Pundit, Madhavi; Ramayandi, Arief; Simba, Patrick Jaime A.; Sorino, Dennis E.; Tan, Sharyl Rose S.

Working Paper

Monitoring business cycle fluctuations in Asia

ADB Economics Working Paper Series, No. 766

Provided in Cooperation with:
Asian Development Bank (ADB), Manila

Suggested Citation: Pundit, Madhavi; Ramayandi, Arief; Simba, Patrick Jaime A.; Sorino, Dennis E.; Tan, Sharyl Rose S. (2025) : Monitoring business cycle fluctuations in Asia, ADB Economics Working Paper Series, No. 766, Asian Development Bank (ADB), Manila, https://doi.org/10.22617/WPS250019-2

MONITORING BUSINESS CYCLE FLUCTUATIONS IN ASIA

Madhavi Pundit, Arief Ramayandi, Patrick Jaime A. Simba, Dennis E. Sorino, and Sharyl Rose S. Tan

NO. 766

January 2025

ADB ECONOMICS WORKING PAPER SERIES

ASIAN DEVELOPMENT BANK

ADB

ADB Economics Working Paper Series

Monitoring Business Cycle Fluctuations in Asia

Madhavi Pundit, Arief Ramayandi,
Patrick Jaime A. Simba, Dennis E. Sorino,
and Sharyl Rose S. Tan

No. 766 | January 2025

Madhavi Pundit (mpundit@adb.org) is a senior economist, Dennis E. Sorino (dsorino@adb.org) is a senior economic officer, and Patrick Jaime A. Simba (psimba.consultant@adb.org) and Sharyl Rose S. Tan (srsy.consultant@adb.org) are consultants at the Economic Research and Development Impact Department, Asian Development Bank.
Arief Ramayandi (aramayandi@adbi.org) is a senior research fellow at the Asian Development Bank Institute.

ASIAN DEVELOPMENT BANK

6 ADB Avenue, Mandaluyong City, 1550 Metro Manila, Philippines
Tel +63 2 8632 4444; Fax +63 2 8636 2444
www.adb.org

ISSN 2313-6537 (print), 2313-6545 (PDF)
Publication Stock No. WPS250019-2
DOI: http://dx.doi.org/10.22617/WPS250019-2

Notes:
In this publication, "$" refers to United States dollars unless otherwise indicated.
ADB recognizes "China" as the People's Republic of China.

ABSTRACT

Timely updates of business cycle fluctuations—commonly represented by movements in the output gap—help policymakers make informed decisions on the appropriate course of action. Unfortunately, business cycle assessments often suffer from lags in actual gross domestic product data releases. This paper proposes the use of an Economic Activity Index, developed to monitor monthly business cycle fluctuations in Asia. The index summarizes a selection of updated monthly economic indicators to gauge movements in the output gap. The paper shows that the application of machine learning models substantially improves the ability of the index to track actual fluctuations of the business cycle compared with models constructed using a traditional principal component analysis. Grouping the information used to construct the index into six categories—consumption, investment, trade, government, financial, and the external sector—makes it possible to break down and explain drivers of movements in the business cycle.

Keywords: macroeconomic monitoring, tracking business cycles, economic fluctuations, nowcasting

JEL codes: C32, C63, E32, E37

1. INTRODUCTION

Real gross domestic product (GDP) is a comprehensive and key indicator in tracking economic activity. Unfortunately, in many developing and emerging economies, GDP data are published at low frequency and with a significant lag, meaning timely and complete information is unavailable for real-time interventions and policy setting. This situation often leads to a gap in timely decision making, as effective policy intervention requires a thorough understanding of current conditions and anticipated trends.

Tracking current economic conditions is a challenge for policymakers. Fortunately, despite the inherent delays in GDP data reporting, relevant high-frequency data are often available to provide valuable insights for policymakers to use to gauge the most recent conditions of economic activity. It is possible to exploit this broad range of high-frequency indicators by developing frameworks to predict current movements in aggregate economic activity. This refers to a "nowcasting" exercise, which is the practice of predicting current economic conditions using readily available data.

Nowcasting has become popular because key economic indicators such as GDP growth often suffer from substantial reporting lags. Nowcasting models are now widely used by central banks and other institutions to provide real-time assessments of the economy, enabling timely decisions in policymaking. For example, the Federal Reserve Banks of Atlanta and New York employ a mixed-frequency Dynamic Factor Model (DFM) framework to develop their quarterly GDP growth nowcasting model using high-frequency indicators (Higgins 2014). Their models continuously update US GDP growth estimates as new monthly data become available to reflect the most current economic conditions. A similar framework has been applied in developing Asian economies, by Bragoli and Fosten (2018) for India and Luciani et al. (2018) for Indonesia.[1]

Although nowcasting GDP growth is important for informing policy direction, relying solely on GDP growth figures is insufficient. Policymakers should also recognize the significance of the output gap, which measures the difference between an economy's actual output and its potential at full capacity. This gap indicates whether an economy is overheating or underperforming, guiding the appropriate monetary and fiscal policy direction. By nowcasting the output gap, policymakers can gain a clearer picture of economic health. This involves using timely data to estimate the current state of economic activity relative to its potential.

This paper constructs an index to gauge the movements of the output gap monthly by utilizing multiple indicators to track economic expansions or contractions in selected Asian economies. The index is labeled the Economic Activity Index (EAI) and provides policymakers with a current picture of the GDP growth gap by tracking the location of a business cycle. In consideration of data availability, the EAI is constructed for 12 Asian economies: The People's Republic of China (PRC); Hong Kong, China; India; Indonesia; Kazakhstan; the Republic of Korea (ROK); Malaysia; the Philippines; Singapore; Taipei,China; Thailand; and Viet Nam. Monthly indicators used in the information set represent various categories of activity: private consumption, government

[1] See Cascaldi-Garcia et al. (2023) for a summary.

consumption, aggregate investment, financial investment, trade, and external sector. Given the non-synchronous nature of monthly data releases, the paper forecasts endpoints of necessary series using a standard autoregressive (AR(1)) procedure to avoid presence of “ragged”/“jagged” edges in the set.

The EAI is constructed using principal component analysis (PCA), which combines individual data series into a summary measure of economic activity, tracking GDP growth deviations from its long-term trend. A zero EAI indicates growth at its trend, signaling a business cycle inflection point, while negative and positive values indicate below- or above-trend growth, respectively, reflecting slowdowns or expansions.

To enhance accuracy, the EAI is reconstructed using machine learning (ML) algorithms trained on the same indicators against real GDP growth gaps. ML-based EAI outperforms PCA in accuracy, effectively capturing historical business cycle movements and key events such as the 2008 financial crisis and the 2020 coronavirus disease (COVID-19) pandemic. However, results vary across economies due to differences in data availability.

This paper contributes to the extensive literature on continuous monitoring of economic activity, which generally relies on three main approaches. The first is nowcasting through DFM,[2] built on Stock and Watson’s (1989) work. Various researchers have addressed the challenges faced in applying DFM, such as differing data frequencies and “ragged/jagged" end issues (Mariano and Murasawa 2003; Camacho and Perez-Quiros 2010; and Aruoba, Diebold, and Scotti 2009; among others). The second approach employs PCA to reduce dimensionality, such as in the Chicago Fed National Activity Index (CFNAI) by Evans, Liu, and Pham-Kanter (2002). The third approach utilizes neural networks to predict real-time economic activity using high-frequency indicators, inspired by the Organisation for Economic Co-operation and Development’s (OECD’s) Real-Time Activity Tracker (Woloszko 2020). This method is shown to be effective in modeling economic activity for OECD and G20 economies, particularly during uncertain periods like the COVID-19 pandemic.

Nowcasting has also been used to track business cycles, such as in the Aruoba-Diebold-Scotti (ADS) Index of the Federal Reserve Bank of Philadelphia. This is constructed by employing a state-space model estimated using Bayesian techniques to summarize information from multiple high-frequency series, including indicators on labor, industrial, housing, and financial. By standardizing growth rates around zero, the ADS score indicates average growth at zero, above-average growth with positive values, and below-average growth with negative values.

This paper proceeds as follows: Section 2 defines the EAI target, Section 3 outlines the EAI construction using PCA, Section 4 applies ML algorithms, Section 5 evaluates both models, and Section 6 concludes.

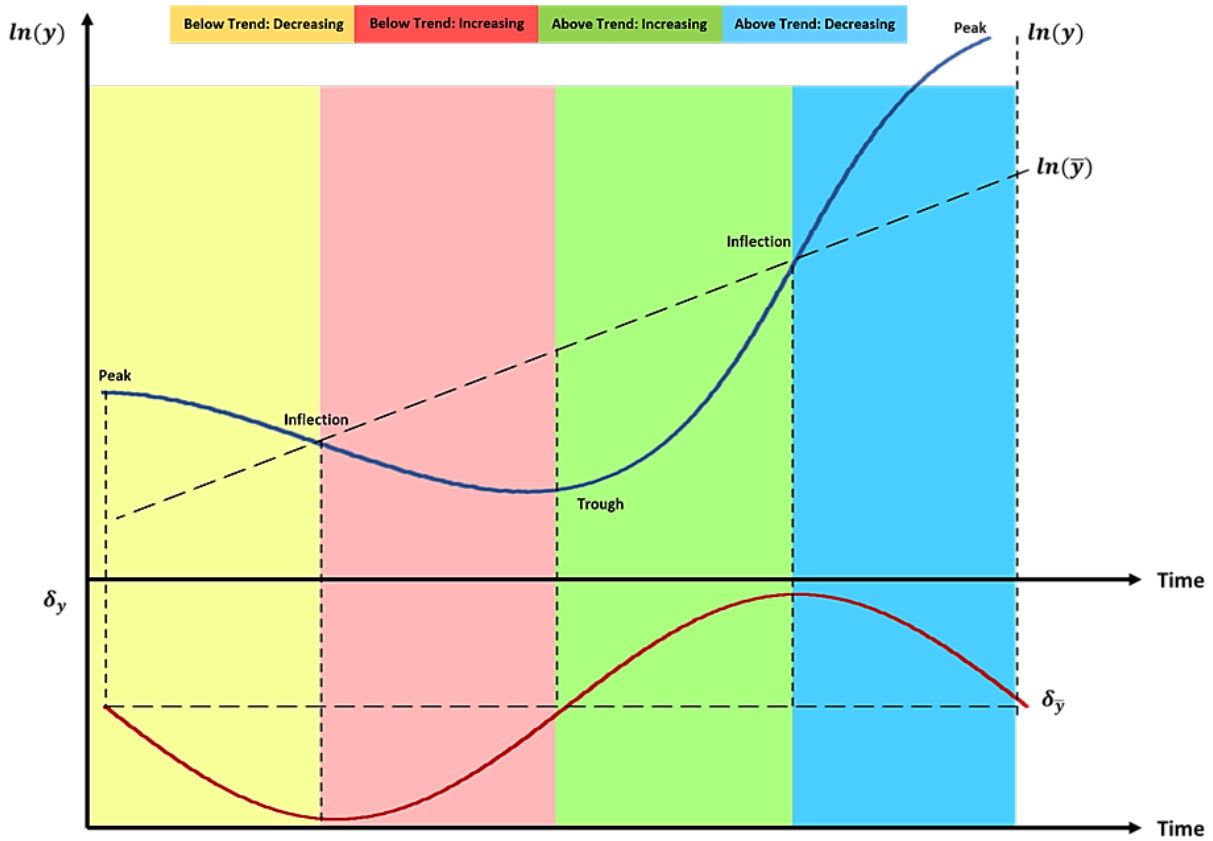
[2] See, for example, Giannone, Reichlin, and Small (2008), Leung et al. (2010), Giannone et al. (2013), Bragoli (2017), Bok et al. (2018), Bragoli and Fosten (2018), and Luciani et al. (2018).

2. GDP GROWTH GAP AND BUSINESS CYCLES

Measuring business cycles—represented by the output gap—is challenging because potential output is unobserved, making it difficult to determine how much actual economic activity deviates from its optimal level. However, the neoclassical growth theory suggests that an economy ultimately converges to a steady state at a constant growth rate, reflecting the long-term balance in the production factors. This theoretical framework suggests that, while potential output may be elusive, the long-term average growth rate can represent a practical proxy for steady-state growth. Consequently, estimating the growth gap becomes more manageable.

Figure 1 illustrates the relationship between the GDP growth gaps and the business cycle. The chart is derived by exploiting the fact that changes in the output gap can be approximated using the difference between the actual growth rate and its long-term average.

Figure 1: The Business Cycle and Growth Gaps



Note: Blue line = actual GDP (in level), representing business cycle; diagonal line = GDP trend (potential output); red line = growth cycle; horizonal line = long-term growth (potential growth rate). The distance between the blue line and its trend is interpreted as the output gap, and the distance between the red line and the long-term average is the growth gap.
Source: Authors' estimates.

The connection between the growth gaps and the business cycle is then defined to be represented by the following four states:

1. When growth is *decreasing and below its average* (the yellow section): the output gap is closing, with GDP (output) declining from its local peak toward its potential, suggesting the economy is heading toward a recession.
2. When growth is *increasing but below its average* (the red section): the output gap is negative and widening as GDP declines further away from its potential during recessions.
3. When growth is *increasing and above its average* (the green section): the output gap is closing from below as the economy is recovering from a recession after hitting its trough.
4. When growth is *decreasing but above its average* (the blue section): the output gap is positive and widening as GDP increases and starts to indicate an overheating.

3. CONSTRUCTING THE ECONOMIC ACTIVITY INDEX: THE PRINCIPAL COMPONENT ANALYSIS APPROACH

To develop a summary statistic that closely tracks periods of economic expansion and contraction by creating a "real-time" statistical measure of coincident economic activity derived from a wide range of high-frequency indicators, the paper begins by following the methodology used by the Chicago Federal Reserve Bank in developing its National Activity Index (Evans, Liu, and Pham-Kanter 2002) to construct the EAI for selected Asian economies. The Chicago Fed National Activity Index (CFNAI) is a weighted average of 85 monthly indicators of economic activity in the US, hand-picked from four categories of data: (i) production and income; (ii) employment, unemployment, and hours; (iii) personal consumption and housing; and (iv) sales, orders, and inventories (Federal Reserve Bank of Chicago, n.d.).

3.1. Data and Transformation

The EAI is constructed for the 12 aforementioned Asian economies with quarterly GDP data. Similarly, in spirit to the CFNAI, a panel dataset of monthly indicators is compiled for each economy to generate the EAI. Unlike the CFNAI, which includes real sector series only, this exercise compiles indicators from all sectors of the economy, partly to compensate for the fewer series available for developing the index as compared with the United States (US). Indicators are selected starting from the year 2000, where available.

The monthly indicators represent various aspects of economic activity, categorized into the following six sectors: (i) private consumption, all personal spending on goods and services; (ii) government consumption, government purchases of goods and services; (iii) aggregate investment, total assets or items acquired to gain appreciation; (iv) financial investment, investments specifically in money; (v) trade, imports and exports; (vi) exogenous international, various international indicators that can affect each Asian economy. The last group of indicators, exogenous international, is included to capture the contribution of external factors affecting the economies. These include the Purchasing Managers' Index (PMI) for developed economies and emerging markets, the US Federal Funds Rate, and commodity prices such as for energy and

agricultural products. Specific indicators selected for each economy are detailed in Tables A1 to A12 in the Appendix.

Once the indicators are collected, pre-processing steps include linking series with multiple base years and removing seasonality when deseasonalized series are not available from the source. A feature of working with a large set of monthly data for economies is the presence of the "ragged"/"jagged" edge problem, caused by non-synchronicity in data releases. This is especially true for Southeast Asian economies such as the Philippines and Viet Nam. To balance the dataset, an autoregressive process (AR(1)) is applied to forecast the indicators to the latest month, as performed by Bok et al. (2018). Once the official update for the indicator is released, the index is re-estimated with the new figures. The series are then checked for stationarity using an augmented Dickey-Fuller test for unit-root process. If a series is non-stationary, it is transformed into growth rates by taking either the year-on-year first difference of the log or the year-on-year percentage change if there are negative values. Indicators that are already in percentage form or are derived from surveys (indexes) are included as levels.

Finally, like the CFNAI, all indicators are rescaled to have mean zero and a standard deviation of one, as follows:

$$x_{ij} = \frac{(z_{ij} - m_z)}{SD_z} \quad (1)$$

where,
x_{ij} = rescaled i^{th} observation of the indicator j
z_{ij} = i^{th} observation of the indicator j
m_z = mean of the indicator
SD_z = standard deviation of the indicator

3.2. Methodology: Principal Component Analysis to Extract Weights

PCA is applied to identify the common factor to overcome the "curse of dimensionality" while maintaining as much statistical variability as possible (Jolliffe and Cadima 2016). This PCA works well when used with large datasets as it can take advantage of their dimensionality. In our case, the first principal component generated by the PCA is used to create the EAI. To demonstrate the methodology, suppose that the principal component index y is based on three indicators, expressed as:

$$y_i = w_1 x_{i1} + w_2 x_{i2} + w_3 x_{i3} \quad (2)$$

where,
x_{ij} = ith observation of indicator j, *j* = 1, 2, 3
w_j = weights generated by the PCA that are assigned to the individual indicators, *j* = 1, 2, 3

To generate these weights, the first principal components are transformed by taking the absolute value and dividing by the sum of the components. The result is a rescaled series of components that sum to 1:

$$w_j = \frac{|p_j|}{\Sigma|p_j|} \qquad (3)$$
$$\Sigma w_j = 1$$

where,
p_j = principal component of the j[th] indicator
w_j = weights from rescaled first principal component of the j[th] indicator

The weights measure the relative importance of the individual indicators in the index (Evans, Liu, and Pham-Kanter 2002). The weights are fixed for each indicator and are not time varying. Using these weights, equation (2) is generalized to create the EAI:

$$EAI_i = \Sigma w_j x_{ij} \qquad (4)$$

with EAI_i defined as the i[th] observation of the EAI.

3.3. Projecting EAI onto Business Cycles

One of the main applications of the estimated EAI is to track periods of economic expansion and slowdown as measured by gap cycles. This is done by projecting the EAI onto the GDP growth gaps, generated by demeaning GDP growth rates. The projection is carried out through a modified version of equation (1), considering the mean and standard deviation of the desired gap series.

$$PEAI_i = m_T + (EAI_i - m_A) * \frac{SD_T}{SD_A} \qquad (5)$$

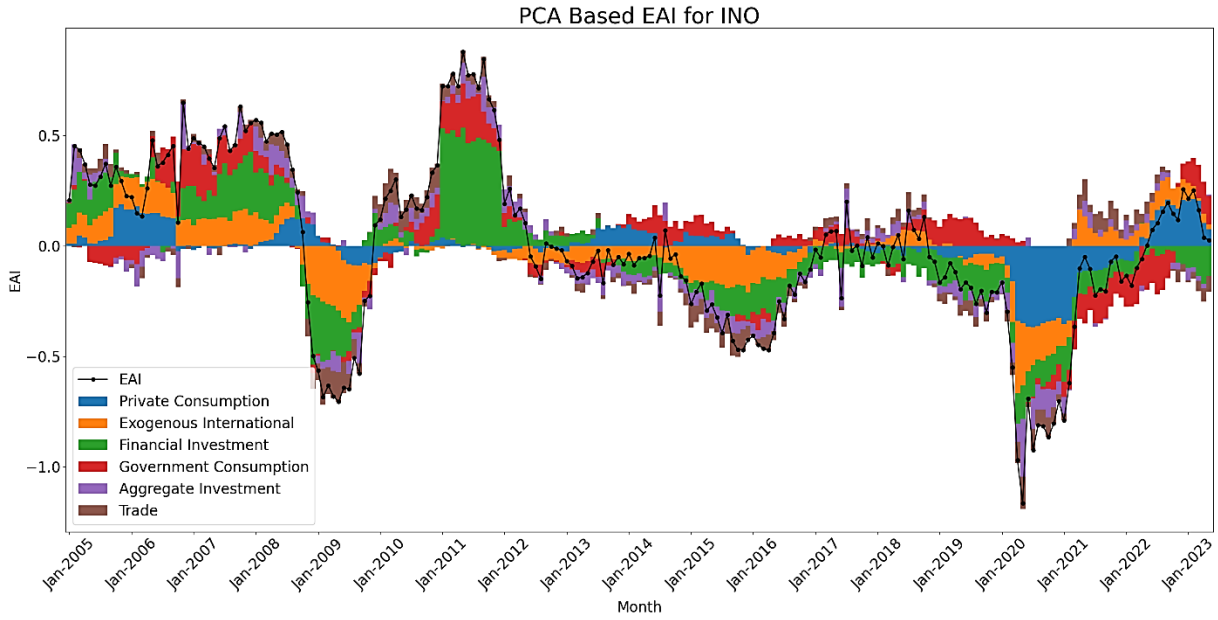
where,
$PEAI_i$ = projected i[th] observation of the Economic Activity Index
m_T = target mean
SD_T = target standard deviation
m_A = actual mean of the index
SD_A = actual standard deviation of the index

3.4. Features of the PCA-Based EAI and Its Components: Results and Interpretation

The PCA is applied to monthly data starting from the year 2000 until 2019 for each economy to produce a summary indicator (EAI) that incorporates all relevant information from a broad range of economic activity.

For each economy, the EAI plot closely resembles the growth gap fluctuations. The EAI provides an objective and timely measure of economic activity drawn from different categories of activity on a monthly frequency. While the EAI captures various relevant information in a single series that reflects characteristics of the aggregate business cycle of an economy, its construction methodology allows for the representation of components that constitute the index. Therefore, it allows for the identification of factors driving the movements of the EAI.

Figure 2: Indonesia's Economic Activity Index



EAI = Economic Activity Index, PCA = principal component analysis, INO = Indonesia.
Source: Authors' estimates.

The EAI is interpreted simply as a summary statistic that tracks the aggregate growth cycle. Figure 2 showcases results for Indonesia, where the EAI mimics cyclical growth movements, including expansions, slowdowns, and turning points. Taken by itself, the interpretation of the index is as follows: by construction, the EAI has an average value of zero and a standard deviation of one. Since many of the indicators are transformed into growth rates, the EAI reflects the deviation of the economic growth rate from its average. Accordingly, an EAI value of zero can be associated with an economy growing at its steady state, a positive value shows an above-average rate of growth, and a negative value shows below average.

Figures A1 to A12 in the Appendix[3] show the computed EAI and its respective components in other Asian economies. The components, represented by colored bars inside the EAI line, are based on indicators drawn from the following categories of economic activity: (i) private consumption; (ii) government consumption; (iii) aggregate investment; (iv) financial investment; (v) trade; and (vi) exogenous international. Weights are computed for individual series, fixed over time, during the PCA and are aggregated according to their assigned economic category. Grouping indicators into categories provides an intuitive interpretation of which sector of the

[3] The Appendix is available at http://dx.doi.org/10.22617/WPS250019-2.

economy is driving the movements in the EAI. The larger the area, the higher the weight of that group during the period. We can consider the case of Hong Kong, China; the ROK; and Thailand, where the weights of the components are clearly visible. It is easy to notice the changing magnitude of the effects of each component across time. This is, however, not true for all the economies; such an example is the Philippines, where it is difficult to present a consistent story. Categories that contributed positively and ones that dragged down EAI every month, as well as the sectors that improved or deteriorated over the previous month, can be seen in most economies though.

The figures also illustrate the volatility of higher-frequency data. Given that the initial estimates of the EAI did not discriminate against the selection of any of the indicators, the results ended with very noisy charts. That is, there may be indicators included that merely contribute noise to the index. Economies like Hong Kong, China; Indonesia; the ROK; Thailand; Singapore; and Taipei,China, however, look cleaner than others. It could be the case that, for these economies, the grouped indicators are reflective of the nature of the economy. It is also interesting to consider the general effect of the exogenous international variables on the economies as this is the only group of indicators that is common for all. This group again contains price indices of agricultural products and oil products, the US Federal Funds Rate, and the PMI.

3.5. Selecting the Most Relevant Indicators

The initial version of the EAI uses all available monthly indicators for each economy. This section discusses techniques for indicator selection that could improve the accuracy of the index for tracking the business cycle. While this indicator selection technique is applicable to all economies, an element of judgment will be applied in the selection process in each case, as elaborated below.

In the first stage, monthly data that feed into the construction of the index are compiled from various sources for available years and transformed. Further treatment is applied to the data as needed. Series with different base years are spliced to have a common base year. The latest base year is always used as the rule of thumb. Series with obvious seasonal patterns are adjusted to remove seasonality using the X-13 autoregressive integrated moving average (ARIMA) package in EViews. All series in levels are converted into log values. Variables already in percentages or rates of change are retained for the estimation. An autoregressive of order one—AR(1)—procedure is applied to complete the series in the dataset to the latest observation month to deal with ragged ends. Finally, all indicators are rescaled to have a mean zero and a standard deviation of one.

To identify the relevant indicators for inclusion in the EAI using the PCA method outlined earlier, three selection methods are used. The first is a correlation method, whereby the correlation of each indicator with the GDP growth rate is computed using Pearson's correlation. Indicators with less than 50% correlation are removed. The other two methods are model averaging techniques—Bayesian model averaging (BMA) and weighted-average least squares (WALS).

Model averaging is used as an automatic variable selection technique to arrive at a set of indicators that serve as a benchmark. These averaging techniques compute weights given the results of the estimates. The focus indicators are selected based on the correlation of each indicator against GDP growth rate using Pearson's correlation. Only those indicators with a significance level of 0.5% and above are included as focus indicators. The remaining indicators are used as auxiliary indicators. The focus regressors are always included, while auxiliary regressors are less certain about the inclusion. To determine which of the auxiliary indicators to use, we rely on the t-statistic computed via both BMA and WALS estimations. Auxiliary indicators with a t-statistic greater than one are considered to have a significant effect on the estimation and are included in the list of selected indicators.

Among the automatic variable selection processes to construct the EAI, the method that best tracks GDP gap cycles is identified. To improve the fit, an ad hoc examination—using economy-specific knowledge—of the existing set of indicators is conducted.

For the PRC, the monthly data for estimating the EAI started in 2000. Data starting after 2000 are excluded. These include the Macroeconomic Climate Index, which starts in 2015; the urban unemployment rate, which starts in 2017; the New Economy Index, which starts in 2017; the Caixin China Composite PMI, the Caixin China General Manufacturing PMI, and the Caixin China General Services PMI, which start in 2018; and the Consumer Confidence Index for income, which starts in 2016. For the PRC, BMA produced the EAI with the best fit to track growth cycles. An ad hoc examination—using economy-specific knowledge—of the existing set of indicators was conducted. The following indicators were added: M2 money supply, which represents the amount of liquidity in circulation; the household saving deposits rate, which reflects the cost of money; and the profits of industrial enterprises, which reflect business conditions. These frequently watched indicators by Bloomberg were added to improve the EAI's accuracy in tracking the target gap cycles. The international commodity price of soybeans, preselected by the WALS method, was added as well. The following indicators were added to improve the EAI's performance, following Giannone, Agrippino, and Modugno (2013): total passengers carried, which reflects the dynamics of tourism; residential floor space sold, which reflects the dynamics of the real estate sector; loans, which reflects credit expansion; and the suppliers' delivery time PMI and new export orders PMI, which reflect business trends in the manufacturing sector.

For Hong Kong, China, the monthly data for estimating the EAI started in 2000. Data starting after 2000 are excluded. These include data on market capitalization, which start in 2004; the Residential Mortgage Survey, which measures the number of loan applications, which starts in 2005; and the survey of usable floor area of completed new buildings, which starts in 2003. The WALS method produced the EAI with the best fit to track growth cycles for Hong Kong, China. To improve the fit, an ad hoc examination—using economy-specific knowledge—of the existing set of indicators was conducted. We removed indicators that did not closely co-move with the growth cycle, such as the unemployment rate. This improved the EAI's performance in tracking the growth cycle. We added additional international indicators, such as the US Industrial Production Index, to capture the dependence of Hong Kong, China on the external sector, and the hotel

occupancy rate, to reflect the large role of tourism in the economy. This improved the EAI's performance. Both variables were preselected using the BMA method.

For Indonesia, the monthly data for estimating the EAI started in 2006. Data starting after 2006 are excluded. These include Sharia Monetary Operation, which starts in 2008; price expectations in the next 12 months, household income allocation for consumption, household income allocation for loan repayments, and household income allocation for savings, which start in 2012; and outstanding loans of micro, small, and medium-sized enterprises, which start in 2011. For Indonesia, the WALS method produced the EAI with the best fit to track the growth cycles. To improve the fit, an ad hoc examination—using economy-specific knowledge—of the existing set of indicators was conducted. The following indicators were added: domestic consumption of cement and M1 money supply, based on the set of indicators used in Luciani et al. (2018), which further improved the EAI's overall performance; loan approvals of commercial and rural banks, to reflect the role of credit flows; and tourist arrivals, to reflect the impact of tourism. The latter two indicators were preselected using the BMA method.

For the ROK, the monthly data for estimating the EAI start in 2000. Data starting after 2000 are excluded. These include household lending by depository corporations and domestic credit card usage, which start in 2003; central government debt, which starts in 2012; foreign purchases of stocks and bonds, which start in 2010; and house sales prices, which start in 2011. For the ROK, the WALS method produced the EAI with the best fit in tracking the growth cycles. Indicators that did not closely co-move with the growth cycle, such as the unemployment rate, were removed; this improved the EAI's performance in tracking the growth cycle. The US Industrial Production Index and the US Federal Funds Rate were added to capture the external effect on the economy. Both indicators were preselected using the growth correlation method.

For Malaysia, the monthly data for estimating the EAI started in 2005. Data starting after 2005 are excluded. These include the overnight policy rate, which starts in 2008; and the Malaysia Industrial Production for Construction Related Cluster, which starts in 2007. WALS produced the EAI with the best fit in tracking the growth cycles. To improve the fit, M1 money supply and net official reserves, both preselected using the BMA method, were added, as they improved the EAI's performance.

For the Philippines, the monthly data for estimating the EAI started in 2003. Data starting after 2003 are excluded. These include the Chamber of Automotive Manufacturers of the Philippines sales of passenger and commercial vehicles and the PMI, which started in 2006; and the overnight reverse repo rate, the balance of payments in domestic direct investment, and the financial account, which started in 2005. The BMA method produced the EAI with the best fit in tracking the growth cycles. To improve the fit, the following indicators were added: government expenditure, to reflect the government's role in the economy, preselected based on the growth correlation method; total rice stock inventory, to capture the state of household and commercial supplies of rice; and visitor arrivals, to reflect the impact of tourism on the economy. The latter two indicators were preselected by the WALS method.

For Singapore, the monthly data for estimating an EAI start in 2000. Data starting after 2000 are excluded. These include the central government cash surplus/deficit, which starts in 2011; the Singapore Overnight Rate Average, which starts in 2005; the Domestic Liquidity Indicator, which starts in 2002; and outbound departures of Singapore residents, which start in 2011. The WALS method produced the EAI with the best fit in tracking the growth cycles. To improve the fit, the following indicators were added: industrial production for manufacturing chemicals and chemical products, and industrial production for manufacturing petrochemicals, to reflect the role of the domestic petrochemical industry in driving the economy; and tourist arrivals and the hotel occupancy rate, to capture the growing significance of tourism in Singapore. These indicators were preselected using the BMA method, which improved the performance of Singapore's EAI.

For Thailand, the monthly data for estimating an EAI started in 2005. Data starting after 2005 are excluded. These include the Industrial Production Index, the Finished Goods Inventory Index, and data on new job applicants and job vacancies, which start in 2011; and the headline manufacturing PMI, which starts in 2017. The growth correlation method produced the EAI with the best fit in tracking the growth cycles. Government expenditure, monthly tourist arrivals, and broad money were added as they improved the EAI's performance in tracking cycles. These indicators were preselected using the WALS method.

For Viet Nam, the monthly data for estimating the EAI start in 2009. Data starting after 2009 are excluded. These include vehicle sales in units, which start in 2013; the Labor Employed Index for Industrial Enterprises and deposits of economic entities at credit institutions, which both start in 2012; and the headline manufacturing PMI, which starts in 2017. The growth correlation method produced the EAI with the best fit in tracking the growth cycles. To improve the fit, the following indicators were added: rice exports, to reflect the role of the decline in agriculture, which accounted for the growth downturns in the second quarter of 2012, the first quarter of 2013, the fourth quarter of 2014, the first and fourth quarters of 2016, and the second quarter of 2017; foreign direct investment, to reflect its increasing role in the economy; retail sales for accommodation, food, and beverage services, and international visitors, to account for the growing importance of services and tourism in Viet Nam's economy; and industrial production for textiles from polyester, to reflect the important role of the garment and textiles industry. All these additional indicators were preselected using the WALS method.

4. PIVOTING TO A MACHINE LEARNING-BASED METHODOLOGY

The PCA-based EAI effectively reflects business cycle movements in selected Asian economies, offering a straightforward way to interpret trends. However, as an aggregation of real indicators rather than precise predictions, it falls short in accurately estimating GDP growth gaps. A more robust methodology could enhance predictive accuracy and better capture business cycle dynamics. Machine learning (ML), a subfield of artificial intelligence, allows systems to "learn and improve from experience" (Sarker 2021). By training ML models on economic indicators used in PCA, data from each period can generate a single EAI value that more accurately represents the GDP growth gap for an economy.

4.1. Comparison of Traditional Methods and Machine Learning

Traditional nowcasting methods, such as PCA, have limitations in predictive power due to strict assumptions about linear relationships and finite learning capacities, making them less effective in modeling nonlinear dynamics or adapting to changing economic conditions.

In contrast, machine learning (ML) excels with large-scale data and advanced computational tools, enabling complex pattern recognition and real-time processing for accurate predictions. The ML workflow includes preprocessing, feature engineering, model training, hyperparameter tuning, and evaluation. Unlike traditional methods, ML requires minimal assumptions about functional forms, automatically identifying relationships and improving accuracy. By incorporating exogenous variables alongside target data, ML enhances prediction accuracy with broader datasets (Makridakis, Spiliotis, and Assimakopoulos 2013).

4.2. Machine Learning in Economics

Significant research has been devoted to exploring the potential of ML in economic forecasting. Hinds et al. (2017) demonstrate the use of Elastic Net, an extension of linear regression that introduces a penalty term to address overfitting, in nowcasting the UK's quarterly GDP growth rate. The model utilizes 28 economic indicators such as the services index, the London Interbank Offered Rate, and tax receipts. It can pick up changes in actual GDP growth direction 94% of the time.

Yoon (2020) and Padmawar, Pawar, and Karande (2021) applied models such as Gradient Boosting Machines and Random Forest to nowcast the GDP growth of Japan and India, respectively. They find that ML models significantly outperform benchmarks in prediction accuracy, supporting the application of ML algorithms in macroeconomic forecasting.

In summary, ML offers more accurate predictions by providing flexibility in modeling, advanced pattern recognition, and efficiency in handling high-dimensional datasets. Employing the techniques discussed in the existing literature will allow timely and accurate predictions of the GDP growth gap.

4.3. Machine Learning-Based Economic Activity Index

4.3.1. Machine learning process

Regression models predict continuous values, while classification models assign class labels. In machine learning, datasets are typically divided into training and testing sets, often in an 80-20 ratio. The training set is used to develop the model, and the testing set assesses its accuracy using metrics such as R-squared for regression or other error measures.

4.3.2. Machine learning algorithms

ML models range from simple linear models to more complex boosting methods, each with unique strengths and applications. ML models can be categorized into the following:

1. Linear models model the relationship between input features and the target variable by fitting a linear equation. They are simple, interpretable, and efficient in training and prediction, but suitable only for linearly separable data. These include Linear Regression, Stochastic Gradient Descent, Kernel Ridge, Elastic Net, and Bayesian Ridge.

2. Boosting models combine multiple weak models sequentially to create a stronger overall model by focusing on correcting the errors of the previous one. They are effective in capturing complex patterns. These include Random Forest, Gradient Boosting, XGBoost, and Light Gradient Boosting Machine.

3. Support Vector Machines find the optimal hyperplane to separate data points of different classes by maximizing the margin between them. This is effective for high-dimensional data but computationally expensive.

The 10 machine learning models employed for each economy are listed below:

1. Linear Regression computes the linear relationship between a dependent variable and one or more independent features. The goal is to find the best linear equation that can predict the value of the dependent variable based on the independent variables. The mathematical equation is found below:

$$y_t = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon$$

   where,
   y is the target variable
   $x_1 \ldots x_n$ are the input features
   β_0 is the intercept term
   $\beta_1 \ldots \beta_n$ are the coefficients corresponding to each input feature
   ε is the error term

2. Stochastic Gradient Descent. Gradient descent is an iterative algorithm that starts from each point of a loss function and travels down until it reaches its lowest point. Stochastic Gradient Descent introduces randomness by picking one data point at each iteration, making the method more efficient than standard gradient descent. An update is performed iteratively for each training example in the dataset to minimize the loss function. It is typically repeated until the model converges to an optimal set of parameters. The Stochastic Gradient Descent update step for linear regression is:

$$\theta := \theta - \alpha(\hat{y_i} - y_i)x_i$$

where,
θ is the vector of model parameters
α is the learning rate

3. Kernel Ridge uses a kernel function to calculate weights. This model is used when there are too many data for a traditional linear model. The Kernel Ridge problem combines ridge regression with the kernel trick to enable nonlinear relationships and can be expressed as:

$$\alpha = (K + \lambda I)^{-1}\mathrm{y}$$

where,
α is the n×1 vector of dual coefficients
K is the n x n kernel matrix
λ is the regularization parameter

4. Elastic Net combines the features of lasso and ridge regression to deal with problems of multicollinearity and overfitting for high-dimensional datasets.

$$\min_{\beta}\left(\frac{1}{2N}\sum_{i=1}^{N}(y_i - \mathbf{x}_i^{\top}\beta)^2 + \lambda_1\|\beta\|_1 + \lambda_2\|\beta\|_2^2\right)$$

where,
y_i is the observed response
$\mathbf{x}_i$ is the predictor vector
β is the coefficient vector
$\|\beta\|_1$ is the L1 norm (sum of the absolute values of the coefficients)
$\|\beta\|_2^2$ is the L2 norm squared (sum of the squared coefficients)
λ_1 and λ_2 are regularization parameters that control the contribution of the L1 and L2 penalties, respectively

5. Bayesian Ridge is a linear regression model that uses probability distributions instead of point estimates to create the linear equation. The output is generated from a normal distribution (Gaussian) characterized by a mean and variance:

$$\mathbf{y} \mid \mathbf{X}, \beta, \sigma^2 \sim \mathcal{N}(\mathbf{X}\beta, \sigma^2\mathbf{I})$$

where,
y is the vector of observed responses
X is the matrix of predictors
β is the vector of coefficients

6. Random Forest is an ensemble learning technique in ML that combines multiple decision trees, each trained on a separate subset of data, to control overfitting and improve model

accuracy. Every tree in the forest is generated using a subset of the training data and a randomly chosen set of attributes. The final forecast is created by averaging or majority voting the different trees' outputs during prediction, improving the model overall stability and robustness. Mathematically, the final prediction can be calculated as:

$$\hat{y} = \frac{1}{n} \Sigma_{i=1}^{n} y_i$$

where,
n is the number of decision trees in the Random Forest
y_i is the prediction of the i^{th} decision tree

7. Gradient Boosting is an ensemble learning technique that sequentially combines multiple modest models, often represented by decision trees. In each iteration, the algorithm enhances the model by creating new trees to address the errors from the previous prediction rounds. This involves updating the target value using gradient descent. The final forecast is arrived at by summing the predictions from all trees. The objective function is given by:

$$\mathcal{L} = \sum_{i=1}^{n} L(y_i, \hat{y}_i) + \Omega(f)$$

where,
$L(y_i, \hat{y}_i)$ is the loss function for the i^{th} instance
$\Omega(f)$ is a regularization term to penalize model complexity

Gradient Boosting builds an additive model of the form:

$$\hat{y}^{(t)}(\mathbf{x}) = \hat{y}^{(t-1)}(\mathbf{x}) + \nu f_t(\mathbf{x})$$

where,
$\hat{y}^{(t)}(\mathbf{x})$ is the prediction at iteration t
$f_t(\mathbf{x})$ is the new base learner added at iteration t
ν is the learning rate

The new base learner is trained to fit the negative gradient:

$$f_t = \arg\min_{f} \sum_{i=1}^{n} (g_i^{(t)} - f(\mathbf{x}_i))^2$$

After training the new base learner, the model is updated as follows:

$$\hat{y}^{(t)}(\mathbf{x}) = \hat{y}^{(t-1)}(\mathbf{x}) + \nu f_t(\mathbf{x})$$

This process is repeated iteratively until the final prediction results in the sum of the initial prediction and all the updates from each iteration:

$$\hat{y}(\mathbf{x}) = \hat{y}^{(0)} + \nu \sum_{t=1}^{T} f_t(\mathbf{x})$$

8. XGBoost applies the gradient boosting algorithm with advanced regularization aimed at reducing overfitting. This model generally performs better and is more efficient than other complex models. Similar to gradient boosting, the objective function of XGBoost, which combines a loss function with a regularization term, is given by:

$$\sum_{i=1}^{N} L(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k)$$

where the first term represents the loss function while the second term represents the regularization term, which penalizes the complexity of the trees in the ensemble. This term prevents overfitting and typically includes:

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2$$

where,
γ is the regularization parameter for the number of leaves T (complexity of the tree)
λ is the regularization parameter for the L2 norm of the weights w_j (tree leaf values)

In XGBoost, a new tree is added to the model in each iteration to improve performance. To find the best split for a decision tree, XGBoost optimizes the following objective:

$$\text{Gain} = \frac{1}{2}\left[\frac{(G_L^2)}{H_L + \lambda} + \frac{(G_R^2)}{H_R + \lambda} - \frac{(G^2)}{H + \lambda}\right] - \gamma$$

where,
G is the sum of gradients (first-order derivatives) for the samples in the node
H is the sum of Hessians (second-order derivatives) for the samples in the node
G_L and H_L are the gradients and Hessians for the left split
G_R and H_R are the gradients and Hessians for the right split
γ is the regularization parameter for the number of leaves

9. Light Gradient Boosting Machine is another gradient boosting-based method that uses decision trees split on the leaf level instead of the tree level to increase efficiency. Its algorithm is largely similar to XGBoost but it uses histogram-based learning to convert continuous features into discrete bins, which speeds up computation.

10. The Support Vector Machines algorithm aims to find the hyperplane that passes through as many data points as possible within a certain distance to minimize prediction error while computing predictions. It is effective in high-dimensional spaces and nonlinear data but can be computationally intensive, especially with large datasets. The objective function is given by:

$$\min_{\mathbf{w},b}\left(\frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{n}\xi_i\right)$$

where,
$\mathbf{w}$ is the weight vector perpendicular to the hyperplane
b is the bias term
ξ_i are the slack variables that allow some misclassification
C is the regularization parameter that controls the tradeoff between maximizing the margin and minimizing classification errors

4.3.3. Time series machine learning

While traditional ML-based regression models are trained on a portion of data taken from a randomized set of static, independent features and their corresponding target variables, time series data are more dynamic because each row or set of variables is indexed based on a date/time attribute. Therefore, when applying ML models to time series data, we would like to train the models on past data points to predict future data (Hyndman and Athanasopoulos 2018).

Generally, regression models forecast or predict a target/forecast value y based on its relationship with one or more time series, predictor variables x. In its simplest form, this relationship can be modeled linearly, such as in basic multiple linear regression:

$$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \ldots + \beta_k x_{k,t} + \varepsilon_t$$

where y is the target variable and $x_1, \ldots, x_k$ are the predictor variables. The coefficients $\beta_1, \ldots, \beta_k$, meanwhile, represent the degree of importance of each independent variable, measuring the marginal effect of each predictor. Taking the above equation into account, one needs only to plug in the values of $x_{1,t}, \ldots, x_{k,t}$ for $t = 1, \ldots, T$ and set the error term ε_t to zero to produce a future prediction at time t: $\widehat{y_t}$.

$$\hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1 x_{1,t} + \hat{\beta}_2 x_{2,t} + \ldots + \hat{\beta}_k x_{k,t}$$

The multiple linear regression model is the simplest algorithm used in predictive ML. These models are usually manufactured as a litmus test against higher level models given their intuitive nature and lightness in terms of computational requirements. The simple principles for linear regression are used for other ML models covered in the following sections as well, with varying degrees of complexity.

The predictive capabilities of the models presented in this paper should not be confused with forecasting models, however. There are two types of forecasts associated with and that can be produced by time series ML models, depending on data availability at the time of prediction: ex-ante and ex-post forecasts. Ex-ante forecasts are generated ahead of the last available data point, using information that is available in advance. This can be done by using other forecasting methods such as exponential smoothing or ARIMA models to extend the reach of each predictor before using the base model to generate predictions.

On the other hand, ex-post forecasts are produced with the use of later information. The machine learning model in this case is only generating predictions using observed predictors. A data set containing features up until Q4 of this year can be used to build a model trained up until Q3, to predict a target value for Q4. The models detailed in the following sections of this paper were built to produce hybrid forecasts using elements from both ex-ante and ex-post approaches. We consider this a hybrid methodology because the models utilize at least one known predictor for their latest predictions, filling in any missing features for that period using ARIMA forecasts.

4.3.4. Time series process

The data transformation steps taken to prepare the monthly indicators and actual GDP figures of each economy remain mostly constant between the PCA-based and the ML-based methodologies. One addition to the ML method has to do with preparing the target data variable for model training and evaluation. Once the quarterly GDP levels are converted first into year-on-year growth and then into a quarterly GDP growth gap, linear interpolation is carried out to fill in the GDP growth gap per month. Moreover, the addition of monthly training points allows the models access to more training data, as well as a better idea of the trend movements of real values. Because the connecting monthly values between real quarterly values are merely interpolated, however, the results that the model provides for those months cannot be evaluated as actual GDP gap predictions. Therefore, final model performance evaluations are conducted only for the points of time for which real, quarterly GDP gap values are available.

Given the nature of time series data, in which a chronological time-based index is assigned to each set of predictors and target variable, one cannot simply separate an arbitrarily sized set, randomly taken from the full data, to train the model and then test it on the remaining data. Instead, the training and testing process is organized to simulate how the model will act in the real world.

As an initial training set, the first 50% of available data is removed from the overall dataset. This collection of predictors and targets is used to train the ML model. Once the model is trained and ready for computation, it is used to predict the next target variable chronologically before repeating the process for the remaining data by continuously adding the predicted values to the training set.

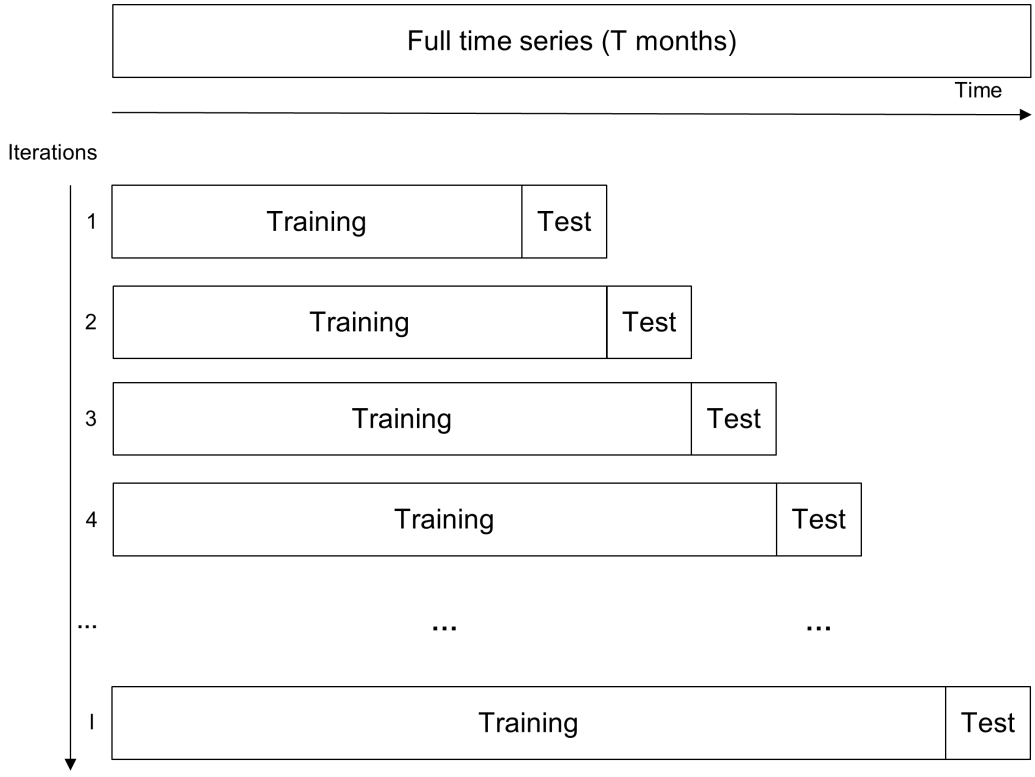
The model in this case is trained to predict the GDP growth gap using monthly economic indicators. If monthly data are available from January 2000 to January 2024, the model is first trained on data from January of 2000 to 2012. The following steps break down the modeling process employed:

1. That model is used to predict the GDP growth gap in February of 2012 using the monthly indicators of February 2012.
2. The predicted value for February 2012 is saved for evaluation.
3. The actual value for February 2012, along with the independent variables used to make the previous prediction, is added to the end of the training set. This means that the training set has grown to include data from January 2000 to February 2012.
4. The previous steps are repeated for the next month, March 2012 and so on, until the final data row in January 2024 is reached.

Once the latter 50% of data has its own predictions, these can be evaluated against actual values to determine the overall performance of the model.

The process used to build the ML models in this paper can simulate how the index is used monthly as each prediction is being made with only past information as a basis. Further, as time goes on, the model is updated to include all past data, making it more sensitive to predictor fluctuations in the present. In essence, the training dataset increases in size to encompass all available information as it becomes available to predict for the future, as Figure 3 shows.

Figure 3: Iterative Time Series Machine Learning Model-Building Process



Source: Authors.

4.3.5. Economy model selection process

To build the best possible EAI-producing models, the time series process for ML detailed in the previous section is conducted using 10 different regression model algorithms for each of the selected economies. Once predictions are calculated and collated for each model, they can be compared against actual GDP growth gap values using metrics such as R-squared—a measure used to determine the goodness of fit of a model by explaining the variation in the dependent variables, the root mean squared error (RMSE)—the average difference between predicted and actual values, and the trend match rate—a metric used specifically in this study that measures the percentage of instances where the business cycle trend of the predictions matches that of actual figures. The results for each of the metrics are compared for all models and economy, with the best performing chosen for the EAI computation.

4.3.6. Model selection results

After running all ML model algorithms for each economy's data, these models performed the best based on all performance metrics (R-squared, RMSE, trend match) used for evaluation.

- ***PRC; Hong Kong, China; Indonesia; Singapore:*** gradient boosting regression
- ***India:*** stochastic gradient descent regression
- ***Kazakhstan, Republic of Korea, Philippines, Thailand, Viet Nam:*** XGBoost regression
- ***Malaysia:*** Bayesian ridge regression
- ***Taipei,China:*** support vector machine

4.4. Interpreting Model Results (Shapley Values)

Results derived from ML algorithms can, at times, be extremely accurate—to a point where they may seem unbelievably so, most often because of the high degree of complexity of ML and deep learning models. This, however, is a double-edged sword, as the complexity of these models hinders the overall interpretability of the results. Researchers have proposed various methods to help with the interpretation of the models' predictions, such as qualitative evaluation, a human-based approach wherein the tester must experiment with a model's inputs and outputs until they are able to evaluate how closely a model acts to regular human reasoning (Alangari et al. 2023). This method, however, is more easily affected by biases, making it less popular. On the other hand, there has been an abundance of proposed, computational methods of interpreting models, though it is often unclear how these methods can be linked, or which methods are better than others, as Lundberg and Lee (2017) discuss.

Lundberg and Lee (2017) propose a computational method to numerically interpret ML models called SHAP (SHapley Additive exPlanations). SHAP or, more specifically, SHAP values, is calculated using an "explanation model", which is an interpretable explanation of the original ML model—whether it is a simple linear model or a more complex one such as an ensemble or deep model.

$$g(z') = \phi_0 + \sum_{i=1}^{M} \phi_i z_i'$$

where,
$z' \in \{0,1\}^M$, M is the number of simplified input features and $\phi_i \in R$

This explanation model is used in at least six other model interpretation or additive feature attribution methods:

1. **Local Interpretable Model-Agnostic Explanations** (LIME) explains a single prediction at a time by training random data points sampled from a training dataset's distribution and weighted based on the distance from the reference feature point being explained, with the feature importance coefficients of the most important features becoming the basis for interpretation (Visani, Bagli, and Chesani 2022).
2. **Deep Learning Important FeaTures (DeepLIFT)** is a method for decomposing the predictions of neural networks by back propagating the contributions of all neurons to each feature of the input to assign "contribution scores" (Shrikumar, Greenside, and Kundaje 2019).
3. **Layer-Wise Relevance Propagation** explains and scales deep neural networks by propagating the prediction backward using a set of designed propagation rules (Montavon et al. 2019).

The last three interpretation methods that employ the explanation model technique all relate to Shapley values and, thus, use principles and equations from cooperative game theory to compute interpretations. These are classic ***Shapley value estimation***, ***Shapley sampling values***, and ***quantitative input influence***. Cooperative game theory refers to a sort of game or situation in which groups of players or coalitions sign binding contracts determining actions and payoffs (Maskin 2016). When building an ML model, we form a game, where the data features inputted into the model are the players and the output prediction is the payoff. The Shapley value is computed by taking the average marginal contribution of each feature by continuously retraining the model on all feature subsets $S \subseteq F$, where F is the set of all features. To compute the effect of each feature, a model $f_{S\cup\{i\}}$ is trained with a certain feature included; in the meantime, another model f_S is trained without the feature. The predictions for the two models are compared to determine the effect. This process is repeated for each feature and set of data in the testing dataset, allowing for the final computation of Shapley values, the weighted average of all possible differences.

$$\phi_i = \sum_{S\subseteq F\setminus\{i\}} \frac{|S|!\,(|F| - |S| - 1)!}{|F|!} [f_{S\cup\{i\}}(x_{S\cup\{i\}}) - f_S(x_s)]$$

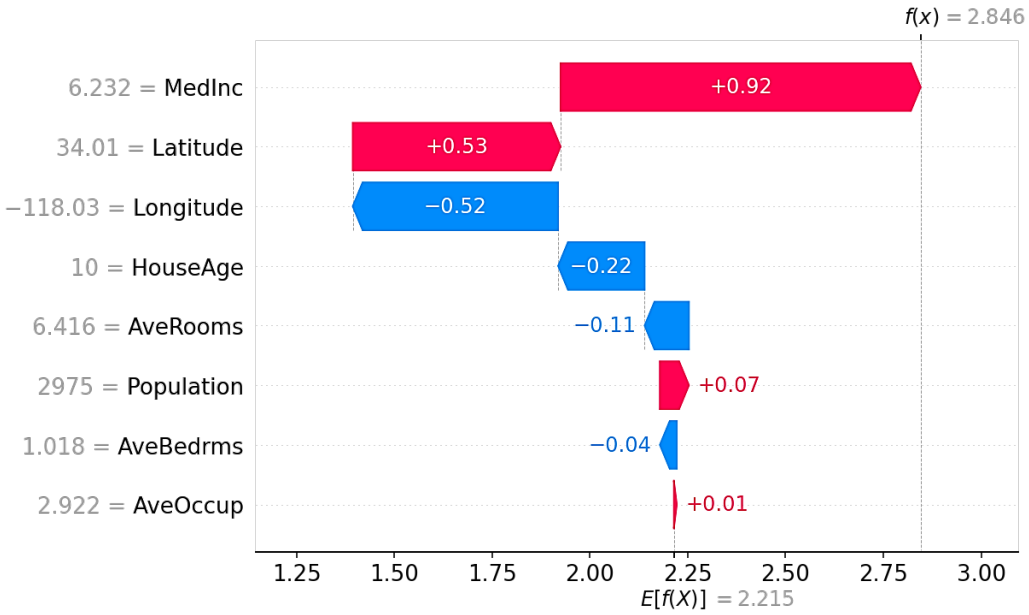
By applying sampling approximations to the preceding equation and then estimating the effect of removing a variable from a model by integrating over samples in the training set, Shapley

sampling values are used to explain a model. On the other hand, quantitative input influence uses a similar methodology while having a broader framework.

SHAP, the unified measure of feature importance proposed by Lundberg and Lee, combines principles and methods from popular additive feature attribution methods such as LIME, DeepLIFT, and layer-wise relevance propagation with the main ideas behind Shapley values. Given the additive nature of Shapley values, each value corresponds to a single feature for one period, and all the values add up to the difference between the final prediction and the average across all outputs. We can see what kind of information SHAP values can give us in the waterfall plot below, taken from the official documentation for SHAP.

Figure 4 depicts a single output using a model to predict housing prices with the features shown on the y-axis. $E[f(x)]$ is the expected prediction or the overall average output and $f(x)$ is the actual prediction for the housing price in a single time period. We can see that median income (MedInc) pushes the prediction up from average the most, followed by latitude, population, and average household members (AveOccup), respectively, while longitude, median age (HouseAge), and average rooms and bedrooms pull the final prediction below the average. It is clear from the information we have that the median income of people in the area, together with the location of the house, contributes the most to the final price.

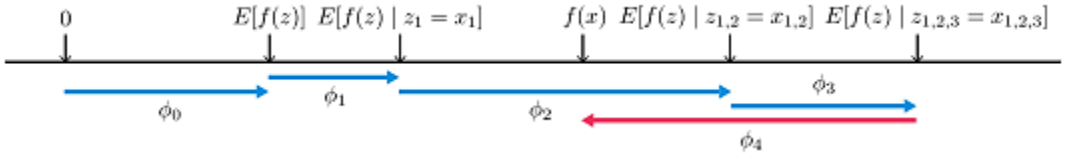
Figure 4: SHAP Values Example—Housing Prices Prediction



SHAP = SHapley Additive exPlanations.
Source: Lundberg and Lee (2017).

With SHAP, we can determine approximate importance measures associated with each feature and for every model prediction in our own models. We are also able to visualize how the data inputs either push up or pull down the results to arrive at a final predicted output.

Figure 5: Interpretation of SHAP Values

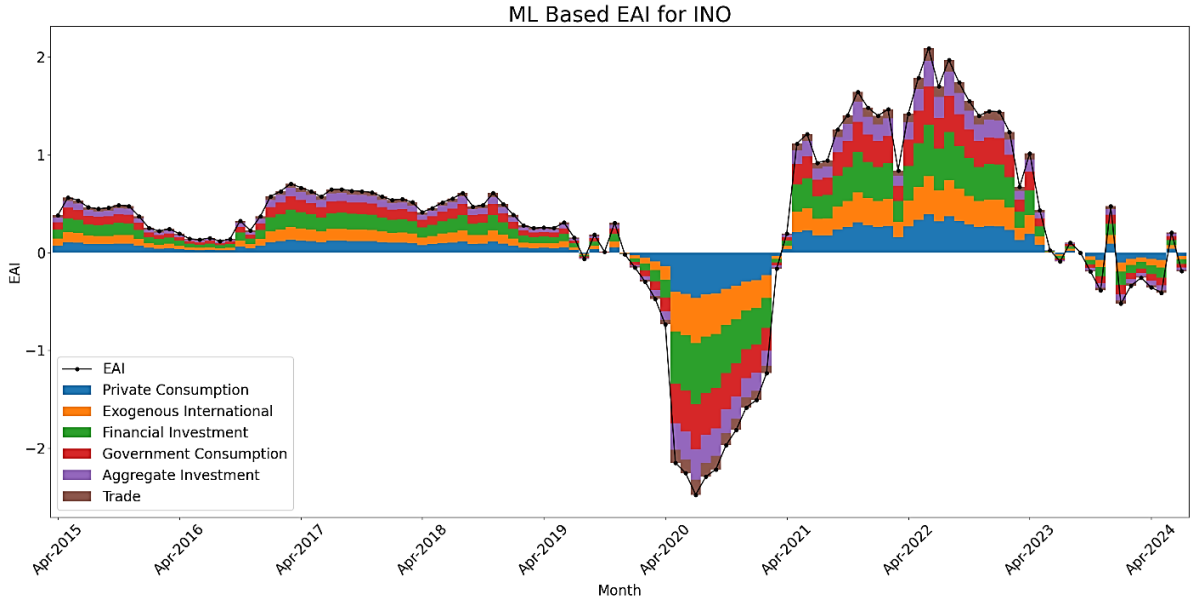


SHAP = SHapley Additive exPlanations.
Source: Lundberg and Lee (2017).

While the initial version of the EAI utilizes weights derived from the first component of a PCA to determine the overall effect of each of the six economic sectors on the index value for per period, the ML-based version employs SHAP values. For each model developed, feature-specific SHAP values were computed for every new GDP growth gap prediction, then summed with the expected growth gap, before the proportion of that value against the final predicted growth gap was computed. This gives us the degree of importance of each feature, which are then aggregated into one of the six specified economic sectors. Using these values, we can approximate the effect of each economic sector on the state of each economy's business cycle at certain times, allowing for more informed policy decisions.

Figure 6 details an example from Indonesia's economy, showing the various categories' importance for each predicted EAI value over the past 5 years (see Figures B1 to B12 in the Appendix for other economies). We can see that the lowest point in the economic downturn predicted by the EAI model in July 2020 resulted mostly from changes in consumption, financial, and the external sectors, with the eventual rise mid-2021 attributed to the same economic sectors, with the inclusion of government expenditures.

Figure 6: Indonesia's Machine Learning Economic Activity Index over Time with Economic Components



EAI = Economic Activity Index, ML = machine learning, INO = Indonesia.
Source: Authors' estimates.

4.5. Machine Learning-Based Methodology: Most Recent Trends

A quick look at comparison charts between the actual GDP growth gap movements of each economy over time versus the predicted EAI for each clearly shows a close relationship. This is especially true for periods in which the growth gap fluctuates close to the zero-trend line. One very similar behavioral trend that can be seen across models is that they tend to underestimate large upturns and downturns. For all economies covered in this paper, while the large economic contraction and following expansion caused by the global COVID-19 pandemic from around early 2020 until 2022 for most economies would have been predicted for, using the data available at that time, the overall severity would not. This may be addressed by using other, less traditional, indicators such as consumer sentiment or mobility, as these may be more telling when it comes to unexpected shocks.

With the most recent predictions being computed on 15 July 2024, real GDP growth gap data for the first quarter of 2024 are available for all economies aside from Kazakhstan and Hong Kong, China, which have the GDP data available only up to the fourth quarter of 2023. With this, EAI predictions have been computed for each economy up to the second quarter of 2024, giving us a small picture of upcoming growth trends.

The outlook for Asian economies in the second quarter of 2024 shows a mixed trend in GDP growth gap predictions: Kazakhstan and Taipei,China are predicted to be above their long-term trend and slowing. Malaysia and Thailand are likewise predicted to be above their long-term trend—but accelerating. Conversely, Indonesia, the ROK, the Philippines, Singapore, and Viet Nam are predicted to be below trend but accelerating, whereas the PRC; Hong Kong, China; and India are predicted to be below trend and slowing.

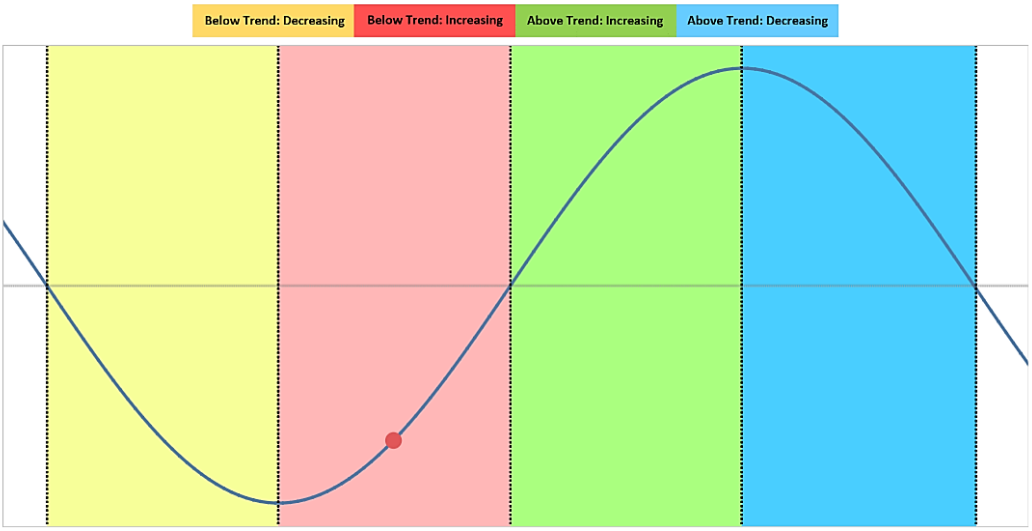
Most economies are predicted to stay within the same phase of the business cycle from Q1 to Q2: Indonesia, the ROK, Malaysia, Singapore, and Viet Nam. Other economies are predicted to move to the following phase in the business cycle: India is expected to cool down to below its long-term trend following a period of above-trend growth. Kazakhstan and Taipei,China are expected to slow down but remain above trend. The Philippines is expected to remain below its long-term trend but should start accelerating. In contrast, the PRC is predicted to revert to a decreasing and below long-term growth trend.

Two economies deviate from the usual phases of the business cycle. Thailand is classified as below trend and increasing in Q1 but by Q2 it had shifted to above trend and increasing. This strong growth was particularly evident in April and June, with investment and trade making the largest contributions, followed by private consumption and financial investment. In contrast, Hong Kong, China transitioned from above trend and increasing in Q4 2023 (latest actual GDP gap) to below trend and decreasing in Q1 and Q2 of 2024. While all sectors contributed to this decline, the main factors were private consumption, trade, and financial investment.

5. ECONOMIC ACTIVITY INDEX RESULTS: PRINCIPAL COMPONENT ANALYSIS VERSUS MACHINE LEARNING

As previously mentioned, the initial methodology used to create the EAI relies heavily on PCA being conducted on selected monthly economic indicators for each economy. The EAI produced using the PCA is projected onto business cycle frequencies and GDP growth rates in hopes that researchers and policymakers can get an idea of upcoming trends in the economy, in that they can see the point in the business cycle in which the economy is currently, and where it will be in the upcoming months or quarters. To depict the phase in which the economy resides at certain times, each real GDP growth gap value, as well as EAI value that is projected onto business cycle frequencies, is assigned to one of four areas in the business cycle: increasing or decreasing below or above the trend line. Figure 7 depicts an economy that is currently increasing below the business cycle trend.

Figure 7: Generic Business Cycle with Outlined Sections



Source: Authors.

The ML-based methodology to track the EAI does not require a projection onto business cycle frequencies or GDP growth levels because models are trained using already computed real GDP growth gap data. As a result, the EAI values produced by the newer models can be compared with actual values both in total value/magnitude and in the position in which the economy sits at certain times. This is in contrast with the PCA-produced results, which can be accurately examined only through the position of the economy and not through actual growth gap values.

Here, we break down the overall performance of both PCA-based and ML-based EAI values produced by the new models as compared with actual GDP growth gap computed data per economy using R-squared and RMSE. We also examine the performance of each model as it pertains to the position of the economy in the business cycle at different times. This metric is called the trend match rate, indicating the rate at which the position of the economy on the

business cycle based on the EAI matches that of the position based on real GDP growth gap data.

The table below presents the overall performance of the ML- and PCA-based EAI methodologies. For each economy, the R-squared and RMSE values were computed by comparing output results for each EAI value against its corresponding actual GDP growth gap. The results were calculated only for months in which the ML- and PCA-based EAI, as well as the actual GDP gap, were available.

Table: Metric Comparison for PCA- and Machine Learning-Based EAI Predictions

Economy	R-Squared		RMSE		Trend Match %	
	ML	PCA	ML	PCA	ML	PCA
People's Republic of China	0.64	0.46	0.32	0.40	0.65	0.74
Hong Kong, China	0.54	0.51	0.62	0.65	0.48	0.64
India	0.72	0.78	1.08	0.96	0.67	0.72
Indonesia	0.78	-0.06	0.45	0.99	0.73	0.27
Kazakhstan	0.77	-2.02	0.20	0.74	0.59	0.26
Republic of Korea	0.75	-0.08	0.28	0.59	0.60	0.60
Malaysia	0.71	0.67	0.57	0.60	0.57	0.67
Philippines	0.89	0.73	0.48	0.74	0.58	0.58
Singapore	0.69	0.63	0.55	0.61	0.62	0.59
Taipei,China	0.66	0.14	0.39	0.63	0.65	0.53
Thailand	0.86	0.72	0.37	0.52	0.80	0.70
Viet Nam	0.84	-2.01	0.16	0.72	0.70	0.60

ML = machine learning, PCA = principal component analysis, RMSE = root mean squared error.
Source: Authors' estimates.

While the PCA-based approach to producing EAI figures is not expected to exactly mirror actual GDP growth gap values, there is still merit in comparing and evaluating them. As expected, R-squared results for the PCA method are relatively low for most economies. Hair and Alamer (2022) suggest that, in the realm of social science research, an R-squared 0.00 to 0.10, 0.11 to 0.30, 0.30 to 0.50, and greater than 0.50 can constitute weak, modest, moderate, and strong explanatory power, respectively. A strong R-squared, meanwhile, can be considered high-performing only if there are no issues with spurious causation or multi-collinearity (Ozili 2023), both of which are assumed to be mitigated by the removal of indicators that are linearly related to others, as well as the initial selection of each indicator by subject matter experts. Based on this principle, from the economies that are examined in this paper, the PCA method produced strong results for five: Hong Kong, China (R-squared = 0.51), India (0.78), Malaysia (0.67), the Philippines (0.73), and Thailand (0.72). The remaining economies have R-squared results at or below 0.46, with Kazakhstan and Viet Nam dropping below -2, indicating extreme cases of weak explanatory power, which can clearly be seen in their respective comparison charts.

Evaluating the ML-based models when producing EAI values against the real GDP growth gap gives us more promising results. The EAI models trained for each economy, once evaluated, produced R-squared values indicating strong explanatory power for every economy covered in this study. Most of the models achieved an R-squared at or above 0.70, with the average for all economies being at 0.74. The only outlier in this case is India, wherein the PCA method of

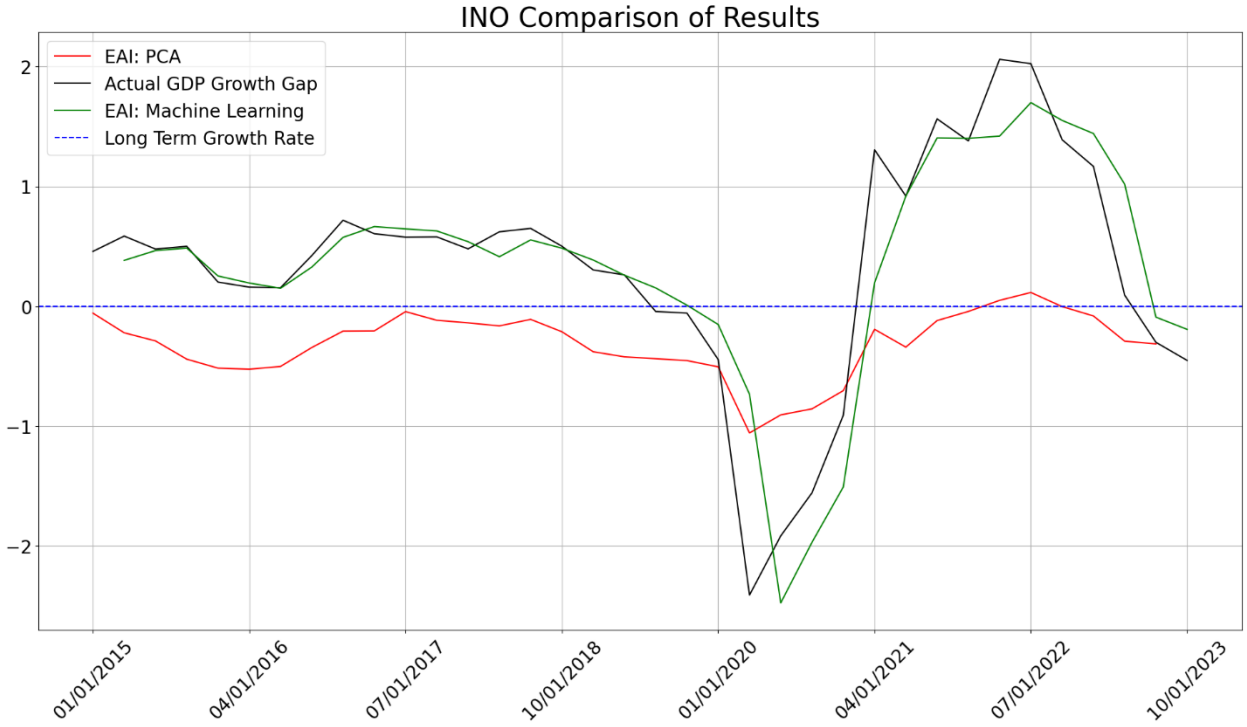
generating the EAI remained a better indicator of activity, with an R-squared value of 0.78 compared with the ML method value of 0.72.

In terms of RMSE, a similar trend can be seen such that economies with higher R-squared values generally have lower RMSE figures. However, it is difficult to take RMSE at face value owing to its high dependence on the average magnitude of the EAI. We can, however, see a sharp decrease in RMSE between the two methodologies for most economies, with the ML method resulting in all economies aside from India being below 0.62, for an average of 0.46 across the board. On the other hand, the PCA EAI averaged an RMSE of 0.68, with only India's figure being lower than its ML counterpart.

The trend match rate, which is the accuracy of the methodology when tracking the position of the economy on the business cycle at certain periods of time, is the most important metric to examine for the case of PCA because the activity index is built to mirror business cycle movements instead of GDP levels. It can be seen in the full results of the table that trend match accuracy is more promising than the results of the other two metrics for the PCA method. The PRC; Hong Kong, China; India; and Malaysia had better trend match scores, while the ROK and the Philippines matched the scores of the ML results. For economies in which the trend match scores of the ML method beat out the PCA method, the score discrepancy is larger on average.

Finally, we can examine the modeling capabilities of each method more clearly through a visual comparison of results. The charts (Figures C1 to C12 in the Appendix) depict the movements of each method's computed EAI over time against the actual GDP growth gap. The chart for Indonesia can be seen below as an example, with the black line depicting the actual GDP growth gap, the green the ML-based EAI, and the red the PCA-based EAI. Each chart begins at a static point in January 2015 and ends around the end of 2023 and shows the long-term growth rate, which is where potential and actual growth meet, depicted by the blue dashed line. Some charts, like those of Indonesia and Viet Nam, do not have EAI ML values dating back to 2015, owing to a lack of indicators going back far enough to train the models and produce values by then. Similarly, PCA EAI values end slightly before ML EAI values, as we transitioned to the ML method and ceased the computation of the PCA-based EAI by then.

Figure 8: Comparison Chart for EAI (PCA and ML) Versus Actual GDP Growth Gap for Indonesia



EAI = Economic Activity Index, GDP = gross domestic product, ML = machine learning, PCA = principal component analysis, INO = Indonesia.
Source: Authors' estimates.

Overall, while the PCA can give us some information about the movements of business cycles for several Asian economies, even going as far as to almost mimic the movements of actual GDP growth gap levels, the indexing method is still quite weak and has a lot of room for improvement. The ML methodology of producing EAIs, meanwhile, proves much more powerful and informative in most cases compared with the relatively simple PCA method, with improvements in most metrics across the board.

6. CONCLUSION

In this paper, we develop an Economic Activity Index that closely tracks periods of economic expansion and contraction within the business cycles of selected Asian economies. We applied PCA and various ML models with the monthly indicators for the PRC; Hong Kong, China; India; Indonesia; the Republic of Korea; Malaysia; the Philippines; Singapore; Taipei,China; Thailand; and Viet Nam to generate a monthly summary statistic that we call the Economic Activity Index, or EAI. To show that this can track periods of economic expansion and contraction, we evaluated the generated results to the respective cycles of the selected economies.

We show that the methodology generated a summary statistic that was able to mimic business cycle activity in these economies. The index also retains granularity, linking the aggregate index with individual macroeconomic series. Next, we also compare the overall performance of the

PCA- and ML-derived EAI figures using metrics such as R-squared and RMSE against actual growth values. We can conclude that the ML methods performed better than PCA for most cases, though they can still be improved upon when it comes to lingering lag issues and data availability.

Further research could explore the use of more consistent, high-frequency indicators that remain uniform in terms of publishing times between economies. Data derived from Google Trends, for example, like the OECD's weekly tracker, could be explored for the Asian economies discussed in this paper. Moreover, other ML or neural network-based algorithms can be utilized, while also incorporating hyperparameter tuning to further improve the performance of the economy models in the future.

REFERENCES

Alangari, Nourah, Mohamed El Bachir Menai, Hassan Mathkour, and Ibrahim Almosallam. 2023. "Exploring Evaluation Methods for Interpretable Machine Learning: A Survey." *Information* 14 (8): 469. https://doi.org/10.3390/info14080469

Aruoba, S. Borağan, Francis Diebold, and Chiara Scotti. 2009. "Real-Time Measurement of Business Conditions." *Journal of Business & Economic Statistics* 27 (4): 417–427.

Bok, Brandyn, Daniele Caratelli, Domenico Giannone, Argia Sbordone, and Andrea Tambalotti. 2018. "Macroeconomic Nowcasting and Forecasting with Big Data." *Annual Review of Economics* 10: 615–643.

Bragoli, Daniela. 2017. "Now-casting the Japanese Economy." *International Journal of Forecasting* 33 (2): 390–402.

Bragoli, Daniela, and Jack Fosten. 2018. "Nowcasting Indian GDP." *Oxford Bulletin of Economics and Statistics* 80 (2): 259–282.

Camacho, Maximo, and Gabriel Perez-Quiros. 2010. "Introducing the Euro-Sting: Short-Term Indicator of Euro Area Growth." *Journal of Applied Econometrics* 25 (4): 663–694.

Cascaldi-Garcia, Danilo, Matteo Luciani, and Michele Modugno, M. 2023. "Lessons from Nowcasting GDP Across the World." International Finance Discussion Paper 1385. Washington, DC: Board of the Governors of the Federal Reserve System.

D'Agostino, Antonello, Domenico Giannone, Michele Lenza, and Michele Modugno. 2015. "Nowcasting Business Cycles: A Bayesian Approach to Dynamic Heterogeneous Factor Models." Finance and Economics Discussion Series. Washington, DC: Board of the Governors of the Federal Reserve System.

Evans, Charles, Chin Te Liu, and Genevieve Pham-Kanter. 2002. "The 2001 Recession and the Chicago Fed National Activity Index: Identifying Business Cycle Turning Points". *Economic Perspectives, Federal Reserve Bank of Chicago* 26, Third Quarter: 26–43.

Federal Reserve Bank of Chicago. n.d. "Chicago Fed National Activity Index (CFNAI)." www.chicagofed.org/research/data/cfnai/current-data

Giannone, Domenico, Lucrezia Reichlin, and David Small. 2008. "Nowcasting: The Real-Time Informational Content of Macroeconomic Data." *Journal of Monetary Economics* 55 (4): 665–676.

Giannone, Domenico, Silvia Miranda Agrippino, and Michele Modugno. 2013. "Nowcasting China Real GDP." CIRANO. www.cirano.qc.ca/conferences/public/pdf/realtime2013/Papers/Giannone_Miranda.pdf

Giannone, Domenico, Lucrezia Reichlin, Marta Bańbura, and Michele Modugno. 2013. "Now-casting and the Real-Time Data Flow." Working Paper 1564. Frankfurt. European Central Bank.

Hair, Joseph, and Abdullah Alamer. 2022. “Partial Least Squares Structural Equation Modeling (PLS-SEM) in Second Language and Education Research: Guidelines Using an Applied Example.” *Research Methods in Applied Linguistics* 1 (3): 100027. https://doi.org/10.1016/j.rmal.2022.100027

Higgins, Patrick. 2014. “GDPNow a model for GDP ‘nowcasting’.” Working Paper. Atlanta, GA: Federal Reserve Bank of Atlanta.

Hinds, Sam, Lucy Rimmington, Hugh Dance, Jonathan Gillham, Andrew Sentence, and John Hawksworth. 2017. “A Machine Learning Approach to Estimating Current GDP Growth.” *UK Economic Outlook* (July). www.pwc.co.uk/economic-services/ukeo/pwcukeo-section4-nowcasting-july-2017.pdf

Hyndman, Rob, and George Athanasopoulos. 2021. *Forecasting: Principles and Practice* (2nd ed.). Melbourne: OTexts.

Jolliffe, Ian, and Jorge Cadima. 2016. “Principal Component Analysis: A Review and Recent Developments.” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374 (2065), 20150202. https://doi.org/10.1098/rsta.2015.0202

Leung, Charles, Kenneth, Chow, Matthew Yiu, and Dickson Tam. 2010. “House Market in Chinese Cities: Dynamic Modeling, In-Sampling Fitting and Out-of-Sample Forecasting.” *Journal of the Asian Real Estate Society* 14 (1): 85–117.

Luciani, Matteo, Madhavi Pundit, Arief Ramayandi, and Giovanni Veronese. 2018. “Nowcasting Indonesia.” *Empirical Economics* 55 (2):597–619.

Lundberg, Scott, and Su-In Lee. 2017. “A Unified Approach to Interpreting Model Predictions.” *Proceedings of the 31st International Conference on Neural Information Processing Systems*: 4768–4777.

Makridakis, Spyros, Evangelos Spiliotis, and Vassilios Assimakopoulos. 2018. “Statistical and Machine Learning Forecasting Methods: Concerns and Ways Forward.” *PLOS ONE* 13 (3). https://doi.org/10.1371/journal.pone.0194889

Mariano, Roberto, and Yasutomo Murasawa. 2003. “A New Coincident Index of Business Cycles Based on Monthly and Quarterly Series.” *Journal of Applied Econometrics* 18 (4): 427–443.

Maskin, Eric. 2016. “How Can Cooperative Game Theory Be Made More Relevant to Economics? An Open Problem.” In John Forbes Nash Jr, and Michael Rassias (eds). *Open Problems in Mathematics*: 347–350.

Montavon, Grégoire, Alexander Binder, Sebastian Lapuschkin, Wojciech Samek, and Klaus-Robert Muller. 2019. “Layer-Wise Relevance Propagation: An Overview.” *Springer Link* 11700: 193–209.

Ozili, Peterson. 2022. “The Acceptable R-square in Empirical Modelling for Social Science Research.” *Social Research Methodology and Publishing Results*. https://ssrn.com/abstract=4128165 or http://dx.doi.org/10.2139/ssrn.4128165

Padmawar, Vaishnavi, Pradnya Pawar, and Akshit Karande. 2021. “Gross Domestic Product Prediction Using Machine Learning.” *International Research Journal of Engineering and Technology* 8 (6): 2817–2821.

Sarker, Iqbal. 2021. “Machine Learning: Algorithms, Real-World Applications and Research Directions.” *SN Computer Science* 2 (3). https://doi.org/10.1007/s42979-021-00592-x

Shrikumar, Avanti, Peyton Greenside, and Anshul Kundaje. 2017. “Learning Important Features Through Propagating Activation Differences.” *Proceedings of the 34th International Conference on Machine Learning*: 3145–3153.

Stock, James H., and Mark W. Watson. 1989. “New Indexes of Coincident and Leading Economic Indicators.” In *NBER Macroeconomics Annual 1989, Volume 4*, edited by Olivier J. Blanchard and Stanley Fischer, 351–409. Cambridge, MA: National Bureau of Economic Research.

Stock, James, and Mark Watson. 2016. “Dynamic Factor Models, Factor-Augmented Vector Autoregressions, and Structural Vector Autoregressions in Macroeconomics.” *Handbook of Macroeconomics* 2: 415–525.

Visani, Giorgio, Enrico Bagli, and Federico Chesani. 2020. “Optilime: Optimized LIME Explanations for Diagnostic Computer Algorithms.” arXiv:2006.05714

Woloszko, N. 2020. “Tracking Activity in Real Time with Google Trends.” Economics Department Working Papers 1634. Paris: Organisation for Economic Co-operation and Development.

Yoon, Jaehyun. 2020. “Forecasting of Real GDP Growth Using Machine Learning Models: Gradient Boosting and Random Forest Approach.” *Computational Economics* 57 (1): 247–265.

Monitoring Business Cycle Fluctuations in Asia

Timely insights on business cycle fluctuations are necessary to guide effective policy decisions, but delays in official data releases pose challenges. This paper introduces an Economic Activity Index to monitor monthly business cycles in selected Asian economies using updated economic indicators across six categories: consumption, investment, trade, government, financial, and external sectors. It shows that machine learning algorithms accurately track output gap movements, offering a robust tool for monitoring economic fluctuations.

About the Asian Development Bank

ADB is committed to achieving a prosperous, inclusive, resilient, and sustainable Asia and the Pacific, while sustaining its efforts to eradicate extreme poverty. Established in 1966, it is owned by 69 members —49 from the region. Its main instruments for helping its developing member countries are policy dialogue, loans, equity investments, guarantees, grants, and technical assistance.

ADB

ASIAN DEVELOPMENT BANK
6 ADB Avenue, Mandaluyong City
1550 Metro Manila, Philippines
www.adb.org