NOVA IMS Information Management School

MEGI

Mestrado em Estatística e Gestão de Informação

Master Program in Statistics and Information Management

# Understanding Geographical Disposition of MIDAS Forecasting Accuracy in European Countries

Bruno Roovers de Avelar Esteves Borges

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa

**NOVA Information Management School**
**Instituto Superior de Estatística e Gestão de Informação**
Universidade Nova de Lisboa

# UNDERSTANDING GEOGRAPHICAL DISPOSITION OF MIDAS FORECASTING ACCURACY IN EUROPEAN COUNTRIES

by

Bruno Roovers de Avelar Esteves Borges

Dissertation presented as partial requirement for obtaining the Master's degree in Statistics and Information Management , with a specialization in Analysis and Information Management.

**Advisors:** Bruno Damásio & Flávio Pinheiro

November 2022

# ABSTRACT

The measuring of economic activity has long been a time consuming process which often results in official aggregated statistics being published with some delay in relation to the end of the observation period. In order to assess where the forecasting of country wide economic activity can be the most useful, we propose an exercise on which the forecasts root mean squared errors of several European countries' GDP are taken trough a MIDAS regression on a rolling estimation window. The results of this exercise should allow for a comparison on the regional dispositions of the forecasting quality and assess which countries show the highest gains from the use of MIDAS methodology in comparason to the ARIMA univariate process. We find that the gains of using the MIDAS methodology in relation to the ARIMA process are heterogenous between countries. Even though the countries appear to have distinct potential for the use of MIDAS models, the geographical sub-region division into Center, Eastern, Northern and Southern European countries does not seem to produce good separation groups.



# KEYWORDS

MIDAS; Forecasting Quality; Economic Activity; European Coutries;

# INDEX

# LIST OF FIGURES

# LIST OF EQUATIONS

# LIST OF ACRONIMS AND ABREVIATIONS

| | |
|---|---|
| **ARIMA** | Autoregressive Integrated Moving Average |
| **BIC** | Bayesian Information Criterion |
| **GDP** | Gross Domestic Product |
| **MAD** | Mean Average Deviation |
| **MIDAS** | Mixed-data Sampling |
| **NLS** | Nonlinear Least Squares |
| **RMSE** | Root Mean Squared Error |
| **U-MIDAS** | Unrestricted Mixed Data Sampling |

# 1. INTRODUCTION

In this dissertation it is discussed the reliability of the MIDAS (Mixed-data Sampling) methodology for the forecasting of 14 European Countries GDP (Gross Domestic Product) by comparing the two step ahead out of sample root mean squared error (forwardly known as RMSE) with the forecast for the same horizon of the ARIMA (Autoregressive Integrated Moving Average) process. This study aims to assess up to which degree can the geographical location affects the MIDAS relative gains.

The datasets and the methodology were carefully selected so that they could be extended to the entirety of the country applications in order to minimize the likelihood of creating biased results. These differences in performance gains would then be consequences of circumstancial parameter and starting weights selection as well as the inevitable creation of a prone data snooping environment. It was found that the gains of using the MIDAS methodology in relation to the ARIMA process are heterogenous between countries. Even though the countries appear to have distinct potential for the use of MIDAS models, the geographical sub-region divisions into Center, Eastern, Northern and Southern European countries does not seem to produce good separation groups hinting that the geographical region is not the deciding factor in the success of the MIDAS application. With the works developed it is expected hope to contribute to the discussion of the degree to which geographical, economical and cultural diferences moderate the MIDAS forecasting accuracy and the relative gains thereof.

## 1.1. RESEARCH GAP

From the literature gathered it has become clear that there isn't extensive research developed on the topic for the scale that it is being proposed in this article, most of the literature review collected focuses on assessing forecasting quality at national level for longer than the than two quarters ahead forecasting using a wide range of combinations for the estimation methods, performance assessment methodology, datasets, assumptions made on statistical and economical side.

## 1.2. RELEVANCE OF THE TOPIC

With the increase for the need to have national statistics that represent the reality closer to the object of observation, the more and diverse approaches have been developed Lehmann & Wohlrabe (2014). New methodologies have been developed to allow for the estimation of these figures even before the end of the observation period such as MIDAS (Mixed-data sampling) regressions. As diverse as the approaches studied can be, it should be relevant to acknowledge the possibility that the use of these methods can have statistically significant dissimilarities in its accuracy depending on the applications.

Throughout this research one have come to terms with the existing trade-off between the depth and granularity that could be given to the forecasting quality analysis and the availability

and scalability of the exercise proposed. It has come to our attention that even though the data availablility has had a tremendous positive evolution in the past years, it is not feasible to produce a study of this nature considering a regional unit smaller than a country for an entire continent. Additionally, given the ever-present interest of society as a whole to have accurate economic activity figures at national level it was decided to perform the study at this level.

The comprehension of the model's ability to forecast and nowcast the economic activity for a given country is relevant for the future replication of studies and forecasting methodologies for different geographies. Finally, this research intends to assess MIDAS estimation quality for several European countries using datasets of similar information amongst them.

Should the resulting exercise manage to differentiate and lump together countries by their forecasting quality, the article could contribute to the basis of subsequent research to deepen the knowledge on the underlying reasons that determine the quality of forecast for any given country.

The following segments will consit of the Literature Review where the research developed so far on the subject and tangent topics of this dissertation. Section 3 will give an encapsulating view of the methodology followed that allowed the achievement of the results whose analysis will be taken in the susquent section. The dissertation ends with the conclusions where the findings, limitations and future research are shared.

## 2. LITERATURE REVIEW

This section aims to give an holistic cover of the collected research documents and main ideas discussed within that were considered for the decisions made in the methodology of this study. The research published in this field can be characterized as having a wide combination of methodologies from the estimation methods, the processes through which the quality of the estimation is assessed and the types of data selected. With this into consideration it was decided to separate the review on the aforementioned methodological axis. This segmentation cannot be done orthogonaly for all the contents discussed below as context of the uses needs to be provided.

### 2.1. General Methodology Review

In order to discuss the state-of-the art methodologies it would be appropriate to begin with the research done by Lehmann & Wohlrabe (2014). This article systematizes the existing literature at the time on regional economic forecasting. Nonetheless, it is mostly useful in detecting the most appropropriate estimation models for different forecasting horizons.

The authors also make crucial remarks on several axes of the most noticeable research published until then. The data scarce environment is a recurrent topic and as such, many researchers that have focused in the optimization of particular regions forecasting have used spatial panel data models. Furthermore, this problem has also contrained the development of structural models despite the relevancy of these models in the modeling of interdependicies of the features and constructs. On the models used, the authors have realized the relative under representation of factor models and the absence of density or interval forecasts in the existing literature, being the most common estimation done with point forecasts. Finally, when it comes to the validation of the models, the most common method would be to assess the relative Root Mean Square Forecast Error measured as both in and out of sample and compare it with the univariate auto-regressive procedure.

The forecasting horizons was also addressed in this research article, the majority of the studies gathered have have forecasted one step ahead, which in the context of the respective articles could be raging from one quarter up to one year. The quantity and relevancy of research that concerns with higher horizons dimishes exponentially after that horizon for regional forecasting of the GDP. One final and important remark raised in this article was the relative absence of literature developed thus far on the estimation and forecasting on economic cycles.

### 2.2. Estimation Methods

The selection of MIDAS and U-MIDAS (Unrestriceted Mixed Data Sampling), the unrestricted version of the MIDAS methodology models first proposed in Ghysels, Santa-Clara, & Valkanov (2004). When this methodology was first proposed it was pledged by the authors to have extensive applicability for finance and economics fields. Since then, countless research articles have been developed using these ideas as background and derivations of the methodology thereof. It was not until Ghysels, Sinko & Valkanov (2007) that it was seen in a unified article a comparison and application of several lag structures specifications for the parsimonious parameterization of a MIDAS regression. Besides, these applications do not explicitly model the dynamics of the dependent variables, in its place the MIDAS directly forecasts the interest variable with current or lagged indicators.

One of the crucial concept that needs to be discussed is the Expontential Almon Distributed Lag as it is one of the functional constraints available in the literature for the parsimonious estimation of the MIDAS. It is based on the Almon lags first introduced by the author of the same name in Almon (1965). In Ghysels, Sinko & Valkanov (2007), the authors apply both lags distributions specifications to a Volatility forecasting problem and the ICAPM model (Intemporal Capital Asset Pricing Model) as proposed by Merton (1973). In the first application, the authors resort to the Mean Average Deviation (MAD) as claimed to be a more robust measeure for the goodness of fit in the presence of heteroscedasticity. For the estimation of the volatility measures, the results show that there is no significant difference between the estimations that use one of the two lag distributions, that is, the Beta distribution and the Expontential Almon. Furthermore, the results also show that the Beta distribution performs better than the Almon Distributed lag for higher frequency estimation of the coefficients that relate the Expected Return to Market Volatility and Almon showed an advantage for estimations for lower frequency of the coefficient.

Despite the concise objective being related to the MIDAS modelling framework, it can not be overlooked some other interesting approaches collected in the literature for forecasting economic activity at the reagional level. Models such as Dynamic and Static Factor models can deal with great amount of features (large cross-section when comparing to temporal dimensionality). As stated in Lehmann & Wohlrabe (2015), the great advantage of these models in comparison to the standard econometrics models, besides the satisfatory forecast, is the parsimonious way in which they can be estimated without concerning themselves with the inherent uncertainty about parameters estimates which would be the case when estimating the models with heavly correlated independent variables.

Another approach to modelling of economic activity seen with relative frequency in the reviewd literature was the use of Bridge Regressions. Due to its simplicity of application these models have been widely used by central banks in nowcasting and forecasting of short horizons of economic activity as concluded by Schumacher (2014). These models were first developed in the context of chemometrics by Frank & Friedman (1993) where they compared what they claim to be common estimation methods in that field, the Partial Least Squares and

the Principal Components Regression with other statistical methods such as ordinary least squares, variable subset selection, the ridge regression and the bridge regression. On one side, the bridge regressions share features with the ridge regressions as both these models belong to the penalized regressions family, however the bridge regressions can fit in to the same model families as the MIDAS regressions in the sense that they are both distributed lags models that can use of mixed frequency data.

Although the Bridge and MIDAS regressions share many attributes there are relevant differences among these them. While the MIDAS regressions perform direct multi-step nowcasting, the bridge regressions equations are based on iterated forecasts. They also diverge in the incorporation of higher frequency variables in the model's body, for instance, the MIDAS regressions uses functional lag polynomials to parsimoniously estimate the coefficients and the Bridge regressions are partiall fixed due to the its time aggregation feature. Finally, MIDAS equations can consider indicators of higher frequency for the quarter of the nowcast while the Bridge regressions cannot. Additionally, there is no direct superiority between the two models as the performance differences can be related with indicators chosen and the evaluation methods as referred by Schumacher (2014). The authors of this article also applied both models and the results seem to suggest that the pooling of the models nowcasts and forecasts tends to have better performance stability.

Throughout the review performed it was common to see the univariate auto-regressive process as well as the full ARIMA in itself being used as the benchmark for the estimation of the MIDAS model's forecasting quality. It was also noticed in the literature consulted the contruction of statistical tests based on a comparative measure between the loss function of the MIDAS model and the benchmark to assess the significance to which it can be rejectes the hypothesis of equal predictive ability.

### 2.3. Estimation Window

In literature it is usual to have a simple horizontal or bootstrap split of the time series between the training and validation data. On the first set, the models are trained and used to determine the estimations of the parameters as well as in-sample accuracy measures, usually RMSE for the out of sample is calculated. The latter set, often to provide external validity of the model. The methods used throughout the literature for the creation of the datasets containing goodness of fit measures and the subsequent evaluations are combinations extensely diverse and seldom do they repeat themselves from one research article to another.

The use of estimation windows in forecasting is not concerned with these features as each one of the windows can preconize the existence of either a training and validation set. The use of an estimation window for the models and the resulting error estimation both in and out of sample has been considered due to the its increased robustness in the face of structural breaks, given that using the entire series for the estimation of the parameters would increase the likelihood of producing a biased estimation as noted in Pesaran & Timmermann (2007). Additionally, there is the possibility that the quality of forecasting at one or susbquent moments in time is not representative of the real accuracy of the model. The use of either recursive or rolling estimation windows becomes even more pertinent in the context of the institutional changes in social reality introduced by the Covid-19 pandemic.

When it comes to the the literature developed that allows the better understanding of the optimal estimation window size, the works developed by Clark & McCracken (2009) become useful. These authors have researched this thematic in the context of structural breaks in the time series for rolling and recursive estimation windows, emphasizing the Bias-Variance trade-off. As explained by the authors, if the earliest data available does not follow the same data generation process then it may lead to biased estimators which accumulate in inefficiency for the out of sample errors. On the other side, reducing the sample size will increase parameters variance. One of the main findings of this study is that the optimal estimation window is weakly decreasing in the magnitude of the break in the time series.

Additional literature collected showed that it is also possible to construct tests robust to the choice of window size for the models that use either rolling or recursive estimation windows like the research of Rossi & Inoue (2012). In particular, the authors concluded that the methodology was useful, by applying it to the evaluation exchange rate models predictive ability.

### 2.4. Performance Assessment

The model's performance assessment is a topic that manages to distinguish several of the researches developed so far and therefore merited a distinct segment from the estimation methods and the estimation windows. The assessment of the models performance between the literature gathered covers a lot of different methodologies and even though some of them do not contribute directly to the goal proposed in this article, they are valid contributions for future research in this thematic and allows to have considered the differences between the methods and the context behind their application.

In regards to the model performance assessment measurements, the overwhelming majority the articles collected have resorted to the RMSE or the MSE obtained from the in and out of sample residual estimation exercise. Some non negligible differences can be seen in the

research gathered namely for testing the statistical significance of the residuals, the use of Diebold-Marinano test or others that allow for a more fair testing when there are more than two competing models is present in a good part of the consulted articles.

On the use and comparison of predictive capabilities as proposed by Diebold and Mariano, the test was proposed in Diebold & Mariano (2002). This test has been widly used in the time series related literature and contributed extensively to derivations and modifications of the test's purpose thereafter. The aim of the methodology proposed in the article is to test for equal predictive ability, in part the test has had great success due to its quality of having a wide variety of loss functions that can have non Normally distributed forecast errors. These tests can also be used in serially and contemporaneously correlated errors. Surprisingly, in a more recently published article, Diebold (2015), the author denounces the abuses of the applications of the Diebold-Mariano test. The author claims that it was not his intention for the test to aid in the comparison of the in-sample loss functions of two given models.

Other tests proposed as in Lehmann & Wohlrabe (2015), where the authors tests if a given model performs better than an entire collection of competing models by using the Superior Predictive Ability test. The biggest contribute of this test is the adequacy of its use given that they intend to compare more than two models and a pairwise testing method would lead to the problem of data snooping. This approach followed was first proposed by Hansen (2005) as a modified version from White (2000) given that White's method suffers from the inclusion of poor or irrelevant models, this test also requires a rolling window estimation. Should the null hypothesis be rejected it can be accepted the alternative one where there is enough evidence to support the claim that there is at least one model that out-performs the benchmark model. To get a set of the best models, the authors then use the MCS algorithm proposed by Hansen et al (2011).

Other approaches can be seen in Rünstler, & Sédillot (2003) and Sédillot & Pain (2003) where the research focus not only in the point forecast of economic activity but also in the acceleration or decelaration of the growth tendency. In this case the authors resorted to the forecasting encompassing tests of the Diebold-Mariano modified by Harvey, Leybourne & Newbold (1997).

It is also common in the literature to find that the assessment of the comparative quality of the model proposed is done by an analysis of the out of sample loss function ratio, usually the RMSE, between the model of interest that is being studied and the pre-selected benchmark model. The benchmark model more often than not is the AR process of the response variable but in some cases it can be similar derivations of the interst model, in this case the MIDAS methodology, or even other estimations of the same model when the estimation consist of iteratable methods such as the Nonlinear Least Squares (forwardly known as NLS).

### 2.5. DATA CONSIDERATIONS

For the variable selection process, there have been interesting approaches developed in the literature collected so far and it was observed a wide variety of methods that tackle the topic of variable selection and data availability. Starting with Rünstler & Sédillot (2003) where the Bridge Regressions variables were selected through a Stepwise process for which the intermediate step was kept as way examine the roubustness of the findings against ou-of-sample foreacasting performance and later on the forecast based on incomplete datasets.

On the topic of dealing with data accessibility where the decision of usage of certain variables that are constrained by it's availability for the period and geography, there has been methodos developed as a work around that consists of disagreggating the variables into more granular time units. The operacionalization of these concepts can be seen in Lehmann & Wohlrabe (2015) where the authors used two methods for the disagregatrion of the annual regional GDP of three east german regions into the respective quarterly regional GDP. The authors use the method proposed in Chow & Lin (1971) for disaggregation through a single regressand linear equation where the low frequency variable is predicted for the higher frequency, naturally the regressand has an higher frequency than the response variable and the regression can only fitted for the periods where the observations coincide. There must also exist a strong pre-validated relationship between the regressand and the variable to be converted to the higher frequency. The other method used has greater applications for the desagreggation of the interest variable according to the authors, the extrapolation method consists of determining the interest variable proportion that has be realized at each higher frequency period by observing the proportion of a higher frequency variable for the same periods. As an example, the authors use the proportion of all turnovers in manufacturing sector that occurred in a given quarter in relation to the entire year for the specific region to assign the proportion of the annual GDP that has occurred in that same period. Still in Lehmann & Wohlrabe (2015) it is clear the approach to have a multifaceted and comprehensive dataset as the variables used in the modelling exercise contain information from distinct sources. The authors grouped the variables by source and contents as macro-economical, prices, sectorial, surveys and compound indicators data at regional, national and international scope.

There are other interesting approaches to variable selection as it is well known the capacity of these types of models to leverage on high-frequency data at which ever could be the frequency ratio, that is, the amount of observations of the high-frequency variable in relation to the low-frequency one. In St Aubyn (2020) the authors used electricity consumption to forecast and nowcast Portuguese economic activity with exciting results when it comes to the significance of the variable as well as the quality of the estimation. Other interesting uses of high-frequency data can be seen in Eraslan & Götz (2021) where they use flights data with a weekly frequency to forecast the quarterly economic activity, given the high

flow of this information and it's completeness, the authors had at their disposal data regarding passangers, cargo, amounts of flights landed and taken off from and to a given country.

Recently, there has been a focus in the literature on the institutional changes that the global economy has gone through in the last couple of years in the context of a world wide pandemic as shown in Jardet & Meunier (2022). The indexes that aimed at producing reliable measures to attest for the intensity of propagation and the health impacts were widely available with diverse granularity even to the most disaggregate regional level. Another quality of these type of variables besides the geographical granularity is the high-frequency on which they were produced. They can be easily found at weekly perdiocity and even in some cases at daily frequency. These indicators were used in the literature either to assess the impact of the pandemic on the response variables or to account for these exact institutional changes in the models, usually related to the occupancy of hospitals and other measures that could indicate the level of stress being put on healthcare resources.

Other uses namely on labor market conditions and economic activity were seen. Even though these applications had distinct level of success, there is not extense literature on the use of composite indexes that explore the degree to which the confinement and other measures taken by the governments to contain the spread of the virus. Indexes such as the Oxford COVID-19 Government Response Tracker's Stringency Index were used in Cross, Ng & Scuffham (2020) whose results showed little promise in relation to its predictive content for economic activity, but the same can not be said when they are applied to forecast the desease's propagation rates where the index was more succeseful in its forecast.

# 3. METHODOLOGY

The methodology for the creation of the results seen in this study will consider three different segments alike the structure in the Literature Review. One related with the data acquisition that will be done mostly on auditing modality consulting and transforming of mostly structured data, other with the estimation methods and the latest concerning the techniques for the assessment of the model validity. Some general considerations regarding the methodology need to be made to contextualize the contents of the segments that will go in further deph on the respective methods.

The majority of the cases studied consider the research that focuses on GDP nowcasting or forecasting of periods not further than a semester ahead and use a mixed-frequency dataset. The articles selected also use direct forecasting of the interest variable instead of the also common estimation by components of the GDP.

On the process taken for the estimation and the comparison of the results in the purpose of this study to correctly assess the quality of the forecast, it has been considered besides the MIDAS residuals, the ARIMA ones. This will help distinguish the quality of the information in the explanatory variables for the forecasting exercise present in the explanatory variables. By creating a ratio between RMSE of the MIDAS forecasts and the ARIMA ones it is expected to more easily detect the relative effectiveness of the MIDAS model in comparison to the competing model. The forecasts and the out of sample errors will be calculated in the periods that procede the rolling windows.

## 3.1. DATASET

The structure of the datasets collected has been collected and structured in it's two possible dimensions: the datasets can have different granularity when it comes to the observation interval – monthly or quarterly intervals. The data can also be divided by the geographical area of aggregation, that is, by region, country or global, naturally the use of the global variables is common for the several countries estimation, these include mostly the financial indexes measured through the closing prices as well as the composite indicators for the global economy. As a way to operacionalize the estimation of MIDAS using these datasets, the regions were grouped by country and the global datasets applicable to every country-wide model.

The countries initially considered for this article was the whole set of the 27 EU countries, it was only later that, due to the accumulation of data constrigencies for some of the countries, that the scope has been reduced. The decision to exclude some of the countries that did not have the full dataset as an alternative to performing the estimation and the subsequent analysis for these countries using a majorly different dataset, was justified to preserve the comparability of the forecasts across the countries included. It is worth noticing that although only the final models accuracy measure will contribute to the comparison of forecasting ability

by country, the use of the same independent variables for all the estimates will drastically reduce the risk that the difference in results is originated by the difference in the datasets. Consequently this will increase the quality analysis of the explanatory variables predictive content on each of the countries and the degree to which the relevant variations of these variables are captured by the MIDAS process. It was also for this reason that during the creation and obtaining of the datasets it was avoided to source from national statistics and given preference to systems that compile data for several regions and countries under the same methodology or one that has contained differences such as European wide repositories. Nonetheless, there was an effort to capture a geographically, economic and cultural representative sample of the European countries which resulted in the following list of 14 countries – Austria, Belgium, Denmark, Estonia, France, Hungary, Italy, Ireland, Lithuania, Netherlands, Poland, Portugal, Spain, Sweden.

The selection of the explanatory variable datasets was done taking into account the existing literature for cases similar to the use given in this dissertation. The process to gather and select variables was heavily inspired by the use of a comprensive multifaceted dataset as in the case of Lehmann & Wohlrabe (2015). Given the approach to build a comprehensive dataset, it was considered the inclusion of other composite indicators. For this set, the composite indicators compiled and used in Baumeister, & Guérin (2021) were sourced - World Industrial Production Index (WIP); Global Steel Production; Kilian Index; Real Commodity Price Factor; Global Economic Conditions (GECON).

Furthermore, and to further explore the readly available use of high-frequency variables, other coincident indicators were considered for the estimation exercise namely electricity consumption or the flight passengers data. This information can have particular relevance and contribute to the models accuracy at stages where severe institutional changes take place which is the case of those introduced in the context of the Covid-19 pandemic. To advise for these changes and to justify part of the dramatic variations in GDP occurred in this time, the additional boolean variable for the covid period was considered defined between March 2020 and June 2022 as the literature gathered did not find statistically significant relations between government imposed lockdowns and economic activity as measured through the e Oxford COVID-19 Government Response Tracker's Stringency Index Cross, Ng & Scuffham (2020). Furthermore, the description of the full dataset considered can be seen in detail in the annex 1.

### 3.2. Estimation Methods

For the MIDAS estimations, some choices had to be made regarding the functional constraints to apply, the weights decisions for the functional constraints, the parameterization of the exponential almon lag and the starting weights for NLS. Given the intent of this article to allow the comparison of the performance of these models as directly as possible throughout the countries considered, it was given priority to the estimation methodology generalization to all the geographies, while simultaneously taking into consideration the uses

and choices that have been done in the literature so far. By giving priority to these aspects there is a natural trade-off of the research's goal requirements to have the most similar model's features across the nationalities with the individual quality of the forecasting methods. Provided this point, it is desirable to minimize the chances that the model's performance is affected by different model's structure, therefore it is with that much more confidence that it can be claimed that the difference in the loss function's evolution is not caused by the luck picking the parameters that maximize the forecasting accuracy for the each country. It was decided instead to go with the combination of hyperparameters that would minimize globally the loss function on the entire dataset for the tested combinations. Due to convergence issues in some of the window estimations for some countries that represent for less than 2% of all estimations performed and alternative MIDAS model was considered with the exact same variables, parameter selection, estimation methodology and functional constraints as the original one apart from the initial weights selected. After experimenting with these two versions of the model it is not clear whether this initial weight has impact on the countries quality of estimation.

On a more formal but not new disposition of the MIDAS structure considered for the exercise takes the functional equation displayed in *Equation 1*:

$$y_{t+h} = \beta_0 + \lambda y_t + \beta_1 B\left(L^{\frac{1}{3}}; \theta; k; p\right) x_{t-w}^{m} + \varepsilon_{t+h} \quad (1)$$

Where the $y$ denotes the outcome variable, while the subscript $t+h$ the period of the forecasting, $\beta_0$ is the intercept and the $\lambda$ represents the coefficient of the autoregressive term $y_t$. Furthermore, the $w = T_x - T_y$ denotes a negative fraction representing the difference between the time that the independent explanatory variables are observed and the moment at which the interest variable is observed in relation to the high-frequency periodicity. As an example, to mention any monthly observation occurred one month before the end of the quarter it could be recognized $w = \frac{1}{3}$. The representation for the noise in the model at the same frequency observation is the $\varepsilon_{t+h}$ parcel. While $x^m$ represents the explanatory of index $m$, the B function represents a lag polynomial defined as follows:

$$B\left(L^{\frac{1}{3}}; \theta; k; p\right) = \sum_{k=0}^{K} \psi(k; \theta; p)\, L^{k/3} \quad (2)$$

In this specification of the lags polynomial, the reference to the $\psi$ function is made as a representation of the functional constraint applied to the model. The U-MIDAS specification would not contemplate this function and the simple lagged polynomial would be used. The $L^{k/3}$ factor of the equation represents the high-frequency lag operator in a way that $x_{t-1/3}^{m} = L^{1/3} x_t^m$.

The use of the functional constraints chosen in the application of this methodology as the Almon Exponential due to its popularity in the literature even though there is not always a

consensus in it's relative superiority to other functional contraints such as the Beta distribution. Below, the definition of this function:

$$\psi(k;\theta;p) = \frac{\exp\left(\sum_{p=1}^{P}\theta_p k^p\right)}{\sum_{j=0}^{J}\exp\left(\sum_{p=1}^{P}\theta_p k^p\right)} \quad (3)$$

By construction, the weights add up to one as $J$ represents the maximum lag order of the independent variable. Furthermore, $p$ is a natural number whose value is manually determined which determines the liberty to the approximate any of distributed lags' coefficients for variable $k$. The selection of this parameter must be taken with caution because the higher the $p$ in relation to the amount of lags of the variable considered in the model represented by $J$, the more agile and nonlinear the relation between the variable's distributed lag coefficients can be. On the other side, the smaller the difference between $J$ and $p$ the less degrees of freedom are available for this estimation. Therefore, by constraining the distributed lags coefficients to have a more linear relationship amongst themselves hence providing greater degrees of freedom seems to be the most conservative approach to this selection and the one adopted in this exercise.

The forecasting exercise should comprehend a rolling estimation window that comprehends five years, the rolling estimation method was chosen to allow for a more robust estimation of the true out of sample RMSE in comparison to the separation by the estimation data (training) and the forecasting data (out of sample). In the selection of the window size in itself one must be careful due to the potential trade-off between bias and variance due structural breaks in the series as explained by Clark & McCracken (2009) where they propose a convex combination of recursive and rolling forecasts. In other words, by considering an estimation window too large there may arise severe differences between the in-sample and out of sample RMSE, however, if it is too small the estimators may be inefficient which would themselves result in a poor estimates. Given that there is no strict method for the definition of the window in the literature it was decided the 5 year mark as a mid-point of the cases collected in the literature.

Furthermore, the standard ARIMA(p,d,q) model will serve as a benchmark for the quality of the overall information for the specific region and time frame. Besides serving as a benchmark for the individual series, using the univariate process will allow to more easily identify the causes of some deeps or surges in estimation accuracy. The parameterization of the ARIMA will be done for the same windows as the model of interest so that the out of sample RMSE relate to the same data. Finally, the benchmark ARIMA model selection process is done through the Bayesian Information Criterion.

As attested in the early analysis of the collected variables, some of them showed a very noticeable trend just by looking into the plots. This is not unexpected in the slightest for these type of data, and so, to have a quantifiable and validated standard test for this aspect, the Augmented Dickey-Fuller test for unit root was performed. The results are presented in the annex 2 and their discussion will be done hereunder in the appropriate results segment.

Even though the Augmented Dickey-Fuller serves to test for a unit root in a time series similarly to the standard Dickey-Fuller, the augmented version of the test assumes that there is auto-correlation higher than the first order. Furthermore, for the determination of the maximum auto-correlation lag it was used the suggested upper bound on the rate at which the number of lags should be forced to increase with the sample size. Below an operationalization of this concept where $Z$ is the rounded maximum order of the lags to be included in the model and let $i$ define the length of the series vector

$$Z \approx \sqrt[3]{i-1} \tag{4}$$

$$Z \in \mathbb{N}$$

For the countries interest variable series who's Augmented Dickey-Fuller null hypothesis test was not rejected, and so the failure to reject the hypothesis that the series is Non-Stationary, it was included a trend variable of the same length as the quarterly variables series in the estimation windows in the estimation of the model. On the Contrary, for the series who's Augmented Dickey-Fuller test null hypothesis was rejected, no further considerations for the trend were made.

### 3.3. Model Validation and Error measuring

For the evaluation of the forecasting quality the key analyzed measure is the out of sample RMSE. The period is from the first quarter of 2017 to the end of the first semester in 2021. Below the formula for this result:

$$RMSE = \sqrt{\sum_{t=1}^{T}(y_t - \hat{y}_t)^2} \tag{5}$$

Where $y_t$ and $\hat{y}_t$represent respectively the actual value and the forecasted one at period $t$ of low frequency. The calculation of the forecasting horizon will be the 2-period ahead.

The discussion and development of the testing and validation of the models performance will always be concerned with maintaining robustness in the face of shocks and other factors that may moderate the relationship between the variables over time. Furthermore, to reach decisive results the Diebold-Mariano test for superior predictive ability will be used between the benchmark and the interest model for the entire collection of the out of sample residuals. This test will provide greater insights to the regional disposition of using MIDAS methodology in comparison to the univariate process. In the two sided test the alternative hypothesis consists in rejecting the null hypothesis in favor of the alternative where it can be accepted that the methods have different forecasting abilities, the one sided test aims at rejecting the null in favor of the alternative hypothesis on which the MIDAS model has a better predictability than the one produced by the univariate process. Finally, this test will consider 95% threshold for the confidence level.

Besides these results it is important to understand the evolution of the loss function and the RMSE ratios throughout the series for the individual countries as well as an aggregated view. Another key analysis is the assessment of the global RMSE out of sample average and standard deviation so that more considerations of the MIDAS process as a whole can be made. For instance, in a case where the average has a considerable variation but the standard deviation is unaltered it can be deduced that the generality of the models had a change in the proportion inversely to the RMSE's movement.

In order to help the assessment, the overall analysis and data visualization of the forecasting quality results, the countries analysis have been mostly compared by their geographical proximity. The first group can be described as the Southern Europe group whose members are Italy, Portugal and Spain, the Eastern Europe group – Estonia, Hungary, Lithuania and Poland, the Central Europe group – Austria, Belgium, France and the Netherlands and the remaining group, Northern Europe which concerns Sweden, Denmark and the perhaps out of place addition of Ireland that can be justified by the closer proximity to the group than any other.

# 4. RESULTS

## 4.1. STATIONARITY OF THE GDP

In this section it will be discussed the results of the Augmented Dickey-Fuller tests done on the interest variable, the GDP. As mentioned in the methodology section, the execution of the test considers the entire series available for the country at the extraction time of the data and the confidence level interval considered for the test is that of 95%.

The results of the test clearly show that there is no room for the rejection of the null hypothesis except for the Austrian series which is rejected a p-value of less than 0,01. The French series also comes close to having the same result as the Austrian one however it falls short with a p-value of ≈ 0,06. The next closest series to be rejected is Estonia with a p-value of 0,14. It can be questionable if the choice of the Augmented Dickey-Fuller instead of the standard one would have changed the outcome of some of these results. By comparing both test results, the French one is the only series close enough to have the test's decision changed, the standard Dickey-Fuller test resulted in a p-value smaller than 0,01. The summary table that contains the maximum lag order used, the test statistic value and the p-value of the Augmented Dickey-Fuller test can be visited in the annex 2.

## 4.2. EVOLUTION OF THE RMSE RATIO

The first aspect that comes to the attention when assessing of the evolution of RMSE ratio between the MIDAS forecast and the ARIMA one is the appearence of some peaks of the ratio for some of the countries at some point in time. These apparently spurious events make it hard to visualize the series evolution and compare them with other countries, nonetheless the graphical representation uses a logarithmic scale for an easier visualization of this evolution that can be seen in the lowest part of this segment in Figure 1.

For the Italian case, in the third quarter of 2017 and even though both the benchmarks and the MIDAS RMSE had climbed significantly, it was not proportionate enough to keep the ratio stable given that the univariate process tripled its RMSE and the MIDAS procedure had it multiplied it by 461. This surge is not carried out in the trend as the MIDAS RMSE quickly comes back to normal levels while the ARIMA ones remain high for some time after. The next important case to be discussed is the Estonian one which, in comparison to it's neighboring countries has a surge in two particular moments of the series - one in 2018 and the other one prolongs from the end of 2019 until the end of the series, contrarily to the previous case, in these two instances of the time series, the RMSE variations are significantly different and so while the benchmark gains accuracy, the MIDAS process seems to have a severe dip in its forecasting quality.

The remaining outliers are not as nearly as severe as the ones just covered but they still stand out in relation to their neighbors statistics. The Belgium ratio reaches the 100 mark in dramatic fashion given that in the previous observation, at 2017 Q1, the MIDAS observed

RMSE was more than 1000 times smaller than the ARIMA model. Furthermore, this ratio in particular climbs dramatically from the 1st quarter of 2020 until the end of the series. The final case worthy of mentioning is the Swedish case where the observed ratios indicate that the model has had a better performance than it's ARIMA competitor for most of the series apart from the incident observed in the first quarter of 2020 where the loss in accuracy of the MIDAS model is contrasted by a variation in the Oposite side by its competitor.

Besides the discussion on the outliers done above, it is pertinent to assess the moments where the observed RMSE ratio gives the advantage to the MIDAS, that is, when the ratio is below one. This analysis does not replace the results of the Diebold-Mariano tests for Superior Predictive Ability that are discussed in the appropriate section below even though the exploration of this ratio may me and indicator to the test's results. When looking into the country groups that appear to have shown the greatest gains from using the MIDAS the two groups that stand out are the Eastern and Northern group where the average seldomly situates above the unit mark. When the average climbs, it is mostly related to the deterioration of the forecasting quality in one of the group's countries rather than a common behavior in the entire group. This can be seen in the Eastern European group in the beginning of 2020 where average increases not representing necessarily the quality of the group which apart from the Estonian case has a downward trend.

Finally, while in the overall set of countries there is no evidence of a trend, by performing a regional disaggregation it might not appear that way, such as the Southern Europe case where there is a clear negative trend for all its members. This trend becomes clearer when looking into the group's average RMSE path namely in the the latest stages of the serie for which the greatest contributions seem to have been the RMSE decline of Italian and Portuguese series when comparing to their values at the start of the series. The Central Europe group can be characterized by its ratio's variation and the performance disparity between the member countries at any given moment of the series, this does not hold true however in the final year of the series where the ratio climbs consistently for all its members.

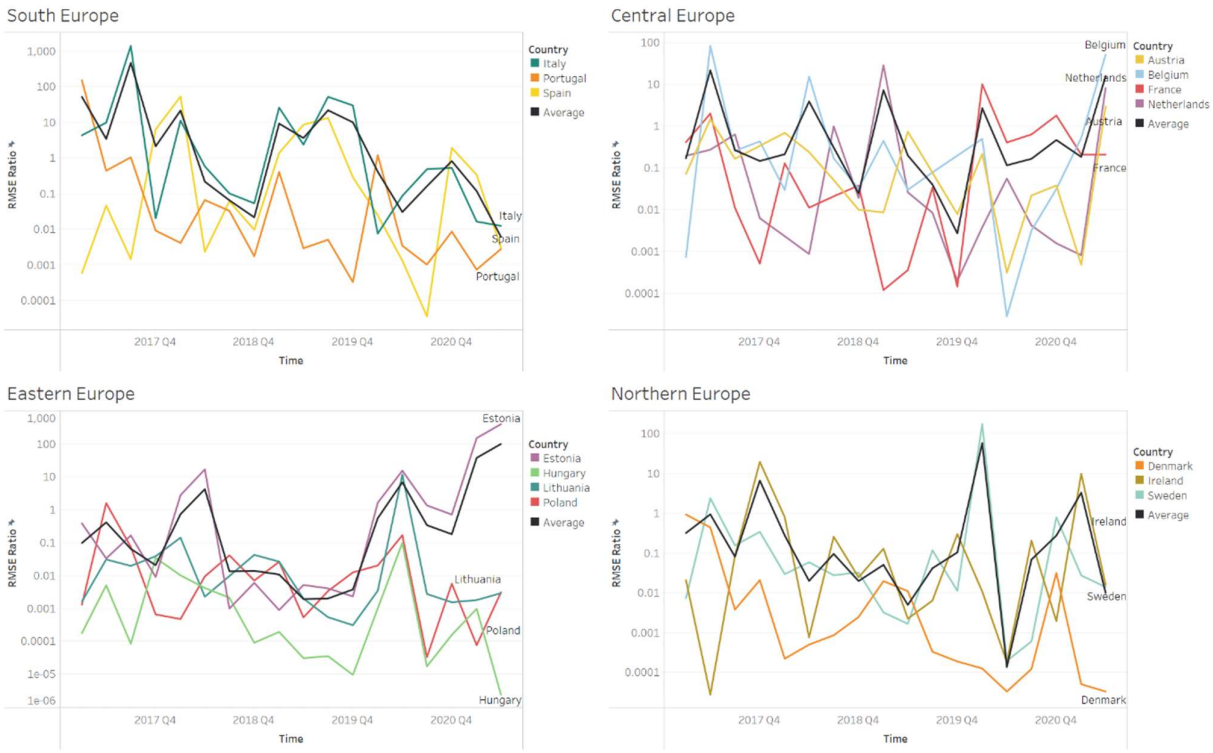


*Figure 1.*

*Ratio between the MIDAS and ARIMA RMSE by Geographical region.*

### 4.3. EVOLUTION OF THE MIDAS FORECASTING QUALITY

The above segment has helped in the analysis of the inter and intra group differences for the interest model in comparison to the benchmark model. Nonetheless, it is not given enough attention to the analysis of the global differences between the MIDAS model's at any moment in time. For this purpose, this segment is introduced to analyze the RMSE's evolution where the average and country wide variance have been captured for any of the out of sample estimation periods. By looking into these statistics it is expected to provide better context for the periods where the model's quality has suffered the most and where it was the most differentiated.

Considering the graphical representation in Figure 2, it becomes clear that there is a positive association between the two measures. This can indicate that the average RMSE was affected by one or more countries models that have deviated heavily from the mean instead of a generalized trend. For the first spike in 2017 Q4 although half of the countries have had a negative variation in relation to the previous quarter, Hungary stands out by contributing the most to this variation given that in gross RMSE variations is of +11 035 327, 57. It may be quickly noted that this variation does not translate to the increase of the ratio in the same proportion, this is due to bad performance of its ARIMA competitor whose RMSE quarterly

variation was of +204 536 875,99. Yet and the despite the ARIMA variation being higher in absolute values the growth of the ratio is caused by the relative growth of the MIDAS RMSE (114990%) in relation to the ARIMA one (182%).

For the next significant deviation occurred in the first quarter of 2019, the divergence's main responsible the Netherlands (+3 642 503,21) in which case it is already reflected in the ratio change as the ARIMA despite weakening its performance, did not climb in the same fashion as its competitor. The final peak occurred in the first quarter of 2020 and relates to Swedish case also covered in the previous segment given that the variation of the MIDAS model was not accompanied by it's competitor which in fact gained performance. The full list with the absolute variations disaggregated by country can be seen in the annex tables 3, 4 and 5.

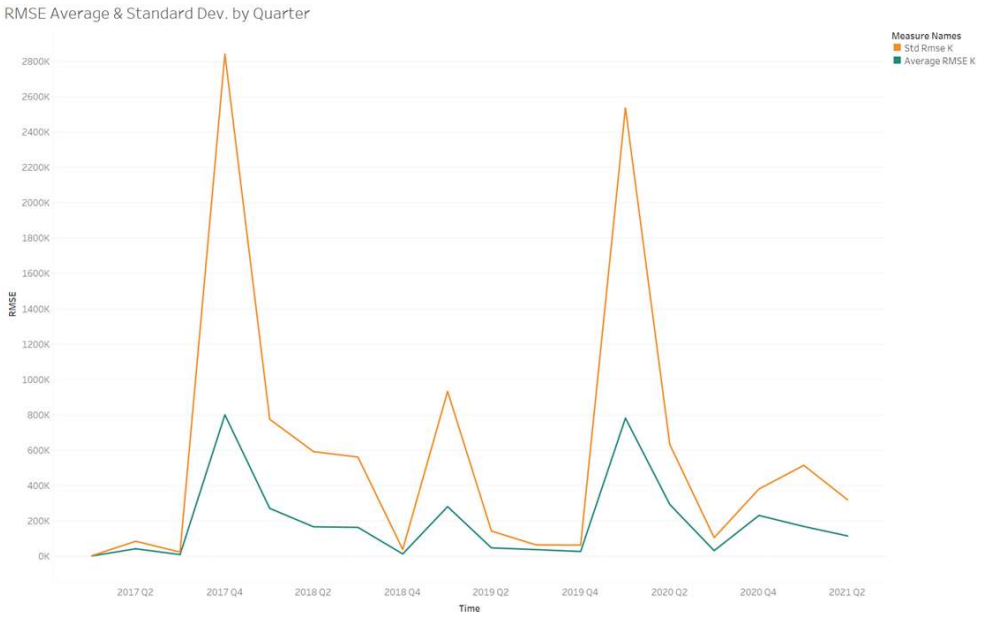


*Figure 2.*

*RMSE Quarterly Average and Population's Standard Deviation.*

### 4.4. DIEBOLD-MARIANO TEST

In regards to the Diebold-Mariano for the Superior of Superior Predictive ability test it is considered the entirety of the out of sample residuals and is a one sided given that the desirable outcome of the test for the superiority of the MIDAS in relation to the benchmark model.

For the majority of the countries considered, the results seem to suggest that there is only enough evidence at the 5% significance level for the rejection of the null for a constrained number of countries. On one side, there are the countries that clearly outperform the benchmark model – Denmark, Ireland, Lithuania, Hungary and Poland. Even though there is a small number of countries that managed to pass this threshold in the generality of the out of

sample testing at the 5% significance level, the majority of the countries fall under the 10% mark, and so adding to the previous list it can be included Austria, France, Italy, and Portugal. The remaining set can still be separated into two groups, the ones where the p-value is below the 35% significance which includes Belgium, Spain and Sweden, the other group of countries – Estonia and the Netherlands would have been closer to the rejection of the null hypothesis if the alternative were the acceptance that the univariate process has a superior predictive ability with p-values surrounding the 87% and 84% levels.

With these results there is no clear trend on the regional disposition of the relative forecasting accuracy of the MIDAS model in relation to the univariate process. The p-values relative to the Diebold-Mariano can be visited in the annex 6 for further detail.

## 5. CONCLUSIONS

In the aftermath of the results analysis and in the light of the MIDAS methodology regional disposition forecasting accuracy assessment, it can be said that there is enough evidence to support the idea that MIDAS accuracy has not been equaly competitive in relation to the ARIMA process for all the countries considered. Furthermore, the countries that did not present the greatest gains from using the MIDAS were also heavily influenced by the significant deterioration of the forecasting quality at moments in time where their ARIMA counterpart stayed relatively stable. The Estonian, Swedish and Dutch cases are the best examples where, due to these performance outliers more often than not, they came really far from rejecting the null hypothesis in the Diebol-Mariano test that would give some statistical significance to the acceptance of the MIDAS model as the more accurate process. On the other side, the Polish, Hungarian and Lithuanian series have had the biggest margin for the rejection of equal predictability.

Apart from these considerations, when the regional groups behavior are considered it is not clear that the relative performance has been favored in one group of countries in relation to another. These results have without margin of doubt been affected by the respective outliers seen in each of the group. This does not mean that there were no interesting patterns or conclusions arising from this grouping as they are geographically bound and therefore immutable apart from the inclusion of the Irish series in the Northern group. While the Southern group showed a consistent decline in the average RMSE ratio, the Central group should be characterized by it's variability and the likeliness of finding the average RMSE ratio near the unitary threshold. The Eastern region was the most consistent when looking at its intra group variance and would have probably been the only group to distinctly show benefits for the use of the MIDAS regression if it was not for the worsening of the Estonian forecasting quality at the end of the series.

When taking into account the pooling of the MIDAS RMSE, the periods of greater average RMSE are also the periods where the models forecasting accuracy differ the most. After carefully analyzing these moments there can be distinctly identified the main country contributors that caused both the average increase as well as the increased standard deviation.

In this exercise some limitations can be deduced from the comparison of the MIDAS process exclusively to the best ARIMA model selected through the BIC (Bayesian Information Criterion) as opposed to other non-univariate models. Besides, this study focuses on the comparative forecasting ability and dismisses the in sample RMSE as well as the estimates evolution.

In the holistic view of the objectives proposed and carried out in this dissertation, it can be concluded that there has been value added in directing future research to the deepening of the insights of the underlying reasons that cause the differences in the estimation quality

of the MIDAS process. The comparison of the MIDAS gains can and should be extended other competing models such as the Dynamic Factor Models, Bridge Regressions among others. Furthermore, the experimentation with a different and more complete dataset is all the more interesting and possibly introduce further stability to the MIDAS RMSE.

## BIBLIOGRAPHY


Almon, S. (1965). The distributed lag between capital appropriations and expenditures. Econometrica: Journal of the Econometric Society, 178-196.

Baffigi, A., Golinelli, R., & Parigi, G. (2004). Bridge models to forecast the euro area GDP. International Journal of forecasting, 20(3), 447-460.

Baumeister, C., & Guérin, P. (2021). A comparison of monthly global indicators for forecasting growth. International Journal of Forecasting, 37(3), 1276-1295.

Bhattacharyya, D. K. (1999). On the economic rationale of estimating the hidden economy. *The Economic Journal*, *109*(456), F348-F359.

Blinov, S. (2017). *Economic Forecasting Based on the Relationship between GDP and Real Money Supply*. University Library of Munich, Germany.

Brautzsch, H. U., & Ludwig, U. (2002). *Vierteljährliche Entstehungsrechnung des Bruttoinlandsprodukts für Ostdeutschland: Sektorale Bruttowertschöpfung* (No. 164/2002). IWH Discussion Papers.

Chow, G. C., & Lin, A. L. (1971). Best linear unbiased interpolation, distribution, and extrapolation of time series by related series. *The review of Economics and Statistics*, 372-375.

Clark, T. E., & McCracken, M. W. (2009). Improving forecast accuracy by combining recursive and rolling forecasts. International Economic Review, 50(2), 363-395.

Clements, M. P., & Galvão, A. B. (2009). Forecasting US output growth using leading indicators: An appraisal using MIDAS models. Journal of Applied Econometrics, 24(7), 1187-1206.

Clemenys, M.P. & Galvão A. (2011). Macroeconomic Forecasting with Mixed Frequency Data: Forecasting Us Output Growth and Inflation, Journal of Business and Economic Statistics, 26(4), 546-554.

Coutiño, A. (2005). On the use of high-frequency economic information to anticipate the current quarter GDP: A study case for Mexico. *Journal of Policy Modeling*, *27*(3), 327-344.

Cross, M., Ng, S. K., & Scuffham, P. (2020). Trading health for wealth: The effect of COVID-19 response stringency. *International Journal of Environmental Research and Public Health*, *17*(23), 8725.

Diebold, F. X. (2015). Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of Diebold–Mariano tests. *Journal of Business & Economic Statistics*, *33*(1), 1-1.

Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. Journal of Business & economic statistics, 20(1), 134-144.

Duarte, C. (2014). *Autoregressive augmentation of MIDAS regressions* (No. w201401).

Eraslan, S., & Götz, T. (2021). An unconventional weekly economic activity index for Germany. *Economics Letters*, *204*, 109881.

Frank, L. E., & Friedman, J. H. (1993). A statistical view of some chemometrics regression tools. *Technometrics*, *35*(2), 109-135.

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2004). The MIDAS touch: Mixed data sampling regression models.

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

Hansen, P. R. (2005). A Test for Superior Predictive Ability. Journal of Business and Economic Statistics, 23 (4), 365–380

Harvey, D., Leybourne, S., & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of forecasting*, *13*(2), 281-291.

Jardet, C., & Meunier, B. (2022). Nowcasting world GDP growth with high-frequency data. *Journal of Forecasting*.

Jiang, Y., Guo, Y., & Zhang, Y. (2017). Forecasting China's GDP growth using dynamic factors and mixed-frequency data. *Economic Modelling*, *66*, 132-138.

Lehmann, R., & Wohlrabe, K. (2014). Regional economic forecasting: state-of-the-art methodology and future challenges.

Lehmann, R., & Wohlrabe, K. (2015). Forecasting GDP at the regional level with many predictors. German Economic Review, 16(2), 226-254.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica: Journal of the Econometric Society*, 867-887.

Mikosch H. & Solanko L., (2019). Forecasting Quarterly Russian GDP Growth with Mixed-Frequency Data, Russian Journal of Money and Finance, Bank of Russia, 78(1), 19-35.

Mikosch, H., & Zhang, Y. (2014). Forecasting Chinese GDP growth with mixed frequency data: Which indicators to look at?. KOF Working Papers, 359.

Pesaran, M. H., & Timmermann, A. (2007). Selection of estimation window in the presence of breaks. *Journal of Econometrics*, *137*(1), 134-161.

Rossi, B., & Inoue, A. (2012). Out-of-sample forecast tests robust to the choice of window size. *Journal of Business & Economic Statistics*, *30*(3), 432-453.

Rünstler, G., & Sédillot, F. (2003). *Short-term estimates of euro area real GDP by means of monthly data* (No. 276). ECB working paper.

Sédillot, F., & Pain, N. (2003). Indicator models of real GDP growth in selected OECD countries.

Schumacher, C. (2014). MIDAS and bridge equations. *Available at SSRN 2797010*.

St Aubyn, M. (2020). Covid 19 and loss of production–an estimate for Portugal from electricity consumption (No. 01/2020). Portuguese Public Finance Council.

Stock, J. H., & W Watson, M. (2003). Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature*, *41*(3), 788-829.

Stock, J. H., & Watson, M. W. (1989). New indexes of coincident and leading economic indicators. *NBER macroeconomics annual*, *4*, 351-394.

White, H. (2000). A Reality Check for Data Snooping. Econometrica, 68 (5), 1097–1126.

## APPENDIX

### ANNEX 1

## *Full Dataset*

| *Code* | *Meaning* | *Frequency* | *Geographical Classification* |
|---|---|---|---|
| *gdp_full* | GDP at Current Prices by Country | Quarterly | Country |
| *encon_full* | Electric Energy Consumption by Country | Monthly | Country |
| *global_indicators* | Global Indicators | Monthly | World |
| *industrial_production* | Industrial Production index by country | Monthly | Country |
| *euro_stoxx* | Euro_stoxx closing price index | Monthly | N/A |
| *m_mass* | Monetary Mass Aggreagates (M1, M2, M3) | Monthly | N/A |
| *hpi* | House Price index by Country | Quarterly | Country |
| *motorcars* | Land vehicles purchase orders | Monthly | Europe |
| *rate_unemploy* | Unemployment Rates by Country | Quarterly | Country |
| *rates_eur* | Cambial Exchange with 42 other currencies (base EUR) | Monthly | N/A |
| *smi* | SMI index closing price | Monthly | N/A |
| *cac* | CAC index closing price | Monthly | N/A |
| *dax* | DAX index closing price | Monthly | N/A |
| *ibex* | IBEX index closing price | Monthly | N/A |
| *ftse* | FTSE closing price index | Monthly | N/A |
| *brent* | Closing price of the brent benchmark | | |
| *flights* | Number of Passangers landed by Country | Monthly | Country |

**ANNEX 2**

| *Country* | *Lag Order* | *Test Statistic* | *P-Value* |
|---|---|---|---|
| *Austria* | 4 | -4,2854 | 0,01 |
| *Belgium* | 3 | -2,4448 | 0,3977 |
| *Denmark* | 4 | -1,9919 | 0,58 |
| *Estonia* | 4 | -3,0503 | 0,141 |
| *France* | 5 | -3,3961 | 0,05702 |
| *Hungary* | 4 | 1,183 | 0,99 |
| *Ireland* | 4 | -0,093927 | 0,99 |
| *Italy* | 4 | -2,035 | 0,5621 |
| *Netherlands* | 4 | -1,932 | 0,6048 |
| *Lithuania* | 4 | -2,1023 | 0,5342 |
| *Poland* | 4 | 1,3404 | 0,99 |
| *Portugal* | 4 | -2,3576 | 0,4283 |
| *Spain* | 4 | -1,9021 | 0,6173 |
| *Sweden* | 4 | -1,3042 | 0,8654 |

Results of the Augmented Dickey-Fuller test for the countries on the left. In this table it can be seen the maximum significance level allowed for the rejection of the null for each of the series. This tests alternative hypothesis is that the series is a stationary process.

**ANNEX 3**

## *Quarterly variation of RMSE by country – 1st Peak (2017 Q4)*

| *Country* | *MIDAS RMSE Quarterly Variation* | *ARIMA RMSE Quarterly Variation* |
|---|---|---|
| *Belgium* | 4,17 | 41,44 |
| *Denmark* | 2 335,33 | 72 669,83 |
| *Estonia* | - 10,26 | 138,34 |
| *Ireland* | 4 273,10 | - 1 512,21 |
| *Spain* | 144 001,98 | 12 913,49 |
| *Italy* | - 90 230,68 | 25 582,92 |
| *France* | - 824,80 | 62 778,27 |
| *Lithuania* | 4,32 | - 75,68 |
| *Hungary* | 11 035 327,57 | 204 536 875,99 |
| *Netherlands* | - 2 337,10 | 4 078,27 |
| *Austria* | 37 361,52 | 13 591,32 |
| *Poland* | - 30 591,09 | - 59 473,45 |
| *Portugal* | - 430,23 | 1 020,86 |
| *Sweden* | 5 910,35 | 13 634,09 |

Table containing the absolute variations of the forecasting RMSE for the MIDAS model and ARIMA at 2017 Q4. In this table it can be seen how the individual countries contributed to the variation in the global scope at the specific moment in time where the MIDAS RMSE explodes.

**ANNEX 4**

## *Quarterly variation of RMSE by country - 2nd Peak (2019 Q1)*

| ***Country*** | ***MIDAS RMSE Variation*** | ***ARIMA RMSE Variation*** |
|---|---|---|
| *Belgium* | 1 450,43 | 913,96 |
| *Denmark* | 12 534,30 | 431 806,89 |
| *Estonia* | - 3,37 | - 184,84 |
| *Ireland* | 398,08 | - 2 372,70 |
| *Spain* | 78 099,64 | 910,06 |
| *Italy* | 10 954,51 | - 18 424,33 |
| *France* | - 17 531,27 | - 64 275,85 |
| *Lithuania* | - 33,26 | - 293,69 |
| *Hungary* | - 37 806,04 | - 491 036 359,33 |
| *Netherlands* | 3 642 503,21 | 86 023,87 |
| *Austria* | - 38,69 | - 3 444,74 |
| *Poland* | 57 016,43 | 942 558,76 |
| *Portugal* | 800,39 | - 3 712,43 |
| *Sweden* | - 1 349,80 | 635 366,05 |

Table containing the absolute variations of the forecasting RMSE for the MIDAS model and ARIMA at 2019 Q1. In this table it can be seen how the individual countries contributed to the variation in the global scope at the specific moment in time where the MIDAS RMSE explodes.

**ANNEX 5**

## *Quarterly variation of RMSE by country – 3rd Peak (2020 Q1)*

| *Country* | *MIDAS RMSE Variation* | *ARIMA RMSE Variation* |
|---|---|---|
| *Belgium* | 654,50 | 1 040,26 |
| *Denmark* | - 9,20 | 326 293,50 |
| *Estonia* | 275,07 | - 748,39 |
| *Ireland* | - 36,86 | 1 712,09 |
| *Spain* | - 16 492,14 | 50 716,48 |
| *Italy* | - 232 016,49 | 1 035 595,23 |
| *France* | 823 642,61 | - 542 787,09 |
| *Lithuania* | 5,50 | - 1 087,36 |
| *Hungary* | 79 924,90 | - 2 380 769 046,29 |
| *Netherlands* | 415,25 | - 7 506,74 |
| *Austria* | 1 631,78 | 3 879,56 |
| *Poland* | 40 807,04 | 243 857,52 |
| *Portugal* | 515,02 | - 8 459,15 |
| *Sweden* | 9 892 643,47 | - 79 544,14 |

Table containing the absolute variations of the forecasting RMSE for the MIDAS model and ARIMA at 2020 Q1. In this table it can be seen how the individual countries contributed to the variation in the global scope at the specific moment in time where the MIDAS RMSE explodes.

**ANNEX 6**

## *Diebold-Mariano test Results*

| ***Country*** | ***P-value MIDAS > ARIMA*** |
|---|---|
| *Belgium* | 0,145717 |
| *Denmark* | 0,003010 |
| *Estonia* | 0,866961 |
| *Ireland* | 0,015622 |
| *Spain* | 0,177464 |
| *France* | 0,051840 |
| *Italy* | 0,081076 |
| *Lithuania* | 0,002453 |
| *Hungary* | 0,000455 |
| *Netherlands* | 0,842892 |
| *Austria* | 0,068940 |
| *Poland* | 0,000126 |
| *Portugal* | 0,053840% |
| *Sweden* | 0,346423 |

Results of the Diebold-Mariano test for the countries on the left-most column's series. In this table it can be seen the maximum significance level allowed for the rejection of the null for each of the series. This tests alternative hypothesis is that the MIDAS has a superior predictive ability in relation to the ARIMA process.