RESEARCH ARTICLE

Nowcasting Philippine Economic Growth Using MIDAS Regression

Cesar C. Rufino
De La Salle University, Manila, Philippines
cesar.rufino@dlsu.edu.ph

One of the most anticipated data releases of the Philippine statistical system is the quarterly real gross domestic product. This all-important variable provides the basis of establishing the economic performance of the country on a year-on-year basis. Official publication of this statistic, however, comes at a significant delay of up to two months, upsetting the planning function of various economic stakeholders. Under this backdrop, data scientists coined the term "nowcasting" which refers to the prediction of the present, the very near future, and the very recent past based on information provided by available data that are sampled at higher frequencies (monthly, weekly, daily, etc.). This study aims to demonstrate the viability of using a state-of-the-art technique called MIDAS (Mixed Data Sampling) regression to solve the mixed frequency problem in implementing the nowcasting of the country's economic growth. Different variants of the MIDAS model are estimated using quarterly Real GDP data and monthly data on inflation, industrial production, and Philippine Stock Exchange index. These models are empirically compared against each other and the models traditionally used by forecasters in the context of mixed frequency. The results indicate the relative superiority of the MIDAS framework in accurately predicting the growth trajectory of the economy using information from high-frequency economic indicators.

***Keywords*:** Nowcasting, MIDAS Regression, Mixed Frequency Problem, Temporal Aggregation, Ragged Edge Problem, Bridge Equations

JEL Classification: C13, C52, C53, E17

Nowcasting has been a buzzword in the current economic forecasting literature. It refers to the prediction of the present, the very near future, and the very recent past (Giannone, Reichlin & Small, 2008), which has a lot of decision making and planning implications. Its relevance to economic planning lies in the fact that the most important indicators of economic health (gross domestic product and its components – personal consumption expenditure, gross domestic capital formation, government expenditure, etc.) are sampled and published quarterly with substantial publication delays of even up to two months; thus, upsetting the planning activities of various stakeholders of the economy (the central bank, legislators, fiscal planners, financial and business firms, and others who are immensely affected by the business cycle). On the other hand, many variables sampled at higher frequencies (monthly, weekly, daily, etc.) that are known to carry predictive information on future economic growth (such as industrial production, inflation, monetary aggregates, interest rates, and stock market index) are already available and the useful information they carry can be extracted to the fullest, even before the final quarterly indicators are released. The central objective of the nowcasting research is in developing models and procedures that will make this information extraction process as effective and as reliable as possible.

Relationships of variables in economics, finance, and other fields are traditionally modeled as a form of regression equations or systems of equations, wherein the variables are sampled in the same frequency. When any or all of the regressors is/are in higher frequency, the usual recourse, called temporal aggregation approach, is to temporally aggregate these variables, usually in terms of their sums or averages to conform with the sampling frequency of the regressand; thus, synchronizing the data sampling of the left hand and right hand side variables to that of the lower frequency regressand, making the analysis viable. Although computationally convenient, this recourse of solving the mixed frequency problem does not extract predictable information from the more frequently sampled regressors because of information loss and possible misspecification errors induced by the process of aggregation might compromise the forecast quality.

An alternative option, called the individual coefficient approach, the extraction of hidden information in the higher frequency regressors may be possible if the model is augmented by the individual components of the regressors, each with its own coefficient to be estimated. For example, if the regressand is quarterly and the regressor has m components (that is, m periods in a quarter, m = 3 if the regressor is monthly, m = 66 if the regressor is daily, etc.) of this variable. This will effectively introduce a multiplier for each component, which may be interpreted as the component's marginal contribution to the regressand during the specific quarter. This option is unappealing because of parameter proliferation (with a consequent loss in degrees of freedom), especially if m becomes large. In the temporal aggregation option, the multipliers effectively all equal to 1/m, when the aggregation scheme is averaging.

The MIDAS regression approach represents an intuitively appealing middle ground between the two options discussed above. The MIDAS approach, introduced by Ghysels, Santa-Clara, and Valkanov (2004) and was further developed by Andreou, Ghysels, and Kourtellos (2010), allows for non-equal weights (multipliers) for the components that are parsimoniously reparametrized through a weighing scheme anchored on the innovative use of lag polynomials. The way lag polynomials are employed in defining the weighing scheme for the multiplier represents a specific MIDAS regression model. For a good brief introduction to MIDAS, see Armesto, Engemann, and Owyang (2010).

In this study, MIDAS regression models are estimated and matched against models traditionally used in dealing with the "mixed frequency" problem.

Technical Specifications of the Models

Suppose the mixed frequency model under consideration is given as follows:

$$Y_t^L = \sum_{i=1}^{q} \beta_i W_{t-i}^L + \lambda f\left(\gamma, X_{j,t}^H\right) + \varepsilon_t \qquad (1)$$

where,

Y_t^L is the dependent variable sampled at low frequency;

W_t^L is the set of regressors sampled at the same (low) frequency as the regressand (possibly including lags of *Y* for autoregressive or ARDL form);

$X_{j,t}^H$ is the set of regressors sampled at a higher frequency;

β_i, λ, and γ are the parameters to be estimated; and

f *(.)* is a function translating the higher frequency data into the low frequency *Y.*

The following are the models that will be used in the study, each of which has a different way of translating the higher frequency data into their low frequency form, through their specific choice of *f (.).*

Model 1: Temporal Aggregation

A conventional way to address mixed frequency samples is to use some type of aggregation, perhaps summing or taking the average of the high-frequency data that occur between samples of the lower-frequency variable (Clements and Galvão. 2008). For example, we can take a simple average:

$$\boldsymbol{X_t^L} = \frac{1}{m}\sum_{j=0}^{m-1} X_{t-j}^H \quad (2)$$

and carry out the following regression:

$$Y_t^L = \sum_{i=1}^{q} \beta_i W_{t-i}^L + \boldsymbol{\lambda X_t^L} + u_t \quad (3)$$

where m – is the number of periods in the higher frequency corresponding to a single period in the lower frequency and X_t^H is the high-frequency observation corresponding to the last observation in period t. As mentioned previously, the problem with this is that it assumes that the slope coefficients on each high-frequency observation of X are equal.

Model 2: VAR Forecasting Model

The forecasting capability of vector auto regressive (VAR) models offers another way of using available high-frequency predictors in forecasting low-frequency target variables. This is done by first converting the high-frequency variables into the sampling frequency of the target variable. After which, an unrestricted VAR model (Sims, 1982) is constructed featuring the target variable and the time aggregated predictors, forming the vector. Forecasts are then made out-of-sample for all of the variables in the vector which are all considered endogenous. The focus of interest in this exercise is the forecast for the target variable.

Model 3: Bridge Equation

Another intuitive alternative in using higher-frequency data (e.g., monthly) to forecast lower frequency series (e.g., quarterly) would be to estimate a "bridge equation." This method uses popular forecasting models (such as VARs, ARIMA, and exponential smoothing) for each of the high-frequency indicators. These models are then used to provide forecasts for the missing higher-frequency values (Baffigi, Golinelli, & Parigi, 2004; Diron, 2008). The forecasts are then aggregated to provide estimates of the quarterly values of the regressors of the bridge equation. A bridge equation is nothing but a low-frequency (quarterly) regression with the aggregated (quarterly) forecasts of high-frequency (monthly) regressors. A bridge equation can be written as:

$$Y_t^L = \beta_0 + \sum_{i=1}^{j} \beta_i(L) X_t^L + u_t \quad (4)$$

where X_t^L are the selected high-frequency indicators (forecasted from the high-frequency VAR, for example) aggregated at low frequency. The lag polynomial $\beta_i(L)$ embeds the parameters of the model for each relevant lags of each regressor. Many central banks use the bridge equation model in coming up with advance releases of important statistics (e.g., Bencivelli, Marcellino, & Moretti, 2012; Rünstler & Sédillot, 2003; Zheng & Rossiter, 2006). Ingenito and Trehan (1996) used bridge equations to nowcast the U.S. real GDP based on nonfarm payrolls, industrial production, and real retail sales.

MIDAS (Mixed Data Sampling) Regression

The key feature of MIDAS regression models is the use of a parsimonious and data-driven weighting scheme:

$$Y_t^L = \sum_{i=1}^{q} \beta_i W_{t-i}^L + \lambda \sum_{j=0}^{m-1} w_{t-j}(\gamma) X_{t-j}^H + u_t \quad (5)$$

where *w(.)* is a weighting function that transforms high-frequency data into low-frequency data.

MIDAS estimation offers a number of different weighting functions/schemes which define a specific MIDAS regression model

Model 4: The Almon or PDL MIDAS

MIDAS regression shares some features with distributed lag models. In particular, one parametrization used is the Almon lag weighting (also known as polynomial distributed lag; Almon, 1965), which is widely used in classical distributed lag modeling. The weighting scheme can be written as follows:

$$Y_t^L = \sum_{i=1}^{q} \beta_i W_{t-i}^L + \sum_{i=1}^{p} \gamma_i \sum_{j=0}^{k} j^{i-1} X_{t-j}^H + u_t \quad (6)$$

where k is the chosen number of lags (which may be longer or shorter than m) and p is the order of the polynomial.

Notice that the number of coefficients to be estimated depends on the polynomial order *(p)* and not on the number of lags *(k)* chosen (Ghysels et al., 2004; Ghysels, Sinko, & Valkanov, 2007).

Model 5: Beta Weighting MIDAS

An alternative method is based on the following beta function:

$$Y_t^L = \sum_{i=1}^{q} \beta_i W_{t-i}^L + \lambda \sum_{j=0}^{k} Z_{j,t} + u_t \tag{7}$$

where,

$$Z_{j,t} = \left(\frac{{\psi_j}^{\gamma_1 - 1} (1-\psi_j)^{\gamma_2 - 1}}{\sum_{i=0}^{k} {\psi_i}^{\gamma_1 - 1} (1-\psi_i)^{\gamma_2 - 1}} + \gamma_3 \right) X_{t-j}^H \tag{8}$$

where $\psi_j = \frac{j-1}{k-1}$.

This function involves the estimation of three parameters, but it can be restricted by imposing either: $\gamma_1 = 1$, or $\gamma_3 = 0$, or $\gamma_1 = 1$, and $\gamma_3 = 0$. The number of parameters estimated can, therefore, be 1, 2, or 3 (depending on the types of restrictions we impose). Notice also that with this weighting scheme, the number of parameters do not increase with the number of lags, but the estimation involves a highly non-linear estimation procedure (Ghysels, Rubia, & Valkanov, 2009).

Model 6: Step Weighting MIDAS

Perhaps the simplest weighting scheme is a step function, where the distributed lag pattern is approximated by a number of discrete steps. The step weighting can be written as:

$$Y_t^L = \sum_{i=1}^{q} \beta_i W_{t-i}^L + \sum_{j=0}^{k} \phi_{t-j} X_{t-j}^H + u_t \tag{9}$$

where, $\phi_j = \gamma_k$, k is a number of lags (k may be longer or shorter than m), $k = \frac{j}{\eta}$, and η is the number of steps.

Step-weighing lowers the number of estimated coefficients because it restricts consecutive lags to have the same coefficient (Forsberg & Ghysels 2007). For example, if k=12 and η =4, the first four lags have the same coefficient, the next four lags have the same coefficient, and so on, all the way up to k=12.

Model 7: U-MIDAS

U-MIDAS or unrestricted MIDAS is appropriate if the differences in sampling frequencies are small (say, monthly and quarterly data). When the difference in sampling frequencies between the regressand and the regressors is large, distributed lag functions are typically employed to model dynamics avoiding parameter proliferation. Introduced by Foroni, Marcellino, and Schumacher (2012), U-MIDAS does not depend on any specific functional lag polynomial since in most macroeconomic applications differences in sampling frequencies are often small, usually quarterly-monthly. In such a case, it might not be necessary to employ distributed lag functions, and parameters can be estimated by OLS. Koenig, Dolmas, and Piger (2003) already proposed U-MIDAS in the context of real-time estimation. In essence, the U-MIDAS model can be written as:

$$Y_t^L = \sum_{i=1}^{q} \beta_i W_{t-i}^L + \sum_{j=0}^{m-1} \gamma_{t-j} X_{t-j}^H + u_t \tag{10}$$

where a different slope coefficient for each high-frequency lag is estimated.

Estimating the Models

The empirical counterparts of Models 1 to 7 are constructed as part of the tasks completed in this study. All of the operational models are estimated using Eviews 10 software, which so far is the only commercial software available that supports estimation of MIDAS regression (although there is an R package midasr; see Ghysels, Kvedaras, & Zemlys, 2016; as well as a MATLAB Toolbox, see Ghysels, 2013). All data to be used—quarterly, monthly, and daily statistics—are accessed through Philippine Statistics Authority, Bangko Sentral ng Pilipinas and Philippine Stock Exchange websites. The following variables over the period 2002–2016 comprise the database of the study:

Quarterly (2002q1-2016q4) – Economic Growth (year-on-year continuously compounded growth of Seasonally Adjusted Gross Domestic Product, in real terms) - the regressand, computed as:

$ecogrowth_t = 400 * 1og(rgdp_t / rgdp_{t-1})\%$

Where $rgdp_t$ = Seasonally adjusted real Gross Domestic Product for quarter *t*. Monthly (2002m1 to 2016m12):

- Inflation: $infl_t = 100 * 1og(cpi_t / cpi_{t-1})\%$
- Growth of Industrial Production: ipg_t - $100 * 1og(ip_t / ip_{t-1})\%$

- PSEI Return: $pseig_t = 100*1og(PSEI_t / PSEI_{t-1})\%$
- Interest Rate: IR_t = 91 days T-Bills Return during month t
- Exchange Rate (Peso to US Dollar) Return: $erg_t = 100*1og(er_t / er_{t-1})\%$

The motivation underlying the use of these higher frequency variables as drivers of economic growth in this study is anchored, not only on their ready availability and intuitive appeal but also on implicit theoretical considerations. To gauge the empirical relevance of these variables to economic growth, a number of analytical procedures and statistical tests are implemented, highlighting on the establishment of the presence of long-run cointegrating relationship(s) of the above relevant variables to economic growth. It is difficult to base entirely to a-priori theoretical considerations the choice of the predictors of economic growth because of the asymmetrical nature of the sampling frequencies of the predictors and the target variable that may distort the underlying data generating process.

Results

Preliminary analysis of the quarterly correlation matrix of the variables reveals the potential of some of the temporally aggregated monthly variables as significant growth drivers. As shown in Table 1, inflation, growth of industrial production, and possibly stock index returns (pseig) produce relatively high contemporaneous correlation with economic growth.

In the above matrix, each cell exhibits the sample correlation coefficient between the row variable and the column variable, together with the p-value of the test for zero correlation. The first column is quite revealing as it indicates the variables with significant correlation with economic growth—inflation, industrial production growth, and PSE returns. Despite their asymmetric sampling frequency with economic growth, these variables may be considered as the key explanatory variables for growth as they may be carrying predictive contents.

All of the variables in all regressions are stationary, as evidenced by the results of individual unit root tests shown in Table 2. These results potentially prevent the occurrence of spurious regressions. This conjecture will be empirically validated through the conduct of cointegration assessment using appropriate procedure.

To empirically demonstrate the presence of long-run relationship(s) among the three key variables and economic growth, testing for cointegration is necessary. Since all variables are integrated of order 0 (i.e., *I (0)*), the Johansen cointegration tests cannot be used; instead, I used the Pesaran Bound test for cointegration under the ARDL approach. Table 3 confirms the presence of a long-run equilibrium relationship between economic growth and its postulated determinants.

Table 1. *Correlation Matrix of the Potential Explanatory Variables for Mixed Frequency Regressions of Economic Growth*

Correlation p-value	Econ Growth	FX Returns	Inflation	IP Growth	Interest Rate
FX Returns	0.008518	1.000000			
	0.9485	-----			
Inflation	-0.286533	0.077610	1.000000		
	0.0264	0.5556	-----		
IP Growth	0.283804	0.074094	-0.148502	1.000000	
	0.0280	0.5737	0.2575	-----	
Interest Rate	-0.174246	-0.066392	0.404472	0.022381	1.000000
	0.1830	0.6143	0.0013	0.8652	-----
PSE Returns	0.202960	-0.468602	-0.150008	0.005459	0.071218
	0.0497	0.0002	0.2526	0.9670	0.5887

Table 2. Stationarity and Unit Root Tests of the Key Variables

Test Statistic	Eco. Growth	Inflation	IP Growth	PSE Returns
KPSS *Statistic*[1]	0.208503 (p>0.10)	0.296980 (p>0.10)	0.180490 (p>0.10)	0.058695 (p>0.10)
ADF *Statistic*[2]	-7.09498 (p<0.000)	-3.94489 (p<0.000)	-8.91786 (p<0.000)	-6.01192 (p<0.000)
PP *Statistic*[2]	-7.09148 (p<0.000)	-2.40887 (p<0.000)	-12.01738 (p<0.000)	-6.025076 (p<0.000)
Order of Integration	I(0)	I(0)	I(0)	I(0)

[1]H_0 : Variable is stationary [2]H_0 : Variable has a Unit root

Table 3. *ARDL Bound Test for the Presence of Cointegration*

Bounds Test for Cointegration	Null Hypothesis: No cointegrating relationships exist			
Test Statistic	Value	Significance Level	I(0)	I(1)
F-statistic	21.89180***	10%	2.72	3.77
		5%	3.23	4.35
		2.5%	3.69	4.89
		1%	4.29	5.61

EC = ecogrowth - (-0.0551*infl + 0.2776*ipg + 0.5095*pseig)

Long Run Coefficients

Variable	Coefficient	Std. Error	t-Statistic	p-value
Inflation	-0.055089	0.272754	-0.201973	0.1411
IP Growth	0.277642	0.157034	1.768031	0.0860
PSE Returns	0.509460	0.162470	3.135730	0.0035

***significant at 0.01 level

ARDL Forms of the Models

It is expected that the effects of the predictors to economic growth is not instantaneous. The explanatory contributions of the regressors are manifested in the target variable with a lag; hence, the auto regressive distributed lag (ARDL) is an appropriate specification of the relationship. However, the central problem is in the determination of the optimal lag of all variables in the model. Different lag configurations for the variables constitute different ARDL models from which I are going to select the optimal specification. I adopted the procedure of model selection based on the Akaike information criterion (AIC).

Out of a total of 500 ARDL models evaluated, the top 20 of these models with the smallest AIC scores are shown in Figure 1. The best among them is the ARDL (1, 4, 0, 1)—autoregressive order is 1, and the distributed lag orders for inflation, industrial production growth, and PSE returns are 4, 0, and 1, respectively.

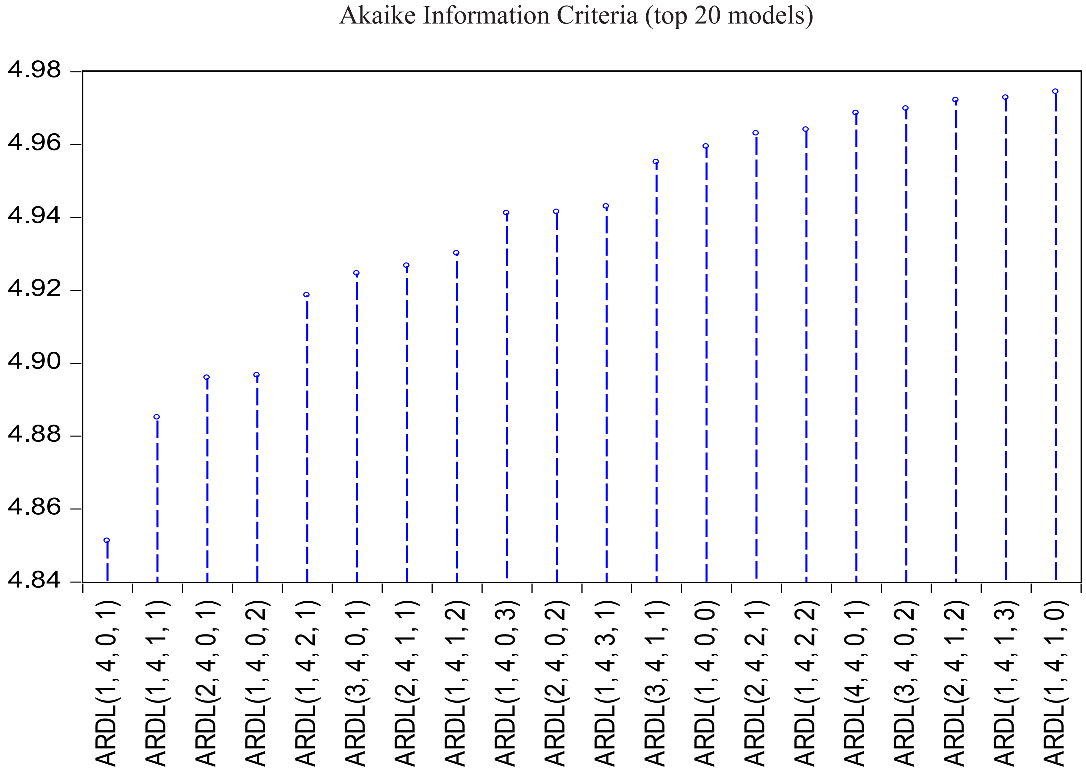


Figure 1. Top 20 ARDL models using the Akaike information criterion.

To implement Model 1 as an ARDL (1, 4, 0, 1) with time aggregated monthly regressors into quarterly frequency, I used a recent inclusion in the Eviews 9.5 suites of commands—the ARDL estimation. The results are presented in Table 4.

Performing the different diagnostic procedures on this model, the following are noted: No residual autocorrelation up to the 4th order, no heteroscedasticity, Ramsey-RESET confirms correct specification, and no structural change. It is important to check for structural change within the sample horizon as its presence will affect the quality of the forecasts. Presented in Table 5 is the result of the Quandt-Andrews Unknown Breakpoint Test for structural change. The test confirms the absence of structural change during the sample period.

Empirical Comparison of the Models Out-of-Sample Forecasting Performance

The different models considered in this study are estimated, tested, and empirically compared as to their capability to effectively track, out of sample, the actual growth data. Presented in Table 6 are the results of this comparison, first, based on each model's ability to encompass the forecasting ability of the other models in the upper panel, and second, their scores on the different evaluation statistics.

The obvious winner in this out-of-sample forecast comparison is a MIDAS model – the step-weighing MIDAS (Model 6). Not only that Model 6 encompasses the other models, but it also obliterated all other competing models in all evaluation criteria, except on the (mean absolute percentage error (MAPE) criterion. Model 2 or the VAR model consistently placed 2nd in all criteria, except MAPE where it ranked 1st. The outstanding performance of the MIDAS model with respect to the RMSE, considered as the benchmark criterion in forecasting, accentuates its superiority as it is the only model with RMSE of less than 2.0. Incidentally, estimation for Model 5, the beta function weighing MIDAS failed to converge.

Summary and Conclusion

As of late, the mixed frequency and rugged-edge problems in economic forecasting and structural analysis have been attracting a considerable following in the literature. This is true most especially among policy makers and planners who are hard pressed in

Table 4. *Estimated ARDL (1, 4, 0, 1)*

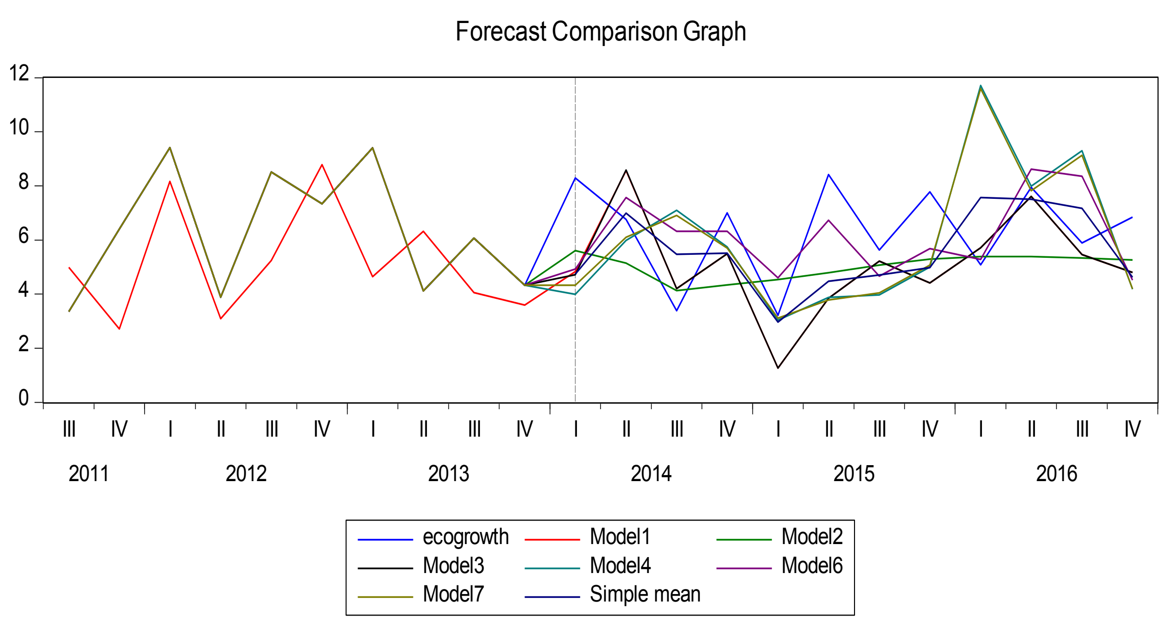
Variable	Coefficient	Std. Error	t-Statistic	Prob.*
ECOGROWTH(-1)	-0.150920	0.141117	-1.069474	0.2924
INFL	-0.306245	0.493605	-0.620424	0.5391
INFL(-1)	1.824895	0.783209	2.330024	0.0259
INFL(-2)	-1.881618	0.888671	-2.117339	0.0416
INFL(-3)	-1.125012	0.916205	-1.227904	0.2279
INFL(-4)	1.424577	0.502055	2.837489	0.0076
IPG	0.319544	0.168774	1.893325	0.0669
PSEIG	0.269582	0.127310	2.117521	0.0416
PSEIG(-1)	0.316767	0.133321	2.375974	0.0233
C	5.380362	1.857334	2.896821	0.0065
R-squared	0.620667	Mean dependent var		5.189676
Adjusted R-squared	0.520256	S.D. dependent var		3.580807
S.E. of regression	2.480194	Akaike info criterion		4.851267
Sum squared resid	209.1463	Schwarz criterion		5.256765
Log likelihood	-96.72788	Hannan-Quinn criter.		5.001645
F-statistic	6.181235	Durbin-Watson stat		2.135071
Prob(F-statistic)	0.000040			
Dependent Variable: ECOGROWTH				
Method: ARDL				
Included observations: 44 after adjustments				
Maximum dependent lags: 4 (Automatic selection)				
Model selection method: Akaike info criterion (AIC)				
Dynamic regressors (4 lags, automatic): INFL IPG PSEIG				
Number of models evaluated: 500				
Selected Model: ARDL(1, 4, 0, 1)				
**Note: p-values and any subsequent tests do not account for model selection.*				

Table 5. *Quandt-Andrews Unknown Breakpoint Test for Structural Change*

Statistic	Value	p-value.
Maximum LR F-statistic (2009Q2)	3.042404	0.1331
Maximum Wald F-statistic (2009Q2)	15.21202	0.1331
Exp LR F-statistic	0.594499	0.4925
Exp Wald F-statistic	5.127762	0.1064
Ave LR F-statistic	1.001204	0.4255
Ave Wald F-statistic	5.006022	0.4255
Null Hypothesis: No breakpoints within 15% trimmed data		
Varying regressors: All equation variables		
Equation Sample: 2002Q4 2013Q4		
Test Sample: 2004Q3 2012Q2		
Number of breaks compared: 32		
Note: probabilities calculated using Hansen's (1997) method		

Table 6. *Forecast Evaluation*

Evaluation sample: 2014Q1 2016Q4						
Encompassing Tests						
Null hypothesis: Forecast of model i includes all information contained in others						
Forecast	**F-stat**	**p-value**				
Model1	2.380380	0.1605				
Model2	1.280371	0.3810				
Model3	2.333357	0.1659				
Model4	10.79025	0.0059				
Model6	1.797065	0.2477				
Model7	9.044966	0.0092				
Evaluation statistics						
Forecast	**RMSE**	**MAE**	**MAPE**	**SMAPE**	**Theil U1**	**Theil U2**
Model1	2.224218	1.777170	27.77141	33.75438	0.185607	0.747321
Model2	2.007592	1.725676	26.10817	29.20138	0.172795	0.699839
Model3	2.239918	1.787617	27.89750	33.94568	0.187012	0.747411
Model4	3.270208	2.662702	45.27516	41.92192	0.250342	0.866209
Model6	1.898047	1.633489	28.71923	27.35359	0.147565	0.437027
Model7	3.183358	2.584310	43.73610	40.59814	0.244464	0.866315
Simple mean	2.171492	1.808920	28.73343	30.26478	0.176487	0.674793



making an updated assessment of the performance of the economy under limited and, at times, missing or incomplete information. Most important data releases related to economic growth are normally done quarterly (e.g., gross domestic product and its components in the national accounts). Moreover, these releases often come with substantial publication delays (which cause the so-called "ragged-edge problem" —missing values for some of the variables, especially at the end of the sample—whereas other equally important statistics are reported more frequently are already available, even before the publication gaps are filled.

These problems of mixed frequency, ragged edge, and asynchronous data availability motivate this study that aimed to demonstrate the viability of using the MIDAS regression modeling—a "state-of-the-art" approach capable of generating "nowcasts" of the country's economic growth. In this study, seven forecasting models, including four variants of the MIDAS model, were estimated, tested, and empirically evaluated for their out-of-sample forecasting performance over a forecast horizon of 12 quarters (2014q1 to 2016q4). Model estimation is over the period 2002q1 to 2013q3 setting aside the remaining available data for out-of-sample forecast evaluation.

The results indicate the outstanding performance of a variant of the MIDAS model, which is the step-weighing MIDAS in practically all of the evaluation criteria. This demonstration led me to conclude the feasibility of using the MIDAS approach in nowcasting the year-on-year quarterly economic growth of the Philippine economy to address the long-standing mixed frequency and ragged edge problems in producing timely and useful forecasts of extremely important indicators.

References

Almon, S. (1965). The distributed lag between capital appropriations and net expenditures. *Econometrica*, *33*, 178–196.

Andreou, E., Ghysels, E., & Kourtellos, A. (2010). Regression models with mixed sampling frequencies. *Journal of Econometrics*, *158*(2), 246–261.

Baffigi, A., Golinelli, R., & Parigi, G. (2004). Bridge models to forecast the Euro area GDP. *International Journal of Forecasting, 20*(3), 447–460.

Bencivelli, L., Marcellino, M., & Moretti, G. L. (2012). *Selecting predictors by Bayesian model averaging in bridge models* (Manuscript). Rome, Italy: Banca d'Italia.

Armesto, M. T., Engemann, K. M., & Owyang, M. T. (2010). Forecasting with mixed frequencies. *Federal Reserve Bank of St Louis Review, 92*(6), 521–536.

Clements, M., & Galvão, A. (2008). Macroeconomic forecasting with mixed-frequency data: Forecasting output growth in the United States. *Journal of Business & Economic Statistics, 26*, 546–554.

Diron, M. (2008). Short-term forecasts of euro area real GDP growth: An assessment of real time performance based on vintage data. *Journal of Forecasting, 27*(5), 371–390.

Forsberg, L., & Ghysels, E. (2007). Why do absolute returns predict volatility so well? *Journal of Financial Econometrics*, *5*(1), 31–67.

Giannone, D., Reichlin, L., & Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, *55*(4), 665–676.

Foroni, C., Marcellino, M., & Schumacher, C. (2012). *U-MIDAS: MIDAS regressions with unrestricted lag polynomials* (CEPR Discussion Paper No. 8828). London : CEPR.

Ghysels, E. (2013). MATLAB toolbox for mixed sampling frequency data analysis using MIDAS regression models. Retrieved from http://www.mathworks.com/matlabcentral/fileexchange/45150-midas-regressionGhysels, E., Rubia, A., & Valkanov, R. (2009). Multi-period forecasts of volatility: Direct, iterated and mixed-data approaches (Working Paper 236). Chapel Hill, University of North Carolina.

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2004). *The MIDAS touch: Mixed data sampling regression models* (CIRANO Working Papers 2004s-20). Montreal, Canada: CIRANO.

Ghysels, E., Sinko, A., & Valkanov, R. (2007). MIDAS regressions: Further results and new directions. *Econometric Reviews, 26*(1), 53–90.

Ghysels, E., Kvedaras, V., Zemlys, V. (2016). Mixed frequency data sampling regression models: The R package midasr. Retrieved from https://CRAN.R-project.org/package=midasrGhysels, E., and Wright, J. (2006), "Forecasting Professional Forecasters," Journal of Business and Economic Statistics 27(4): 504-516

Ingenito, R., & Trehan, B. (1996). Using monthly data to predict quarterly output. Federal *Reserve Bank of San Francisco Economic Review, 3*, 3–11.

Koenig, E., Dolmas, S., & Piger, J. (2003). The use and abuse of real-time data in economic forecasting. *Review of Economics and Statistics, 85*(3), 618–628.

Rünstler, G., & Sédillot, F. (2003). *Short-term estimates of Euro area real GDP by means of monthly data* (Working Paper 276). Frankfurt, Germany: European Central Bank, Sims, C. (1982). Macroeconomics and reality. *Econometrica, 48*(1), 1–48.

Zheng, I., & Rossiter, J. (2006). *Using monthly indicators to predict quarterly GDP* (Working Paper 2006-26). Ottawa, ON: Bank of Canada.