Towards a Usable Set of Leading Indicators for Hong Kong

William Chow[1]
Economics and Business Facilitation Unit, Financial Secretary's Office
Hong Kong Special Administrative Region Government, HONG KONG, CHINA
Corresponding author: William Chow, e-mail: william_chow@fso.gov.hk

Abstracts

This article considers the prospect of compiling leading economic indicators (LEI) that would help predict the Hong Kong economy, especially the turning points. While the pursuit of forecasting accuracy is comprehensible from both an intellectual and practical point of view, the most important role of LEI should be to shorten the recognition lag and implementation lag so that timely and effective policies can be better managed. In fact, forecasting in the context of LEI requires knowing *a priori* future values of the components which, in general, would not be possible without relying on other forecasting models. Thus, we will not emphasize out-of-sample accuracy of the compiled LEIs, but will focus on the timeliness of the warning signals. The component variables are selected using a combination of time domain and frequency domain analysis. Time disaggregation and seasonal adjustment are applied to the chosen variables, when appropriate, to yield smoothed monthly series. Three different approaches of compilation – the Conference Board type, a Dynamic Factor model, and a Neural Network model – are assessed using the dataset. Our results showed that the Neural Network specification delivers relatively the most precise and reliable signals as compared to the other two approaches.

Keywords: leading indicators, neural network model, Hong Kong

1. Introduction

This paper considers the functionality of various leading economic indicators (LEI) and the prospect of applying them in predicting economic upturns and downturns in Hong Kong. The methodologies surveyed are the OECD/Conference Board approach (CB), the Common Factor approach (COM) and the Neural Network approach (NN). Except for the autoregressive class of models, many LEIs are in itself inadequate tools for forecasting as there is a need to impute future values of the LEI components and this is often done with the aid of auxiliary statistical models. The forecasting accuracy of the LEIs thus depends in part on the accuracy of these toolkits. As such, our evaluation will not be based on the "point forecasts" of these indicators but on the ability of turning point detection which is often the prime focus of government agencies and statistics bureaus.

2. Data and Compilation

A total of 21 economic indicators (see Table 1) are selected for the compilation. Most of the variables are available on a monthly basis, and the two quarterly components are disaggregated into monthly series to give the raw data needed. These are then seasonally adjusted, and are evaluated individually in terms of their cross-correlations and dynamic correlations (Croux et al., 2001) with the disaggregated Real GDP series.

---

[1] The views and analysis expressed in the paper are those of the author and do not necessarily represent the views of the Economic Analysis and Business Facilitation Unit.

The dynamic correlations are essentially correlations of two data series at different frequencies. In our context, we try to identify the largest correlation over the medium to low frequencies to avoid spurious short run linkages. The largest dynamic correlations of the variables over all frequencies are stated in Table 1.

Table 1. Components used to compile leading economic indicators

Variables	Dynamic Correlation	Variables	Dynamic Correlation
Gross Domestic Fixed Capital Formation	0.7224	Retained Imports – Capital Goods	0.7682
Hang Seng Index	0.7991	Retained Imports – Consumer Goods	0.7578
M3	0.4240	Exports to Mainland China	0.8632
Retail Sales (Value)	0.8900	Total Loans	0.5628
Retail Sales (Volume)	0.9102	Total Deposits	0.3798
Electricity Consumption	-0.2113	Buildings with Consent to Comment Work – 8 qtr MA	0.4082
Air Cargo	0.8167	U.S. Orders PMI	0.4183
Sales & Purchase Agreements (Number)	0.4195	Inventory	0.5343
Sales & Purchase Agreements (Value)	0.5393	HIBOR 1 month – 12 month Spread	-0.2755
Inbound Tourists	0.5063	OECD Leading Indicator	0.6088
Retained Imports	0.8665		

It is not reasonable to expect all component series to exhibit leading property over RGDP. As long as they are not clear laggards to RGDP, they are included in the dataset. In fact, our exercise includes all of the 21 variables in the compilation as the differences from leaving out a subset are negligible. The final sample runs from April, 1997 to September, 2010.

Marcellino (2006) offers a concise review of the CB and COM approaches. The former is a non-model based approach widely received and practiced by governments and international organizations. Essentially, this is a fixed weight averaging scheme of the underlying components with the weights being inversely proportional to their variances. The COM approach in this paper belongs to the class of Factor Models promulgated by Stock and Watson (1989). The idea is to decompose a set of variables into a sum of common component(s) and idiosyncratic components. Notationally, we can write:

$$x_{it} = \lambda'_i f_t + u_{it} \qquad (1)$$

where λ_i is a *k* dimensional vector of factor loadings, f_t is a *k* dimensional vector of common factors for all the $i = 1, \cdots, 21$ components, and u_{it} is the idiosyncratic component of the *i*-th series. All the items on the right hand side of the equation above are unknown which is what makes the estimation complicated. In our study, we restricted *k* to 1 for simplicity. There are a few ways to solve for the common factor f_t, and the algorithm of Bai and Ng (2002) is adopted.

The last method we considered is the NN. This is a data-mining technique which prides itself in the area of pattern recognition and forecasting. It mimics the structure of biological neural systems and processes information via a group of artificial neurons. One can think of it as an input-output network with hidden layers of neurons that take on weighted sums of inputs. It learns from the gaps/errors between the desired output and the computed output and refines its estimates until certain criteria are met. The model considered in this paper is a multi-layer feedforward NN, reminiscent of the one of Qi (2001). It can be expressed as:

$$f(X) = g\left(\sum_{j=1}^{N_2} \beta_j g\left(\sum_{k=1}^{N_1} \alpha_{kj} x_k + \theta_j^1\right) + \theta_j^2\right) \qquad (2)$$

Here, x_k are the inputs (the 21 constituent components) and $f(X)$ is the output (the LEI). The function g is a logistic transfer function $g(z) = 1/(1 + e^{-z})$. In brief, the N_1 inputs are weighted by α_{kj}, aggregated into binary values by the logistic transfer function and corrected for the bias θ_j^1 before being fed to the N_2 neurons in the layer. The composite signal is then passed onto a layer with N_2 neurons where they are weighted by β_j, aggregated and corrected for bias θ_j^2 in a similar fashion. As in Qi's paper, the Levenberg-Marquardt algorithm[2] is used together with the criterion of minimizing mean squared errors. The parameters are set at: $N_1 = 21,\ N_2 = 4$. Unlike standard regression models, focus will not be on the interpretation of the inter-locking weights, but on how well the generated output $f(X)$ matches the reference series RGDP. A major difference between the neural network model here and Qi's is that we attempt to match $f(X_{t-6})$ to y_t instead of matching contemporary values. In a way, we use brute force to amplify any leading property the *X*s may have in them. A 3-month exponentially weighted moving average is applied to both the COM and NN leading indicators.

3. Performance Evaluation

We denote the compiled LEIs as CLI-CB, CLI-COM and CLI-NN. Figure 1 illustrates the resulting LEIs as compared to RGDP. The shaded regions in the diagrams represent the contraction periods derived using the Harding-Pagan (2002) algorithm. As mentioned in the introduction, a crucial function of leading indicators is to detect turning points of economic activities. There are *ad hoc* detection methods as well as model-based detection methods that serve this purpose. For models with an explicitly defined probabilistic structure like in Bayesian models, the latter will be straightforward. In our case, *ad hoc* methods seem more applicable. Specifically, we consider (i) the conventional three month rule – 3 consecutive monthly declines in the LEI signify an economic downturn; and (ii) the 4/7 rule – 4 monthly declines in the past 7 months as a signal of downturn. Of the three LEIs, CLI-NN exhibits the earliest lead time in generating warning signals overall. It also has the least false signals, but unfortunately, it also missed the mild recession in the SARS period. In those cases where it gave correct warnings, the signals started either on or before the peak dates. There is much less consistency in the signals generated by the other two LEIs as far as the 3-month rule is concerned. Regarding the 4/7 rule, the performances are mixed for the three LEIs with no clear winners.

We next compare the growth cycles implied by the various LEIs and those of RGDP. Growth cycles are deviations from long term growth trend, and are used to measure the momentum of the underlying economy[3]. For leading indicators, it would be most ideal to have procyclical movements between these cycles and those of the reference series, and with a lead time. Figure 2 plots the growth cycles of the LEIs and of RGDP. They are derived using the bandpass filter, see Christiano and Fitzgerald (2003), tuned to shut down signals of frequency less than 18 months and more than 96 months.

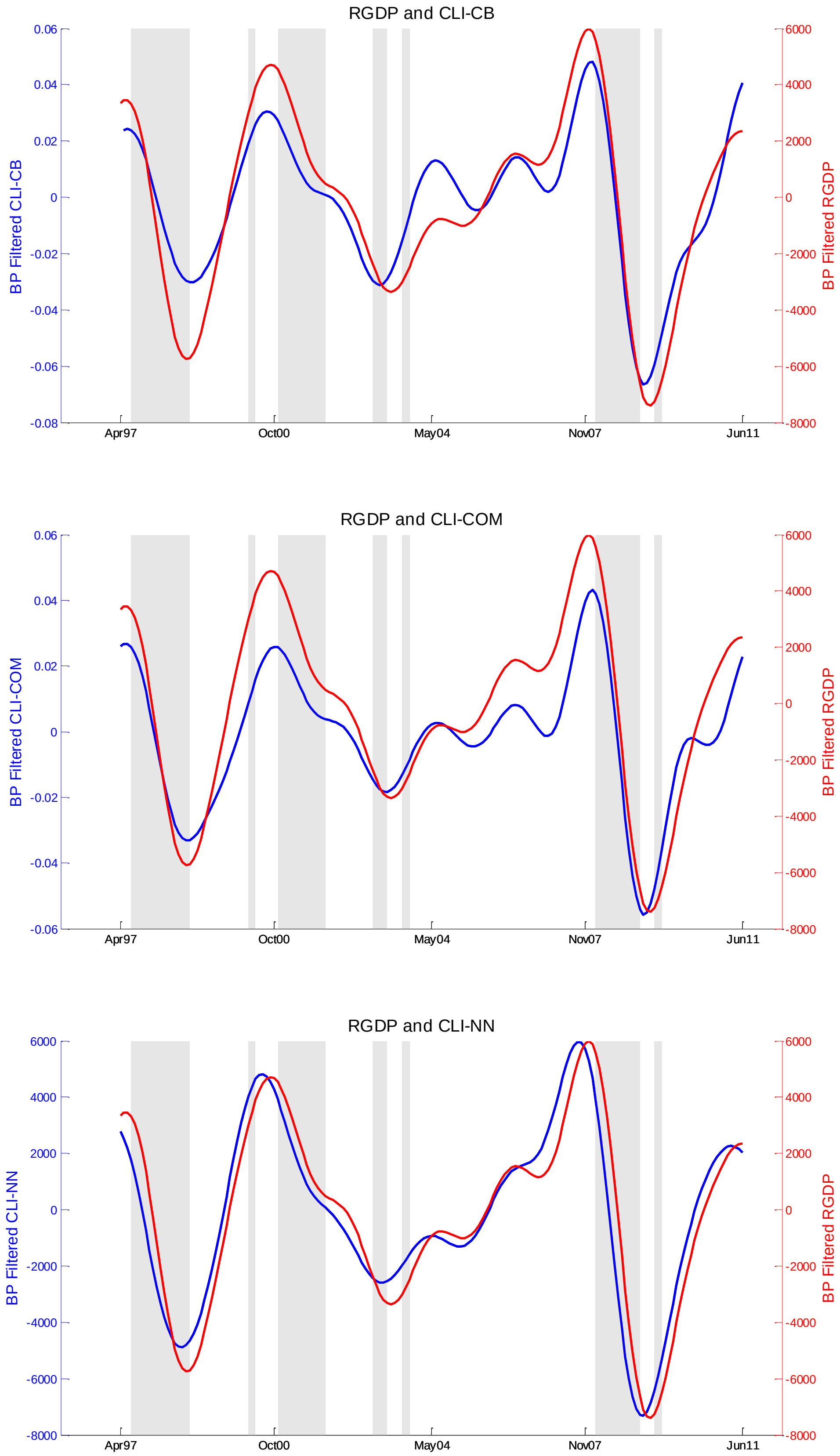
---

[2] This is a hybrid form of optimization technique which interpolates between the Gauss-Newton and the gradient descent methods.

[3] These may or may not perfectly coincide with actual cycles.

Figure 2: RGDP and the various compiled LEIs

Figure 3: Growth cycles of RGDP and the compiled LEIs



The filtered RGDP is marked in red in these diagrams. It can be seen that there were three peaks in these growth cycles – in mid 1997, in the second half of 2000 and 2007. The growth cycles of all three LEIs have up- and down-swings largely matching those of RGDP. However, only CLI-NN shows a clear lead over the RGDP cycle. An analysis of the cross-correlations of the growth rates of RGDP and the LEIs showed that CLI-NN leads RGDP by about 5-6 months. Unfortunately, when it comes to policy-making, not all these lead time can be exploited as we need 3 months to confirm a signal because of the lag in data compilation and reporting.

Conclusions

This article reviews the usefulness of leading indicators compiled based on the Conference Board approach, the Factor Model approach and the Neural Network approach. Performance of the compiled indices are assessed and compared in terms of the timeliness and the accuracy of the signals generated. It is found that the Neural Network approach has a slight edge over the other two methods when judged by the performance benchmarks. The longest lead time we found is about 5-6 months but the actual usable time prior to a growth reversal is much less than that when issues like delay in data reporting is taken into account.

References

Bai, J. and Ng, S. (2002) "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70, 191-221.

Christiano, L.J. and Fitzgerald, T.J. (2003) "The Band Pass Filter," *International Economic Review*, 44(2), 435-465.

Croux, C., Forni, M. and Reichlin, L. (2001) "A Measure of Comovement for Economic Variables: Theory and Empirics," *The Review of Economics and Statistics*, 83, 232-241.

Harding, D. and Pagan, A. (2002) "Dissecting the cycle: a methodological investigation," *Journal of Monetary Economics*, 49(2), 365-381.

Marcellino, M. (2006) "Leading Indicators: What have we learned?" in G. Elliott, C.W.J. Granger and A. Timmermann (eds.) *Handbook of Economic Forecasting*, Elsevier, 879-960.

Qi, M. (2001) "Predicting US recessions with leading indicators via neural network models," *International Journal of Forecasting*, 17, 383-401.

Stock, J.H. and Watson, M.W. (1989) "New Indexes of Coincident and Leading Economic Indicators," in O.J. Blanchard and S. Fischer (eds.) *NBER Macroeconomic Annual*, MIT Press.