 569-583.

Reserve Bank of India, (1998), "Report of the Working Group on Money Supply ", Reserve Bank of India, Mumbai.

Reserve Bank of India, (2002), "Report of the Working Group of Economic Indicators ", Reserve Bank of India, Mumbai.

Reserve Bank of India, (2007), "Report of the Technical Advisory Group on Development of Leading Indicators for the Indian Economy", Reserve Bank of India, Mumbai.

Stock, J. H., and Watson, M.W., (2002), "Macroeconomic Forecasting Using Diffusion Indexes," *Journal of Business & Economic Statistics*, 20, 147–162.