Research Article

Nowcasting the Transportation and Accommodation Sectors Growth using the Google Trends Index

Dhiar Niken Larasati[1], Nucke Widowati Kusumo Projo[2,*]

1   Dir of Statistical Analysis and Development, Statistics Indonesia, Jl. Dr. Sutomo 6-8 Jakarta 10710 Indonesia
2   Politeknik Statistika STIS, Jl. Otto Iskandardinata No. 64 C, Jakarta Timur 13330

*   *Corresponding author: nucke@stis.ac.id*

Abstract: This research aims to assess the possibility of the daily and weekly Google Trends Index (GTI) to predict the quarterly GDP growth. The U-MIDAS approach is utilized because it allows using of daily and weekly basis data to forecast quarterly indicators without aggregating them onto a quarterly basis hence it does not eliminate useful information on the daily and weekly data. This research uses quarterly GDP for the transportation sector and the accommodation and restaurant sector which are considered potential industries for the future of Indonesia's economy. The result shows that the daily basis GTI can effectively predict the quarterly GDP growth better than the weekly basis GTI based on the RSE scores.

Keywords: alternative data, big data, GDP, MIDAS, mixed frequency data

Introduction

The evident characteristic of economic development in many countries is the existence of structural transformation [1]. A structural transformation is a change in an economy from agricultural sectors which is labor-intensive activities to manufacturing and services sectors which are skill-intensive activities [2]–[4]. The evidence of countries experiencing a structural transformation appears in the shifting of the sector's contribution to GDP or the movement of the laborers ([2], [5].

Indonesia is most likely experiencing a structural transformation. It is evidenced by the contribution of service sectors to total GDP continues to increase in the past 10 years. Its contribution is around 46,2 percent to the total GDP in 2021, increasing from 41,76 percent in 2010. Two crucial service sectors for Indonesia's economy are the transportation sector and accommodation and restaurants. The average growth of the transportation sector in the past 10 years is 7.3 percent while the accommodation and restaurants sector is 5.7 percent, higher than the average growth of the agriculture (3.5 percent) and manufacturing sector (4.4 percent) which are the two sectors with the largest contribution to GDP. The labor absorption in the transportation and accommodation and restaurants sectors in 2022 also shows an increase of 16.6 for transportation and 54.1 percent for accommodation and restaurants from 2010, strengthening the existence of structural transformation in Indonesia. These sectors also play an important role in tourism development in Indonesia which is considered a potential industry for the future of Indonesia's economy [6]. Accordingly, forecasting the growth of the transportation sector and accommodation and food services sector will be prominent in providing insight related to the development of the tourism sector in Indonesia which is essential for both the government and private sectors.

Most forecasting methods usually require data with the same frequencies. However, some datasets available at higher frequencies may provide prospectively fruitful information for forecasting which will most likely fail to be applied by those forecasting methods [7], [8]. This concern leads to the possibility of using mixed-frequency data for forecasting known as Mixed Data Sampling (MIDAS) regression [7].

Most studies of MIDAS implementation in Indonesia and around the world use monthly socioeconomic data to forecast the quarterly GDP [8]–[11]. However, in the era of internet proliferation, data can come from various sources with various frequencies and various forms which are later known as big data or alternative data. Some of the alternative data do not rule out the possibility of being strongly associated with the GDP. Nakazawa [12] has proven the possibility of MIDAS implementation using monthly basis alternative data namely the google trends index and METIPOS to predict Japan's GDP. Nevertheless, based on the literature review carried out so far, this research has not found a study that implements daily and weekly basis alternative data to predict the GDP, particularly in Indonesia. Therefore, this research contributes by offering the possibility of using the daily and weekly Google Trends Index to predict Indonesia's Quarterly GDP.

The rest of this research is arranged as follows: the next section describes the current literature on MIDAS Regression and Google Trends, followed by data and methodology. Next, the result of this research is presented. Finally, the two last sections provide some discussion and conclusion.

Methods

Mixed Data Sampling (MIDAS) Regression and Google Trend

MIDAS is first introduced by Ghysels et.al [13]. The objective is to utilize high frequencies data to forecast lower frequencies data without aggregating them into lower frequencies basis. The rationale for this idea is that many researchers experience the condition where the explanatory variables are in high frequencies, while the response variable is in low frequency. Transforming high-frequency data to lower-frequency data may induce deprivation of information [8]. Some recent studies have applied MIDAS to forecast economic indicators, particularly GDP. Barsoum and Stankiewicz [8] performed MIDAS to predict the U.S GDP growth with monthly macroeconomic and financial variables. Subsequent to Barsoum, Bilgin et.al [9] utilized MIDAS to forecast the U.S quarterly GDP using monthly CPI and unemployment rates. Another study was done in Thailand [10] who performed MIDAS regression using monthly indicators that are export growth, unemployment rates, and stock index. There were also some studies of MIDAS implementation to forecast GDP in Indonesia. Utari [14] used the monthly agricultural export data and JCI to forecast quarterly Indonesia's GDP.

Google Trends is a feature from Google that provides information concerning search requests of specific topics made to the Google search engine [15]. It provides real-time and archived daily, weekly, and monthly data by geographic location and category from 2004 [16], [17]. Because of its characteristics, it's considered one of the alternative data that is powerful to be used for predicting, forecasting, moreover for nowcasting. Google Trends has been used dispersedly for forecasting in many sectors. In the business world, the evidence of the study by Wijnhoven and Plant [18] showed that Google Trends gave a better prediction for car sales than social media sentiment did. Another study was done by Ahmed et.al [19] who applied machine learning to analyze whether human behavior captured by Google Trends affects stock markets in Pakistan. In the economic aspect, [20] constructed a machine-learning algorithm to estimate real-time GDP growth in 46 countries using Google Trends. The last study shows evidence of the performance of Google Trends for nowcasting.

There are also some studies regarding the utilization of Google Trends to forecast tourism indicators. [21] applied Random Forest to predict tourism the number of inbound tourists in China using Google Trends. [22] Performed time series regression and ARIMAX model to forecast room occupancy rates using google trends. The study showed that the ARIMAX model gave the best performance based on the RMSE score. Besides the ARIMAX model, a MIDAS approach also has been applied to predicting the tourism indicator from google trends. The result of the study [23] showed that the use of weekly Google Trends in the MIDAS model gave a better performance in predicting monthly tourist arrivals than excluding the Google Trends in the model.

Data

This research uses two datasets, the first is the quarterly GDP growth. This research uses the year-on-year (y-on-y) growth rates of Indonesia's quarterly GDP for the transportation sector and the accommodation and restaurants sector from the second quarter of 2012 to the first quarter of 2022. This data portrays the change of value-added in the transportation sector and accommodation and restaurant

sector in a quarter compared to the same quarter in the previous year. This data is regularly produced and published by Badan Pusat Statistik, the National Statistical Office of Indonesia. Henceforth, this research uses the "GDP sector H" term to represent the quarterly GDP for the transportation sector and the "GDP sector I" term to represent the quarterly GDP for the accommodation and restaurants sector.

The second data set is the Google Trends Index (GTI). The Google Trends gives open access to the public to obtain Google Trends Index (GTI) data. The public can acquire the data by downloading it from the official website or by performing some queries to retrieve it using programming languages. This research applies a package from the R language called "GtrendsR" [24] to perform queries for acquiring information from Google Trends. To get better information on Google Trends regarding the transportation sector and accommodation and restaurants sector from Google Trends, this research utilizes categories instead of specific keywords for the queries [25]. The category "Transportation & Logistics" is used for acquiring information regarding the transportation sector. While the category "Hotels & Accommodations" is used to acquire information regarding the accommodation and restaurant sectors. This research performs queries to obtain the GTI data from 01 January 2012 to 30 June 2022 and calculate the year-on-year growth rate for the index.

Disaggregating Monthly GTI to daily GTI

GTI is an index obtained by normalizing the number of specific search requests to the Google search engine over the selected period [16]. The normalized index of Google Trends is ranged between 0 and 100. A value of 100 indicates the highest number of searches for a certain keyword in a certain period. An index value of 50 indicates the number of searches for a keyword is half of the highest number of searches for the given period. A value of 0 indicates not enough data available/no searches for that keyword. Mathematically, the google trends normalized index can be written as follows:

$$GTI_{ip} = \frac{KW_{ip}}{max(KW)_p} \times 100\% \tag{1}$$

Where GTI_{ip} is the Google Trends Index at time point i in period p, KW_{ip} is the number of searches for a specific keyword at time point i in period p. $max(KW)_p$ is the highest number of searches for a specific keyword at period p. According to formula 1, it should be noticed that GTI is a relative index so different periods return a different index for a specific time point and equivalent keywords.

Google Trends provides different frequencies of the GTI depending on the selected period. Table 1 shows the detail of GTI frequency and the selected period.

Table 1. Period and Frequency of GTI

Period	Frequency of GTI
1 hour	Every minute
4 hours	Every minute
1 day	Every 8 minutes
7 days	Hourly
30 days	Daily
90 days	Daily
12 months	Weekly
5 years	Weekly
More than 5 years	Monthly

Because this research uses growth rates of Quarterly GDP data from 2012 to 2022, the GTI used in this research should meet the period of the data. Regarding Table 1, Google Trends can only provide monthly GTI to meet the period used in this research. Accordingly, this research performs a method introduced in [26] to disaggregate monthly GTI from a large period to daily and weekly GTI.

First, this research retrieves the GTI from January 2012 to June 2022 which returns monthly GTI relative to the selected period. Then, this research also retrieves the GTI for every month in the period which returns daily GTI relative to the selected month. To obtain daily GTI that is comparable over months in the period, this research applies a formula as follows:

$$DGTI_{m_i d_j} = DGTIr_{m_i d_j} \times MGTI_{m_i} \times 100\% \tag{2}$$

Where $DGTI_{m_i d_j}$ is comparable daily GTI over months in the selected period for month i and day j, $⟦$ $DGTIr_{m_i d_j}$ is daily GTI relative to month i for day j, $⟦MGTI_{m_i}$ is monthly GTI for month i. in the selected period. For weekly GTI, this research performs aggregation of daily GTI to weekly GTI using the median score in a respective week. The median score is used thus the aggregation results are not affected by extreme values.

Unit Root Test

This research performs Augmented Dickey-Fuller (ADF) and Phillip-Perron (PP) as unit root tests to check the stationarity of the datasets [27] . The ADF test checks the existence of a unit root from the data using the OLS estimator of an autoregressive model [28] as follows,

$$Y_t = \rho Y_{(t-1)} + \varepsilon_t \tag{3}$$

The existence of a unit root indicates the dataset is not stationer. From equation 3, The Y_t is considered stationary time series if $|\rho| < 1$ and not stationer if $|\rho| = 1$. Accordingly, the time series dataset can be decided as stationer if it meets the alternative hypothesis of ADF test.

PP test is proposed by Phillip and Perron that permits to check of the presents of a unit root for dependent and heterogeneously data using a nonparametric approach [29]. PP test is an improvement of the ADF test which allows the distribution of error. The time series data set is considered stationer if satisfy the alternative hypothesis of the PP test which is $|\rho| < 1$.

Cross Correlation Function (CCF)

The cross-correlation function is used to evaluate the association between two time series datasets [30]. The concept of cross-correlation is similar to cross-variance in statistics. Cross-correlation aims to measure the similarity of two signals to obtain features by comparing the less informative signal with the more informative signal [31]. The cross-correlation function formula can be written as follow,

$$\rho_{xy}(\tau) = \frac{\gamma_{xy}(\tau)}{\sigma_x \sigma_y} \tag{4}$$

$$\text{given } \gamma_{xy}(\tau) = E\big[(X_t - \mu_x)(Y_{t+\tau} - \mu_y)\big]$$

where σ_x and σ_y are the standard deviations of X_t t and Y_t, while μ_x and μ_y are the means of X_t and Y_t.

Mixed Data Sampling (MIDAS) Regression

This research performs an unrestricted MIDAS regression or U-MIDAS. U-MIDAS is a type of MIDAS regression that does not consider the lag polynomials distribution function, or in other words without constraint on the parameters [32], [33]. U-MIDAS parameters is estimated using Ordinary Least Square. U-MIDAS regression can be written as follows,

$$\alpha(B)y_t = \beta(L)'x_{t,0} + \varepsilon_t \tag{5}$$

Where $\alpha(z) = 1 - \sum_{j=1}^{p} \alpha_j z^j$ and

$$x_{t,0} := \big(x_{tm0}^{(0)}, \dots, x_{tmi}^{(i)}, \dots, x_{tmh}^{(h)}\big)',$$

$$\beta(z) = \sum_{j=0}^{k} \beta_j z^j,$$

$$\beta_{j=} \big(\beta_j^{(0)}, \dots, \beta_j^{(i)}, \dots, \beta_j^{(h)}\big)',$$

$$L^j x_{t,0} := x_{t,j} = \big(L^j x_{tm0}^{(0)}, \dots, L^j x_{tmi}^{(i)}, \dots, L^j x_{tmh}^{(h)}\big)$$

The MIDAS estimation in this research is performed using the R package called "midasr".

Model Evaluation

This research utilizes the residual standard error (RSE) score to evaluate the MIDAS model. RSE is one of the goodness-of-fit tests to appraise whether a model fits a dataset [34]. The RSE score can be measured as follows,

$$RSE = \sqrt{\frac{\sum(Y - Y_{est})^2}{df}} \tag{6}$$

where Y-Y_est is residuals of the model and df is degrees of freedom.

Results and Discussions

Overview of longitudinal patterns of the GDP sector H, the GDP sector I, and Google Trends Index

Figure 1 provides the comparison of the daily and weekly GTI time series for the category "Transportation and Logistics" with the GDP sector H. The daily and weekly GTI time series show similarities to the GDP time series pattern.
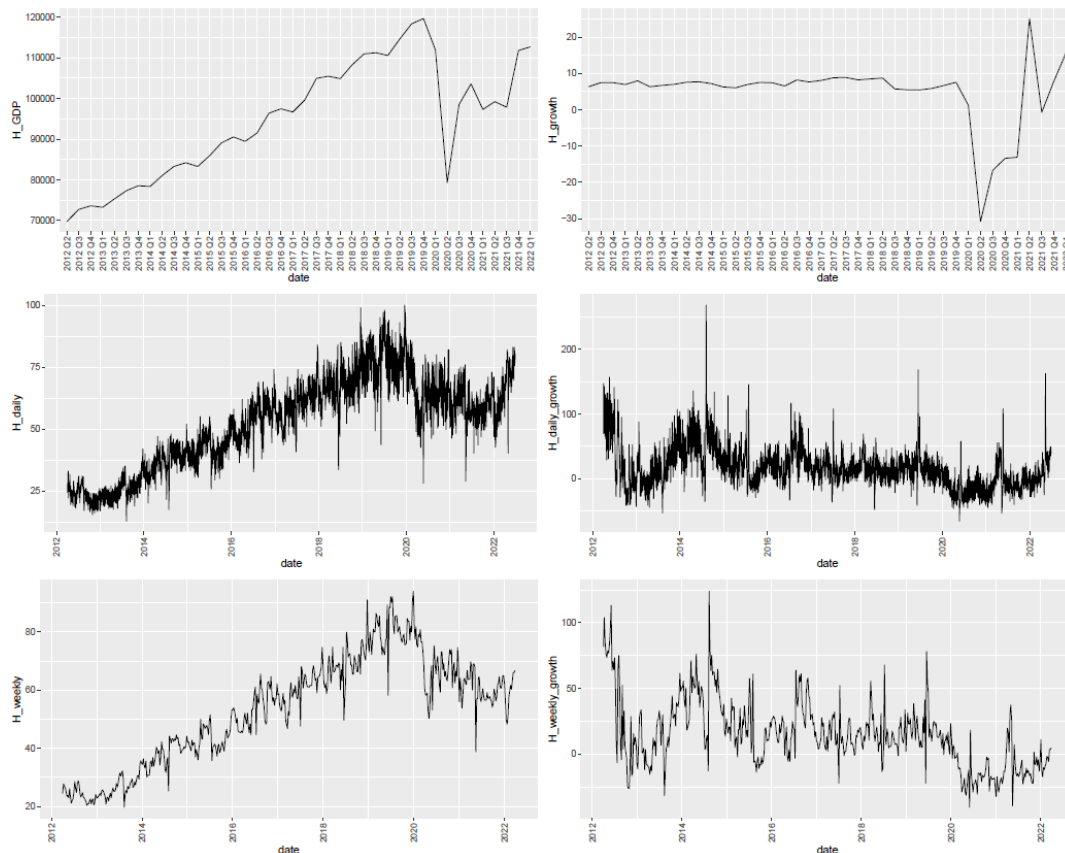

Figure 1. The daily and weekly GTI time series for the "Transportation and Logistics" category from 2012 to 2022 shows an identical pattern to the GDP Sector H for both levels and growth.

The search requests made to the Google search engine regarding transportation and logistics increased from 2012 to early 2020, decreased significantly in mid-2020, showed a relatively stagnant trend until the end of 2021, and later started to increase in 2022. An identical pattern shows in GDP sector H. Looking at the similar trends between the GDP sector H and the GTI category "Transportation and Logistics" indicate the potential of GTI to capture the activities in the transportation sector in Indonesia.

The level shift that occurred in mid-2020 was caused by the Covid-19 outbreak that disrupted most economic activities. It's clearly shown by the growth rates of GDP sector H which is relatively constant but suddenly drop in the second quarter of 2020.

Figure 2 shows the pattern of the GDP sector I that have a similar condition to the pattern of GDP sector H. The search requests made daily and weekly to the Google search engine regarding hotels and accommodation apparently can effectively describe the activities in the accommodation sector.

According to the delineation of Figure 1 and Figure 2, this research believes that the daily and weekly GTI regarding transportation and logistics as well as hotels and accommodations can be useful to nowcast the GDP Sector H and the GDP Sector I respectively.
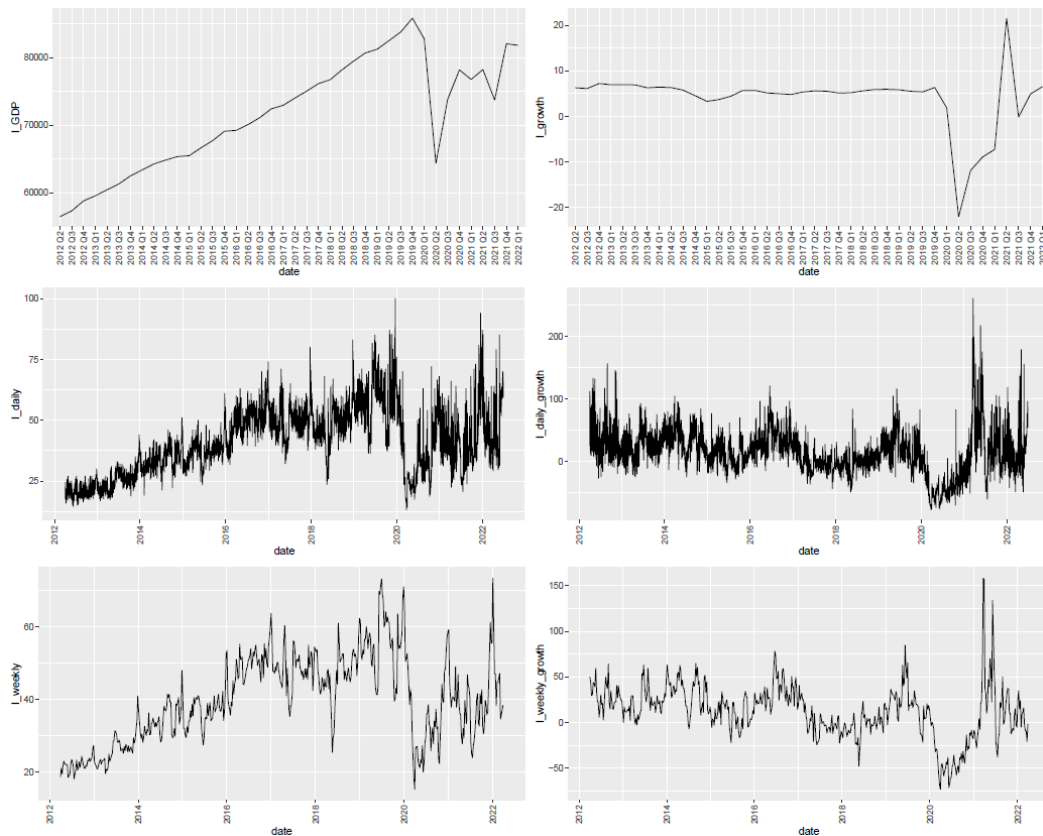

Figure 2. The daily and weekly GTI time series for the "Hotels and Accommodations" category from 2012 to 2022 shows an identical pattern to the GDP Sector I for both levels and growth.

To confirm that all the datasets are eligible to perform MIDAS regression, the datasets should meet the stationarity assumption [35]. Table 2 provides the result of the unit root test for all datasets used in this research. The unit root tests show that all datasets are significantly stationary on alpha 0.05 which is considered a satisfactory result. According to this result, all datasets are appropriate for the models.

Table 2. Result of Stationarity Test

Variables	p-value	
	ADF	Phillip Perron
Growth rates of The GDP sector H	<0.01	<0.01
Growth rates of daily GTI for The GDP sector H	<0.01	<0.01
Growth rates of weekly GTI for The GDP sector H	<0.01	<0.01
Growth rates of The GDP sector I	<0.01	<0.01
Growth rates of daily GTI for The GDP sector I	<0.01	<0.01
Growth rates of weekly GTI for The GDP sector I	0.016	<0.01

This research then performs a cross-correlation function to determine whether the growth rates of the GTI are considered a lagging, coincidence, or leading indicator for the growth rates of the GDP. Because the two datasets have different frequencies, the growth rates of the GTI should be aggregated to the quarterly index therefore the cross-correlation function can be applied. Figure 3 provides the result of the cross-correlation function of the GDP and the GTI.
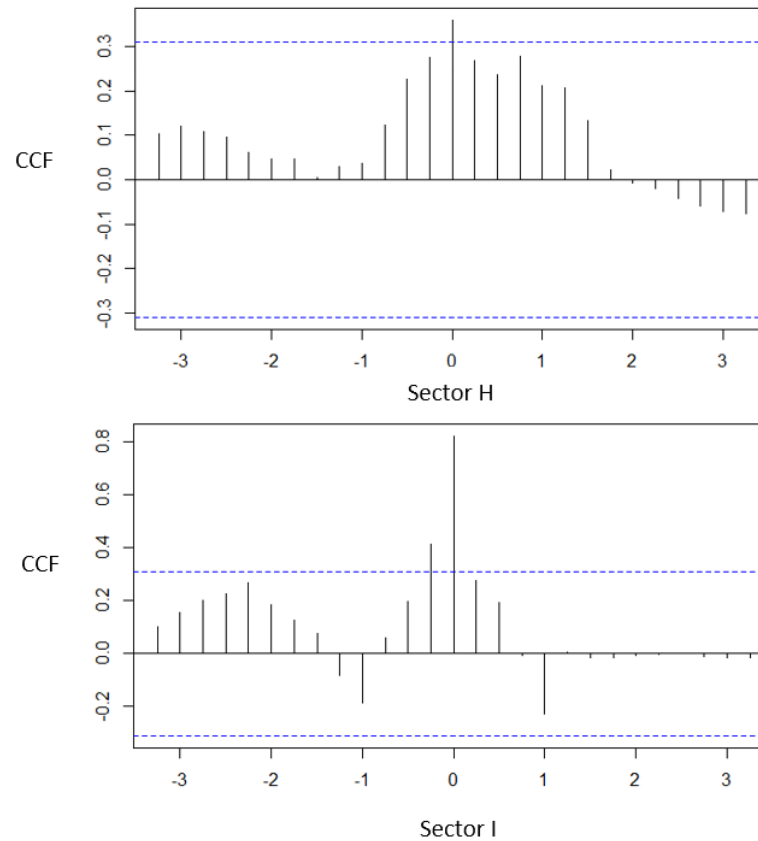

Figure 3. The cross-correlation of the growth rates of the quarterly GDP Sector H and Sector I with the growth rates of the quarterly GTI for "transportation and logistics" and "hotels and accommodations", respectively. The growth rates of the quarterly GTI for both sectors have the highest correlation with the growth rates of the quarterly GDP in lag 0, indicating that the GTI coincides with the GDP.

From Figure 3, this research finds that the growth rates of the quarterly GDP from 2012 to 2022 coincide with the growth rates of the quarterly GTI for the same period. It means that the growth rates of daily and weekly GTI in a quarter will have the highest correlation with the growth rates of the GDP in that quarter if the growth of the quarterly GDP is released on the first day of the next quarter. This is in accordance with the statement of Varian [17] who claims that google trends may be fruitful in predicting the present rather than the future.

MIDAS Regression to nowcast GDP

Based on the cross-correlation result (see Figure 3), this research focuses on the daily and weekly GTI growth in a quarter to determine how many lags effectively describe the GDP growth rates in each corresponding quarter. Therefore, this study performs a U-MIDAS regression for all possible combinations of the daily and weekly GTI lag counts to check which combination that gives the best result. The model with the lowest RSE score is considered the best model.

<div align="center">**Table 3.** Best Model Based on RSE Score</div>

Model	Covid Var	Lag (prior to the last day of a Quarter)	RSE
Sector H – daily GTI	No	Day 7-36	2.29
Sector H – weekly GTI	No	Week 0-12	5.82
Sector H – daily GTI	Yes	Day 6-35	2,27
Sector H – weekly GTI	Yes	Week 0-11	5,91
Sector I – daily GTI	No	Day 57-85	1.98
Sector I – weekly GTI	No	Week 3-11	3.47
Sector I – daily GTI	Yes	Day 57-86	1,28
Sector I – weekly GTI	Yes	Week 3-11	3.40

Table 3 shows the result of the model selection. Considering the change in the GDP trends caused by the Covid-19 outbreak (see Figure 1 and Figure 2), this research also performs simulation by including the Covid dummy variables to explain the shifting in the trends.

Figure 4. Plots of the GDP growth for Sector H and Sector I and their fitted values of MIDAS model using the daily GTI with and without Covid dummy variable. All models fit well with the actual values.

This research finds that the daily GTI growth rate for the "transport and logistics" effectively predicts the GDP growth rate for Sector H if it is in the last month of each corresponding quarter. For instance, on the MIDAS model without Covid dummy variables, the daily GTI for "transportation and logistics" from day 7 to day 36 prior to the last day of a quarter will give the best prediction on the GDP growth for sector H for that quarter. While including the covid dummy variable, the MIDAS model using

the daily GTI from day 6 to day 35 prior to the last day of a quarter will give the best prediction on the GDP growth for that quarter.

On the other hand, according to the RSE score, the best MIDAS model to predict the GDP growth for the sector I use information on the daily GTI growth in the first month of each corresponding quarter. For instance, on the MIDAS model without Covid dummy variables, the daily GTI for "hotels and accommodations" from day 57 to day 85 prior to the last day of a quarter will give the best prediction on the GDP growth for sector I for that quarter. While including the covid dummy variable, the MIDAS model using the daily GTI from day 57 to day 86 prior to the last day of a quarter will give the best prediction on the GDP growth for that quarter.

For the MIDAS model with weekly GTI growth rates, this research finds that most weeks of a quarter effectively predict GDP growth rates for that quarter. The weekly GTI growth rates for "transportations and logistics" from week 11 or 12 to the last week of a quarter effectively predict the GDP growth rates for sector H. While the weekly GTI growth rates for "hotels and accommodations" from week 11 to week 3 prior to the last day of a quarter effectively predict the GDP growth rates for sector I.
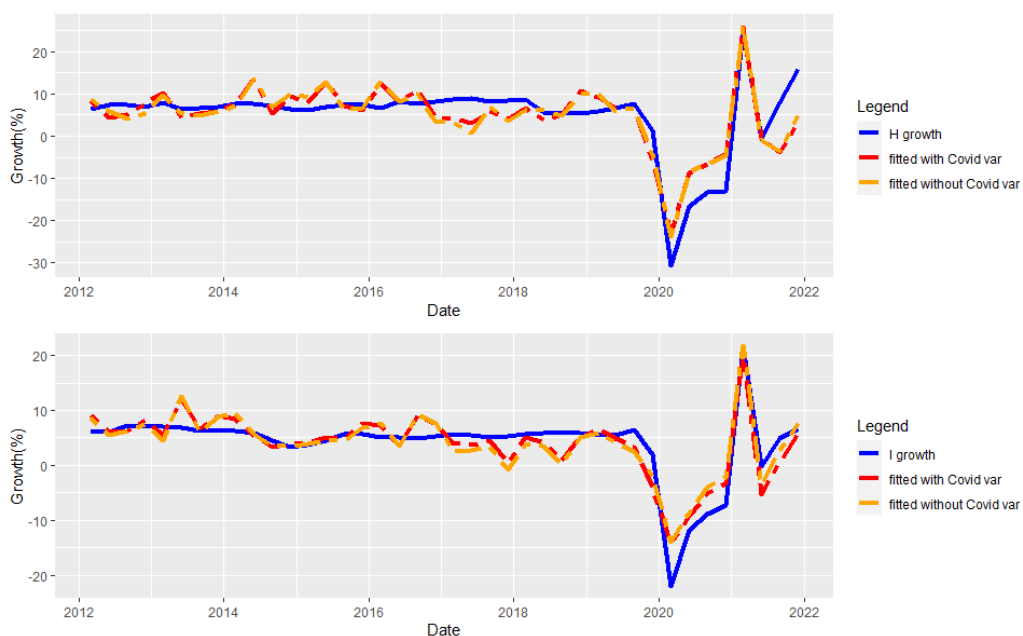

Figure 5. Plots of the GDP growth for Sector H and Sector I and their fitted values of MIDAS model using the weekly GTI with and without Covid dummy variable. All models are not very good at predicting the GDP growth rates.

Figure 4 and Figure 5 present visual descriptions of the goodness of the models compared with the actual values. The Covid dummy variables show an insignificant impact on the prediction for all models (see Table 3).

Discussions

This research finds that the MIDAS model using the daily Google Trends Index gives a better result to predict the quarterly GDP growth. This is evident by the RSE scores for the daily MIDAS model which are lower than the RSE scores of the weekly MIDAS hotel. It strengthens with the plot of fitted values and the actual values (see Figure 4 and Figure 5). Apparently, the statement from Foroni [32]which said that the U-MIDAS model will be more fruitful if applied to datasets with small frequency differences is not applicable to the condition in this research. The provisional presumption is this research gain an advantage from utilizing a dataset that has an identical pattern to the responding dataset (see Figure 1

and Figure 2). However, conducting further exploration by comparing it to the restricted MIDAS model and performing other model evaluations such as monte Carlo simulations and AIC is necessary to obtain factual explanations.

Conclusion

This research has proven the possibility of the use of daily and weekly alternative data to nowcast the GDP. The nowcasting is performed using the U-MIDAS model and the explanatory variable used for the model is the daily and weekly Google Trends Index (GTI). The models provide satisfactory results based on the RSE score. The evidence in this research shows that the U-MIDAS model using the daily GTI is better to predict the current GDP than the U-MIDAS model using the weekly GTI.

Acknowledgement

References

[1]   E. M. Bah, Structural Transformation Paths Across Countries, Emerg. Mark. Finance Trade., 47 (2011) 5–19.

[2]   E. Zahrotul Awaliyyah, S.-E. Chen, and R. Anindita, Analysis of Structural Transformation of Labor from Agriculture to Non-Agriculture in Asia, Agric. Soc. Econ. J., 20(4) (2020) 335–341.

[3]   L. Schlogl and A. Sumner, Rethinking international development series Disrupted Development and the Future of Inequality in the Age of Automation. [Online]. Available: http://www.palgrave.com/gp/series/14501

[4]   UN-Habitat, Structural transformation in developing countries: Cross Regional Analysis, (2016).

[5]   K. Sen, Structural transformation around the world: Patterns and drivers, Asian Dev. Rev., 36(2) (2019) 1–31.

[6]   A. Haryana, Economic and Welfare Impacts of Indonesia's Tourism Sector, Jurnal Perencanaan Pembangunan: The Indonesian Journal of Development Planning, 4(3) (2020) 300–311.

[7]   E. Ghysels, A. Sinko, and R. Valkanov, MIDAS Regressions: Further Results and New Directions, Econom Rev., 26(1) (2007) 53–90.

[8]   F. Barsoum and S. Stankiewicz, Forecasting GDP growth using mixed-frequency models with switching regimes, Int. J. Forecast, 31(1) (2015) 33–50.

[9]   D. Bilgin, D. Jankovic, and A. Lam, MIDAS regression using inflation and unemployment to predict GDP, 2018. Accessed: Aug. 18, 2022. [Online]. Available: http://mathstat.carleton.ca/~smills/2017-18/STAT4601-5703/Research%20Projects/2018%20Submissions/JankovicLamBilgin/Time%20Series.pdf

[10]   T. and S. S. Kingnetr Natthaphat and Tungtrakul, Forecasting GDP Growth in Thailand with Different Leading Indicators Using MIDAS Regression Models, in Robustness in Econometrics, S. and H. V.-N. Kreinovich Vladik and Sriboonchitta, Ed. Cham: Springer International Publishing, (2017) 511–521.

[11]   D. T. Utari and D. Rosadi, Forecasting Indonesian GDP growth using mixed data sampling (MIDAS) regression, 2018.

[12]   T. Nakazawa, Constructing GDP Nowcasting Models Using Alternative Data Constructing GDP Nowcasting Models Using Alternative Data, 2022.

[13]   A. Ghysels, E. Santa-Clara, P. Valkanov, E. Ghysels, P. Santa-Clara, and R. Valkanov, UCLA Recent Work Title The MIDAS Touch: Mixed Data Sampling Regression Models Permalink https://escholarship.org/uc/item/9mf223rs Publication Date The MIDAS Touch: Mixed Data Sampling Regression Models, [Online]. Available: https://escholarship.org/uc/item/9mf223rs

[14]   D. T. Utari and H. Ilma, Comparison of methods for mixed data sampling (MIDAS) regression models to forecast Indonesian GDP using agricultural exports, in AIP Conference Proceedings, 2021(1) (2018).

[15]   Google, FAQ about Google Trends Data - trends help, https://support.google.com/trends/answer/4365533?hl=en&ref_topic=6248052 (accessed Aug. 15, 2022).

[16]   A. Mavragani and G. Ochoa, Google trends in infodemiology and infoveillance: Methodology framework, JMIR Public Health and Surveillance, 5(2) (2019).

[17]   H. Choi and H. A. L. Varian, Predicting the Present with Google Trends, Economic Record, 88 (2012) 2–9.

[18] F. Wijnhoven, Sentiment Analysis and Google Trends Data for Predicting Car Sales Review spam View project Artificial intelligence for business decisions View project, 2017. [Online]. Available: http://aisel.aisnet.org/icis2017/DataScience/Presentations/1.

[19] F. Ahmed and M. Muzammil, Financial Market Prediction using Google Trends, 2017. [Online]. Available: www.ijacsa.thesai.org

[20] N. Woloszko, OECD Economics Department working papers, Tracking activity in real time with Google Trends, 2020. [Online]. Available: https://www.oecd-ilibrary.org/economics/oecd-economics-department-working-papers_18151973

[21] Y. Feng, G. Li, X. Sun, and J. Li, Forecasting the number of inbound tourists with Google Trends," Procedia Comput. Sci., 162 (2019) 628–633.

[22] I. Ayuningtyas, I. Wirawati, Nowcasting tingkat penghunian kamar hotel menggunakan google trends, Seminar Nasional Official Statistic 2020, Politeknik Statistik STIS.

[23] T. Havranek and A. Zeynalov, Forecasting tourist arrivals: Google Trends meets mixed-frequency data, Tour. Econ., 27(1) (2021) 129–148.

[24] P. Massicotte and D. Eddelbuettel, Perform and display google trends queries [R package gtrendsr version 1.4.6], 2022. Accessed: Aug. 15, 2022. [Online]. Available: https://cran.microsoft.com/snapshot/2020-06 20/web/packages/gtrendsR/index.html

[25] P. Trasborg, Google trends categories, 2007. https://github.com/pat310/google-trends-api/wiki/Google-Trends-Categories

[26] F. B, Google trends: How to acquire daily data for broad time frames, https://medium.com/@bewerunge.franz/google-trends-how-to-acquire-daily-data-for-broad-time-frames-b6c6dfe200e6 (accessed Aug. 18, 2022).

[27] D. A. Dickey and W. A. Fuller, Distribution of the Estimators for Autoregressive Time Series with a Unit Root," J. Am. Stat. Assoc., 74(366a) (1979) 427–431.

[28] E. Paparoditis and D. N. Politis, The asymptotic size and power of the augmented Dickey–Fuller test for a unit root, Econom Rev., 37(9) (2018) 955–973.

[29] P. C. B. Philips and P. Perron, Testing for a unit root in time series regression, Biometrika, 75(2) (1988) 335–346.

[30] T. R. Derrick and J. M. Thomas, Time-Series Analysis: The cross-correlation function, in Innovative Analyses of Human Movement, N. Stergiou, Ed. Champaign, Illinois: Human Kinetics Publishers, 2004, 189–205.

[31] A. E. Usoro, Some basic properties of cross-correlation functions of n-dimensional vector time series, J. Stat. Economet. Meth., 4(1) (2015) 63–71.

[32] C. Foroni, M. Marcellino, and C. Schumacher, U-Midas: Midas Regressions with Unrestricted Lag Polynomials, 2012. Accessed: Aug. 18, 2022. [Online]. Available: https://ssrn.com/abstract=2013819

[33] E. Ghysels, V. Kvedaras, and V. Zemlys, Mixed Frequency Data Sampling Regression Models: The R Package midasr, J. Stat. Softw., 72(4) (2016) 1–35.

[34] Zach, How to interpret residual standard error, 2021. https://www.statology.org/how-to-interpret-residual-standard-error/ (accessed Aug. 18, 2022).

[35] K. Karagoz and S. Ergun, Forecasting Monthly Inflation: A MIDAS Regression Application for Turkey, 2020.