FUNDAÇÃO GETULIO VARGAS
ESCOLA DE ECONOMIA DE SÃO PAULO

JULIA LADEIRA FERREIRA

TRANSMUTING UNEQUALLY SPACED DATA:
A MIDAS REGRESSION TOUCH TO FORECAST REAL GDP GROWTH IN BRAZIL

SÃO PAULO
2020

JULIA LADEIRA FERREIRA

TRANSMUTING UNEQUALLY SPACED DATA:

A MIDAS REGRESSION TOUCH TO FORECAST REAL GDP GROWTH IN BRAZIL

Dissertação apresentada à Escola de Economia de São Paulo da Fundação Getulio Vargas, como requisito para obtenção do título de Mestre em Economia

Área de Concentração: Macroeconomia Financeira

Orientador: Prof. Dr. Pedro Luiz Valls Pereira

SÃO PAULO

2020

Ferreira, Julia Ladeira.

Transmuting unequally spaced data : a MIDAS regression touch to forecast real GDP growth in Brazil / Julia Ladeira Ferreira. - 2020.

43 f.

Orientador: Pedro L. Valls Pereira.

Dissertação (mestrado profissional MPFE) – Fundação Getulio Vargas, Escola de Economia de São Paulo.

1. Produto interno bruto - Brasil. 2. Previsão econômica. 3. Modelos econométricos. 4. Macroeconomia. I. Pereira, Pedro L. Valls. II. Dissertação (mestrado profissional MPFE) – Escola de Economia de São Paulo. III. Fundação Getulio Vargas. IV. Título.

CDU 330.55(81)

Ficha Catalográfica elaborada por: Isabele Oliveira dos Santos Garcia CRB SP-010191/O

Biblioteca Karl A. Boedecker da Fundação Getulio Vargas - SP

JULIA LADEIRA FERREIRA

TRANSMUTING UNEQUALLY SPACED DATA:

A MIDAS REGRESSION TOUCH TO FORECAST REAL GDP GROWTH IN BRAZIL

Dissertação apresentada à Escola de Economia de São Paulo da Fundação Getulio Vargas, como requisito para obtenção do título de Mestre em Economia

Área de Concentração: Macroeconomia Financeira

Data de Aprovação: 16/12/2020

Banca Examinadora:

Prof. Dr. Pedro Luiz Valls Pereira
(Orientador)
FGV-EESP

Prof. Dr. Emerson Fernandes Marçal
FGV-EESP

Prof. Dr. Diogo de Prince Mendonça
UNIFESP

To my grandpa's support during my education

To my grandma's bliss during my childhood

ACKNOWLEDGEMENTS

First and foremost, I express my gratitude and admiration for my advisor Prof. Dr. Pedro Luiz Valls Pereira of the Escola de Economia de São Paulo at Fundação Getúlio Vargas. In the absence of his supervision and patience, this research would not have been accomplished.

Likewise, I am beyond thankful for Prof. Paulo Ney of Berkeley University support and guidance during all my MSc program in an infinitude of ways.

This work reflects a myriad of influences beyond the MSc two year time span. I am particularly grateful for the foundations received at the undergraduate program in Economics at PUC-Rio. Surprisingly, I identified that precious debugging lessons, applied in my R script, came from the same institution as I observed my mom and her fellows coding during their graduate program. Besides, I am grateful for Laurie and Roger Anderson's kindness to shape my English in the 90s and how, after 25 years, they still do it.

Beyond all, I especially recognize my husband's extreme dedication to our family. Without his everyday commitment, I wouldn't have enjoyed the precious time needed for research. Moreover, I particularly appreciate his unfaltering encouragement as we were crossing north to Michigan in the pandemic outbreak. Equally, I heartfully value our kid's resilience and joyfulness during this journey.

Finally, I appreciate the example set by my parents that learning and inventiveness are a continuous and unbounded path.

“In God we trust, all others must bring data.”

W. Edwards Deming

ABSTRACT

Unequally spaced data poses a dilemma on how to aggregate high-frequency variables to model a low-frequency variable. To tackle this quandary, this work proposes to apply MI(xed) DA(ta) S(ampling) (MIDAS), which allows the independent and dependent variables to be sampled at various and different frequencies, to forecast the real GDP growth in Brazil using macroeconomic data. The results show that the restricted polynomial MIDAS specification can outperform the AR(1) for out of the sample recursively estimated nowcasts. Moreover, IBC-BR restricted lag polynomial based MIDAS showcase the best performance under all the computed metrics for evaluation. Not only did the restricted IBC-Br MIDAS outperform the benchmark, but it also beat the U-MIDAS. Fortuitously, the cumulative MSE ratio revealed that between 2014Q3 until the end of 2015, the quotient for the monetary base MIDAS model continuously declined. While this behavior might not be related to the "fiscal pedaling", its trend contributes to the economic policy narrative during those years.

KEYWORDS: MIDAS. Economic Forecasting. Econometric Models. GDP. Macroeconomics.

RESUMO

Dados espaçados desigualmente impõem um dilema sobre como agregar variáveis de alta frequência. Este trabalho propõe a aplicação de MI(xed) DA(ta) S(ampling) (MIDAS), que permite modelar variáveis independentes e dependentes com diferentes frequências. Esse trabalho utiliza essa abordagem para prever o crescimento real do PIB no Brasil com séries macroeconômicos. Os resultados mostram que é possível superar a acurácia das previsões fora da amostra do AR(1) com a especificação polinomial recursivamente estimada. Dentre todos os regressores, o IBC-Br apresentou a melhor performance. O modelo com IBC-Br não apenas ultrapassou o desempenho do benchmark, mas também apresentou uma performance melhor do que o U-MIDAS. Por fim, o índice MSE acumulado revelou que, entre 2014Q3 e o final de 2015, o quociente para o modelo MIDAS da base monetária declinou continuamente. Embora esse comportamento possa não estar relacionado à "pedalada fiscal", sua tendência contribui para a narrativa da política econômica durante esses anos.

PALAVRAS CHAVE: MIDAS. Produto interno bruto - Brasil. Previsão econômica. Modelos econométricos. Macroeconomia.

LIST OF FIGURES

LIST OF TABLES

SUMMARY

1 INTRODUCTION

Unequally spaced data engenders a dilemma on how to aggregate high-frequency variables to model a low-frequency variable. As an example, gross domestic product is measured in quarters while most of the activity indicators are released monthly. Aggregating the high-frequency to obtain a balanced dataset would solve this puzzle, but yielding the cost of precious loss of information. This problem can be especially challenging for researchers trying to forecast economic growth in Brazil since the country does not publish as many indicators of activity as in other OECD countries. In confront with fewer activity statistics released, any leak of content might compromise the forecast precision. Furthermore, since a flat-weighting aggregation conveys a convolution of the relationships among the variables, as remarked by Marcelino (1999), a more engineered method might be conveyed.

Other than simply aggregating the data, Foroni and Marcellino (2013) survey the approaches to deal with unbalanced spaced data, including bridge equations, MIDAS, and MF-VAR. Notably, the secular stagnation faced by the Brazilian economy increases the importance to identify a turning point for both policy-makers and investors. This requires the continuing development of forecast and nowcast econometric methods. Remarkably, the promising results obtained by MIDAS forecast by different researches suggest that this mixed regression might also compliment policy-makers and investors toolkit.

Mixed Data Sampling framework, introduced by Ghysels, Santa-Clara, and Valkanov (2004) and Ghysels, Sinko, and Valkanov (2007), allows the independent and dependent variables to be sampled at various and different frequencies. In this way, quarterly real GDP growth can be modeled by monthly indicators of activity and daily financial asset prices. The inclusion of a distributed lag polynomial enables many lags orders to be included by few parameters. This release of the degrees of freedom from the parsimonious model is exceptionally accommodating when dealing with shorter time series. Every researcher familiar with Brazilian statistics is well aware that dealing with local data poses this additional challenge, especially in quarterly released data. Although any scheme of weightings can be designed for each regressor, the prevalent parametric MIDAS relies on an Almon or Beta lag distribution. Furthermore, this mixed frequency model can be combined with classical econometric approaches, embedding considerably high modeling flexibility and providing a family of MIDAS estimators. To this extent, MIDAS regression can model complicated dynamic relationships in a versatile method without parameters proliferation.

In particular, this research proposes to analyze the forecast results obtained for quarterly seasonally adjusted real GDP growth in Brazil with a MIDAS regression versus an autoregressive model of order 1 using a set of monthly indicators. This comparison relies on the mean-squared errors (MSE) ratio between the two models and on the Diebold and Mariano test. Besides, this paper computes the cumulative MSE fraction to uncloak the possible presence of an outlier pattern during 2014Q2 and 2019Q3. Moreover, this empirical application is built recursively and simulates a real-time GDP nowcast. Although Zuanazzi and Ziegelmann (2014) work had already applied MIDAS framework to forecast Brazilian GDP, their research focused on financial asset prices, and it did not include relevant macroeconomic indicators such as IBC-Br (Brazilian monthly GDP released by Brazilian Central Bank), industrial capacity, retail sales, and monetary aggregates.

The main result of this research is that the restricted MIDAS specification using a normalized exponential Almon lag polynomial carries the potential to outperform the autoregressive model of order 1. Additionally, IBC-Br, as an independent variable, generates the single best nowcast for quarterly seasonally adjusted real GDP growth in Brazil. Besides, for the best performance regressor, the restricted MIDAS beats U-MIDAS, indicating that, even for macroeconomic applications with a small difference between the dependant and independent frequencies, the polynomial inclusion might enhance the nowcast performance. As stated by Foroni, Marcellino, and Schumacher (2015), the choice between the two frameworks requires an empirical approach and might even vary depending on the data set.

This dissertation's findings relate to the results in the MIDAS effervescent literature in three aspects. First, it supports that empirical MIDAS-based implementation detains the ability to beat linear time series forecasts and enhance the predictability of the GDP forecast. Second, it confirms the relevance of the election of regressors and the predominance of statistical facts for the MIDAS framework, as emphasized by Ghyseld et al. (2018). Third, it validates the relevance of the monthly GDP index computed by Central Banks to estimate economic growth. While in Lebouef and Morel (2014), Eurocoin was one of the best indicators to predict the economic performance for the euro-area, this dissertation suggests that IBC-Br yields the best performance nowcasts for the Brazilian economy.

Finally, the cumulative MSE ratio disclosed an incidental finding. Between 2014Q3, until the end of 2015, during the "fiscal pedaling"maneuvers, the cumulative MSE ratio for the monetary base continuously declined. Whereas this trend might not be intrinsically affiliated with the budgetary machinations, it conveys a sturdy historical narrative for the economic policies implemented during those years and contains implications for future research in macroeconomics.

The remainder of this dissertation is crafted as follows. Section 2. describes the Mixed Data Sampling framework and some of its main literature empirical applications in macroeconomics. Section 3. presents the MIDAS specifications, the macroeconomic dataset, and the metrics used to evaluate the models. Section 4. provides the empirical results on the nowcasting for quarterly GDP growth in Brazil with a set of monthly regressors. The final part recapitulates the findings and concludes.

2 LITERATURE REVIEW

2.1 MIDAS Framework

MIDAS is a novel econometric method introduced by Ghysels, Santa-Clara, and Valkanov (2004) and further developments in Ghysels, Sinko, and Valkanov (2007). Their methodology allows the independent and dependent variables to be sampled at various and different frequencies without extensive parameter proliferation by the inclusion of a distributed lag polynomial. Considering the Mixed Data Sampling and the distributed lag regressions resemblance, Ghysels, Santa-Clara, and Valkanov (2004), derived the MIDAS estimator properties in comparison with the latter.

Ghysels et al. (2004) show that MIDAS consistency is similar to lag regression's and that finer sampling of data eventually eliminates discretization bias. Additionally, MIDAS is always more efficient than flat-weighted aggregation. Accordingly, as argued by Ghysels and Marcellino (2018) and Marcelino (1999), temporal aggregation alters the kinetics of data, which compromises the dynamics of the generated model and key econometric features. Furthermore, under certain circumstances, Ghysels et al. (2004) show that its asymptotic efficiency is comparable to lag regressions efficiency.

Ghysels et al. (2007) present several finite and infinite polynomial specifications. The parameterization is one of the central MIDAS components as it weights the dependant variable and selects the lag length included in the model. Notably, the flexibility of this MIDAS feature enables practically any shape of weighting scheme. Moreover, the total lags, which is merely data-driven, does not influence the number of estimated coefficients. Thereby, it would be unjustified to use an information criteria as applied for ARMA or ARDL models to select the lag order as the penalty function will remain constant Ghysels et al. (2018).

Ghysels, Rubia, and Valkanov (2009) complement Ghysels et al. (2007) discussion by including linear, hyperbolic, and geometric schemes as an option to the parameterization. Besides these popular schemes, Ghysels et al. (2007) and Ghysels et al. (2018) also discuss U-MIDAS and MIDAS with step functions features. Although this econometric method accepts a variety of polynomials specifications, the prevalent parametric applications, as described in the next subsection, rely mostly on an Almon or Beta lag distribution.

2.2 MIDAS Applications

Whilst Ghysels (2004) method envisioned its primer use in finance, (risk-return tradeoff and volatility prediction), recent literature suggests that this regression also provides a powerful technique to forecast and nowcast macroeconomic data (GDP, inflation, and fiscal data). As an example, Clements and Galvão (2008) use MIDAS-AR to predict real output growth in the US. Their results suggest that MIDAS-AR always outperforms MF-DL and is preferred to ADL-F when the horizon is not an integer multiple of quarters.

Similarly, Marcellino and Schumacher (2010) infer, for GDP in German, that factor-MIDAS performs better than factor models if nowcasting is not quarterly. In the same direction, Zuanazzi and Ziegelmann (2014) reason that, for GDP in Brazil, MIDAS and UMIDAS yields better performance than ARMA, especially when inside the quarter. These first three articles derive comparable results: when information for the quarter is missing, MIDAS achieves better results than classical econometric models.

Foroni, Marcellino, and Schumacher (2015) compared the results for GDP nowcasting in the US and Euro area with MIDAS and U-MIDAS and proposed that the choice between the two frameworks might require an empirical approach. Likewise, Kuzin, Marcellino, and Schumacher (2011) theorize that, for nowcasting GDP in the Euro Area, the choice between MIDAS and MF-VAR depends on the time-horizon. When incorporating daily financial assets and factors, Andreou, Ghysels, and Kourtellos (2013) identify that US GDP forecast combinations using FADL-MIDAS regression beat traditional models and benchmarks (RW, AR, FAR, ADL, and FADL).

Although most of the literature is concentrated in GDP, recent works input that MIDAS framework potentially improves monthly inflation prediction accuracy when using daily generated data. Monteforte and Moretti (2013) identify that mixed frequency regression using daily financial indicators diminishes real-time inflation forecast errors in the Euro Area when compared to VAR predictions. Breitung and Roling (2015) describe that the inclusion of daily commodities price index can improve the inflation forecast in German, especially with a non-parametric MIDAS. Finally, Li, Shang, and Wang (2015) suggest that MIDAS framework outperforms ARIMA when daily google search data is used to forecast inflation in China.

For fiscal data, although fewer studies are available, they also indicate that MIDAS has the potential to improve forecasts. Ghysels and Ozkan (2015) suggest that combining forecasts of ADL-MIDAS constructed with single predictors outperforms AR and combinations of ADL for fiscal receipts and expenditures. Paredes, Pedregal, and Pérez (2014) conclude that MIDAS beats U-MIDAS and flat-weighted forecasts for USA budgetary data.

Finally, through Monte Carlo simulations, Foroni, Marcellino, and Schumacher (2011) illustrate that when sample differences between the dependent and independent variable is small, the U-MIDAS performance is similar to MIDAS. Ghysels et al. (2018) argue that sampling difference is often lower in macroeconomics variables than in finance. These conjectures ratify the empirical approach of estimating both MIDAS and U-MIDAS in macroeconomic applications.

2.3 MIDAS: the Role of Variables

Ghysels et al. (2018) highlight that the selection of variables is particularly relevant and might vary over time. Moreover, statistical facts might dominate since MIDAS is not a standard macroeconomic model. As an example, Marcellino and Schumacher's (2010) regressions evidence the preeminent role of survey data and industrial production in predicting German GDP under Factor-MIDAS. Accordingly, Kuzin et al. (2009) show that survey data and industrial price index outperform a selection of 23 variables to forecast GDP in the Euro Area with AR-MIDAS and MF-VAR.

Similarly, survey data, industrial production, and the Eurocoin display the best performance among 72 indicators when forecasting GDP in the Euro area using U-MIDAS in Lebouef and Morel (2014). For Japan, the same authors highlight that statistics on consumption perform better than 74 other variables. Kuzin et al. (2009), Marcellino and Schumacher (2010) and Lebouf et Morel (2014) results reinforce the role of survey data in predicting short-term real GDP as highlighted in Godbout and Jacob (2010) for the Euro Area, Japan, United Kingdom, China, and the World Economy.

When predicting real GDP in Brasil using MIDAS and U-MIDAS, Zuanazzi and Ziegelmann (2014) identify that Ibovespa, industrial production, and exports are the best variables among 16 indicators. Those variables used by Zuanazzi and Ziegelmann (2014) do not comprise survey data, consumption, and IBC-Br (which closely relates to the Eurocoin). Finally, when predicting US GDP under FADL-MIDAS, Andreou et al. (2013) confirm the power of the financial asset to predict output as pointed by Stock and Watson (2003).

3 PROPOSED METHODOLOGY AND DATABASE

3.1 Proposed Methodology: Restricted MIDAS

This work is a straightforward application of MIDAS regression introduced by Ghysels et al. (2004) and further developments in Ghysels et al. (2007). Moreover, the forecast of the real quarterly GDP growth in Brazil, Y_{t_q}, using monthly activity and financial indicators, X_{t_m}, will rely on their shrewd specification. In the trivial matrix form, MIDAS is specified as:

$$Y_t = \beta_0 + \beta_1 B\left(L^{1/m};\theta\right) X_t^{(m)} + \varepsilon_t \tag{1}$$

for $t = 1, \ldots, T$, indexing the low-frequency regressand and error ε_t, and for $t_m = 1, 2, 3, \ldots, T_m$ indexing the high-frequency regressors with $T_m \geq T_m = 3T$. Since GDP is quarterly and the explanatory variables are monthly, $m = 3$ and $T_m = 3T$ if all the monthly indicators are available for the last trimester.

$B\left(L^{1/m}, \theta\right)$ represents the polynomial identified by $\sum_{k=0}^{K} b(k;\theta)L^{k/m}$ in which $L^{1/m}X_t^{(m)} = X_{t-k/m}^{(m)}$ is the monthly lag operator which generates a multi-dimensional vector for each regressor i, such as:

$$X^i = \begin{bmatrix} x^i_{(k+1)/m} & \cdots & x^i_{1/m} \\ \vdots & \vdots & \vdots \\ x^i_n & \cdots & x^i_{n-k/m} \end{bmatrix}$$

This vector is weighted by $b(k;\theta)$ which is a function of a small-dimensional vector of parameters θ. Without it, the model would face an undesired proliferation of parameters. Moreover, it selects the lag-order, subjected to the underlying data. Similarly to the distributed lag literature, the proposed parsimonious polynomial specifications to weight the lags is the exponential Almon, as its properties mitigate the undesired multicollinearity.

$$b(k, \boldsymbol{\theta}) = \frac{\exp\left(\theta_1 k + \theta_2 k^2\right)}{\sum_{k=0}^{K} \exp\left(\theta_1 k + \theta_2 k^2\right)} \tag{2}$$

In an Almon lag exponential polynomial specification, the data determines the shape of the lag scheme. As an example, when $\theta_1 = \theta_2 = 0$, this specification coincides with a flat-weighing aggregation. Moreover, a faster decline of the regressor's weights will reduce the number of lags included in the model. Although, the exponential lag function $B\left(L^{1/m}, \theta\right)$ allows for great flexibility in a single-equation approach, it requires a non-linear least square estimation.

Without the inclusion of a polynomial $B\left(L^{1/m}, \theta\right)$, the data would either have to be aggregate a priori, or it would lead to overparameterization. The first extreme outcome is an undesirable loss of information as unbalanced datasets are traditionally circumvented by reducing its frequency with flat-weighing aggregation. This would be similar to impose ad-hoc restrictions to the $B\left(L^{1/m}, \theta\right)$ polynomial before the model estimation. In the absence of both the polynomial and data aggregation, every $X^{(m)}_{t-k/m}$ lag included in the model would require an extra parameter.

Therefore, the $B\left(L^{1/m}, \theta\right)$ polynomial solves the dilemma between including more information and estimating an additional coefficient. The parameterization enables large K to approximate the impulse response function. No matter the size of K, the number of estimated coefficient is always 3 for each regressor. This is particularly accommodating if many lags are required or if the difference between the high and low-frequency data is substantial. For macroeconomic applications that do not include financial data, the lags are usually the restricting feature as the published frequency is often either monthly or quarterly.

3.2 Benchmarks: AR(1) and U-MIDAS

Although the restricted MIDAS offers a compelling framework, only an empirical approach can determine if its application enhances forecast accuracy. In that manner, the primary purpose of this study is to evaluate if the restricted MIDAS specification outperforms the AR(1) when used to estimate the GDP growth in Brazil. Massimiliano Marcellino (2005) argues that, even for sophisticated econometric models for GDP growth, the linear time series models is still a hard benchmark to be beaten. From this standpoint, the chosen reference was the AR(1). This study employed midasr package in R studio for the estimations.

Additionally, to evaluate if the restrictions imposed by $B\left(L^{1/m},\theta\right)$ polynomial were able to capture the underlying data generating process, it is insightful to compare MIDAS with a model without constraints on the weights. Foroni, Marcellino, and Schumacher (2011) denoted this model as the unrestricted MIDAS (U-MIDAS) as follows:

$$Y_t = \beta_0 + \sum_{k=0}^{K} \beta_{k+1} L^{\frac{k}{m}} X_t^{(m)} + \varepsilon_t \tag{3}$$

The above specification is analogous to equation (1) except that it replaces the $B\left(L^{1/m},\theta\right)$ polynomial for K parameters. This substitution enables the estimation via ordinary least squares, which is not feasible in MIDAS. The disadvantage is that, except when $K \leq 3$, U-MIDAS requires the estimation of more parameters than the restricted framework. This comparison provides discerning guidance about the contribution of the $B\left(L^{1/m},\theta\right)$ polynomial in solving the dilemma between more information or an additional coefficient.

3.3 Empirical MIDAS

The database comprises the quarterly GDP index used as the low-frequency dependant variable and 18 monthly time series used as the high-frequency regressors over the sample period from 2003M1 until 2019M9. The regressors comprise industry statistics, consumption indicators, and monetary aggregates downloaded from the Central Bank of Brazil and Instituto Brasileiro de Geografia e Estatística (IBGE) in 2020M107 as detailed in the appendix. All nineteen series employed are originally published with seasonal adjustment.

To the level log series, the standard Augmented Dickey-Fuller (ADF) and Kwiatkowski – Phillips – Schmidt – Shintest (KPSS) were performed. The results indicate that the series are non-stationary. As the ADF is a left-tail test, the positive results found in the equation with no intercept and trend are discarded, and the unit root null hypothesis not rejected. To obtain stationarity, the log difference was applied to all series values. On the transformed series, the tests were repeated. Table 1 summarizes the statistics for both tests.

The first training set comprises eleven years ending in 2014Q1. The training set is recursively expanded, and both MIDAS regressions and the autoregressive reestimated

Table 1 – Augmented Dickey-Fuller and KPSS Test

Code		ADF none	ADF intercept	ADF intercept trend	KPSS
PIM g	level	NA	-2.3615	-2.3313	
	dif	-16.3299***	-16.3037***	-16.4633***	0.0347
PIM e	level	NA	-2.6177*	-3.5038**	
	dif	-8.8057***	-8.8388***	-8.97***	0.03581
PIM t	level	NA	-2.1206	-2.2888	
	dif	-17.0553***	-17.023***	-17.1675***	0.0537
PIM bk	level	NA	-2.3762	-2.2629	
	dif	-19.6377***	-19.6174***	-19.7548***	0.0353
PIM int	level	NA	-2.0289	-2.3405	
	dif	-14.238***	-14.2022***	-14.2921***	0.0313
PIM bc	level	NA	-2.8061*	-2.4945	
	dif	-19.4257***	-19.4227***	-19.5522***	0.081
PIM bcd	level	NA	-3.1449**	-3.0995	
	dif	-16.924***	-16.9031***	-16.9386***	0.0814
PIM bcnd	level	NA	-2.8866**	-2.4761	
	dif	-21.0511***	-21.0443***	-21.1686***	0.1001
PMC	level	NA	-3.3139**	-0.4514	
	dif	-3.2262***	-14.9721***	-16.0511***	0.1469**
PMC a	level	NA	-2.5601	-1.1964	
	dif	-17.9187***	-18.5224***	-18.8543***	0.1156
PMC a comb	level	NA	-1.2469	-0.8235	
	dif	-14.0824***	-14.0467***	-14.0746***	0.1427*
PMC a alim	level	NA	-2.4994	-0.6354	
	dif	-17.0535***	-18.2207***	-18.6113***	0.1049
PMC a veic	level	NA	-2.1646	-1.7808	
	dif	-19.7883***	-19.8758***	-19.9392***	0.1001
IBC-BR	level	NA	-2.8514*	-1.1861	
	dif	-7.3861***	-7.7222***	-13.3564***	0.0762
Cap Inst	level	NA	-1.2849	-2.4529	
	dif	-9.9447***	-9.9271***	-9.9607***	0.0476
BM	level	NA	-2.1537	-0.5742	
	dif	-20.1061***	-14.0021***	-14.2718***	0.1032
M1	level	NA	-3.9804***	-1.13526	
	dif	-2.7681***	-16.5869***	-17.5687***	0.0713
M2	level	NA	-2.2671	-1.1987	
	dif	-1.8422*	-3.6664***	-4.2179***	0.1169

* 10%, ** 5%, and *** 1% significance levels.

for every one of the 21 subsequent quarters. Since this study includes 18 regressors, this comprises a total of 396 restricted MIDAS estimations. For each estimation, a one-step-ahead forecast is computed, simulating a real-time estimative. This approach resembles a nowcast, although it lacks the multiple data revision that later undoubtedly modified the published series. Moreover, Bernanke and Boivin (2003) and Schumacher and Breitung (2008) advocate that data revisions would not influence forecast precision significantly.

Unlike the restricted MIDAS, where the selection of the lags and weights scheme is endogenous, the unrestricted specification requires the choice of the lag scheme. For this dissertation, U-MIDAS is recursively estimated with 3, 4, 5, and 6 lags, comprising a total of 1584 estimations.

3.4 Model Evaluation: MSE and Modified DM Test

To compare the parametric MIDAS nowcasts with the AR(1) benchmark, this paper computes the relative mean-squared error (MSE) between the two model out-of-the sample forecasts with MIDAS on the numerator. The evaluation between MIDAS and U-MIDAS will rely on the same metric.

Additionally, this paper will graph the parametric MIDAS cumulative relative mean-squared error (MSE) for every quarter as calculated as shown:

$$r^{t,i} = \frac{MSE_{t,i}^{MIDAS}}{MSE_{t,i}^{AR}} = \frac{\sum_{t=44}^{T}(GDP_t - \widehat{GDP_{t,i}^{MIDAS}})^2}{\sum_{t=44}^{T}(GDP_t - \widehat{GDP_{t,i}^{AR}})^2} \tag{4}$$

for t= 44,...,66 representing the recursively reestimated estimations periods from both the restricted MIDAS and the autoregressive, ranging from 2014Q2 to 2019Q3 and i indexing the 18 regressors. The cumulative rate can uncover outliers and unveil if a particular regressor staged an unprecedented role during that time interval in forecasting the GDP. Distinct time series might undertake unusual significance during certain moments.

Although the MSE is a widely used criterion, it does not evaluate if the forecast precision derived from the MIDAS and U-MIDAS is statistically different than the AR(1) prediction. To this extent, this study implements the modified test proposed by Harvey, Leybourne, and Newbold (1997) test with a loss function power equal to 1. The null hypothesis is that the two methods yield the same forecast accuracy as described:

$$H_0 : E\left[l\left(\varepsilon^i_{t+h|t}\right)\right] - E\left[l\left(\varepsilon^{AR}_{t+h|t}\right)\right] = 0 \tag{5}$$

$l(\cdot)$ represents the quadratic lost function applied on $\varepsilon^i_{t+h|t}$, the i-th regressor model prediction error, and on $\varepsilon^{AR}_{t+h|t}$, the benchmark prediction error. Additionally, the alternative hypothesis is that the AR(1) is less accurate than MIDAS.

3.5 Forecast Combination

Finally, forecast combination for each set of restricted and unrestricted is implemented to produce an overall forecast. Bates and Granger (1969) suggest that combining forecast can dramatically improve accuracy. In addition to the equal-weighted combination as in Clements and Galvão (2008), this work computes the inverse MSE weighted as Kuzin, Marcellino, and Schumacher (2009), described below, providing a total of ten forecast combinations.

$$w_i^f = \frac{\frac{1}{MSE_i^f}}{\sum_{i=i}^{18} \frac{1}{MSE_i^f}} \tag{6}$$

where w_i^f stands for each regressor's weight at a given f framework. For each overall forecast, model evaluation relies on overall forecast MSE and the Modified DM Test against the AR(1).

4 EMPIRICAL RESULTS

This section summarizes the main MIDAS results vis à vis with the benchmark. The focus is on the comparable results as the primary purpose of this work is to evaluate if MIDAS can enhance the AR(1) forecast accuracy. The first part presents the MSE ratio and the modified Diebold and Mariano test. The second part describes the weights structure for every regressor. The thirds part compares MIDAS and U-MIDAS results under information criteria, MSE and DM-test, and the final section summarizes this work's contributions to the literature.

4.1 Restricted MIDAS Evaluation vis à vis the Benchmark

Table 2 reports the performance for the 18 recursive models for horizon h=1 vis à vis the autoregressive of order 1. The second column displays the relative mean-squared error (MSE) between the two model forecasts where the benchmark is on the denominator. In that way, an MSE fraction below 1 suggests that MIDAS outperformed the benchmark. The two other columns present the modified Diebold and Mariano test

Table 2 – Model Evaluation

		DM test	
MIDAS Series Code	**MSE ratio**	**Statistic**	**P-value**
IBC-Br	0.2846	-1.8666***	0.0383
PMC	0.5346	-1.0949	0.1432
PMC a	0.7032	-0.6	0.2776
PMC a veic	1.0473	0.0911	0.5358
BM	1.1454	0.3196	0.6237
PIM g	1.2621	0.6055	0.7241
Cap Inst	1.3067	0.7627	0.7727
PIM bcnd	1.3155	0.9939	0.8339
PIM t	1.3639	0.7766	0.7767
PMC a comb	1.4113	1.5701	0.9339
PIM int	1.5053	2.5909	0.9912
PMC a alim	1.5635	1.2977	0.8954
M1	1.6357	1.1789	0.8738
PIM bk	1.8111	1.4021	0.9118
M2	1.8972	2.1529	0.9781
PIM bcd	2.7472	3.4159	0.9986
PIM bc	2.9863	2.1633	0.9786
PIM e	3.294	3.7598	0.9993

proposed by Harvey, Leybourne, and Newbold (1997) statistics and p-value. The table summarizes the results with ascending ratio for the cumulative MSE quotient ending in 2019Q3. Accordingly, the top line showcases the model with the utmost performance. An MSE fraction below 1 suggests that MIDAS outperformed the benchmark.

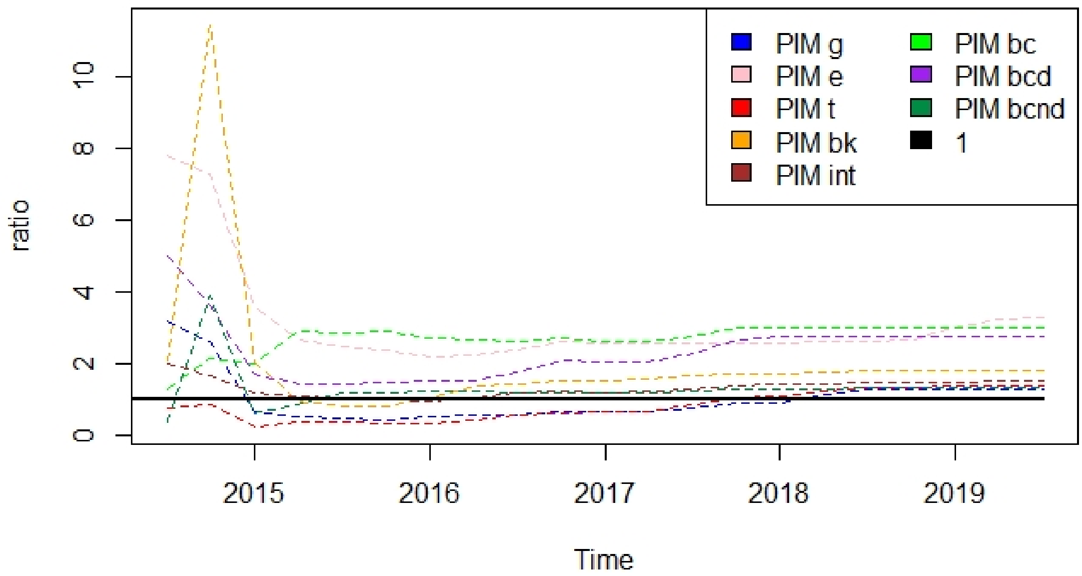
Under the MSE ratio criteria (Table 2 - column 2), the restricted MIDAS specification with a normalized exponential Almon lag polynomial beats the benchmark when employing IBC-Br, retail sales, and amplified retail sales as regressors. The result that MIDAS based specification holds the ability to outperform linear time series forecasts and improve GDP forecast accuracy is analogous to the findings in the mentioned literature. Unfortunately, Brazil does not release the same spectrum of activity statistics as OECD countries. Therefore, the number of MIDAS based specifications that enhance forecast precision is substantially reduced when compared with the developed countries' literature.

The results in Table 2 emphasizes Ghysels et al. (2018) argument that the selection of variables is particularly relevant in non-standard macroeconomic models. The literature supports that while MIDAS regression can improve the forecast performance, it remarkably depends on the selected independent variable. Besides it, strictly theoretical considerations would not be able to ascertain which regressor performs better for each country. While survey data, industrial production, retail sales, and activity indicators seem to be superior when predicting GDP growth, only an empirical approach can determine which will improve forecast accuracy for each country. Furthermore, this result might also depend on the family of MIDAS employed.

To evaluate if the two models forecast deviation is statistically significant, Table 2 presents the modified Diebold and Mariano test. For that extend, the only regressor that provides a statistically superior MIDAS specification is the IBC-Br. This result endorses the Central Bank's effort to estimate a monthly economic activity index. Their index, combined with the MIDAS framework, can improve GDP growth nowcast for the Brazilian economy. In the literature, this result compares with Lebouef and Morel (2014) that includes Eurocoin as one of the preferred indicators to predict GDP in the Euro Area.

However, Table 2 only provides a static measure. To examine if there were

Figure 1 – MSE Cumulative Ratio - Industrial Production



any indicators accuracy shift when forecasting GDP, graphs 1 to 3 present the cumulative MSE ratio for the 18 monthly indicators. The first graph focus on the evolution of the industrial production cumulative MSE ratio. From 2015 to 2017, the industrial production and the intermediate goods industrial production indexes outperformed the benchmark. Nonetheless, there is a continuous loss of predictability power from 2015 onward. This trend is coherent with the shrinking industrial sector contribution to the Brazilian economy.

The second graph exhibit the MSE cumulative ratio for the retail sales indicators. Although the coefficient remained below 1 for almost all the test set for the retail sales and the amplified retail sales, the modified Diebold and Mariano test displayed in Table 2 does not support that there is a statistical difference between its forecast and the benchmark.

Finally, other than just reinforcing the standing IBC-Br ability to nowcast the GDP growth in Brazil, the third graph reveals a fortuitous finding. From 2014Q3 until the end of 2015, the cumulative MSE fraction for the monetary base consistently decreased. While this coefficient behavior might not be intrinsically related to the fiscal pedaling, its course provides an additional narrative to the economic policies during those years as the

Figure 2 – MSE Cumulative Ratio - Retail Sales

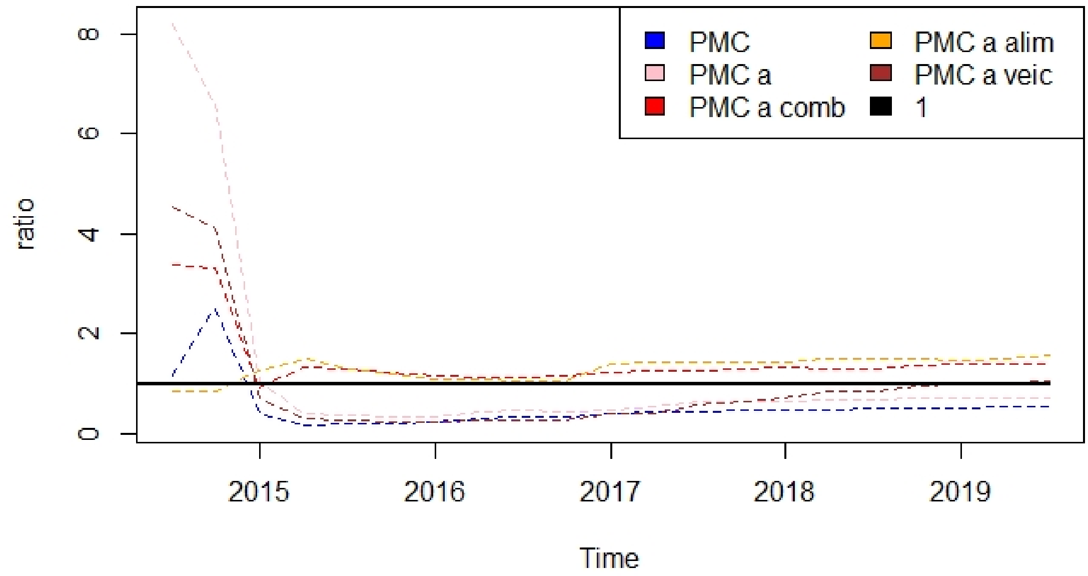
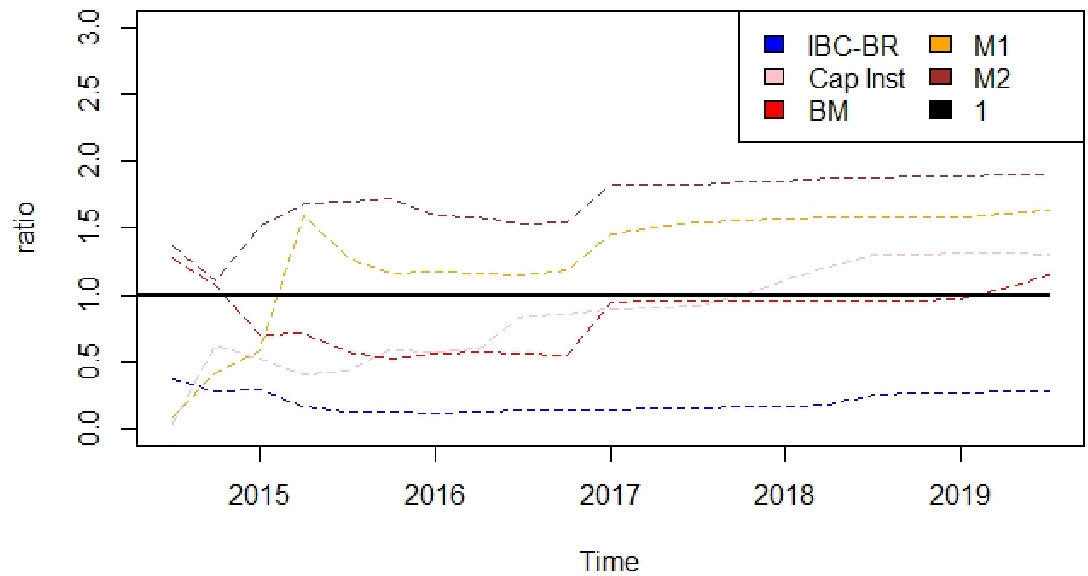


Figure 3 – MSE Cumulative Ratio - IBC-Br, Ind. Cap. and Mon. Aggregates



money supply infiltrated the Brazilian domestic growth.

4.2 Restricted MIDAS Weighting Structure

One advantage of the restricted MIDAS approach is that it can include many lags without the proliferation of parameters. Therefore, if the weights decay very fast, the inclusion of the the $B\left(L^{1/m}, \theta\right)$ polynomial would be unnecessary, and a simpler specification as the U-MIDAS could yield better results.

Graphs 4 to 6 present the weighs structure for all the 18 monthly regressors for the last estimation ending in 2019Q3. Firstly, the majority of the regressor's weights follows the expected format. It peaks on the first lags and then decays smoothly with almost no mass after the 12th lag.

Other than for the monetary aggregate M2, there is mass beyond after the third lag for all models. This suggests that the $B\left(L^{1/m}, \theta\right)$ might tackle the dilemma between including more information and estimating an additional coefficient. Apparently, the U-MIDAS does not resemble to be a competitive alternative if it were to include all the lags with mass, as shown in the graphs.

Comparing graphs 4 and 5, retail sales tend to have longer lags than industrial production. Additionally, the industrial production weights seem to peak earlier than

Figure 4 – Restricted MIDAS Weighting Scheme - Industrial Production

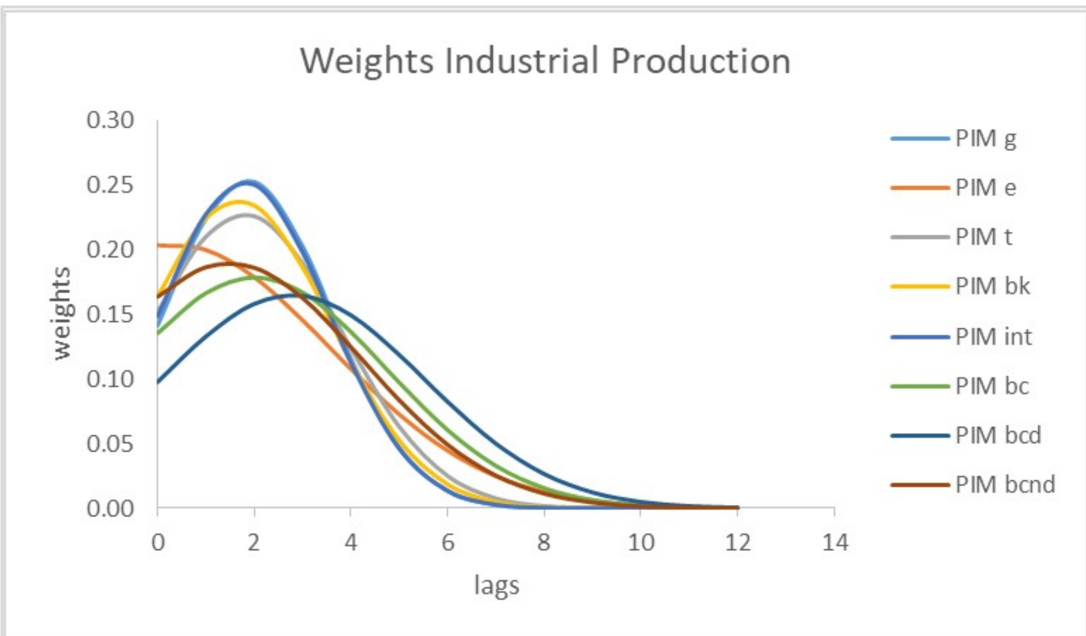
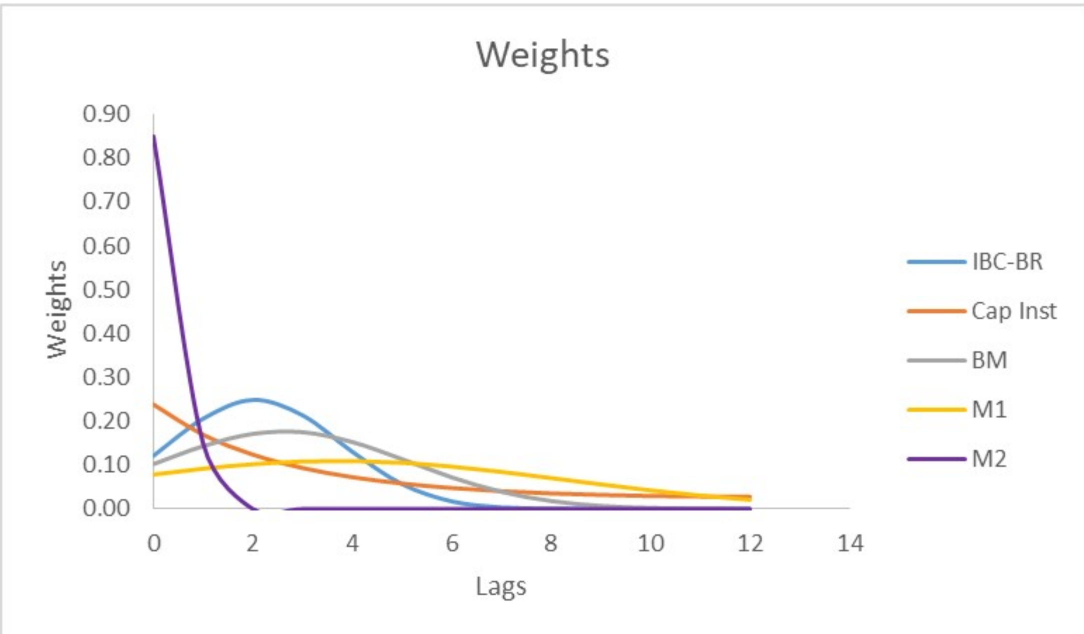


Figure 5 – Restricted MIDAS Weighting Scheme - Retail Sales

Figure 6 – Restricted MIDAS Weighting Scheme - Miscellaneous



Retail Sales. This is especially true if the two best performance retail sales indicators are taken into account. Finally, for the best performance model, IBC-BR, there is almost no mass after the 6th month.

4.3 Restricted MIDAS Evaluation vis à vis the Unrestricted MIDAS

This subsection focus on the evaluation between the MIDAS and U-MIDAS performance for out-of-the sample nowcasts estimated recursively. Tables 4 and 5 present the information criteria for all the estimated models for the last training set. Under the Akaike (AIC) and the Bayesian (BIC) information criteria, MIDAS provides minimum values for all the regressors. This is compatible with the expected role of the $B\left(L^{1/m}, \theta\right)$ polynomial. Reinforcing the previous findings, IBC-Br under the lag polynomial MIDAS schema outperforms all other models.

Table 5 displays the MSE ratio for all the 18 regressors. Although the model with the Almon polynomial model does not beat the U-MIDAS for all lags schema, the restricted framework dominates for the best performance nowcasts. Additionally, only IBC-BR, retail sales, and amplified retail sales based models delivered MSE ratios below

Table 3 – AIC: Restricted x Unrestricted MIDAS

	MIDAS R	U-MIDAS			
AR(1) 140.4541		**lag =3**	**lag=4**	**lags=5**	**lags=6**
PIM g	116.5617	143.6192	122.5151	124.4961	122.6408
PIM e	184.9912	199.453	194.259	195.9164	194.8946
PIM t	127.0035	151.3444	135.492	136.9447	135.4835
PIM bk	160.3073	175.3113	169.5251	170.9191	171.3512
PIM int	141.8944	155.333	149.0384	150.8985	147.3926
PIM bc	133.0089	172.5796	156.6494	150.1798	144.0442
PIM bcd	145.4853	184.6853	165.6175	165.0773	161.6951
PIM bcnd	162.2392	177.9015	177.9984	168.9318	166.7735
PMC	164.2862	172.5953	172.4136	174.3766	174.8401
PMC a	148.7338	190.3552	158.113	158.9917	157.1038
PMC a comb	188.6096	200.1877	198.1462	199.9313	198.6452
PMC a alim	192.0418	207.7954	209.7193	211.6399	208.0108
PMC a ve	152.9964	202.8991	177.1637	177.4841	170.7487
IBC-Br	111.9542 *	138.2252	121.4885	120.9627	119.9881
Cap Inst	156.9365	165.832	167.7701	168.0278	167.2907
BM	176.3259	203.6252	205.5218	202.2676	186.7489
M1	180.4461	190.3582	191.8959	193.7547	189.0778
M2	201.3198	210.4929	210.8736	205.5611	201.5705

Table 4 – BIC: Restricted x Unrestricted MIDAS

	MIDAS R	U-MIDAS			
AR(1) 144.0225		**lag =3**	**lag=4**	**lags=5**	**lags=6**
PIM g	127.1161	156.5725	137.6273	141.7672	141.929
PIM e	195.5456	212.4063	209.3712	213.1875	214.1828
PIM t	137.5579	164.2977	150.6042	154.2157	154.7717
PIM bk	170.8617	188.2646	184.6373	188.1902	190.6394
PIM int	152.4488	168.2863	164.1506	168.1696	166.6808
PIM bc	143.5633	185.5329	171.7616	167.4508	163.3324
PIM bcd	156.0397	197.6386	180.7297	182.3484	180.9833
PIM bcnd	172.7936	190.8548	193.1106	186.2029	186.0617
PMC	174.8406	185.5486	187.5258	191.6476	194.1283
PMC a	159.2882	203.3085	173.2252	176.2628	176.392
PMC a comb	199.164	213.141	213.2583	217.2023	217.9335
PMC a alim	202.5962	220.7487	224.8315	228.911	227.299
PMC a veic	163.5507	215.8524	192.2759	194.7551	190.0369
IBC-Br	122.5086 *	151.1785	136.6007	138.2337	139.2763
Cap Inst	167.4909	178.7853	182.8823	185.2988	186.5789
BM	186.8803	216.5785	220.634	219.5387	206.0371
M1	191.0005	203.3115	207.0081	211.0258	208.366
M2	211.8741	223.4462	225.9857	222.8322	220.8587

1. For these three best performance models, solely for the amplified retail sales did the restricted version outperform the unrestricted. Foremost, IBC-Br based MIDAS achieved the minimum MSE within all the models, standing out as the utmost out-of-sample forecast. The restricted version also outperformed the U-Midas for the 11 out of 15 models that didn't beat the benchmark.

Table 6 presents the modified DM test statistics and p-value for out-of-the sample recursively estimated nowcasts against the AR(1). The results are analogous to the previously reported findings. Only the restricted polynomial IBC-Br based MIDAS is statistically different from AR at 5%. Additionally, with 4 and 5 lags, IBC-Br U-MIDAS is statistically different from AR at 10%. This last observation is compatible with the IBC-BR restricted MIDAS weighting scheme that shows no mass beginning at the 6-th lag, as shown in graph 3.

Although the restricted polynomial PMC and amplified PMC based MIDAS reported MSE below 1 in table 2, the model forecast accuracy is not statistically different from the AR at 10% significance level. Nonetheless, these two regressors displayed negative sign ADF statistics for all the MIDAS models.

Table 5 – MSE ratio: Restricted x Unrestricted MIDAS

	MIDAS R	U-MIDAS			
code		**lag =3**	**lag=4**	**lags=5**	**lags=6**
PIM g	1.2621*	2.6055	1.5427	1.6401	1.5391
PIM e	3.294*	3.4073	3.7091	3.7838	3.7975
PIM t	1.3639*	3.0628	2.1023	2.1681	2.1621
PIM bk	1.8111*	2.2815	1.8972	1.9599	2.1238
PIM int	1.5053*	1.9853	1.5841	1.6158	1.5255
PIM bc	2.9863	3.0496	1.6457	1.3129	1.2318*
PIM bcd	2.7472	3.3146	2.0268	1.9487	1.9063*
PIM bcnd	1.3155	1.9574	1.8552	1.144	1.0831*
PMC	0.5346*	0.6745	0.7001	0.7351	0.8193
PMC a	0.7032	0.9592	0.6748	0.6364	0.5779*
PMC a comb	1.4113*	1.7642	1.5301	1.5907	1.4975
PMC a alim	1.5635*	2.0484	2.192	2.1745	2.0132
PMC a ve	1.0473*	1.8621	1.5309	1.5004	1.1929
IBC-Br	0.2846*	1.0244	0.4503	0.446	0.4493
Cap Inst	1.3067*	1.4872	1.5167	1.5413	1.4922
BM	1.1454*	2.1082	2.1178	2.283	1.4881
M1	1.6357	1.0649	1.0617	1.0513	0.8372*
M2	1.8972*	2.0977	2.4005	2.5716	2.8038

4.4 Forecast Combinations

Table 7 compares forecast combinations originated from the restricted and unrestricted Midas computed in this dissertation. As expected, the forecast combinations achieved better results than the single MIDAS models. Except for the equal-weighted U-Midas with 3 lags, all the forecast combination outperforms the AR(1) under MSE ratio evaluation. These results reinforce that forecast combinations can enhance forecast precision.

Overall, the inverse MSE weight promoted better results than the equal-weighted combination, as proposed by Bates and Granger (1969). The forecast improvement is especially substantial for the U-Midas models. While there was no model that statistically outperformed the AR(1) in the single restricted Midas, only the 3 lags combination forecast under the pooling schema did not outperform the benchmark. Furthermore, only the inverse MSE weight provides a statistically different forecast against the AR(1) at 5% under Modified DM-test. Finally, as in the previous outcomes, the restricted Midas appears to provide better accuracy than the unrestricted framework.

Table 6 – DM test: Restricted and Unrestricted MIDAS x AR

	MIDAS R		**U-MIDAS**			
			lag =3		**lag=4**	
code	**Statistic**	**P-value**	**Statistic**	**P-value**	**Statistic**	**P-value**
PIM g	0.6055	0.7241	1.4378	0.917	0.8308	0.792
PIM e	3.7598	0.9993	4.2978	0.9998	3.8447	0.9994
PIM t	0.7766	0.7767	1.5319	0.9294	1.3442	0.903
PIM bk	1.4021	0.9118	2.2167	0.9808	1.5445	0.9309
PIM int	2.5909	0.9912	2.2762	0.983	2.0183	0.9714
PIM bc	2.1633	0.9786	1.3456	0.9032	1.2853	0.8933
PIM bcd	3.4159	0.9986	1.9809	0.9692	1.3473	0.9035
PIM bcnd	0.9939	0.8339	1.7376	0.9511	1.8048	0.9569
PMC	-1.0949	0.1432	-0.7310	0.2366	-0.6526	0.2607
PMC a	-0.6000	0.2776	-0.1849	0.4275	-0.6565	0.2594
PMC a co	1.5701	0.9339	1.8761	0.9623	2.3116	0.9842
PMC a al	1.2977	0.8954	1.6282	0.9404	1.5943	0.9367
PMC a ve	0.0911	0.5358	2.7875	0.9943	1.4126	0.9134
IBC-Br	-1.8666**	0.0383	0.0341	0.5134	-1.4453*	0.0819
Cap Inst	0.7627	0.7727	1.0890	0.8554	1.1178	0.8615
BM	0.3196	0.6237	2.7819	0.9942	2.7694	0.994
M1	1.1789	0.8738	0.1961	0.5767	0.1814	0.571
M2	2.1529	0.9781	2.3817	0.9863	2.7238	0.9934

	lag =5		**lag=6**	
	Statistic	**P-value**	**Statistic**	**P-value**
PIM g	0.8685	0.8022	0.7956	0.7822
PIM e	3.8425	0.9994	3.913	0.9995
PIM t	1.2997	0.8957	1.3409	0.9025
PIM bk	1.7034	0.948	2.0042	0.9706
PIM int	2.2096	0.9805	1.5875	0.9359
PIM bc	0.7397	0.7659	0.4718	0.6789
PIM bcd	1.3316	0.901	1.2111	0.88
PIM bcnd	0.5538	0.707	0.3057	0.6185
PMC	-0.5682	0.2881	-0.3801	0.3539
PMC a	-0.7464	0.232	-0.8554	0.2012
PMC a co	2.6538	0.9923	2.1897	0.9797
PMC a al	1.605	0.9379	1.2158	0.8809
PMC a ve	1.297	0.8953	0.3964	0.652
IBC-Br	-1.3656*	0.0936	-1.2341	0.1157
Cap Inst	1.0338	0.8432	0.9644	0.8268
BM	2.2405	0.9817	1.1959	0.8771
M1	0.1521	0.5597	-0.4889	0.315
M2	2.6822	0.9928	2.9602	0.9961

* 10%, ** 5%, and *** 1% significance levels.

Table 7 – Forecast Combination - Pooling Equal-Weighted and Inverse MSE

		MSE ratio	**DM-Stat**	**DM-p-value**
R-Midas	pool	0.7266	-1.5056	0.0739*
	pool MSE	0.3871	-2.1183	0.0234**
U-Midas - 3 lags	pool	1.1211	0.5363	0.7012
	pool MSE	0.7801	-1.0358	0.1563
U-Midas - 4 lags	pool	0.7983	-1.0415	0.1550
	pool MSE	0.4981	-1.9106	0.0352**
U-Midas - 5 lags	pool	0.7618	-1.1473	0.1324
	pool MSE	0.4735	-1.9135	0.0350**
U-Midas - 6 lags	pool	0.6337	-1.5852	0.0643*
	pool MSE	0.4000	-2.0225	0.0283**

* 10%, ** 5%, and *** 1% significance levels.

4.5 Contributions

To the author's knowledge, this is the first study that applies MIDAS to nowcast GDP in Brazil using macroeconomic data. Although Zuanazzi and Ziegelmann (2014) had already implemented the MIDAS framework to forecast Brazilian GDP, their work targeted financial asset prices and did not include relevant macroeconomic indicators such as IBC-Br which this research showed that encompass the highest predictive content. Like the vibrant literature in MIDAS, this empirical application supports its implementation, for it could enhance GDP nowcast accuracy in Brazil.

Moreover, this dissertation indicates that the constraints imposed by $B\left(L^{1/m}, \theta\right)$ polynomial are reasonable and can well capture the underlying data generating process as the outperforming IBC-Br based restricted MIDAS beats U-MIDAS. Although Foroni, Marcellino, and Schumacher (2011) illustrate, through Monte Carlo simulations, that when sample differences between the dependent and independent variable is small, the U-MIDAS performance is similar to MIDAS, this work suggests that the restricted MIDAS might be advantageous even for macroeconomic applications with small frequency differences. Furthermore, the forecast combinations outperformed the single MIDAS models. Moreover, the restricted framework outperforms the unrestricted version for forecast combinations for the two chosen weighting schemas.

Finally, this study showcased that the money supply had an unprecedented predictive power to explain the GDP in Brazil from 2014Q3 to the end of 2015. This finding deserves further research as the literature had not fully explored it, and it might correlate with other economic policy choices during that period.

5 CONCLUSION

This empirical study compares if the restricted polynomial MIDAS specification outperforms the AR(1) for out of the sample recursively estimated nowcasts when used to estimate the GDP growth in Brazil with macroeconomic indicators. As in previous literature, the results endorse that the restricted MIDAS can enhance the nowcast accuracy when used with adequate regressors. For the Brazilian economy, IBC-Br restricted lag polynomial based MIDAS showcase the best performance under all the computed metrics for evaluation. Not only did the restricted IBC-Br MIDAS outperform the benchmark, but it also beat the U-MIDAS. This result reinforces that the lag polynomial constraints are reasonable and that, even for small sample differences, the restricted approach can deliver superior nowcasts than the least squared estimated MIDAS. Moreover, the forecast combinations performed better than the single MIDAS models.

Finally, the cumulative MSE ratio graph revealed that between 2014Q3 until the end of 2015, the quotient for the monetary base continuously declined. While this behavior might not be related to the "fiscal pedaling", its trend contributes to the economic policy narrative during those years, potentially offering implications and directions for future investigation. Lastly, the multiple MIDAS with variable selection via automated methods are left to further research.

References

ANDREOU, E.; GHYSELS, E.; KOURTELLOS, A. Should macroeconomic forecasters use daily financial data and how? *Journal of Business & Economic Statistics*, Taylor & Francis, v. 31, n. 2, p. 240–251, 2013.

BATES, J. M.; GRANGER, C. W. J. The combination of forecasts. The OR Society, 1969.

BERNANKE, B. S.; BOIVIN, J. Monetary policy in a data-rich environment. *Journal of Monetary Economics*, Elsevier, v. 50, n. 3, p. 525–546, 2003.

BREITUNG, J.; ROLING, C. Forecasting inflation rates using daily data: A nonparametric midas approach. *Journal of Forecasting*, Wiley Online Library, v. 34, n. 7, p. 588–603, 2015.

CLEMENTS, M. P.; GALVÃO, A. B. Macroeconomic forecasting with mixed-frequency data: Forecasting output growth in the united states. *Journal of Business & Economic Statistics*, Taylor & Francis, v. 26, n. 4, p. 546–554, 2008.

FORONI, C.; MARCELLINO, M.; SCHUMACHER, C. Unrestricted mixed data sampling (midas): Midas regressions with unrestricted lag polynomials. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, Wiley Online Library, v. 178, n. 1, p. 57–82, 2015.

FORONI, C.; MARCELLINO, M. G. A survey of econometric methods for mixed-frequency data. *Available at SSRN 2268912*, 2013.

FORONI, C.; MARCELLINO, M. G.; SCHUMACHER, C. U-midas: Midas regressions with unrestricted lag polynomials. Bundesbank Series 1 Discussion Paper, 2011.

GHYSELS, E.; MARCELLINO, M. *Applied economic forecasting using time series methods*. [S.l.]: Oxford University Press, 2018.

GHYSELS, E.; OZKAN, N. Real-time forecasting of the us federal government budget: A simple mixed frequency data regression approach. *International Journal of Forecasting*, Elsevier, v. 31, n. 4, p. 1009–1020, 2015.

GHYSELS, E.; SANTA-CLARA, P.; VALKANOV, R. The midas touch: Mixed data sampling regression models. 2004.

GHYSELS, E.; SINKO, A.; VALKANOV, R. Midas regressions: Further results and new directions. *Econometric Reviews*, Taylor & Francis, v. 26, n. 1, p. 53–90, 2007.

GHYSELS, E.; VALKANOV, R. I.; SERRANO, A. R. Multi-period forecasts of volatility: Direct, iterated, and mixed-data approaches. In: *EFA 2009 Bergen Meetings Paper*. [S.l.: s.n.], 2009.

GODBOUT, C.; JACOB, J. *Le pouvoir de prévision des indices PMI*. [S.l.], 2010.

HARVEY, D.; LEYBOURNE, S.; NEWBOLD, P. Testing the equality of prediction mean squared errors. *International Journal of forecasting*, Elsevier, v. 13, n. 2, p. 281–291, 1997.

KUZIN, V.; MARCELLINO, M.; SCHUMACHER, C. Midas vs. mixed-frequency var: Nowcasting gdp in the euro area. *International Journal of Forecasting*, Elsevier, v. 27, n. 2, p. 529–542, 2011.

LEBOEUF, M.; MOREL, L. *Forecasting short-term real GDP growth in the euro area and Japan using unrestricted MIDAS regressions.* [S.l.], 2014.

LI, X. et al. A midas modelling framework for chinese inflation index forecast incorporating google search data. *Electronic Commerce Research and Applications*, Elsevier, v. 14, n. 2, p. 112–125, 2015.

MARCELLINO, M. Some consequences of temporal aggregation in empirical analysis. *Journal of Business & Economic Statistics*, Taylor & Francis Group, v. 17, n. 1, p. 129–136, 1999.

MARCELLINO, M.; SCHUMACHER, C. Factor midas for nowcasting and forecasting with ragged-edge data: A model comparison for german gdp. *Oxford Bulletin of Economics and Statistics*, Wiley Online Library, v. 72, n. 4, p. 518–550, 2010.

MONTEFORTE, L.; MORETTI, G. Real-time forecasts of inflation: The role of financial variables. *Journal of Forecasting*, Wiley Online Library, v. 32, n. 1, p. 51–61, 2013.

PAREDES, J.; PEDREGAL, D. J.; PÉREZ, J. J. Fiscal policy analysis in the euro area: Expanding the toolkit. *Journal of Policy Modeling*, Elsevier, v. 36, n. 5, p. 800–823, 2014.

SCHUMACHER, C.; BREITUNG, J. Real-time forecasting of german gdp based on a large factor model with monthly and quarterly data. *International Journal of Forecasting*, Elsevier, v. 24, n. 3, p. 386–398, 2008.

STOCK, J. H.; WATSON, M. W. Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature*, v. 41, n. 3, p. 788–829, 2003.

ZUANAZZI, P.; ZIEGELMANN, F. Previsões para o crescimento do pib trimestral brasileiro com séries financeiras e economicas mensais: Uma aplicação de midas. *Economia Aplicada*, Scielo, v. 18, n. 2, p. 295–318, 2014.

APPENDICES

Downloaded Series from IBGE and Central Bank of Brazil

Table 8 – Monthly Seasonally Adj. - downloaded from IBGE

CODE	DESCRIPTION	TABLE	VARIABLE
PIM g	General industry	Table 3653	1- Indústria geral
PIM e	Extractive industry	Table 3653	2 Indústrias extrativas
PIM t	Manufacturing industry	Table 3653	3 Indústrias de transformação
PIM bk	Capital goods industry	Table 3651	1 Bens de capital
PIM int	Intermediate goods industry	Table 3651	2 Bens intermediários
PIM bc	Consumer goods industry	Table 3651	3 Bens de consumo
PIM bcd	Durable consumer goods industry	Table 3651	31 Bens de consumo duráveis
PIM bcnd	Nondurable consumer goods industry	Table 3651	32 Bens de consumo semiduráveis e não duráveis

Table 9 – Monthly Seasonally Adjusted Retail Sales - downloaded from IBGE

CODE	DESCRIPTION	TABLE	VARIABLE
PMC	Retail sales	Table 3416	
PMC a	Amplified retail sales	Table 3417	
PMC comb	Amplified retail sales - fuels	Table 3419	Combustíveis e lubrificantes
PMC a alim	Amplified retail sales- food, beverages and tobacco	Table 3419	Hipermercados, produtos alimentícios, bebidas e fumo
PMC a veic	Amplified retail sales - vehicles	Table 3419	Veículos, motocicletas, partes e peças

Table10 – Quarterly Seasonally Adj. GDP - downloaded from IBGE

CODE	DESCRIPTION	TABLE	Variable
GDP	GDP - Physical Production	Table 1620	Série encadeada do índice de volume trimestral (Base: média 1995 = 100)

Table11 – Monthly Seasonally Adjusted Series - downloaded from Central Bank of Brazil

CODE	DESCRIPTION	TABLE	VARIABLE
IBC-Br	index of economic activity	24364	Índice de Atividade Econômica do BC (IBC-Br) - com ajuste sazonal
Cap Inst	Capacity utilization	28561	Utilização da capacidade instalada – indústria de trans. (FGV) - Dados dessazonalizados
BM	Monetary Base	27840	BM - Base monetária restrita - sazonalmente ajustado
M1	Monetary Aggregate	27841	Meios de pagamento - M1 - Novo - sazonalmente ajustado
M2	Monetary Aggregate	27842	Meios de pagamento amplos - M2 - Novo - sazonalmente ajustado