

**Use of High-Frequency Indicators compiled by  
the Central Bank of Sri Lanka to enhance  
informed Monetary Policy Decisions**

**By**

**Gunawardhana S R C L, Krishantha R M G S, Maheepala M M J D**

**SAARCFINANCE Seminar and Working Group Meeting  
Male, Maldives November 2, 2022**

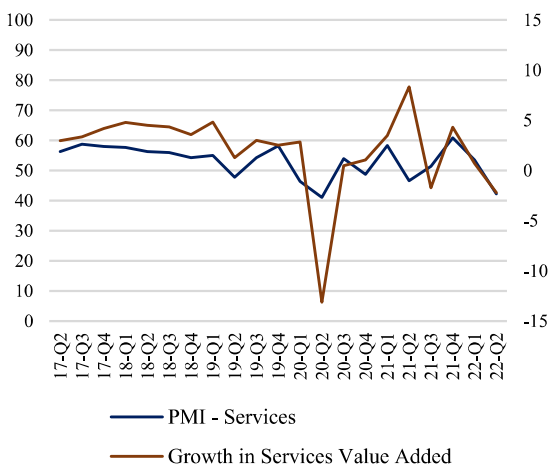
## Motive to High-Frequency Indicators

The necessity of forerunning quality inputs to pre-emptive policy decision-making has become more critical than ever due to the current unprecedented economic dynamism. Hence, the monetary authorities are on a mission for newfangled solutions to tackle those high-magnitude risk factors, including the pandemic-induced fallbacks, inflation tensions provoked by the energy pressures, and the subdued global economic outlook owing to the across globe monetary and fiscal tightening measures. Foreseeing the demand, the Central Bank of Sri Lanka (CBSL) revamped its data collection process years back by implementing an array of technology-driven solutions, including developing high-frequency indicators, which are upgraded from time to time.

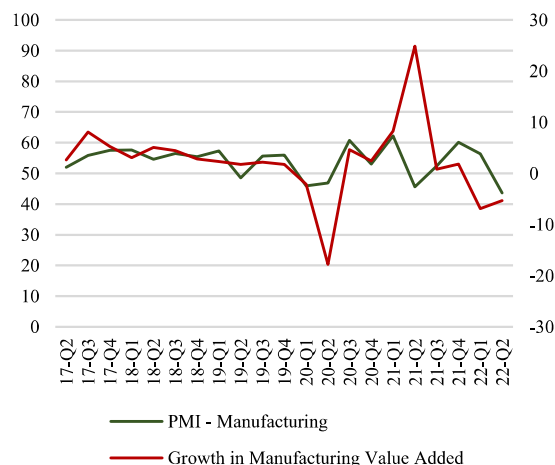
## Tracking the Economy through Regular Business Surveys

The CBSL commenced Purchasing Managers' Index (PMI) surveys in 2015 to generate a timely set of indicators which are available in advance of the release of official statistics. Accordingly, monthly surveys are conducted for the key activities of Manufacturing, Services and Construction to obtain firsthand information on the industry performances and their insights. The headline as well as sub-indices of production/activities, new orders, employment, suppliers' delivery time, stock of purchases, backlog of work, future expectations and price changes are compiled, which serve as leading indicators. The information thus collected is used in advanced analysis on fields such as economic growth, inflation and labour market, assisting policy decisions and near-term forecasting of macroeconomic variables. In addition, the PMI-Manufacturing and Services reports are published by the mid of the following month for the use of analysts and decision-makers to benchmark industry performance and formulate business strategy.

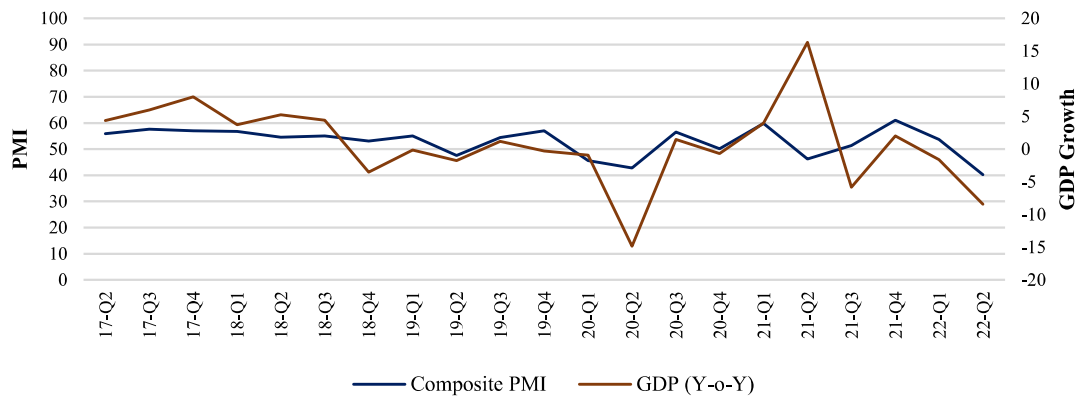
**Figure 1: PMI - Manufacturing vs Growth in Manufacturing Value Added**



**Figure 2: PMI - Services vs Growth in Services Value Added**



**Figure 3: Composite PMI vs GDP Growth**



### **New Avenues for Data Collection**

With the advancement of new technologies and the emergence of the concept of Big Data, the CBSL is urged to use new avenues of information due to constraints in data availability under conventional set-up. Accordingly, alternative high-frequency data sources such as web portals and Google Trends were identified, and indicators were developed in the fields of real estate market, vehicle market and retail sales. In addition, text data analysis was carried out by developing sentiment indices using web-based news articles and analysing social media content available on specific matters to support informed decision-making. Furthermore, experimental work is carried out using data science techniques to explore the possibility of enhancing the current inflation forecasting process.

With the digital expansion, the amount of data available in cyberspace is growing exponentially, enabling access to data in more sophisticated ways. Web Scraping, an automated extraction of web-based information, is revolutionary and can provide advanced, real-time and more granular insight for economic and financial analysis. For central banks, scraping augmented with big data opens the possibility of extracting timelier economic signals, applying new statistical methodologies and thus enhancing economic forecasts to obtain rapid feedback on policy impacts. Such information can aid a better understanding of the policy tools in the national context giving policymakers the ability to refine and/or choose the optimal strategies. Nevertheless, for emerging economies, utilising web scraping is challenging, largely due to lack of updated web content and defective content structures.

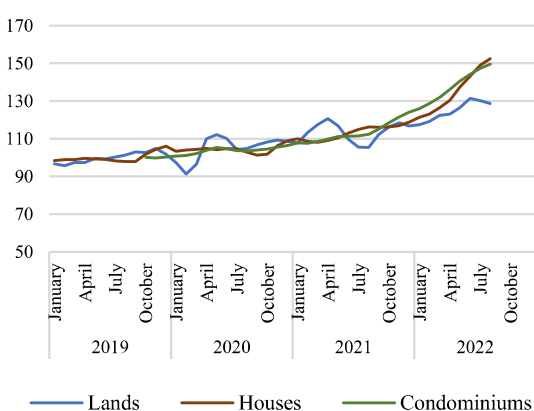
### **Economic and Financial Insights Via Property Price Indices**

Considering the relatively large contribution of the real estate sector to the overall economy, monitoring the trends in real estate property prices is important. However, the biggest challenge in compiling property price indices is the unavailability of required data. For this purpose, property transaction data available at government institutions are the most suitable source, which is currently used by many countries that possess well-established IT platforms for managing official data. However, IT systems available at the government authorities in Sri

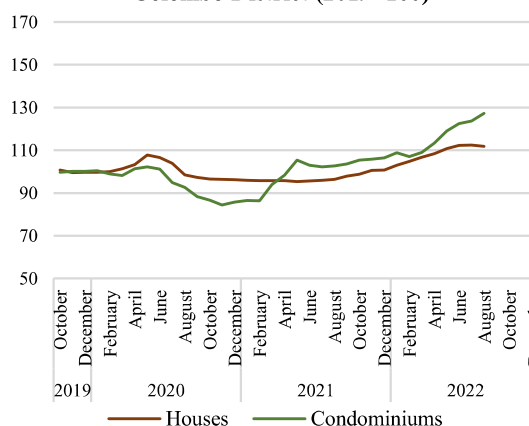
Lanka, such as Land Registries, Provincial Revenue Authorities and Survey Department, are underdeveloped, which has limited the availability of a time-bound transaction data set on a continual basis with required granular level data. Thus, the CBSL explored alternative data sources in compiling property price indices. Surveys with property developers and real estate property sales advertisements published on property websites are such alternative data sources to obtain sales information on properties. Accordingly, the CBSL collected the advertised (asking) prices and other property characteristics available in the sales advertisements published on popular property websites using web scraping techniques. In addition, property rent advertisement information was also collected through web scraping to observe the price movements in the real estate property rental market.

Using the data collected through web advertisements, which is a non-traditional data source, asking price indices for lands, houses and condominiums, and rental indices for houses and condominiums were compiled on a monthly basis. The indices are compiled employing hedonic regression-based rolling window time dummy method, complying with international best practices. These indices are used to observe the price movements in the real estate market as well as an early warning indicator of the economy, particularly to monitor investment trends. The asking price indices for lands, houses and condominiums are reported to the management to assist decision-making, especially on macroprudential surveillance. These indices are published on the CBSL website to provide more reliable information to the general public on the real estate sector. Further, the availability of information would make the industry more transparent, enabling stakeholders such as investors, buyers, developers and banks to make more informed decisions. Meanwhile, the rental indices compiled for houses and condominiums are at the experimental level at the moment and are used only for internal analysis of the Statistics Department of the CBSL. Consequently, it is used for inflation forecasting as an indicator to observe the trend in rental expenditure component in the inflation basket.

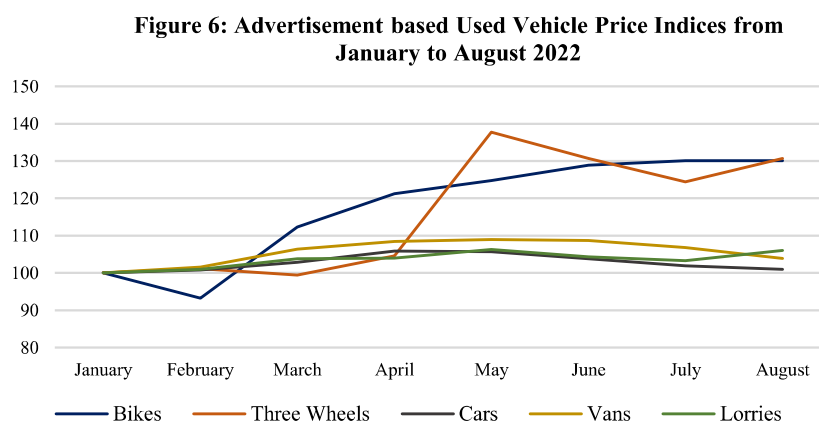
**Figure 4: Asking Price Indices for Lands, Houses and Condominiums in Colombo District (2019=100)**



**Figure 5: Advertisement based Rental Indices for Houses and Condominiums in Colombo District (2019=100)**

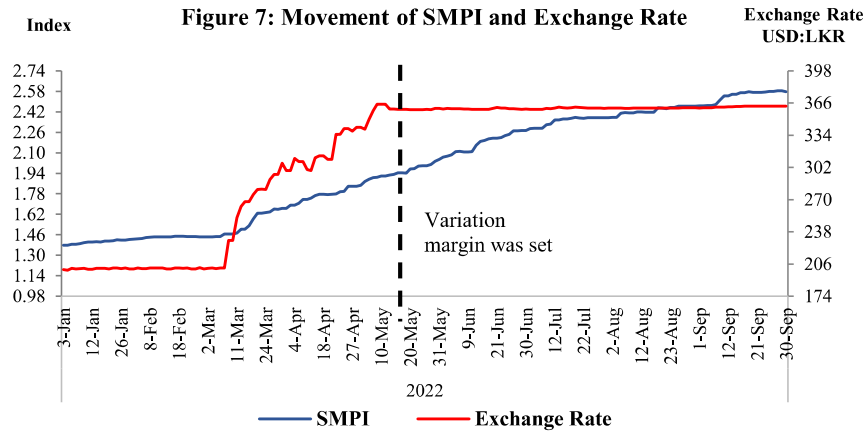


In the meantime, it was observed that there is no specific indicator in Sri Lanka to monitor the price movements in the vehicle market as well. Similar to real estate property data, vehicle transaction data are also unavailable on the official level. Therefore, information available in vehicle advertisements were collected through web scraping on a weekly basis, and relevant price indices were compiled monthly under five categories, namely, motorbikes, three-wheelers, cars, vans and lorries. The same technique used in property price indices, which is hedonic regression-based rolling window time dummy method, was employed in compiling the used vehicle price indices. However, the compilation process is at the experimental level and work are underway to improve the index compilation process. These indices can be used for economic analysis on customer behaviour and expenditure patterns, and also as a proxy indicator to observe the price movements of vehicle component in the inflation basket. Moreover, in the background of suspensions to import vehicles and foreign exchange shortage in Sri Lanka, these used vehicle price indices are useful to analyse vehicle price movements with key economic variables such as exchange rates, inflation and interest rates. Further, considering the significance of credit and lease facilities provided to purchase vehicles, an analysis can be carried out to determine the effects of the vehicle market on financial stability.



## Retail Market Price Effects on Inflation and Exchange Rate

In addition, popular supermarket websites were scraped on a daily basis to collect retail prices of selected consumer goods, with the objective of enhancing the quality and quantity of data used in forecasting the inflation. Accordingly, the price changes of consumer goods in the inflation basket were observed and used to validate the respective price movements collected via other modes. In addition, a daily Supermarket Price Index (SMPI) is compiled using prices of selected imported consumer goods and used in studying the exchange rate pass-through effect. The index was compiled following the Jevons Index methodology to capture the price movements of imported items. The results obtained through the analysis are used for inflation forecasting process and reported to the senior management to be used in monetary policy decision-making as and when needed.



### Sentiment Analysis on Public Perceptions

Further, with the advancement of technology, many Central Banks tend to conduct text data analysis/ text mining to gain insights to support their decision-making process. Text mining is the process of transforming the information available in the form of text into a structured format to identify meaningful patterns and new insights. With the recent advent of computational tools and data science techniques, a range of new methodologies has been used to quantify linguistic content into a form suitable for analysis or modelling. Text data analysis concepts are presently widely used in fields such as marketing, politics and defence to effectively analyse a variety of written information to reveal patterns hidden in information.

The research interest in text mining in the central banking context has also increased over the years. This new technique provides a range of data sources useful for assessing sentiments/expressions, which is otherwise impossible to analyse quantitatively using any other conventional means. Further, the development of technical tools facilitating information retrieval and analysis has resulted in the widespread application of these novel concepts in analysing macroeconomic issues, studying central bank communication and financial stability in particular. The key text data sources for central banks include news articles, social media, supervisory information and various written reports.

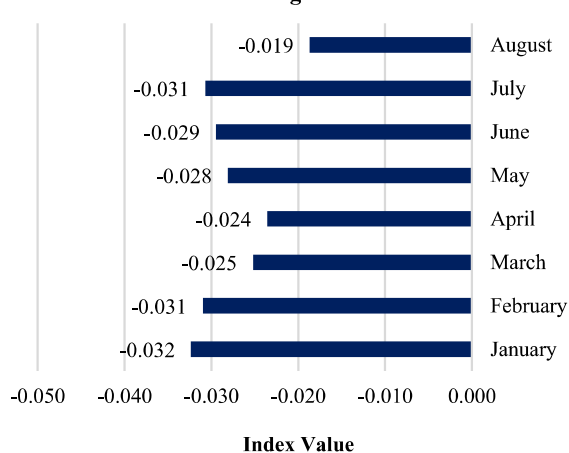
Accordingly, a Sentiment Analysis was initiated to gauge the external opinion expressed through the news articles published related to the activities of the CBSL. Data Science techniques such as Natural Language Processing (NLP) and Machine Learning (ML) were employed for this analysis. Accordingly, data preparation was carried out using NLP and ML models based on four algorithms, namely, Naïve Bayes, Support Vector Machine, Random Forest and Boosting were developed and used to determine the polarity of sentiments of the collected news articles. The analysis was carried out separately for English and Sinhala news articles, and among the alternative models experimented, Random Forest model was identified as the most accurate model. Therefore, using the outputs of the random forest model, the

Sentiment Indices<sup>1</sup> were computed on a monthly basis, which would be used to understand the opinion expressed through the news articles published related to the CBSL. This analysis is also currently carried out on an investigational basis to explore the possibility of gaining insights from text data for informed decision-making.

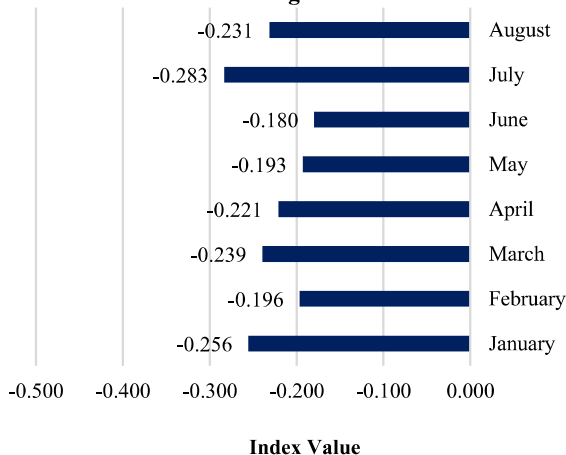
**Table 1: Accuracy of Machine Learning Models Experimented**

Machine Learning Model	Model Accuracy	
	English	Sinhala
Naïve Bayes	64.58%	56.73%
SVM	76.49%	62.60%
Random Forest	85.38%	62.11%
Boosting	63.80%	60.01%

**Figure 8: Sentiment Indices Compiled for English News Articles based on Random Forest Algorithm**



**Figure 9: Sentiment Indices Compiled for Sinhala News Articles based on Random Forest Algorithm**



## Analysing Social Media Content on Workers' Remittances

Another analysis was carried out using text data collected through web scraping social media. Recently, a necessity arose to obtain information on unofficial channels for workers' remittances of Sri Lanka during a short period of time. Thus, time was limited to conduct a survey with relevant parties to get the required information. At this point, the study was carried out based on social media inputs collected through web scraping the comments of relevant YouTube and Facebook contents. The text data collected was analysed to identify the reasons behind the downfall of workers' remittances.

<sup>1</sup> The index was calculated as follows.

$$\text{Index Value} = \frac{\text{No. of Positive Sentences} - \text{No. of Negative Sentences}}{\text{Total no. of Sentences}}$$

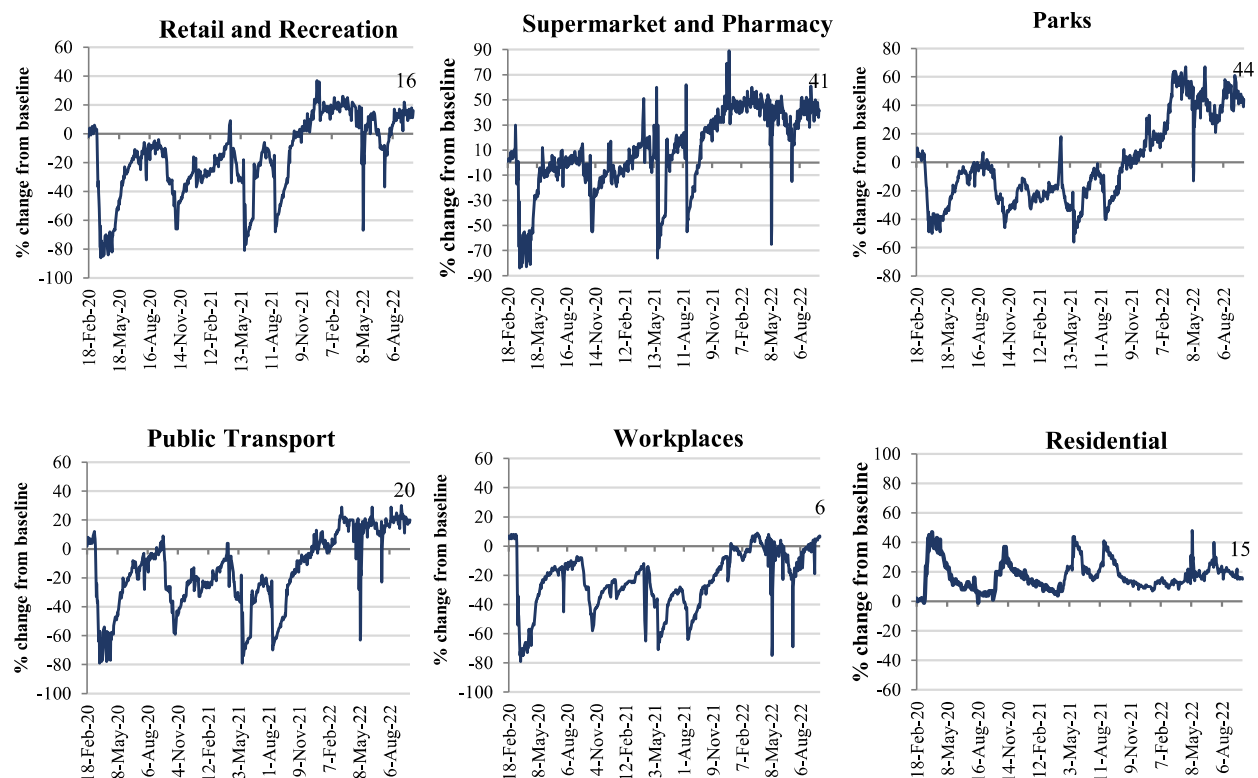
A positive index value indicates an overall positive sentiment for news articles within the specified data set and vice versa.



## Bridging the Data Gap during the Pandemic

During the pandemic time, the Department of Census and Statistics of Sri Lanka (DCS) has delayed publishing Gross Domestic Product (GDP) estimates due to unavailability of survey results and mobility restrictions. Therefore, the CBSL has moved into several high-frequency indicators to capture the economic activities of the country. Google mobility report is one of such high-frequency indicators which is widely used to monitor community movements across six different sectors, namely, Retail and Recreation, Supermarkets and Pharmacy, Parks, Public Transport, Workplaces and Residential which enables to capture economic activities of the country and to examine how fast the economies converge to the pre-COVID levels.

**Figure 10: Google Mobility Trends**



Source: Community Mobility Reports – Google

Note: 1. Excludes weekends and Sri Lankan holidays

2. Google has revised data for all countries from 17 August 2020, onwards. The graphs show the new data series

Further, time to time, air pollution levels in main cities in Sri Lanka were observed using air quality index data compiled by National Building Research Organisation on daily basis, available for PM 2.5 and PM 10 (Particulate Matter (PM) which is a mixture of solid and liquid particles that are suspended in the air. Especially after the COVID-19 pandemic and its impact on the economy, these indicators are used to gauge the level of economic activities compared to the activities during pandemic induced lockdown period in early part of the year to measure



the pace of recovery. These indicators are used for nowcasting GDP based on spread sheet analysis.

In addition, stringency index<sup>2</sup> and highway traffic data also were better early indicators to identify contraction during pandemic periods. These high-frequency data are considered as proxies to capture economic activities in Sri Lanka.

As mentioned above, the CBSL has explored the possibility of using high-frequency data and data science techniques to enhance its functions. These findings have provided valuable inputs to policy formulation of the country.

### **Way Forward**

Since forecasting inflation is a key function carried out by the CBSL to support monetary policy decisions, possible alternative forecasting methods are explored due to the current inflation pressure in the country. Inflation forecasting has been carried out using conventional statistical models over the years, and recently, an experimental study was carried out to forecast inflation using machine learning models to be used as a supportive component. However, the accuracy of the machine learning forecasts was low compared to the conventional model forecasts. Accordingly, work is underway to further expand the data avenues by exploring possible high-frequency data and to enhance inflation forecasting by employing appropriate data science techniques such as deep learning and natural language processing.

In addition, it is expected to improve the text data analysis conducted using news-based data and social media content to obtain information on inflation expectations using text mining in order to gain insights for monetary policy decision-making process.

---

<sup>2</sup> The nine metrics used to calculate the Stringency Index are school closures, workplace closures, cancellation of public events, restrictions on public gatherings, closures of public transport, stay-at-home requirements, public information campaigns, restriction on international movements, and international travel controls. A higher score indicates a stricter response (i.e. 100= strictest response).