Economic Letter 2019/08

Composite economic indicators in Hong Kong¹

Dr HO Chi Pui Economist

July 2019

Abstract

This article evaluates three approaches to construct composite economic indicators: the OECD/Conference Board approach, the common factor approach and the artificial neural network approach. By design, a composite economic indicator constructed by the artificial neural network approach performs the best in detecting recessions and tracking growth cycles, with a lead time of about 4-5 months. However in practice not all this lead time can be exploited to provide early warnings as 3-4 months are required to confirm a recession signal (by the 3-month rule or the 4/7 rule). The OECD/Conference Board CEI and Neural Network CEI are also helpful for improving the accuracy of real GDP forecasts.

香港的綜合經濟指標

摘要

本文評估三種建構綜合經濟指標的方法:經合組織/會議委員會方法、共同因素方法及人工神經網絡方法。通過設計,以人工神經網 絡方法建構的綜合經濟指標在檢測經濟衰退和追踪增長周期方面表 現最佳,領先時間約為4-5個月。然而,由於實踐中需要3-4個月來 確認經濟衰退的信號(通過3個月規則或4/7規則),固並非所有 這些領先時間都可用於提供早期預警。以經合組織/會議委員會及人 工神經網絡方法建構的綜合經濟指標亦有助提高預測實質本地生產 總值的準確性。

The views and analysis expressed in this article are those of the author and do not necessarily represent the views of the Office of the Government Economist.

¹ This article is an update of Dr. William Chow's work on "Leading Indicators of Hong Kong Economic Output" dated February 2011.

I. INTRODUCTION

1. This article considers the functionality of various composite economic indicators (CEIs) and the prospect of using them to predict aggregate economic activity in Hong Kong. CEIs compiled by three methodologies are surveyed: the OECD/Conference Board (CB) approach, the common factor approach, and the artificial neural network approach. The performance of these CEIs in detecting recessions, tracking growth cycles and forecasting real GDP figures is evaluated.

II. BACKGROUND

2. Business cycle indicators and related composite indices are useful tools for analysing business cycles. The National Bureau of Economic Research (NBER) in the United States began to use the indicator approach to analyse and forecast business cycles in the 1930s. From the late 1960s onwards, the U.S. Department of Commerce Bureau of Economic Analysis (BEA) began to publish the business cycle indicators and relevant indices. In 1995-1996, the business cycle indicators programme was privatised and the Conference Board took over the responsibility of publishing the monthly report on business cycle indicators and relevant indices.²

3. For business cycle indicators, over the years, academics and practitioners have classified economic variables according to their co-movements with aggregate economic activity, as proxied by real GDP. In *Table 1*, the business cycle properties of 31 economic variables in Hong Kong have been surveyed.³ This table shows the cyclicality, lead-lag relationship and the maximum cross correlation (in absolute value) of the cyclical component of the concerned variables with the cyclical component of real GDP, based on quarterly data from 1973Q1 to 2018Q4.

² Ozyildrim, A. (2017). "Business Cycle Indicator Approach at the Conference Board." *Handbook on Cyclical Composite Indicators for Business Cycle Analysis*. Luxembourg: Publications Office of the European Union, 225-240.

³ Please refer to Dr. Chi Pui Ho's related work on "Stylised facts on business cycles in Hong Kong" dated 28 June 2019 for details on the methodology.

Variable	Cyclicality	Leads (+) /	Maximum
		lags (-)	cross
		real GDP by	correlation
Labour market variables			
Unemployment rate	Countercyclical	-1 quarter	-0.75
Underemployment rate	Countercyclical	-1 quarter	-0.69
Price variables			
Composite consumer price index (CCPI)	Acyclical		
Quarter to quarter change in CCPI	Procyclical	0 quarter	0.36
Effective exchange rate indices for Hong	Countercyclical	+1 quarter	-0.32
Kong Dollar			
Oil (WTI) price	Procyclical	+1 quarter	0.39
CRB index	Procyclical	0 quarter	0.73
Gold spot	Procyclical	+1 quarter	0.38
Property market variables			
RVD's residential property price index	Procyclical	0 quarter	0.49
RVD's residential property rental index	Procyclical	-1 quarter	0.82
Number of S&P agreements for	Procyclical	+1 quarter	0.36
residential building units			
Financial market variables			
Hang Seng Index (period end)	Procyclical	+1 quarter	0.74
Aggregate balance	Countercyclical	-2 quarters	-0.55
1-month HIBOR	Procyclical	0 quarter	0.51
3-month HIBOR	Procyclical	-1 quarter	0.52
12-month HIBOR	Procyclical	-2 quarters	0.43
HSBC best lending rate	Procyclical	-3 quarters	0.34
Mortgage rate (BLR-based)	Procyclical	-3 quarters	0.33
M1	Acyclical		
M2	Procyclical	0 quarter	0.43
M3	Procyclical	0 quarter	0.44
Loans and advances	Procyclical	-1 quarter	0.63
Total deposits	Procyclical	0 quarter	0.40
Tourism variables			
All visitor arrivals	Procyclical	0 quarter	0.46
Hotel occupancy rate	Procyclical	+1 quarter	0.78
Logistics and transport variables			
Total container throughput	Procyclical	0 quarter	0.23
Air cargo throughput	Procyclical	+2 quarters	0.85
Air passenger traffic	Procyclical	0 quarter	0.52
Average daily cross-boundary vehicles	Procyclical	+2 quarters	0.56
Total number of water-borne arrivals and	Procyclical	0 quarter	0.60
departures			
Total number of road-borne arrivals and	Procyclical	0 quarter	0.48
departures			

Table 1: Business cycle properties of selected economic variables in Hong Kong

Remarks: Data are deseasonalised with the U.S. Census Bureau's X13-ARIMA model and detrended with the Hodrick-Prescott filter. Lead and lags are in quarters, with a plus sign indicating that the variable is a leading variable.

4. While *Table 1* covers the business cycle characteristics of variables across many diverse aspects of economic activity, they are not independent of one another. They reflect underlying business cycles which exert influence throughout the economy. Consequently, the individual variables can be combined into composite economic indicators (CEIs) that utilise more of the information in the data and exhibit less volatility. Ideally, the CEIs would also provide early signals of turning points in business cycles and especially recessions, allowing a timely analysis of the current economic situation. This article surveys three methods to compile CEIs.

III. COMPILATION OF CEIs

5. The component variables to be included in the compilation of CEIs are selected based on a number of criteria (OECD, 2012; Ozyildrim, 2017):⁴

- **Consistent timing**: the component variables should exhibit consistent patterns as leading, coincident or lagging indicators. Monthly series are preferred to quarterly series.
- **Economic significance**: the cyclicality of the component variables should be economically meaningful and logical.
- **Statistical adequacy**: the time series of the component variables should be collected in a statistically reliable way.
- **Timeliness**: the component variables should be timely, being available very soon after the period used to construct the CEI.
- Length and revision: long time series with no breaks and that are not subject to significant revisions are preferred.

6. Excluding the two acyclical variables in *Table 1*, and air cargo throughput which is only available from January 2008, the 28 remaining economic variables fulfil the above criteria and are used to construct the CEIs in this article. Monthly data on these variables has been available since January 1999. The methodologies to construct three types of CEIs are briefly summarised below:

⁴ OECD (2012). "OECD System of Composite Leading Indicators." Available at <u>https://www.oecd.org/sdd/leading-indicators/41629509.pdf</u> (Accessed on 5 June 2019). Ozyildrim, A. (2017). "Business Cycle Indicator Approach at the Conference Board." *Handbook on Cyclical Composite Indicators for Business Cycle Analysis*. Luxembourg: Publications Office of the European Union, 225-240.

III.1. Conference Board/OECD-Type CEI

7. This is the approach commonly adopted by government bureaux and quasigovernmental bodies. It requires relatively simple aggregation procedures to assign weights to the individual economic variables. In this article, weights are assigned to N = 7 leading economic variables in *Table 1* (except air cargo throughput for which data are only available from January 2008). This involves the following main steps:

- (a) Calculate the month-to-month rates of change for each of the X13-ARIMA deseasonalised individual variables (components x_{it}), where i = 1, ..., N. If the component is not in percentage form, the rate of change formula is $r_{it} = 200 \cdot \frac{x_{it} x_{it-1}}{x_{it} + x_{it-1}}$; if the component is in percentage form, the rate of change formula is $r_{it} = x_{it} x_{it-1}$. If a component is countercyclical, its month-to-month change is multiplied by -1 to invert it before adding it to the composite index.
- (b) Assign to each component a weight w_i that equals the normalised inverse standard deviation of the month-to-month changes calculated in step (a):

$$w_i = \left(\frac{1}{s_i}\right) / \sum_{i=1}^N \frac{1}{s_i}$$

where s_i is the standard deviation of the month-to-month changes of x_{it} .

(c) Multiply each component by its corresponding weight and sum up the monthto-month changes of the individual components:

$$C_t = \sum_{i=1}^N w_i r_{it}$$

(d) The weighted sum is then adjusted to match the mean and standard deviation of the month-to-month growth rates of real GDP:

$$\tilde{C}_t = \mu_{real \, GDP \, growth} + \frac{S_{real \, GDP \, growth}}{S_C} (C_t - \mu_C)$$

(e) The CEI is then calculated according to the following method: the initial value is set at $I_1 = 100$ for the first month of the sample period

(January 1999). The index in the second month is $I_t = I_{t-1} \cdot \frac{200 + \tilde{C}_t}{200 - \tilde{C}_t}$, and this formula is used successively to calculate the indices in the following periods.

III.2 Common Factor Based CEI

8. The second approach is to employ the class of factor models developed by Stock and Watson (2002) that decomposes the evolution of a high-dimension observable dataset into a few unobservable common factors that are used in predicting future economic activity.⁵ Technically, this approach reduces a high-dimension vector of observable economic variables to a few latent common factors plus a vector of mean-zero idiosyncratic disturbance components. In this article, the static factor model is applied to N = 28 individual economic variables:

$$X_t = \Lambda \cdot F_t + e_t$$

where X_t is the $N \times 1$ vector of the set of observable deseasonalised economic variables at time t, Λ is the $N \times q$ vector of factor loadings, F_t is a q dimensional vector of common factor at time t, and e_t is the $N \times 1$ vector of idiosyncratic disturbance components at time t.

9. Under the factor model, the correlation matrix of X_t , Σ , is decomposed as follows:

$$\Sigma = \Lambda \Phi \Lambda' + \Psi$$

Assuming uncorrelated common factors, $\Phi = I$. Then, by finding appropriate eigenvectors for the correlation matrix, the factor loading Λ can be computed as the leading eigenvector scaled by the square root of the appropriate eigenvalue. Following Stock and Watson (2002), we then perform the OLS regression as below :

$$y_{t+6} = \beta'_F F_t + \beta'_y y_t + \varepsilon_{t+h}$$

where y_t is deseasonalised real GDP at time t and ε_t is the error term. The common factor based CEI is simply the fitted value of y_t .

⁵ Stock, J.H., & Watson, M.W. (2002). "Forecasting using principal components from a large number of predictors." *Journal of the American Statistical Association*, 97(460), 1167-1179.

III.3. Artificial Neural Network CEI

10. An artificial neural network (ANN) is a data-mining and forecasting method that is based on a self-learning pattern recognition process. An ANN learns from examples and captures subtle relationships within the data even if prior knowledge of relationships among the data are unknown. This modelling approach, with the ability to learn from experience, is valuable for practical problems for which data are readily available but less suited for making theoretical inferences about the underlying processes that generate the data.

11. At its core, an ANN is a machine learning algorithm that mimics biological neural networks for receiving and processing information. It uses successive layers of processing nodes to convert input to output, with each layer relying on inputs from the layer before (or the raw data) to identify higher-level features of the data. An ANN learns from discrepancies between predicted output and observed output until it finds a set of weights on input that minimizes some overall error measures.

12. Applying this concept to the computation of CEIs involves finding the optimal weights of the economic variables used to compute the CEI with a self-learning pattern recognition process. The three-layer feedforward ANN (Qi, 2001), as shown below, is the most widely used model and is adopted in the present study:

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

where f(X) is the target dependent variable, x_k are the deseasonalised component economic variables, g is a logistic function $g(z) = \frac{1}{1+e^{-z}}$, and ε is the error term. The equation says that the N_1 inputs are weighted by α_{kj} , corrected for the bias θ_j^1 and transformed by the logistic function, and then fed into the N_2 processing functions in the second layer. Then the N_2 inputs are weighted by β_j , corrected for the bias θ_j^2 and transformed by the logistic function; this gives the predicted output dependent variable (the third layer). The parameters θ_j^1 , θ_j^2 , α_{kj} , β_j are estimated by the Levenberg-Marquardt algorithm, which is by far the fastest algorithm for moderatesized (up to several hundred free parameters) feedforward ANNs:⁶

$$\operatorname{argmin}_{\left\{\theta_{j}^{1},\theta_{j}^{2},\alpha_{kj},\beta_{j},N_{2}\right\}}\Sigma\varepsilon^{2}$$

⁶ Qi, M. (2001). "Predicting US recessions with leading indicators via neural network models." *International Journal of Forecasting*, 17(3), 383-401.

The data from January 1999 to September 2018 are used as the training data to estimate the parameters, and those from October 2018 to March 2019 are saved as the testing data. The major difference between the neural network CEI in this article and Qi (2001)'s is that the normalised real GDP $f(X_{t+6})$ is matched to $f(X_t)$ plus the error term, that is, brute force has been used to amplify any leading property the Xs may have.⁷

In this paper, the parameter N_1^{\wedge} is set to be 28 (i.e. including all individual economic variables), while $N_2^{\wedge} = 4$, which minimizes the sum of squared errors $\sum \varepsilon^2$. The 3-month simple moving average is then applied to the constructed CEI (which is rebased to 100 at January 1999) to smooth volatility.

IV. PERFORMANCE EVALUATION

13. *Figure 1* shows real GDP and various CEIs from January 1999 to March 2019. There are three ways to evaluate CEI performance: (1) recession detection with ad hoc rules; (2) growth cycle prediction; and (3) forecasting (deseasonalised) real GDP.



Figure 1: Real GDP and various CEIs

Remark: The real GDP figures shown in *Figure 1* are deseasonalised with the X13-ARIMA method. For example, in *Figure 1*, the nearly flat blue line in 2018 means that, after removing the seasonal factors, there was little underlying growth in real GDP.

⁷ For more details on the methodology of ANNs in forecasting business cycles, see Qi (2001).

14. The first application of CEIs is to detect recessions. For the purpose of this assessment, recession periods are defined as two periods of negative growth in deseasonalised real GDP (Wecker, 1979; Harding and Pagan, 2002).⁸ Two ad hoc detection rules are applied to compare the recession detection ability of the three CEIs. The first is the 3-month rule, where three consecutive monthly declines would signify an economic downturn. Another is the 4/7 rule, where four monthly declines of the CEI in the past seven months would signify an economic downturn. *Table 2* shows the performance of the three CEIs in detecting recession episodes since 1999, and *Table 3* shows the false recession alarms produced by the three CEIs over the same timeframe.

Recession period	Detected by CB/OECD CEI?		Detected by Common Factor CEI?		Detected by Neural Network CEI?	
	3-month	4/7 rule	3-month	4/7 rulo	3-month	4/7 rule
Early 2001	Yes	Yes	No	No	Yes	Yes
5	(+3 m)	(+4 m)			(+3 m)	(-1 m)
Mid 2001-early 2002	No	Yes	No	Yes	Yes	Yes
		(+3 m)		(-6 m)	(+3 m)	(+5 m)
Early 2003	No	No	No	Yes	Yes	Yes
				(-4 m)	(+4 m)	(+0 m)
Mid 2008	No	No	Yes	No	Yes	Yes
			(-1 m)		(+3 m)	(+2 m)
Late 2008-early 2009	Yes	Yes	Yes	Yes	Yes	Yes
	(-1 m)	(-2 m)	(-2 m)	(+0 m)	(+7 m)	(+6 m)
Early 2016	Yes	Yes	No	No	Yes	Yes
	(+5 m)	(+6 m)			(+4 m)	(+5 m)
Mid 2018	No	No	No	No	No	No
Late 2018	Yes	Yes	Yes	Yes	Yes	Yes
	(+4 m)	(+3 m)	(-1 m)	(+1 m)	(-1 m)	(+4 m)

Table 2: Recession detection by three CEIs using ad hoc rules

Note: The number in brackets is the number of months between the actual recession and the first recession signal generated by the CEI (a plus sign indicates that the first CEI recession signal led the actual recession). Except for late 2018, all the detection results for the Neural Network CEI shown in this table are within the training period for the model.

⁸ Wecker, W.E. (1979). "Predicting the turning points of a time series." *Journal of Business*, 52(1), 35-50. Harding, D., & Pagan, A. (2002). "Dissecting the cycle: a methodological investigation." *Journal of Monetary Economics*, 49(2), 365-381.

CB/OECD CEI		Common Factor CEI		Neural Network CEI	
3-month	4/7	3-month	4/7	3-month	4/7
rule	rule	rule	rule	rule	rule
	Mid 2013		Late 2011		
	Early 2015				
Total: 0	Total: 2	Total: 0	Total: 1	Total: 0	Total: 0

Table 3: False recession alarms by three CEIs using ad hoc rules

15. Comparing the three CEIs, by design the neural network CEI in general displays the earliest warning signals for economic downturns, with an average leading time of around 4 months. Similarly, it is the only CEI that can provide a leading indication of a large and obvious recession like the 2008-2009 recession. While the neural network CEI seldom produces false recession alarms, it unfortunately also misses the mild recession in mid 2018. The CB/OECD CEI and common factor CEI fail to detect quite a number of recession episodes. Overall, the neural network CEI is the better-performing CEI in terms of providing early warning signals for recessions.

IV.2. Business Cycle Prediction

16. The second application of CEIs is to predict business cycle movements (including predicting the amplitude of booms/recessions in addition to timing). Business cycles are associated with expansions and contractions in economic activity. In particular the analysis based on detrended data, i.e. deviations from the long-term growth trend or the cyclical component of the underlying economy, is referred to as growth cycle analysis. The growth cycles implied by the three CEIs can be compared to the growth cycle of (deasonalised) real GDP. Ideally, the growth cycles implied by the CEIs would replicate the growth cycle of real GDP, but in advance. To explore this, the growth cycles of real GDP and the three CEIs are compiled by applying Christiano and Fitzgerald (2003)'s band-pass filter: the parts of real GDP and the three CEIs that reflect the business cycle are identified to be those with frequencies between 18 and 96 months, and the parts with longer frequencies (i.e. the trend) and

higher frequencies (i.e. the noise) are taken out.⁹ The remaining portions are the growth cycles of real GDP and the three CEIs, as depicted in *Figure 2*.

17. *Figure 2* shows that the growth cycles of the three CEIs (except Common Factor CEI before mid-2000s) possess growth cycles with up- and down-swings largely matched to those of real GDP. Yet, only the neural network CEI cycle shows a clear lead over the real GDP cycle. The same conclusion can be drawn if the cross-correlations of the year-on-year growth rates of the three CEIs and real GDP are plotted (*Figure 3*). In general, the growth cycle of neural network CEI leads growth cycle of real GDP by around 5 months. The growth cycle of Conference Board-CEI leads the growth cycle of real GDP by around 2 month. The growth cycle of Common Factor CEI is coincident with the growth cycle of real GDP. Again the better leading property of the neural network CEI reflects its design of "brutally" matching the CEI to the real GDP six months ahead.

⁹ Christiano, L.J., & Fitzgerald, T.J. (2003). "The band pass filter." International Economic Review, 44(2), 435-465.

















18. The third application of CEIs is to help forecast (deseasonalised) real GDP. It is investigated whether the three CEIs serve this purpose. A simple vector autoregressive model with six lags is employed. Four models are considered: the benchmark model is a univariate AR(6) model with real GDP being the only variable, that is,

$$y_t = \beta_0 + \sum_{i=1}^6 \beta_i y_{t-i} + \varepsilon_i$$

where y is real GDP and ε_i is the error term. Then real GDP is forecast as follows:

$$b_{t+h} = \hat{\beta}_0 + \sum_{i=1}^6 \hat{\beta}_i \, y_{t+h-i}$$

where *b* is predicted real GDP using the benchmark model, $\hat{\beta}_0$ and $\hat{\beta}$ are estimated coefficients. The other three models regress the vector of real GDP and each of the three CEIs on their lagged values, that is,

$$\begin{pmatrix} y_t \\ CEI_t \end{pmatrix} = \begin{pmatrix} \gamma_{10} \\ \gamma_{20} \end{pmatrix} + \sum_{i=1}^6 \begin{pmatrix} \Pi_{11}^i & \Pi_{12}^i \\ \Pi_{21}^i & \Pi_{22}^i \end{pmatrix} \begin{pmatrix} y_{t-i} \\ CEI_{t-i} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

where y is real GDP, *CEI* is the CEI being tested, and $(\varepsilon_{1t} \quad \varepsilon_{2t})'$ is the error term. Then real GDP is forecast as follows :

$$\hat{y}_{t,h} = \hat{\gamma}_{10} + \sum_{i=1}^{6} \widehat{\Pi}_{11}^{i} y_{t+h-i} + \sum_{i=1}^{6} \widehat{\Pi}_{12}^{i} CEI_{t+h-i}$$

where $\hat{y}_{t,h}$ is forecast real GDP using the CEI, *h* is the forecast horizon, and $\hat{\gamma}_{10}$, $\hat{\Pi}_{11}^{i}$ and $\hat{\Pi}_{12}^{i}$ are estimated coefficients.

19. The forecast performance is evaluated and compared by using Theil's U statistic, which is the ratio of the root mean square error of a given model relative to that of the benchmark forecast:

$$U = \frac{\sqrt{\frac{1}{T - T_0} \sum_{t=T_0}^{T} (y_{t+h} - \hat{y}_{t,h})^2}}{\sqrt{\frac{1}{T - T_0} \sum_{t=T_0}^{T} (y_{t+h} - b_{t+h})^2}}$$

where *T* is the full sample size, T_0 is the first out-of-sample observation, *h* is the forecast horizon, *y* is realized real GDP, \hat{y} is forecast real GDP using the CEI, and *b* is predicted real GDP using the benchmark model. If U < 1, the model with CEI outperforms the benchmark model in forecasting real GDP, and vice versa. *Table 4* summarises Theil's U statistics for the three CEI models for forecasting real GDP from October 2018 to March 2019.¹⁰

Horizon	CB/OECD CEI	Common Factor CEI	Neural Network CEI
1-month	0.50	0.77	0.94
2-month	0.40	0.77	0.88
3-month	0.42	0.94	0.88
4-month	0.36	1.04	0.85
5-month	0.39	1.17	0.83
6-month	0.65	1.26	0.83

 Table 4: Theil's U statistics of various CEIs

20. The Theil's U statistics for CB/OECD CEI and Neural Network CEI are all smaller than 1, meaning that these two CEIs would help improve the forecast accuracy for real GDP as compared to the univariate autoregressive model. Note that the Theil's U statistics for the Common Factor CEI were larger than one for forecast horizons of four months or more, meaning that univariate autoregressive model outperformed the model with common factor CEIs as additional regressors for longer forecast horizons. Notwithstanding that the latter model is nested in the former, this remains theoretically possible because the forecasts are out-of-sample forecasts. We also note that the relative performance of the three CEIs in predicting real GDP might change over different forecast periods.

¹⁰ The monthly real GDP figures are obtained by linearly interpolating quarterly real GDP figures and then deseasonalised by applying X13-ARIMA regression. For the three CEI VAR models, real GDP data from January 1999 to September 2018 are used to estimate the VAR models, and the real GDP data from October 2018 to March 2019 are reserved for evaluating forecast accuracy.

V. CONCLUSION

21. This article evaluates three approaches to construct composite economic indicators: the OECD/Conference Board approach, the common factor approach and the neural network approach. By design a composite economic indicator constructed by the neural network approach performs the best in detecting recessions and tracking growth cycles with a lead time of about 4-5 months. However in practice not all this lead time can be exploited to provide early warnings as 3-4 months are required to confirm a recession signal (by the 3-month rule or the 4/7 rule). The CB/OECD CEI and Neural Network CEI are also helpful for improving the accuracy of real GDP forecasts.