CHARLES UNIVERSITY
FACULTY OF SOCIAL SCIENCES
Institute of Economic Studies

GDPNow for the Czech Republic

Master's thesis

Author: Bc. Jan Kutman
Study program: Economics and Finance
Supervisor: prof. PhDr. Tomas Havranek, Ph.D.

Year of defense: 2022

Declaration of Authorship

The author hereby declares that he or she compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and agrees with the thesis being used for study and scientific purposes.

Prague, January 2, 2022

Jan Kutman

Abstract

The gross domestic product (GDP) is an essential measure of the state of economic activity and serves as a crucial tool for policymakers, investors, or businesses. However, the official GDP estimate in the Czech Republic is only available with a lag of approximately 60 days, and the Czech National Bank (CNB) announces its GDP forecast once in each quarter. This thesis focuses on predicting GDP growth in the current quarter, referred to as nowcasting. I employ several methods to nowcast the real GDP growth in the Czech Republic in a pseudo-real-time setting and compare their performance. Additionally, I investigate the possibility of creating an ensemble model by using a weighted average of several nowcasting models. The results suggest that the Dynamic Factor Model (DFM) performs best in the GDP nowcasting task, and its predictive accuracy is comparable with the official CNB nowcast. Furthermore, the model averaging process yields accuracy close to the best individual model while addressing model uncertainty. The GDP nowcast of the DFM will be made available to the public in real-time on a website and updated with a daily frequency.

JEL Classification	F12, F21, F23, H25, H71, H87
Keywords	GDP Nowcasting,Forecasting,Forecast Combination, Dynamic Factor Model
Title	GDPNow for the Czech Republic

Abstrakt

Hrubý domácí produkt (HDP) je základním měřítkem stavu ekonomiky a slouží jako klíčový ukazatel pro zákonodárce, investory nebo podniky. Česká Národní Banka (ČNB) ale zveřejňuje oficiální odhad HDP až přibližně 60 dní po konci daného čtvrtletí, přičemž předpověď budoucího růstu HDP zveřejňuje ČNB vždy jen jednou v každém čtvrtletí. Tato diplomová práce se zaměřuje na predikci HDP v současném čtvrtletí, tzv. *Nowcasting*. Používám několik metod na průběžnou předpověď HDP v České republice a porovnávám jejich výsledky. Zkoumám také možnosti kombinace několika modelů za použití váženého průměrování jejich předpovědí. Výsledky ukazují, že Model Dynamického Faktoru předpovídá růst HDP v České republice nejlépe ze všech použitých modelů a přesnost jeho odhadů je srovnatelná s oficiálními predikcemi

ČNB. Kombinace sedmi nejlepších individuálních modelů dosahuje přesnosti predikcí obdobné s nejlepším individuálním modelem a zároveň řeší problém nejistoty při výběru optimálního modelu. Predikce HDP získané modelem dynamického faktoru plánuji zveřejnit na internetové stránce a denně aktualizovat.

Klasifikace JEL	F12, F21, F23, H25, H71, H87
Klíčová slova	Predikce HDP,Dynamické modelování, Kombinace Předpovědí
Název práce	GDPNow pro Českou Republiku

Acknowledgments

I would like to express my gratitude to my supervisor prof. PhDr. Tomas Havranek, Ph.D.for his help, encouragement, feedback, and guidance throughout the writing of this thesis.

I would also like to thank my family for their support and patience, which they showed while writing the thesis and during my entire studies.

This thesis is part of a project that has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 870245.

Typeset in LaTeXusing the IES Thesis Template.

Bibliographic Record

Kutman, Jan: *GDPNow for the Czech Republic.* Master's thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2022, pages 93. Advisor: prof. PhDr. Tomas Havranek, Ph.D.

Contents

List of Tables viii

List of Figures ix

Acronyms xi

Thesis Proposal xii

1 Introduction 1

2 Literature Review 5

2.1 Macroeconomic nowcasting as a discipline . . . 5

2.2 Empirical Literature about GDP Nowcasting . . . 9

2.3 GDP nowcasting in the Czech Republic . . . 12

2.4 Model Averaging . . . 13

3 Data 17

4 Methodology 20

4.1 Dynamic Factor Model . . . 20

4.1.1 The Model . . . 20

4.1.2 Impact of new data releases . . . 22

4.2 MIDAS . . . 23

4.3 Factor MIDAS . . . 25

4.4 Machine Learning Algorithms . . . 26

4.4.1 Gradient Boosting . . . 26

4.4.2 Random Forest . . . 27

4.4.3 Support Vector Regression (SVR) . . . 27

4.4.4 Parameter Tuning . . . 28

4.5 Naive Benchmarks . . . 30

4.6 Model Combination 30
4.7 Design of the Nowcasting Exercise 31

5 Empirical Results 34
5.1 In-sample properties 34
5.2 Nowcasting performance of individual models 37
5.3 Impact of new data releases 44
5.4 Results of model averaging 47
5.5 The impact of COVID-19 on nowcasting models 51

6 Conclusion 57

Bibliography 62

A Data I

B Further Results VII

List of Tables

5.1 Average RMSE of selected models 41
5.2 Average RMSE of selected models (including the pandemic) . . 53

A.1 List of variables II

B.1 The RMSE of estimated models (excluding covid period) VII
B.2 The RMSE of estimated models (including covid period) VIII

List of Figures

5.1 Factor estimated in 2021 35
5.2 Factor loadings 36
5.3 Random Forest: variable importance 37
5.4 RMSE of selected models 40
5.5 GDP growth predicted by the DFM 42
5.6 RMSE of the DFM model compared with CNB nowcast 43
5.7 The impact of new data releases on 2020Q2 GDP nowcast . . . 45
5.8 The average impact of variables on nowcast adjustments 47
5.9 GDP predicted by the selected models 48
5.10 The RMSE of the ensemble model 50
5.11 GDP predicted by the ensemble model 51
5.12 RMSE of the selected models 52
5.13 GDP growth predicted by the DFM 54
5.14 RMSE of the DFM and the CNB official forecast 55
5.15 RMSE of the model combination 56

A.1 Variables III
A.2 Variables continuation IV
A.3 Variables continuation V
A.4 Variables continuation VI

B.1 Nowcasting error of the DFM and CNB (Q(0) M2 start) IX
B.2 Nowcasting error of MIDAS and CNB (Q(0) M2 start) IX
B.3 Nowcasting error of Factor MIDAS and CNB (Q(0) M2 start) . IX
B.4 Nowcasting error of the SVR and CNB (Q(0) M2 start) X
B.5 Nowcasting error of XGBoost and CNB (Q(0) M2 start) X
B.6 Nowcasting error of the RF and CNB (Q(0) M2 start) X
B.7 Nowcasting error of the ensemle model and CNB (Q(0) M2 start) XI
B.8 Nowcasting error of RW and CNB (Q(0) M2 start) XI

B.9 Nowcasting error of the DFM and CNB (Q(1) M2 start) XI
B.10 Nowcasting error of MIDAS and CNB (Q(1) M2 start) XII
B.11 Nowcasting error of Factor MIDAS and CNB (Q(1) M2 start) . XII
B.12 Nowcasting error of the SVR and CNB (Q(1) M2 start) XII
B.13 Nowcasting error of XGBoost and CNB (Q(1) M2 start) XIII
B.14 Nowcasting error of the RF and CNB (Q(1) M2 start) XIII
B.15 Nowcasting error of the ensemle model and CNB (Q(1) M2 start) XIII
B.16 Nowcasting error of RW and CNB (Q(1) M2 start) XIV

Acronyms

DFM Dynamic Factor Model

MIDAS Mixed Data Sampling Model

ML Machine Learning

RF Random Forest

SVR Support Vector Regression

SVM Support Vector Machine

GDP Gross Domestic Product

VAR Vector Autoregression

SAM System for Averaging Models

ARIMA Autoregressive Moving Average

DSGE Dynamic Stochastic General Equilibrium

AR(n) Autoregressive process of order n

EM Expectation Maximization

XGBoost Extreme Gradient Boosting

RW Random Walk

MA(n) Moving Average of order n

MDI Mean Decrease of Impurity

Master's Thesis Proposal

Author	Bc. Jan Kutman
Supervisor	prof. PhDr. Tomas Havranek, Ph.D.
Proposed topic	GDPNow for the Czech Republic

Motivation The growth rate of real GDP is a crucial metric of the pace of economic activity and thus is an essential indicator for policymakers assessing the current state of the economy. However, the GDP along with other key macroeconomic statistics are typically released with significant lags and are also frequently subject to revisions. Because of that, estimates of GDP growth in the current quarter are very important for policymakers. Enter Nowcasting. Giannone (2008) defines nowcasting as the prediction of the present, the very near future, and the very recent past.

GDP nowcasting models have been gaining importance as a tool to gain timely insight into the real-time state of the economy for central banks and other institutions. Most of the recent studies focused on the nowcasting of the GDP and other macroeconomic metrics use the Dynamic Factor Model framework suggested by Giannone (2008). Many such studies across different countries,i.e., Bragoli and Modugno (2014) for Brazil or Luciani and Ricci (2013) for Norway, conclude that dynamic factor not only outperforms competing statistical models but is also comparable with professional and judgment-adjusted forecasts by central banks. It is important to emphasize that nowcasting predictions are solely a result of mathematical models and are in no way influenced by expert judgment.

There are some nowcasts of GDP in specific countries available in real-time online, most notably the GDPnow nowcast by the Federal Reserve Bank of Atlanta that provides GDP growth of the current quarter estimate in the USA for each day. The GDPnow model is described in detail in Higgins (2014). Using real-time data since 2011, the GDPnow predictions were only slightly inferior to the consensus near-term forecasts from Blue Chip Economic Indicators.

There is no such running estimate available in the Czech Republic. However, Rusnak (2013) examines the performance of the dynamic factor model in nowcasting Czech GDP using vintages of real-time data. He concludes that the accuracy of model-based nowcast is comparable to the official CNB nowcast that is judgment adjusted. The CNB GDP nowcast, however, is only available on the last day of the current quarter. These results and successful GDP nowcasting efforts in many studies globally provide strong motivation for creating real-time GDP nowcast for the Czech Republic. Apart from creating a model to nowcast GDP growth for the Czech Republic, I will also focus on comparing the performance of multiple methods commonly used in nowcasting, and I will examine the proposition of model averaging in nowcasting.

Hypotheses

> Hypothesis #1: GDP growth predictions of selected model successfully compete with CNB judgmental nowcast.
>
> Hypothesis #2: Model averaging improves nowcast accuracy.
>
> Hypothesis #3: : The accuracy of GDP growth nowcast improves substantially through new data releases.

Methodology To build the models to nowcast the GDP growth, I will use a dataset consisting mainly of vintages of real-time data. The reason for that is the fact that a lot of the macroeconomic statistics released are subject to later revisions, while nowcasting must rely on the latest unrevised data. The dataset will consist of a large number of time series covering labor, production, sales, and trade sectors of the Czech Republic. Financial variables and confidence indicators will also be included as well as a number of foreign variables accounting for the fact that the Czech Republic is a small open economy. The collected data will cover the period 2005-2020, and to collect them, I will use mainly the OECD real-time database and the CNB monetary statistics Monthly Bulletin publications. The selection of variables is heavily influenced by existing studies focused on GDP nowcasting, with the most relevant one being Rusnak (2013).

I plan on building and comparing multiple models designed to nowcast GDP growth and comparing their performance with several statistical benchmarks. As done in multiple previous papers, specifically Rusnak (2013), I will employ a Dynamic Factor Model (DMF) framework introduced by Giannone (2008) to nowcast GDP growth in the Czech Republic. The DFM has been shown by many studies(Luciani and Ricci 2013, Bragoli and Modugno 2014) to be superior to other statistical models used for nowcasting because it is equipped to deal with large unbalanced datasets with a mixed frequency that are typical for nowcasting problems. The main assumption behind the DFM is that THE majority of comovements in macroeconomic variables are driven by a few common unobserved factors. The factors are modeled as a VAR process of order p. The principle idea of the model can be expressed by the two following equations.

$$x_t = \mu + \Lambda f_t + \varepsilon_t$$
$$f_t = A_1 f_{t-1} + ... + A_p f_{t-p} + u_t$$

Where x_t is a vector of monthly series transformed into stationary ones (all the independent and dependent variables included in the model), f_t is a vector of r unobserved common factors, Λ is the vector of factor loadings. $A_1...A_p$ are the autoregressive coefficients of the factors.

Other models that I will use for the nowcasting of the GDP and compare their performance with the DFM include the Bayesian Vector Autoregressive model (BVAR), Leading indicator Model, Dynamic Stochastic General Equilibrium model (DSGE), and Error Correction Model (ERM). These are models traditionally used for macroeconomic forecasting and are often compared to DFM in nowcasting literature. Additionally, following the promising results of Richardson et al. (2018) in nowcasting GDP in New Zealand, I will also employ several machine learning algorithms. Namely, the algorithms used will be the K-nearest neighbors regression, boosted trees, support vector machines regression, and neural networks. The statistical benchmarks used will be a simple autoregressive model, a moving average model, and a random walk model.

The accuracy of the used models, measured by root mean squared error (RMSE), will be compared not only among themselves and the naive statistical benchmarks but also with the official CNB nowcast available at the end of the last month of the reference quarter. This nowcast is computed by a set of equations of expenditure components and adjusted by expert judgment. For a model to be considered successful, it does not have to necessarily provide more accurate predictions of GDP growth than the CNB nowcast. Another advantage of the nowcast I will use is its daily availability, not only at one point in time, as is the case for the CNB nowcast. However, the comparison with the judgment-adjusted CNB nowcast is essential for the usefulness of any nowcasting model.

Apart from comparing the predictive performance of the individual models among themselves, I will also examine the possibility of improving nowcast accuracy through weighted averaging of multiple models. I will follow the methodology presented by Bjornland et al. (2012), who derives the model weights in the averaging scheme from the inverse of root-mean-squared errors (RMSE) of the respective models. If the combination of multiple models provides a more accurate forecast than the best-performing individual model, I will use it for the nowcasting exercise.

The examination of the final hypothesis is fairly straightforward. As the selected model will provide a daily estimate of GDP growth, the hypothesis suggests that the root mean square error of the model predictions will diminish throughout the period as more and more relevant data become available.

Finally, the real GDP growth estimates of the selected model will be displayed on a website and updated with a daily frequency.

Expected Contribution The goal of my thesis is to create a model that will nowcast real GDP growth in the Czech Republic in the current quarter for each day. The estimate should be adjusted with every release of relevant data. To the best of my knowledge, there is now such real-time nowcast of Czech GDP growth available. Such readily available nowcast for current period GDP growth should be valuable information for policymakers as well as for anyone whose decision-making is dependent on the current pace of the Czech economy.

There has been a variety of research dedicated to nowcasting of GDP growth in a large number of countries. To predict the GDP growth, a wide range of different models was used, from VAR models (Mccracken et al. 2015), factor models (Giannone et al. 2008, Luciani and Ricci 2013), bridge equation models (Higgins 2014) to machine learning algorithms (Richardson at al. 2018). One of the goals for my thesis is to compare the performance of the most frequently used nowcasting techniques and also to test whether averaging predictions from multiple models can be useful. In the context of nowcasting, Aastveit et al. (2010) already provided promising results of a density combination approach from multiple forecasting models. In forecasting in general, the benefits of model averaging have been found in multiple studies, for instance, Bjornland et al. (2012). It is, therefore, reasonable to expect that to nowcast Czech GDP, model averaging could be a useful tool.

Outline

1. Motivation: Introduction to nowcasting, specifically in the context of GDP growth and other macroeconomic metrics. Explanation of the importance of such models for policymakers.

2. Literature Review: I will review the literature most relevant to my thesis, with special attention being given to studies that are describing real-time GDP nowcasting in numerous countries.

3. Data: Detailed description of all the data sources used for my analysis.

4. Methodology: I will present a number of different models that will be used for GDP growth nowcasting along with the methodology for model averaging

5. Empirical Results: In this section, the results of each empirical model will be presented and compared among themselves as well as with the official CNB projections. The results of a potential model averaging will also be discussed.

6. Concluding Remarks: The last section will conclude how well the selected model nowcasts GDP growth and how relevant the real-time estimate is for policymakers.

Core bibliography

Higgins, P. 2014. GDPNow: A Model for GDP Nowcasting. Federal Reserve Bank of Atlanta - Working Paper 2014-7

Rusnak, M. 2013. Nowcasting Czech GDP in Real Time. CNB Working Paper Series.

Doz, C., Giannone, D., Reichlin, L. 2011.A two-step estimator for large approximate dynamic factor models based on Kalman filtering. Journal of Econometrics - 164(1):188-205

Richardson, A.,Thomas van Florenstein, M.,Vehbi, T. 2018. Nowcasting New Zealand GDP using machine learning algorithms - Reserve Bank of New Zealand

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

Banbura, M., Giannone, D., Reichlin, L. 2010. Nowcasting.Working Paper Series 1275 - European Central Bank

Bragoli, D., Metelli, L., Modugno, M. 2014. The Importance of Updating: Evidence from a Brazilian Nowcasting Mode. FEDS Working Paper No. 2014-94

Luciani, M., Ricci, L. 2013. Nowcasting Norway.International Journal of Central Banking - vol. 10(4),215-248

McCracken, M. W., Owyang,M. T., Sekhposyan,T. 2015. Real-Time Forecasting with a Large, Mixed Frequency, Bayesian VAR.Federal Reserve Bank of St. Louis - Working Papers 2015-30

Bjornland, H. C., Gerdrup, K. R., Jore, A. S., Smith, C., Thorsrud, L. A. 2012. Does Forecast Combination Improve Norges Bank Inflation Forecasts?. Oxford Bulletin of Economics and Statistics - Vol. 74, Issue 2, pp. 163-179

Aastveit, K.A., Trovik, T. 2012. Nowcasting norwegian GDP: the role of asset prices in a small open economy. Empirical Economics, Springer, vol. 42(1), pp. 95-119

Chapter 1

Introduction

The main goal of this thesis is to compare the performance of several methods in the nowcasting of the quarterly Year on Year (YoY) real GDP growth in the Czech Republic. Additionally, the potential improvement reached through a weighted average of predictions of individual models is examined. Finally, the best–performing model will be used to continuously nowcast Czech GDP growth, with the results being updated with every release of relevant data and made available to the public in real-time on a website.

The growth rate of GDP is a crucial indicator of the pace of economic activity and thus is considered essential information for policymakers assessing the current state of the economy. However, the official yearly or quarterly GDP figures are typically released with significant delays and are also very often subject to subsequent revisions. This provides a challenge for policymakers who are in need of timely estimates of GDP and other key macroeconomic indicators. For this reason, forecasts of the GDP, inflation, and other metrics for current and future periods are a necessary tool for policymakers and anyone else whose decision-making is dependent on the current state of the economy. The forecasting of metrics in the current period or periods in the very near future is commonly referred to as Nowcasting. Giannone *et al.* (2008) define nowcasting as the prediction of the present, the very near future, and the very recent past. Nowcasting has been gaining importance in macroeconomics, with the estimates provided by the nowcasting models typically being used by central banks as inputs for their models that produce both short-term and long-term projections.

GDP nowcasting relies on data related to the GDP released with higher frequency and less of a lag. For illustration, the Czech GDP is only available in quarterly frequency, and the first estimate is released approximately six weeks after the end of the reference quarter. However, other metrics related to GDP, such as industrial production, are released monthly and published with a shorter lag. These data sources can be used to derive early estimates of GDP, which have been shown in multiple studies, such as Chernis & Sekkel (2017), to be comparable in accuracy to the GDP estimates provided by central banks that are typically adjusted by expert judgment and are not released continuously. In the Czech Republic, the CNB only releases GDP forecasts once each quarter[1].

From a technical perspective, macroeconomic nowcasting provides a very particular challenge due to the non-standard nature of the available data. This comes from the need to leverage a wide range of the most recently available data published in different frequencies. Additionally, at different points in time, specific indicators may or may not be available. Overall, the datasets typically used for real-time nowcasting are unbalanced towards the end of the sample and published at different frequencies. New methods have been introduced for this econometric task because of the specific nature of the data in nowcasting problems. The most important method explicitly designed for macroeconomic nowcasting is the Dynamic Factor Model (DFM) developed by Giannone *et al.* (2008). This model is set up to deal with unbalanced datasets with ragged edges and mixed frequencies of the variables. The detailed description of the DFM framework and all the other methods used are provided in Chapter 4 of the thesis.

Since its introduction, the DFM has become the most commonly used model for short-term forecasting and real-time nowcasting in macroeconomics. The DFM nowcasting predictions have been shown in many studies to be comparable with the judgment-adjusted forecasts by central banks and superior to simple models that have been used previously for macroeconomic forecasts. The most frequent tasks in macroeconomic nowcasting are the predictions of GDP or inflation, which are key indicators for policymakers who use the real-time estimates as inputs for core models that provide long-term projections and identify potential inflationary pressures.

GDP nowcasting based entirely on results of statistical models, without any subjective adjustment, has been proven in many cases to be comparable or even slightly superior to the forecasts made by central banks and other institutions.

[1]The CNB releases nowcasts of GDP for the previous quarter, current quarter, and following quarters up to 7 quarters ahead around the 5th day of the second month of each quarter

The official forecasts are typically adjusted with expert judgment and released on a quarterly basis, as is the case in the Czech Republic. An indisputable advantage of entirely statistical nowcasting models is their ability to predict GDP growth in real-time based on currently available data. However, only a handful of nowcasting models worldwide have been automatized so that the updated predictions are displayed publicly on a website and updated continuously with new data releases. The main such application is the GDPNow of the Federal Reserve Bank of Atlanta, described by Higgins (2014). GDPNow is a model that nowcasts the real GDP growth of the United States, with the nowcasts being updated daily and published on the Federal Reserve Bank of Atlanta website. I take inspiration from this application in an effort to create a similar model for the Czech Republic and continuously publish the current estimates on a website. This converts a mainly theoretical work into a valuable tool for policymakers and anyone else with interest in the current prediction of the pace of growth of the Czech economy based on the latest data.

As for research regarding the nowcasting of the GDP in the Czech Republic, the most relevant study was conducted by Rusnak (2013), who examined DFM's performance in nowcasting the growth rate of GDP in the current quarter. He found the DFM nowcast predictions to be comparable to the nowcasts of CNB that are adjusted by expert judgment. Therefore, this study provided a strong indication that creating a real-time nowcast of Czech GDP should be achievable in such a manner that the model predictions would be relevant and helpful for policymakers and anyone whose decision-making is influenced by the current state of the Czech Republic economy.

The results reached through empirical analysis in this thesis confirm the success of DFM in GDP nowcasting tasks and find it to be the best performing individual model. The DFM not only beats comfortably the benchmark model for all nowcasting horizons but is also comparable with the official CNB nowcast for longer horizons and beats it for predictions starting from the second month of the nowcasted quarter. Based on an out-of-sample nowcasting exercise through 32 quarters, these results provide strong support for the usefulness of this real-time nowcasting model. Other methods, including MIDAS regression, factor-augmented MIDAS, and several ML algorithms, also outperform the benchmark model. By taking a weighted average of the predictions of individual models, accuracy comparable to the best performing individual model is obtained, while the uncertainty related to the choice of model is mitigated. The results also indicate that the predictive accuracy of all models generally improves with the arrival of new information during the nowcasted quarter, even though the improvements are not substantial for most data releases. The most significant performance improvement can be attributed to the release of GDP for the previous quarter.

The impact of the COVID-19 pandemic on the performance of nowcasting models is also addressed. While all the estimated models beat benchmarks by even larger margins and their relative performance remains similar, the errors increase substantially in absolute terms. Also, during the pandemic-effected period, even the best performing individual model lacked significantly in accuracy compared to the official CNB nowcasts that are adjusted with expert judgment.

This thesis is structured as follows: Chapter 2 provides a summary of the relevant literature, both from a theoretical and empirical point of view. Chapter 3 describes the dataset collected for the empirical analysis and comments on its use in the construction of a number of pseudo-real-time vintages. Chapter 4 presents all models used for GDP nowcasting in this thesis, as well as the approach utilized for model averaging. Additionally, the detailed process of the out-of-sample nowcasting exercise is explained. Chapter 5 reports the performance of all models and also provides interpretation to how new data releases impact the nowcasts. Finally, chapter 6 concludes and comments on the possibilities of future research.

Chapter 2

Literature Review

The existing literature about nowcasting in general and GDP nowcasting specifically is very extensive. This section summarizes the most relevant literature to the research question of this thesis. The existing literature can be divided into three strands. The first is the literature focused on macroeconomic nowcasting in general, describing the fundamental theory and the challenges and methods particular to this field. The second strand of literature is focused on the empirical results of GDP nowcasting in many countries or regions worldwide, using a wide range of methods and providing diverse results. The third strand of literature covered is explicitly dedicated to the nowcasting of Czech GDP. This is very relevant to this thesis in many ways, specifically concerning the data and methods used and the conclusion reached in efforts to predict the GDP of the Czech Republic. As model averaging is employed in this thesis, the fourth subsection of this chapter focuses on the most relevant literature concerning model averaging related to macroeconomic nowcasting.

2.1 Macroeconomic nowcasting as a discipline

Even though short-term forecasting and nowcasting have been of interest in macroeconomics for a long time and have been routinely used by central banks and other institutions, very little academic literature has focused on nowcasting before Giannone *et al.* (2008). The traditional approach to the nowcasting of GDP has relied on bridge equation methods. The principal idea behind this method is that the key expenditure components of GDP are forecasted separately using regression of the quarterly target component on its lags and a small number of preselected monthly predictors (Giannone *et al.* (2009)).

The monthly predictors are converted to the quarterly frequency of the target variables, typically using equal weights. If there are missing values for one or more of the monthly series, these are forecasted using auxiliary models, typically univariate time series models like ARIMA. This approach to GDP nowcasting was introduced and summarized by Klein & Sojo (1989), and other early applications are Trehan (1989) or Parigi & Schlitzer (1995). Most central banks have historically implemented this traditional approach to nowcasting to obtain early insight into the current state of the economy. Another common practice in GDP nowcasting done by policy institutions was to rely on expert judgment or at least adjust the nowcasts produced by the bridge equations or other simple models with some subjective judgment.

The traditional bridge equations approach to GDP nowcasting has several limitations: it cannot handle a large number of predictive variables, and also its ability to deal with missing observations is questionable, as Guagliano & Mantovani (2014) point out. To combat these limitations, Giannone *et al.* (2008) presented the single most influential study in the field of macroeconomic nowcasting. They argue that even though the *bridging* between the quarterly and monthly variables is a crucial part of nowcasting, it requires a more complex process than the one used in bridge equation models or other simple single equation models. They introduce a methodological framework designed to combat many of the issues that the econometric methods traditionally used for forecasting could not handle in the case of real-time nowcasting. The main idea behind this setting is a dynamic factor model (DFM) in a state-space framework. The DFM is designed to capture common movements from a large set of time series variables and synthesize them into a few latent factors. This model is equipped to deal with mixed frequency data and unbalanced datasets towards the end of the sample, a common issue presented in nowcasting. Another advantage of the DFM is that it provides interpretation to the changes of the GDP nowcast with each release of new data. This trackability of changes in the model predictions to specific data releases is of particular interest to policymakers. The DFM model is described in detail in chapter 4. The vast majority of the recent empirical literature focused on GDP nowcasting uses this framework developed by Giannone *et al.* (2008).

Even though since the introduction of DFM by Giannone *et al.* (2008), it became the leading method, the spectrum of models that have been suggested for GDP nowcasting is vast. One of the alternative approaches to the traditional simple forecasting models that gained significant importance is mixed-frequency data sampling (MIDAS). The MIDAS methodology was developed by Ghysels *et al.* (2004) as a forecasting tool specifically for time series variables sampled at different frequencies, which is very common in macroeconomics and finance. The MIDAS regressions allow a variable measured at a specific frequency to be explained by current and lagged values of time-series variables sampled at a higher frequency. A detailed description of the MIDAS model is also included in chapter 4. The early applications of MIDAS were predominantly on forecasting using high-frequency financial data. However, Clements & Galvao (2008) suggested it as an appropriate method for short-term forecasting of GDP and found it to outperform the bridge equations model in nowcasting of quarterly GDP in the USA. Marcellino & Schumacher (2008) introduce a Factor MIDAS model, which combines the MIDAS methodology with a factor model. Here the MIDAS regression is run on the estimated monthly factors as regressors, rather than a group of monthly indicators as would standardly be done in the MIDAS regression. They test this approach on nowcasting German GDP growth and find the results promising.

It is not uncommon for the studies regarding GDP nowcasting to develop specific models designed for a particular country or available data. These approaches often add together some of the models typically used individually. For instance, Higgins (2014) combines the traditional bridge equation approach with a variant of the factor model proposed by Giannone *et al.* (2008) and with the Bayesian VAR model as suggested by Chin & Miller (1996) to nowcast the GDP of the USA. Marcellino & Schumacher (2008) developed a Factor MIDAS model, which combines the MIDAS methodology with a factor model. Factor augmented VAR was also used in certain GDP nowcasting efforts, such as Grui & Lysenko (2017). These examples illustrate that the potential number of methods that can be used for GDP nowcasting is significant, and it is not a goal of this section of the thesis to summarize all of them but rather to present an overview of the most influential ones.

With the recent rise of popularity of Machine Learning (ML) along with ever-increasing computing power, ML has been suggested as an alternative to the wide range of econometric methods traditionally used for nowcasting of crucial macroeconomic variables. The evidence from literature comparing ML with traditional statistical models in forecasting is somewhat inconclusive. While Makridakis *et al.* (2018) find that out-of-sample forecasting performance of selected ML algorithms is inferior to more traditional statistical methods, Chakraborty & Joseph (2017) conclude that ML models outperform traditional statistical approaches in forecasting tasks. An argument favoring using ML in GDP nowcasting specifically is that GDP is potentially connected with a very large number of domestic and international macroeconomic and financial-market statistics. The ML algorithms are typically very well suited for handling datasets with many potential predictors, which is challenging for many traditional statistical methods. On the other hand, most ML algorithms do not offer a sophisticated way of dealing with variables sampled at different frequencies, other than converting the high-frequency variable to the lower frequency through averaging. This results in a loss of information. Another drawback of the use of ML is that it typically lacks trackability[1] in the interpretation of the nowcast, which may be particularly concerning for policymakers for whom the "storytelling" behind the particular forecast may be of significant importance. The paper that uses ML in a setting most similar to this thesis is Richardson *et al.* (2019), who compares the performance of several ML algorithms with more traditional statistical models in nowcasting the GDP of New Zealand. Specifically, the ML algorithms used in this paper include K-nearest neighbor regression, Least-square boosting, Support Vector Machine Regression (SVM), Lasso, and Neural Networks. The results show that, particularly, SVM and Lasso have been able to outperform commonly used statistical models. However, the entire analysis was conducted on data converted to quarterly frequency, which disregards the principal advantage of the DFM to deal with mixed-frequency data.

As far as the data used for GDP nowcasting, most of the existing studies agree on the need to utilize a large number of indicators that together provide as complete information about the entire economy of a particular country or region as possible. For instance, Giannone *et al.* (2009) point out the importance of both the so-called hard data covering production, labor, and trade sectors of the economy and soft data, which includes survey and financial data. All this data is typically specific to the country or area where the GDP is nowcasted. However, Rusnak (2013) argues that in the case of small open economies such as the Czech Republic, it is important also to include a set of international variables providing information both about the economy globally and the most important trading partners.

An essential aspect of the datasets used for GDP nowcasting is the different lags of the release of the monthly variables typically used as predictors. To

[1]The trackability refers to the ability of the model to associate updates of nowcast with new releases of particular variables.

give an example, the index of industrial production of the Czech Republic is released approximately 40 days after the reference month, while the services confidence indicator is available around five days before the end of the reference month. Furthermore, the hard indicators along with the GDP are also typically subject to subsequent revisions, often multiple times. The existing literature is not uniform in dealing with these real-time data features. While most of the recent research focused on evaluating GDP nowcasting models considers the publication lags of different series, the approach towards considering data revisions is split. The papers that consider data revisions and therefore use truly real-time data vintages include Schumacher & Breitung (2008) for Germany and Banbura *et al.* (2013) for the US, among others. On the other hand, Marcellino & Schumacher (2008) or Chernis & Sekkel (2017) are of those studies that ignore the data revisions and use pseudo-real-time data to examine the performance of nowcasting models. Several empirical studies focused on the role of data revisions on forecast accuracy, such as Bernanke & Boivin (2003) or Schumacher & Breitung (2008), suggest that data revisions do not affect forecast accuracy substantially.

The data used for GDP prediction has a significant impact on the methods that are suitable for the nowcasting exercise and vice versa. The traditionally used bridge equation method or other simple single equation models could only handle a limited number of regressors which provided risk that some important predictors might have to be left unused. The introduction of DFM by Giannone *et al.* (2008) allowed a much larger number of variables to be used as predictors, typically at least 20 but often around 100. In the case of ML algorithms, the number of time series that can be potentially used becomes practically unlimited. For instance, Richardson *et al.* (2019) use a dataset consisting of around 550 domestic and international variables to evaluate the performance of several ML algorithms in the nowcasting of New Zealand GDP.

2.2 Empirical Literature about GDP Nowcasting

The empirical literature testing various methods of GDP nowcasting is very extensive. Overall, the creation of GDP nowcasting models has found a significant amount of success. Many studies conclude that the nowcasting models are comparable in accuracy with official central bank forecasts adjusted by expert judgment.

Since Giannone *et al.* (2008) developed and introduced a framework designed specifically for macroeconomic nowcasting, the empirical literature focused on GDP nowcasting mainly became centered around the DFM he developed. Giannone *et al.* (2009) use the framework developed by Giannone *et al.* (2008) to nowcast quarterly GDP and annual inflation in the Eurozone. They do not focus on comparing the model nowcast with official estimates or competing models but instead show how new data releases influence the nowcast during different points of the quarter that is being nowcasted. In particular, they find the release of survey data significantly impacts the nowcast. Giannone *et al.* (2009) also comment on the development of nowcast accuracy during the quarter and point out that the predictions' root mean square error (RMSE) was reduced by 50% on average from the beginning until the end of the reference quarter. This clearly illustrates the trade-off between the timeliness and accuracy of the GDP estimates. Another important takeaway from this application is the comparison of two datasets used for GDP nowcasting. The first one includes only major indicators on the euro area economy, twenty-six in total. Alternatively, a dataset that includes also disaggregated data on industrial production and labor market information is used. This dataset consists of 64 variables in total. The results show that including the disaggregated data does not improve the accuracy of the nowcast. This finding was later supported also by Banbura *et al.* (2010), who show that the benefits of including more than 20-40 variables are modest and that disaggregate information is not beneficial for nowcast accuracy.

Other notable applications of the DFM model for GDP nowcasting include Guagliano & Mantovani (2014), who were the first to apply it to the main eurozone countries: Germany, France, Italy, and Spain. They find that the model predicts the GDP well, both in terms of trend and magnitude. There is growing evidence that the DFM can produce GDP nowcasts that are superior to simple benchmarks but are also often on par with professional GDP nowcasts using expert judgment. These studies include Bragoli *et al.* (2014) for Brazil, Barhoumi *et al.* (2010) for France, or Chernis & Sekkel (2017) for Canada.

The empirical studies examining the ability of several ML algorithms to nowcast GDP are scarce and all very recent; however, the existing evidence is encouraging. Richardson *et al.* (2019) find that ML algorithms generally outperform other methods traditionally used for GDP nowcasting, including DFM, in terms of the accuracy of the forecasts. They also show that ML has the potential to improve official nowcasts of the Reserve Bank of New Zealand. On the other hand, they point out that ML generally lacks the ability to interpret the particular nowcast and attribute its movements to individual data releases, as the DFM can. Also, it is worth noting that these promising results were obtained using a very large dataset consisting of about 600 time series with their frequency ranging from daily to quarterly, with the higher-frequency series being converted to quarterly frequency by taking a simple average. Other encouraging evidence on the ability of ML, namely the nearest neighbor algorithm, is provided by Jonsson (2020) on the case of nowcasting Swedish GDP. Soybilgen & Yazgan (2021) then combine the DFM with several tree-based ML algorithms in a setting where the dynamic factors are first extracted from a large set of financial and macroeconomic variables. These factors are then used for the ML algorithms. Their results show that this synthesized approach generally outperforms the pure DFM.

There has also been limited empirical evidence of the usefulness of model averaging and combination for GDP nowcasting. Aastveit *et al.* (2011) produce a density combination nowcast from a system of VARs, factor models, leading indicator models, and a DSGE model. They show that this approach can deal with the model uncertainty of the individual methods. Especially towards the end of the nowcasted quarter, it outperforms even the best performing individual models in the nowcasting of Norwegian GDP. Chikamatsu *et al.* (2018) also find evidence that there is a gain from combining individual models along with professional forecasts in nowcasting the GDP of Japan. A detailed description of the methods suggested for model combination in the case of macroeconomic nowcasting is provided in chapter 4.

Most of the empirical literature regarding GDP nowcasting only tests the performance of the selected models on past data, which essentially only provides theoretical knowledge about the performance and usefulness of such models. However, some notable cases of GDP nowcasts have been made available in real-time to the public, which allows to examine the performance of the models also on future data with respect to the time of model creation. Another clear benefit of such readily available nowcasts is the ability of anyone interested to see the GDP nowcast at any point in time, as it reflects every new release of relevant data. One of such nowcasts is the GDPNow of the Federal Reserve Bank of Atlanta, which provides a running nowcast of the current quarter real GDP growth in the USA. The nowcasting model, described in detail by Higgins (2014), unites the DFM model proposed by Giannone *et al.* (2008) with the bridge equation approach by aggregating 13 subcomponents that make up the GDP. Since the GDPNow introduction in 2011, the model forecasts using real-time data are slightly less precise than the official consensus near-term GDP forecasts from Blue Chip Economic Indicators. My thesis aims to replicate this running nowcast for the case of the Czech Republic and make it available in real-time to the public.

2.3 GDP nowcasting in the Czech Republic

There have been several studies focused on short-term forecasting of Czech GDP. Horvath (2013) investigated the role of confidence indicators in improving forecasts of several key macroeconomic indicators using a VAR model. The results suggest that the confidence indicators are not particularly important predictors of GDP and other variables describing the state of the economy. Havranek *et al.* (2012) again use a VAR model to examine the impact of financial variables in forecasting GDP on the case of the Czech Republic. The main takeaway from the paper is that financial variables seem to have a notable effect on GDP, and thus, their inclusion improves its forecasting. However, the predictive performance of individual financial variables is not stable over time. Benda & Ruzicka (2007) developed a short-term forecasting method based on the Leading Economic Indicators (LEI) approach and used it to predict Czech GDP growth for the current and following quarter. They find that this model provides relatively precise nowcasts of Czech GDP. Arnostova *et al.* (2011) evaluate the performance of several competing models in short-term forecasting of Czech GDP. The models used include VAR, bridge equations, and DFM or standard principle components model (PCM). They find the PCM model to be the most accurate in forecasting Czech GDP up to 3 quarters ahead. However, it did not outperform the official CNB forecast in the one-quarter ahead forecast.

The drawback of the studies mentioned above is that they largely ignore the real-time nature of data that forecasters have to deal with when tasked with predicting GDP. Most of them do not consider publication lags or data revisions. This makes their relevance for real-time GDP nowcasting questionable. Rusnak (2013) addresses these issues and uses real-time data vintages to examine the performance of the DFM in nowcasting Czech GDP growth over the period 2005-2012. The results show that the performance of this nowcasting model is comparable in accuracy with the official nowcasts of CNB that are adjusted with expert judgment and only released once in each quarter. Another noteworthy takeaway from the paper is the significant importance of foreign variables in nowcasting the GDP, while financial data and confidence indicators did not appear to be important predictors.

Another notable paper focused on nowcasting GDP in the Czech Republic was published by Franta *et al.* (2016). They compare the performance of several mixed-frequency models, including MIDAS, mixed-frequency VAR, and the DFM, in predicting GDP growth over the period 2005-2014. Their results suggest that DFM performs the best overall, and also, the mixed-frequency VAR is superior to benchmarks. Furthermore, they compare the performance of the models with the official CNB forecasts and conclude that CNB nowcasts outperform the purely statistical models for short forecasting horizons. For longer horizons, the mixed-frequency models appear to provide comparable predictions to those of the CNB.

Only very recently, Adam *et al.* (2021) of the CNB introduced a weekly index of Czech economic activity called Rushin which they use to nowcast the GDP growth rate. The creation of the index was motivated by the turbulent economic conditions caused by the COVID-19 pandemic, which saw a lot of the standard nowcasting models decrease in accuracy due to the unprecedented nature of the shocks to economies caused by lockdowns and the complete shutdown of specific sectors. These conditions caused many indicators typically used in nowcasting models to lose their explanatory power. To combat these challenges, Adam *et al.* (2021) follow a recent strand of literature that uses alternative high-frequency indicators such as electricity consumption or air pollution to extract meaningful signals about the pace of the economy. They use a principal component analysis (PCA) on data consisting of a combination of standard macroeconomic variables and several alternative high-frequency indicators. These include Google search trends, highway truck toll mileage, and electricity consumption. Although the index was still a work in progress at the time of writing this thesis[2], the preliminary results show that the predictive model's performance was particularly strong during the COVID crises.

2.4 Model Averaging

One of the main contributions of this thesis to the literature concerning GDP nowcasting is the examination of the performance of a combination of individual nowcasting models. Model or forecast combination have been of interest in

[2] As of January 1st 2022, the Rushin index was already being published weekly and signaled Czech GDP to grow by 0.2% in the 4th quarter of 2021

statistics for a long time. The premise of using a model combination rather than an individual model, which is certainly the more standard approach, is the ever-remaining challenge of model uncertainty. In the case of GDP nowcasting, many different models have been suggested as appropriate, as was highlighted in previous parts of the thesis. This naturally provides motivation to explore the possibility of using a model combination to combat the uncertainty of individual models.

Bates & Granger (1969) provide one of the earliest studies advocating forecast combination. They consider a setting where a decision-maker has access to multiple forecasts and is tasked to make a choice. They show that the weighted average of multiple unbiased forecasts with the weights assigned ex-post always has a variance lower or equal than the lowest variance of the individual model. The optimal weights are unknown beforehand, and Bates & Granger (1969) provide several alternatives for deriving the weights ex-post.

Timmermann (2006) identifies three reasons for the usefulness of forecast combination. The first one is similar to a portfolio diversification argument. If not all relevant information about possible models is observable, they argue that their combination is the best way to exploit information behind the individual forecasts. The second reason Timmermann (2006) described is that there can be unknown structural breaks, which favor one model over the other in some instances. This is very relevant for the GDP nowcasting problematic, as it has been suggested that the predictive performance of different models can change in tranquil periods and times of crisis (see, for example, Proietti & Giovannelli (2021)). A combination of multiple forecasts may, in this case, be more robust to these instabilities than a single forecasting model. In his third argument, Timmermann (2006) finds model combination to solve possible omitted variable bias by the individual models. If the forecasts are subject to different biases, combining the models can result in averaging out these biases and thus improved predictions.

From a practical perspective, a few issues need to be addressed when choosing a process of forecast combination. Firstly, a strategy to assign weights to point forecasts of individual models has to be selected. Additionally, it is not a priori clear what is the optimal number of models to be considered for the forecast combination. Bjornland *et al.* (2012) develop a strategy which they refer to as *System for Averaging Models* (SAM). First, they train a series of models commonly used for macroeconomic forecasting and use them to predict Norwegian inflation for one up to four quarters ahead. The model classes used are autoregressive moving average (ARIMA) models, VAR models, Bayesian estimated VAR models, error correction models, factor models, and DSGE models. Several variants are used for each model type, generally depending on the number of lags of the explanatory variables included. Overall, they estimate around 80 individual models. To derive the model weights for the forecast combination, they use the inverse of the individual models' root mean squared error (RMSE). Subsequently, they consider three options for selecting a number of models to be used for the combination: Top 8 performing individual models (around 10% of all the models), top 16 models, and the combination of all the models. Apart from the weighted average of point forecasts based on the RMSE, they also observe the performance of the simple mean of all models (the weights are equal). Their empirical analysis shows that the optimal number of models to be used for combination is only eight and that averaging the forecasts not only outperforms the benchmark models but also the Norges Bank's own inflation forecasts. This holds for the forecasts up to 1 year ahead.

The empirical literature about the model combination for GDP nowcasting is relatively scarce, however, there have been some encouraging studies. Aastveit *et al.* (2011) use a density combination framework to produce nowcasts of the GDP growth in the US. In this setting, they focus not only on the first moment of the forecast (the point forecast) but also on the underlying probability density, i.e., on the uncertainty of the model. They consider three different model classes suitable for GDP nowcasting: bridge equation models, factor models, and mixed-frequency VAR models. They find that the model combination approach leads to better point forecasts than the individual models. They also argue that the point forecast is not the only relevant measure if the decision maker's loss function is not quadratic. In such a case, it is necessary to consider the characterization of forecast uncertainty. Combining models decreases forecast uncertainty and thus could be preferred even if it would provide less accurate point forecasts than individual models. Proietti & Giovannelli (2021) use model averaging to nowcast monthly GDP in the euro area. They estimate many bivariate models for all possible predictors of the GDP. For real-time GDP nowcasting, they assign weights to the nowcasts of individual models based on the mean squared error of the model. An interesting contribution of this paper is the consideration of different weight assignments based on the phase of the business cycle or the point in time of the nowcasted quarter. They find that if the nowcast is made at the first part of the quarter, the consumer and business survey indicators receive more weight while the role of hard data is more important towards the end of the nowcasted quarter. As far as the business cycle phase, Proietti & Giovannelli (2021) find that soft indicators are more relevant in economic recession as they provide a timely signal of bad economic conditions. In the case of GDP nowcasting in the Czech Republic, Rusnak (2013) comments on the possible forecast combination and shows that combining point forecast of the model developed by him using the DFM framework with the official CNB nowcast improves accuracy.

Chapter 3

Data

The selection of variables is crucial for the performance of the predictive models. It is desirable to select data that is descriptive of the state of the economy while being parsimonious. The choice of variables in this thesis is influenced by existing literature dedicated to GDP nowcasting, with particular attention paid to those empirical studies focused on GDP nowcasting in small open economies such as the Czech Republic. The table of all the variables included in the analysis, along with their source, is available in Table A.1 in the Appendix. The set of predictive variables consists of 26 time series collected at a monthly frequency. These variables can be divided into four basic groups:

- Hard data
- Financial variables
- Survey data
- International (exogenous) variables

While the inclusion of the first three groups is almost unanimously standard across relevant literature[1], the further addition of foreign variables is due to the fact that the Czech Republic is a small open economy with strong connections to the European Union. To account for that, EURIBOR rates, the yields on EU bonds, and the EURO area business climate indicator are included. Additionally, German exports and the IFO indicator of the business climate in Germany are used as Germany is the most important trading partner of the Czech Republic. The logic of using financial data such as stock prices

[1](for examples, see Higgins (2014) or Giannone *et al.* (2009))

or interest rates is that these are forward-looking indicators related to the economy. Another advantage is the fact that these measures are available with no lag, with the exception of money aggregates. I include monthly interest rates, the PX-50 stock index price, exchange rate between CZK and EURO, along with the M2 monetary aggregate.

The usefulness of survey data relies primarily on their timeliness. It is generally even more timely than financial variables as the surveys are usually released several days before the end of the reference month. As Banbura *et al.* (2013) point out, surveys appear to be an important predictor of GDP when hard data is not available in the early part of the particular quarter. However, when hard data such as industrial production or retail sales becomes available later, the importance of surveys in nowcasting the GDP diminishes. The survey data includes confidence indicators related to the industry, trade, services, and construction.

The hard data series included have a direct link with economic activity, and therefore their use for the prediction of GDP is natural. Their disadvantage compared with financial indicators and survey data is that they are generally released with at least one month lag. They are also not forward-looking indicators because they do not reflect the expectations of producers and consumers. However, as Giannone *et al.* (2009) find, the hard data is still the most relevant GDP predictor when available. I include data covering industrial production, labor market, construction, retail, and trade as the hard macroeconomic indicators. Overall, the collected dataset consists of 26 monthly variables and the GDP at a quarterly frequency. The monthly time series span from January of 2005 until August of 2021 when writing this thesis. This results in 203 monthly observations for the variables released without a lag. For most hard data, the release lag only allows 201 monthly observations, the last one being for June of 2021. The quarterly time series of real GDP starts with the first quarter of 2005, and the last released figure is for the second quarter of 2021, which leaves us with 66 observations.

As the goal of the analysis in this thesis is to evaluate the performance of several models for predicting actual YoY growth of GDP in the Czech Republic, it is necessary to construct such datasets that reflect as closely as possible the actual data available at the time of individual nowcasts. I, therefore, use the collected dataset to construct pseudo-real-time data vintages simulating the inflow of data. For each month in which GDP growth is predicted, three vintages are constructed, taking into account the lag in the release of each variable. To give an example, for a data vintage taken on the last day of any given month, the financial and survey data for that month will already be available. However, as most of the hard data is released with a lag of approximately 40 days, the data for the last two months for these variables will not be available. The process of vintage construction and subsequent GDP nowcasting is explained in more detail in the subsection 4.7 of the chapter 4. The constructed vintages are not fully real-time, as data revisions are not considered, hence the pseudo-real-time label.

Regarding data revisions, the existing literature is split between the studies that ignore them and construct pseudo-real-time data vintages as is done in this thesis and those that take revisions into account and use fully real-time data. Pseudo-real-time data is used by Giannone *et al.* (2009) or Marcellino & Schumacher (2008), to name a few, while Guagliano & Mantovani (2014), among others, use real-time data. The importance of data revisions in macroeconomic forecasting has been a focus of several studies, with Bernanke & Boivin (2003) or Schumacher & Breitung (2008) concluding that data revisions do not influence forecasting performance considerably. Another issue to consider is the availability of real-time data itself, which is a necessary condition for them to be used in empirical studies. For the case of the Czech Republic, Rusnak (2013) evaluates the real-time performance of the DFM for nowcasting GDP using data from the OECD real-time database for the variables that are subject to subsequent revisions. I find the use of the OECD real-time database problematic, as, on several occasions, the data published does not correspond to the actual date of release of the figure by the CNB or the Czech Statistical Office. The pseudo-real-time approach is therefore adopted.

The dependent variable in the predictive models, the YoY GDP growth, is the only variable being announced only at quarterly frequency rather than on a monthly basis. This problem of different frequency of the dependent variable from its predictors creates a specific issue. The solutions of the individual models to the mixed frequency are explained in detail in chapter 4. Following the related literature, all the variables are transformed into stationary ones. This is done by taking yearly percentage change[2] for the variables, except for several confidence indicators for which the absolute yearly difference is taken. Before every model estimation, the variables are also standardized to have zero mean and unit variance.

[2]Through empirical results, I found the yearly percentage change of the variables to predict the GDP YoY growth better than monthly percentage change

Chapter 4

Methodology

This thesis employs several methods for GDP nowcasting, ranging from econometric models traditionally used for forecasting in macroeconomics to ML algorithms used only sporadically in this field. Additionally, an approach to model combination is adopted. This chapter is organized followingly: First, the models and methods are presented individually from a theoretical perspective along with a description of the model combination framework. Then the design of the pseudo-real-time nowcasting exercise itself is introduced.

4.1 Dynamic Factor Model

4.1.1 The Model

The DFM developed by Giannone *et al.* (2008) quickly became the most prevalent method for short-term forecasting and nowcasting in macroeconomics. It can extract the most important features from data by assuming the majority of the comovements in the variables are driven by a few common latent factors. In this sense, the DFM is very similar to the principal component analysis. The detailed model is set up in the following way:

First, we specify the model for monthly variables:

$$x_t = \mu + \Delta f_t + \epsilon_t$$

where x_t is a vector of monthly series transformed into stationary ones, f_t is a vector of r unobserved common factors and Λ is the vector of factor loadings. The factors are modeled as:

$$f_t = A_1 f_{t-1} + ... + A_p f_{t-p} + u_t$$

Where $A_1, ..., A_p$ denote a matrix of $r \times r$ autoregressive coefficients for the factors.

Additionally, we assume that the ith idiosyncratic component of the monthly variables follows an AR(1) process:

$$\varepsilon_{i,t} = \alpha_i \varepsilon_{i,t-1} + \varepsilon_{i,t} \ , \ \epsilon_{i,t} \ N(0, \sigma_i^2) \text{ with } E[\epsilon_{i,t}\epsilon_{j,s}] = 0 \text{ for } i \neq j$$

The quarterly GDP time series is incorporated into the model by using the approximation introduced by Mariano & Murasawa (2003). In this setting, the quarterly GDP is expressed as the sum of its unobserved monthly contributions:

$$GDP_t^Q = GDP_t^M + GDP_{t-1}^M + GDP_{t-2}^M$$

We define $Y_t^Q = 100 \times log(GDP_t^Q)$ and $Y_t^M = 100 \times log(GDP_t^M)$. The unobserved monthly growth rate of GDP, $y_t = \Delta Y_t^M$, is assumed to follow identical factor model representation as the monthly variables:

$$y_t = \mu_Q + \Lambda_Q f_t + \varepsilon_t^Q,$$

$$\varepsilon_t^Q = \alpha_Q \varepsilon_{t-1}^Q + e_t^Q, e_t^Q \ i.i.d.N\left(0, \sigma_Q^2\right),$$

with $\Lambda_Q = (\Lambda_{Q,G}, \Lambda_{Q,R})$. To create a link between y_t and the observed GDP data, a partially observed monthly series is constructed:

$$y_t^Q = \begin{cases} Y_t^Q - Y_{t-3}^Q & t = 3, 6, 9, ... \\ \text{unobserved} & \text{otherwise} \end{cases}$$

Finally, using the approximation by Mariano & Murasawa (2003) we get:

$$\begin{aligned} y_t^Q = Y_t^Q - Y_{t-3}^Q &\approx (Y_t^M + Y_{t-1}^M + Y_{t-2}^M) - (Y_{t-3}^M + Y_{t-4}^M + Y_{t-5}^M) \\ &= y_t + 2y_{t-1} + 3y_{t-2} + 2y_{t-3} + y_{t-4}, \qquad t = 3, 6, 9, ... \end{aligned}$$

If we define $\bar{x}_t = (x'_t y_t^Q)'$ and $\bar{\mu} = (\mu', \mu_Q)'$, then the joint model as specified by the previous equations can be cast in a state space representation:

$$\bar{x}_t = \bar{\mu} + Z(\theta)\,\alpha_t$$
$$\alpha_t = T(\theta)\,\alpha_{t-1} + \eta_t,$$

where η_t $i.i.d.$ $N(0, \sum_\eta(\theta))$ and θ is a collection of all the parameters of the model and $Z(\theta)$, $T(\theta)$, $\sum_\eta(\theta)$ matrices depending on the individual parameters.

The model is estimated by maximum likelihood using the expectation-maximization (EM) algorithm. This approach for large datasets has been proposed by Doz *et al.* (2006) and further generalized by Banbura & Modugno (2013) so that the DFM can deal with any pattern of missing observations. Briefly, the process of estimation by EM can be described as an iterative process alternating between the two following steps:

- The conditional expectation of the likelihood function is calculated using the estimates of parameters from the previous iteration
- The parameters are re-estimated by maximizing the likelihood function from the previous step

For a detailed technical survey of the EM estimation for the DFM, see Banbura & Modugno (2013).

There are several possibilities for the specification of the model. These depend on the number of factors used to model the comovements of the indicators, the choice of variables on which the factors are loaded as well as on the number of lags of the factors. I opt to estimate the models with the number of global factors ranging from 1 to 4 and experiment with factors that load only on real and nominal variables, respectively. To model the dynamics of the factor, 1 and 2 lags are used.

4.1.2 Impact of new data releases

Banbura & Modugno (2013) developed a method to quantify the effect of new data releases and revisions on the predictions of the DFM. This methodology allows us to attribute the nowcast revisions to new releases of specific variables. The ability to track the source of nowcast updates is of interest to policymakers, who are often interested in the ”storytelling” behind the nowcasting results.

An essential concept to this method is called *news*, which refers to the unexpected part of a new data release. The revision of a nowcast is due solely to the *news*, as any information available during the former nowcast would be accounted for. Let Ω_v be the information set available at time v and GDP growth at time t as y_t^Q. The new nowcast can then be decomposed into two components:

$$\underbrace{E[y_t^Q|\Omega_{v+1}]}_{\text{new forecast}} = \underbrace{E[y_t^Q|\Omega_v]}_{\text{old forecast}} + \underbrace{E[y_t^Q|I_{v+1}]}_{\text{revision}},$$

where I_{v+1} is a subset of the information set Ω_{v+1} that is orthogonal to all elements of Ω_v. The revision, which is the effect of the *news* is given by:

$$E[y_t^Q|I_{v+1}] = \sum_{j\in J_{v+1}} b_{j,t,v+1} \underbrace{(x_{j,T_{j,v+1}-E[x_{j,T_{j,v+1}}|\Omega_v])}}_{\text{news}}$$

where $b_{j,t,v+1}$ are the weights derived from model estimation, and J denotes the set of new variables. The effect of the release of a variable on the nowcast is defined as the combination of *news* associated with the release of the given variable and its weight. The total nowcast revision is then the sum of effects of all variables.

4.2 MIDAS

Mixed-data sampling (MIDAS) is an econometric regression designed by Ghysels *et al.* (2004) for direct forecasting of time series sampled at different frequencies than the explanatory variables. Even though it was initially used to forecast financial variables, MIDAS quickly found use as a macroeconomic forecasting tool of quarterly GDP. The direct (one equation) approach of MIDAS is what separates it from other mixed-frequency techniques, such as bridge equations or the DFM, where the missing values of the higher frequency indicators are first forecasted.

In the MIDAS framework, the observations of variable sampled at lower frequency are directly related to current and lagged observations of the high-frequency variables. Let's assume a case of GDP growth y_t at a quarterly frequency, where $t = 1, 2, ..., T_y$, with T_y being the final quarter for which GDP growth is available. Then, for simplicity, we assume one stationary monthly indicator x_t^m, where the time index t is an adequate fraction of the corresponding quarter of the GDP growth: $t = 1 - \frac{2}{3}, 1 - \frac{1}{3}, 1, 2 - \frac{2}{3}, ..., T_x - \frac{1}{3}, T_x$. T_x is the month of the last observation of the monthly indicator. Typically, $T_x \geq T_y$ as the monthly indicators are usually available in a more timely manner than GDP.

The goal is to nowcast the growth of GDP y_{t+h} with horizon $h = 1, ..., H$. The MIDAS equation for GDP growth in quarter $t + h$ is then expressed as:

$$y_{t+h} = \beta_0 + \lambda y_t + \beta_1 B(L^{\frac{1}{3}}; \theta) x^m_{t+w} + \varepsilon_{t+h}$$

Where $w = T_x - T_y$ is the time difference of the availability of the monthly variable and the quarterly GDP. The right-hand side is specified in terms of period $t + w$ to condition the nowcast on the information about the monthly indicator available in the current quarter while the availability of GDP lacks. λy_t is an autoregressive term and $B(L^{\frac{1}{3}}; \theta)$ is a lag polynomial

$$B(L^{\frac{1}{3}}; \theta) \ = \sum_{k=0}^{K} b(k; \theta) L^{k/3},$$

where the monthly lag operator is defined as $x^m_{t-1/3} = L^{1/3} x^m_t$. The lag polynomials are usually chosen to avoid parameter proliferation in the case of long lags K of the monthly indicator. So, even if the researcher uses a high number of lags of the high-frequency indicator, only two or three parameters of the polynomial have to be estimated. In my empirical analysis, I use both the exponential Almon lag and the so-called Beta lag, the most commonly used functional form in the existing MIDAS literature. The exponential Almon lag is expressed as:

$$b(k; \theta) \ = \ \frac{exp(\theta_1 k + \theta_2 k^2)}{\sum_{j=0}^{K} exp(\theta_1 j + \theta_2 j^2)},$$

with parameters $\theta = [\theta_1, \theta_2]$. For any given θ, the lag polynomial provides a parsimonious way to account for a large number of monthly indicators.

The Beta lag is defined as:

$$b(k;\ \theta) = \frac{f(\frac{k}{K}, \theta_1, \theta_2)}{\sum_{k=1}^{K} f(\frac{k}{K}, \theta_1, \theta_2)},$$

where $f(\frac{k}{K}, \theta_1, \theta_2) = \frac{k^{\theta_1 - 1}(1-k)^{\theta_2 - 1}\Gamma(\theta_1 + \theta_2)}{\Gamma(\theta_1), \Gamma(\theta_2)}$ and $\Gamma(\theta) = \int_0^\infty e^{-x} x^{\theta - 1} dx$. The parameters θ_1 and θ_2 control the shape of the distribution, which can be very diverse because of the general parametrization.

One of the main features of the MIDAS approach consists in individual parameter estimation for each forecast horizon. The MIDAS parameters for each forecast horizon h are estimated by non-linear least squares (NLS), and the direct forecast is then the conditional expectation

$$\hat{y}_{T_y+h|T_x} = \hat{\beta}_0 + \hat{\lambda} y_{T_y} + \hat{\beta}_1 B(L^{1/3}; \hat{\theta}) x^m_{T_x},$$

where $T_x = T_y + w$ in such a manner that the monthly predictor's latest observations are included in the conditioning set of the projection.

For the empirical analysis, I follow the methodology proposed by Chernis & Sekkel (2017), among others, and first estimate the MIDAS model individually for all the 26 available monthly predictors. The predictions of all these bivariate models are then combined using both equal weights and weights proportional to the inverse of the RMSE of the individual bivariate models.

4.3 Factor MIDAS

The DFM and MIDAS are the two most common approaches to solve the challenge of mixed-frequency datasets with ragged edges in GDP nowcasting. Marcellino & Schumacher (2008) proposed a combination of these two methods. This synthetic methodology was also later used by Gul & Kazdal (2021), among others. I also follow this methodology that merges the two mixed-frequency econometric techniques. It can be summarized in the following two steps:

1. One or more monthly factors are extracted from the large monthly dataset using the DFM methodology explained in the prior subsection.

2. The monthly factors are then used as regressors in the MIDAS model

The summarization of many variables into 1 or 2 factors helps to avoid the process of selection of appropriate variables for the MIDAS regression. I use 1 and 2 two global factors as regressors, respectively, and estimate both of these models using the exponential Almon lag polynomial as well as the beta lag polynomial.

4.4 Machine Learning Algorithms

4.4.1 Gradient Boosting

Gradient boosting is a powerful ML technique that combines a series of individual weak learners to build a single strong predictor. The weak learners are typically decision or regression trees. It can be used for regression, classification, prediction, or other tasks. The model is built sequentially, as is the case in boosting methods.

The algorithm uses an arbitrary differentiable loss function (i.e., least squares or least absolute deviation). First, the initial model (decision tree) is fitted to predict the target variable with the loss function applied to the residuals of the models fitted so far. This iterative process continues until a certain stopping criterion regarding the loss function is reached. The calculated contribution of every tree is based on minimizing the overall error of the composite model. This iterative process can be expressed in totality as:

$$F_m(x) \;=\; F_{m-1}(x) + \eta\Delta_m(x),$$

Where $F_m(x)$ is the newly added model mapping x to the target variable,$F_{m-1}(x)$ is the previous model, $\Delta_m(x)$ is the weak learner and η is the learning rate. Generally, there are two main parameters in GB models. These are the learning rate and the number of trees to be fitted (the number of iterations). The learning rate also called shrinkage, impacts the effect of each additionally fitted tree on the overall model. Lower learning rates help generalize the model and decrease the risk of overfitting. On the other hand, estimating a model with a low learning rate can become very computationally expensive. Additionally, several characteristics of the individual decision trees can be adjusted as parameters. Examples of these tree-specific parameters include the maximum number of terminal nodes or maximum depth allowed for each tree. The limitation of the complexity of the individual models helps avoid the risk of overfitting. The process of finding the appropriate parameters for gradient boosting models, as well as for other ML algorithms, is presented in the subsection 4.4.4.

I use a modern implementation of gradient boosting, XGboost, which stands for Extreme Gradient Boosting. This implementation was designed for computational efficiency in terms of speed and performance. This algorithm has dominated ML competitions in the last several years and has generally proven to outperform other tree-based methods in accuracy and fellow gradient boosting implementations in computational efficiency.

4.4.2 Random Forest

Random forest (RF) is another ensemble ML method. As is the case in boosting models, RF uses many decision or regression trees (weak learners) to construct a single predictive model superior to its components. The main difference between RF and boosting methods is that in the former, the individual regression trees are trained independently, and the predictions are reached through averaging their output.

The RF model process can be summarized in the following steps:

1. Obtain a bootstrapped sample of size N from the training dataset
2. Using the bootstrapped set from the first step, estimate a regression tree considering only a certain portion x of the total p available variables. For each terminal node, the best variable split is determined until minimum node size n_{min} is reached.
3. The first two steps are repeated B times

Bootstrapping and the random selection of variables to be considered for the estimation of each tree help to reduce the dependence between the regression trees and thus avoid overfitting. This is a common issue with a simple decision or regression trees when they often fit the training set well, but their performance deteriorates with added data. The final output of the RF is calculated by averaging the predictions of all the regression trees. Averaging reduces the variance and makes the overall predictive performance more stable.

4.4.3 Support Vector Regression (SVR)

Support Vector Machine (SVM) is one of the most popular ML algorithms for classification tasks; however, recently, it's other variant (SVR) also started to be increasingly used in regression analysis. It is a supervised learning algorithm that is used to predict discrete values. The main difference between SVR and other regression models is that the latter typically try to minimize the error between the actual and the predicted values. In contrast, SVR tries to fit the best line within a given threshold value.

In practice, SVR attempts to find a function $f(x)$ with a maximum deviation of ε from the actual values of the target variable for all the training data. For a linear function, $f(x)$ is specified as:

$$f(x) = (w, x) + b,$$

where w is the weight vector, x is the vector of input variables, and b is the bias. The algorithm aims to find the linear function while ensuring it's smoothness to the maximum possible extend.

However, it is possible that the data is not linearly separable, and therefore no such function exists. In such a case, the algorithm allows for slack variables, ξ_i, that allow the regression errors to occur up to the values ξ_i and ξ_i^*. The cost function then includes the term w to penalize non-smoothness by minimizing the coefficients on the explanatory variables. The goal of the SVR is then minimization given by:

$$min.\ 0.5 \parallel w \parallel^2 + C \sum_{i=1}^{l} (\xi_i + \xi_i^*)$$

$$\text{subject to: } f(x) = \begin{cases} y_i - (w, x_i) - b \leq \varepsilon + \xi_i \\ (w, x_i) + b - y_i \leq \varepsilon + \xi_i^* \end{cases}$$

where C is the regularization parameter and sets the tradeoff between minimizing the errors and penalizing the overfitting of the model.

4.4.4 Parameter Tuning

ML algorithms typically have several parameters that are not known beforehand and affect the model's performance. The three algorithms that we use to nowcast Czech GDP - the SVR, XGBoost, and Random Forest – are no different. For instance, let's consider the XGboost algorithm. As a boosting model, it consists of several regression trees trained sequentially. One class of parameters that can be adjusted are tree-specific parameters. These parameters restrict the complexity of each tree that is fitted, which can reduce overfitting. Overfitting is a problem especially relevant for a smaller dataset. Examples of these parameters include a maximum three-depth or a maximum number of leaves in a tree. The second class of parameters governs the boosting process as a whole. It includes parameters such as the number of trees to be fitted or the rate the model makes corrections with each added tree: the learning rate, also referred to as shrinkage.

The optimal values of all of these parameters depend on the specific dataset used to train the model. Therefore, it is desirable to conduct an analysis before model training to find parameters that would maximize the performance of the given model. This process is referred to as parameter tuning.

I conduct a parameter tuning process fairly popular in ML, which uses randomized grid search with cross-validation. The process can be summarized in the following steps:

1. The parameters of the model that can be tuned are identified along with the range of values that these parameters can typically take on based on existing literature and available empirical applications. For each parameter, we get a list of values that it can take on.

2. A specified number of random combinations of the model parameters are found – so-called randomized grid search. A model is trained on several subsamples of the dataset while its out-of-sample predictive performance is recorded. This resampling procedure is called cross-validation. At the end of this step, the combination of parameters with the best performance is identified.

3. The parameter values found in the previous step are used to construct a more specific grid of parameter values. Typically only a few values around the value found in randomized grid search are considered.

4. Finally, the performance of models with all possible combinations of parameter values is evaluated using cross-validation. The parameter combination with the superior performance is then used for the actual nowcasting model.

The parameter tuning process described above can be computationally expensive, as it usually requires the fitting of several hundreds of models. Furthermore, in our nowcasting exercise, a different model is trained for each nowcasted quarter and each nowcasting horizon – the number of models to be estimated totals 608. The parameter tuning process could, in theory, be employed separately for each model that is fitted in the nowcasting exercise. This would, however, be extremely computationally expensive. I, therefore, elect to only perform the parameter tuning process using the entire dataset from the final data vintage. The parameters found with parameter tuning are then used for models fitted on every data vintage.

4.5 Naive Benchmarks

Apart from the complex econometric models explicitly designed for macroeconomic forecasting and the different ML algorithms, several naive benchmarks are considered for comparison of the nowcasting performance with the main models proposed. Firstly, we consider a random walk (RW) model, where the current quarter GDP growth nowcast is equal to the GDP growth in the previous period:

$$y_t = y_{t-1} + \varepsilon_t$$

The second benchmark model is an autoregressive process of order 2 (AR(2))where GDP growth is regressed on its first and second lag:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \varepsilon_t$$

Additionally, we also consider a moving average model, where the current period nowcast is an average of the last 4 observations (MA(4)):

$$y_t = \frac{1}{4}(y_{t-1} + y_{t-2} + y_{t-3} + y_{t-4}) + \varepsilon_t$$

4.6 Model Combination

The methodology of forecast combination used in this thesis follows the system set up by Bjornland *et al.* (2012), who developed the so-called System for Averaging Models (SAS). This system addresses two main issues related to model averaging. The first is the choice of models to be included in the averaging process, and the second is the derivation of weights of each model's predictions in forecast averaging.

As Bjornland *et al.* (2012), I also estimate models from several different classes used in the relevant literature for GDP nowcasting. These classes include DFM, MIDAS models, Factor MIDAS, ML algorithms, and univariate benchmarks. For each class of forecasting models, several different model specifications are estimated. For illustration, the DFM is estimated using a different number of factors as well as different structure of factor loadings on the variables. For the MIDAS models, bivariate models are estimated for all monthly predictors while using different functional forms of the lag polynomial.

Each of the nowcasting models is estimated for 16 different horizons for quarters ranging from the first quarter of 2012 till the fourth quarter of 2019, as is explained in detail in the section 4.7. The point out-of-sample nowcasts are evaluated for each horizon by calculating the RMSE, which can be interpreted as a measure of the difference between predictions and values actually observed at a later stage. It is defined as:

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

where T is the length of the series of point forecasts, $\hat{y}_t$ is the predicted value of GDP YoY growth, and y_t is the actual GDP growth.

The weight of each model's predictions in the model averaging is then computed as the inverse of the RMSE of the model's forecasts for a particular nowcasting horizon. The horizon-specific weights reflect the possibility that each model's predictive accuracy may differ depending on the forecasting horizon. For example, a model can be relatively strong in a 1-month ahead forecast but be much worse in forecasting GDP growth four months from now.

As Bjornland *et al.* (2012) conclude that using only several best-performing models for the forecast combination is superior to using all the available estimated models, I opt to include the seven best nowcasting models in the averaging process. The decision not to include more specifications of the same model class that only differ in the lag order was also based on the observation that these specifications typically yield nearly identical predictions. The inclusion of more of these specifications would therefore not provide additional information.

4.7 Design of the Nowcasting Exercise

The nowcasting process starts with constructing pseudo-real-time vintages that simulate the data available at the time of the nowcast. The vintages reflect the different delays of the release of the individual variables that result in the so-called ragged end of the datasets available at any point. For each relevant month, three vintages are constructed, the first one on the 10th day of the month, the second one on the 24th day, and finally one at the end of the month. The selection of these days reflects the approximate release calendar of the monthly indicators. For illustration, the vintage on the 24th day of the month includes the newly released confidence indicators that are typically available several days before the end of the reference month. The vintage at the end of the month then mainly reflects the availability of financial variables that are released with no lag. The detailed summary of the publication lags and subsequent pattern of missing observations is summarized in table A.1 in the appendix. Even though the exact dates of releases differ from month to month, this supposed pattern of releases is a very close approximation of the actual data availability.

For each quarter, the first prediction is made on the 10th day of the 3rd month of the quarter prior to the reference quarter. The predictions are then updated with each additional vintage, and the last one is made on the 10th day of the 2nd month of the following quarter, which corresponds to the release of the first (flash) estimate of GDP by the CNB. In total, for each quarter, 16 different predictions are made. I refer to the predictions made before the reference quarter as forecasts, those made during the reference quarter as nowcasts, and those after it as backcasts in the remaining of the thesis. The update of predictions allows to examine the effect of the arrival of new information on each model's nowcast accuracy and evaluate its predictive power for different nowcasting horizons.

The nowcasting process consists of 32 total nowcasting rounds, starting with the first quarter of 2012 and concluding in the final quarter of 2019. I opted not to include the period affected by the Coronavirus pandemic that started in the first quarter of 2020 in the baseline model. This decision was made due to the unprecedented economic conditions during this period, largely caused by politically imposed lockdowns and shutdowns of entire sectors. This could cause bias in the comparison of the accuracy of individual models as the collected sample is relatively short. Nevertheless, all the models with the inclusion of covid period are also estimated and commented on separately in the following section.

All the nowcasting models presented in the preceding section are used to produce 16 nowcasts for each quarter. The predictions are then evaluated by comparing them to the actual GDP YoY growth for the reference quarter. Root mean squared error (RMSE) is then computed separately for each horizon to measure the nowcasting accuracy of the models – one RMSE is calculated only from predictions made on the 10th day of the last month before the reference quarter and so on.

The fact that for each model, the RMSE is calculated separately for every nowcasting horizon allows the model averaging process to be also adjusted for different horizons. This means that if a particular model has relatively high accuracy for longer nowcasting horizons but is not that precise for shorter-term nowcasting or backcasting, the weight assigned to this model's predictions will be higher for the longer horizon and lower for the short-term predictions.

In general, all the nowcasting models are estimated using both an expanding window and a rolling window with a fixed size. The advantage of the expanding window approach is the ability to use all the data available at the time of the forecasts for the model's training to make the out-of-sample prediction. On the other hand, the rolling window can control against potential structural breaks that may cause older data to be less relevant. For every model, the approach yielding lower average RMSE is adopted.

Chapter 5

Empirical Results

This chapter is divided into four sections. First, we take a look at several properties of the models using the last available vintage. The second section provides an analysis of the nowcasting accuracy of all the individual models. As the DFM allows for detailed examination of the effect of the arrival of new information on the nowcasts, the impact of data releases is also examined. The third section addresses the results of model averaging, and the final section replicates the entire analysis with the inclusion of the pandemic period.

The DFM is the best performing model based on an out-of-sample nowcasting exercise over 32 quarters. The DFM comfortably beats the benchmark model for all nowcasting horizons, and for short-term nowcasts and backcasts also provides more accurate predictions than the official GDP nowcasts of the CNB. The rest of the estimated models also perform reasonably well and are superior to the RW benchmark at a 95% significance level. Furthermore, model averaging results in predictive accuracy comparable with the best individual model while decreasing the uncertainty related to the choice of the most appropriate model.

5.1 In-sample properties

Before the nowcasting performance of the individual models is examined using pseudo-real-time vintages, several properties of some of the models are studied using the final data vintage from July 2021.

Starting with the DFM, figure 5.1 shows the estimated monthly factor along with the actual quarterly YoY GDP growth. The factor is extracted using the DFM specification with one global factor with one lag. The majority of the monthly variables along with the GDP have negative loadings on the factor. Figure 5.1 plots the factor multiplied by -1 so that the correlation between the GDP growth and the estimated factor is more apparent. Overall, it seems that the factor is tracking Czech GDP growth well, both in times of expansion and contraction.

Figure 5.1: Factor estimated in 2021

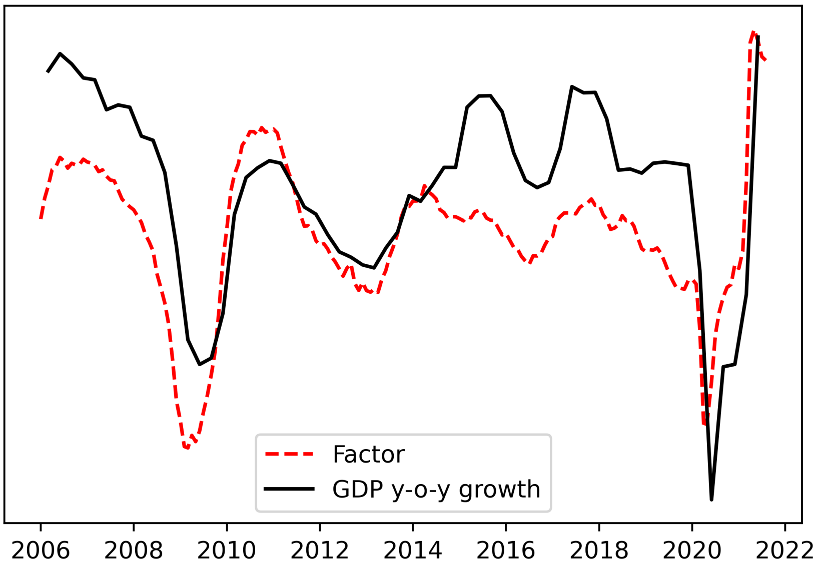
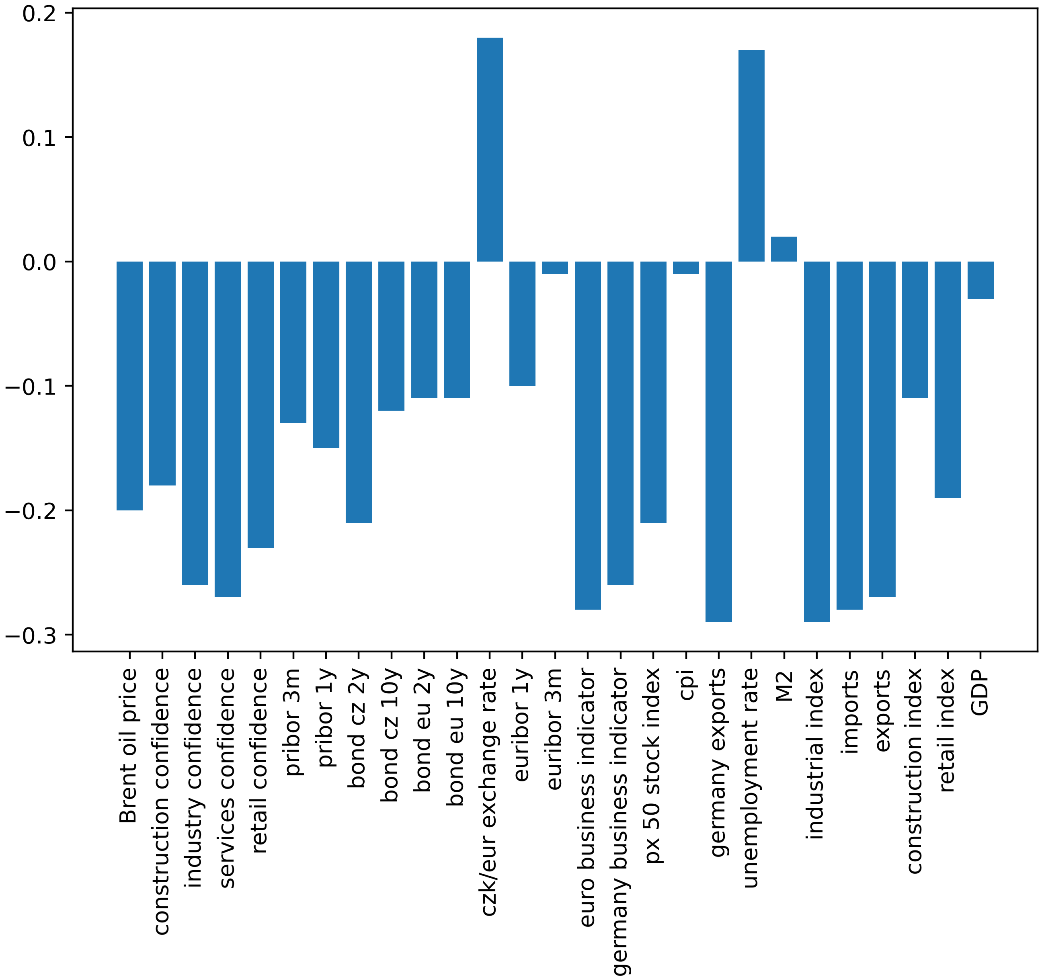


Figure 5.2 reports the estimated loadings of all the explanatory variables on the factor. According to the loadings, all but 3 of the series are procyclical. Unemployment, the CZK/EUR exchange rate, and M2 appear to be countercyclical, even though the estimated loading on M2 is minimal. It is worth noting that the international variables, except for the financial ones, have relatively large loadings, especially the Euro area business indicator and the Germany business indicator. Another group of variables with loadings of large magnitude are the hard series with industrial index, exports, and imports leading the way. Also, some domestic confidence indicators possess only slightly inferior loadings to those of the mentioned hard series. The group of series with the smallest factor loadings are the financial variables. These observations are in line with the findings of Rusnak (2013), who estimated the DFM on a very similar dataset.

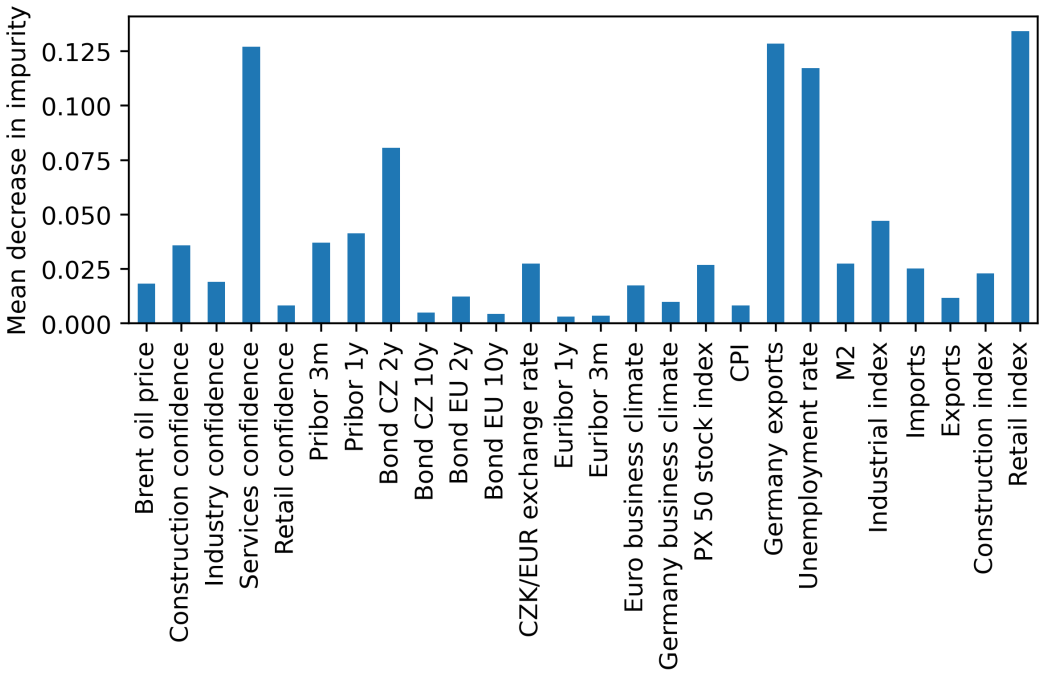
Figure 5.2: Factor loadings



Additionally, the importance of individual variables is examined using the in-sample predictions of one of the ML algorithms, the RF. To evaluate the importance of individual variables for the algorithm's predictions, the mean decrease of impurity (MDI) is considered. To understand the concept of MDI, one must start with the process of RF construction. RF operates by constructing multiple regression trees. Each tree is built by a sequence of nodes, in which the sample is split by a particular variable. An example of such a node can be a simple decision rule: if the construction confidence indicator is higher than 2, then the predicted GDP growth is 2%, and if it is lower than,2 then the prediction is 0.5%. In simple terms, the MDI measures how many times each variable has been used to split the sample, impacting the prediction-making process.

Figure 5.3 reports the MDI of all 26 of the monthly series. The two variables with the highest MDI are the services confidence indicator and the retail index. Another two variables with relatively large importance are German exports and the unemployment rate. These variables represent the groups that have relatively high factor loading in the DFM. Following the trend, the financial variables do not appear to have significant importance in the RF model's predictions, perhaps except for the two-year bond yield of the Czech Republic.

Figure 5.3: Random Forest: variable importance



5.2 Nowcasting performance of individual models

The collection of models that are estimated can be divided into 5 groups:

1. DFM
2. MIDAS
3. Factor MIDAS
4. ML Algorithms
5. Naive benchmarks

For each of the groups, several specifications are estimated. In the first step, the best-performing model for each class is selected. This could be non–trivial in case that the relative performance of two competing models changes for different nowcasting horizons, i.e., model A beats model B for certain nowcasting horizons but model B beats model A in others. In this case, the model with the lower average RMSE of all the 16 nowcasting horizons is selected. Subsequently, the predictive accuracy of the models representing each class is compared between them.

Models with 1 to 4 global factors are estimated for the DFM class. Additionally, a specification with one global factor and a factor that loads only on the real and nominal variables (DFM (real, nominal)) is estimated. These five specifications are estimated with factors modeled with 1 and 2 lags. The two best specifications measured by the RMSE are the one global factor with one lag and the DFM (real, nominal) model with all factors modeled with two lags. The former performs slightly better for longer horizons, while DFM (real, nominal) is superior in predictions starting with the last month of the reference quarter. Overall, DFM (real, nominal) has a slightly lower average RMSE and therefore is selected as the representative model for the DFM class.

The MIDAS models are estimated using the exponential Almon and beta lag polynomials, respectively. The predictions are constructed as a weighted average of point predictions of 26 bivariate models, using all the monthly predictors in our dataset. The process is explained in detail in chapter 4. The predictions are nearly identical for both specifications, with the model with exponential Almon lag polynomial being selected as the best performing MIDAS model based on average RMSE.

The factor augmented MIDAS models estimation consists of 2 steps. First, the monthly dynamic factors are extracted from the dataset identically as in the DFM models. Next, these factors are used as predictors for the MIDAS model. One and two global factors are extracted from the dataset and used as a regressor in the MIDAS regression – again, the factors are modeled with lags of order 1 and 2. The specification with 1 factor modeled with a lag of order 1 performs the best across all nowcasting horizons.

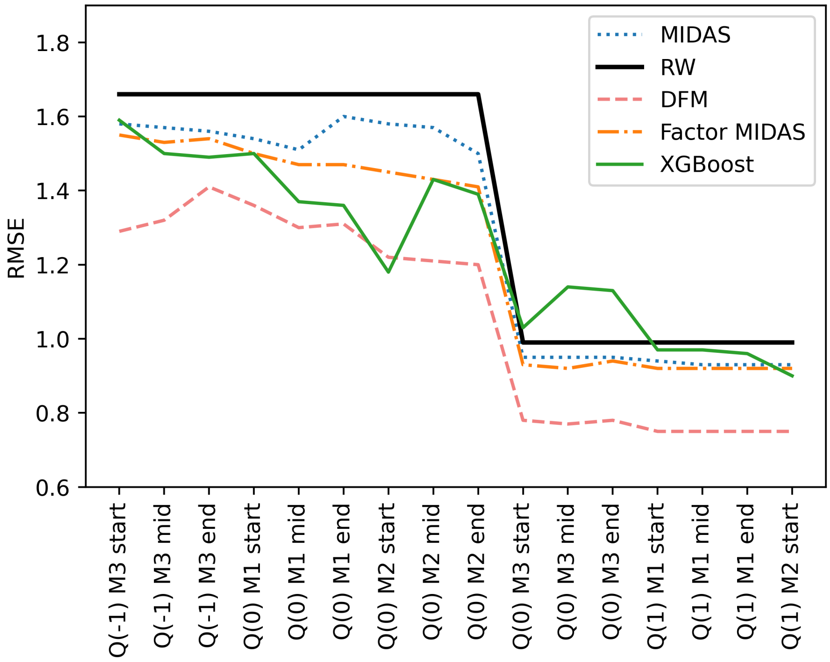
The ML algorithms include two tree-based methods (XGBoost, Random Forest) and Support Vector Regression. Two approaches are tested. In the first one, predictors include all 26 monthly variables from the dataset. The monthly variables are converted to quarterly frequency by taking their mean in the given quarter. In the second approach, following Soybilgen & Yazgan

(2021), dynamic factors are first extracted from the dataset and then used as predictors for the ML models. Again, the monthly factors need to be converted to quarterly frequency. As Soybilgen & Yazgan (2021) suggest, I also opt to use the estimated factor from the last month of the quarter to be considered as the quarterly factor. As I do not dispone with as large a dataset as Soybilgen & Yazgan (2021), only two alternatives are considered for the factor augmented ML models, with 3 and 2 factors, respectively. Before the estimation of each model, its parameters are found using a parameter tuning process which is described extensively in chapter 4. The general conclusion from the comparison of the performance of the ML methods is that the models with all variables as predictors outperform the models with monthly factors as explanatory variables. The two tree-based algorithms yield similar predictive accuracy, both outperforming SVR. XGBoost is the best model overall and therefore is the reference ML model in further analysis.

Finally, three naive benchmarks are considered – Random Walk (RW), Moving Average model using the four last available GDP figures at the time of the forecast (MA(4)), and an Autoregressive model of order 2 (AR(2)). The RW model and AR(2) model clearly outperform MA(4) for all nowcasting horizons. AR(2) is slightly more accurate than RW for shorter nowcasting horizons (predictions starting in the last month before the reference quarter through the 2nd month of the reference quarter), while RW yields better results from the 3rd month of the reference quarter. Overall, RW has the lowest average RMSE from the benchmark models and will serve as a point of comparison for all the other models.

Now that the best-performing model from each class has been selected, their performance across all forecasting horizons can be compared. Figure 5.4 shows the RMSE of the selected models for predictions starting on the 10th day of 1 month prior to the reference quarter and ending on the 10th day of the 1st month after the reference quarter when the CNB releases the first estimate of the GDP.

Figure 5.4: RMSE of selected models



The DFM, MIDAS, and the synthetic factor augmented MIDAS model beat the best performing benchmark model, the RW, in every nowcasting horizon. The selected ML algorithm, XGBoost, provides better predictions than the benchmark, measured by RMSE, in all but three nowcasting horizons. However, XGBoost is the best performing model overall for predictions made on the 10th day of the second month of the nowcasted quarter. Another interesting observation is that the factor-augmented MIDAS model beats the simple MIDAS regression in all horizons but fails to beat the linear DFM model.

The Diebold – Mariano test to statistically evaluate the significance of differences in the predictive abilities of forecasting models is performed for every pair of estimated models. The results indicate that all the DFM, MIDAS, MIDAS-DFM, and XGBoost are superior predictive models than the benchmark (RW) at a 99% significance level. Furthermore, the DFM is superior to XGBoost at a 95% significance level. Other than that, the results do not indicate any statistically significant differences between the competing models. This can be attributed to the relatively small sample of observations for the baseline model.

Overall, we can observe that the performance of all the models generally improves as new data arrives, although this does not hold for every data release. The most significant decrease of the RMSE for all the models occurs at the beginning of the third month of the reference quarter. This corresponds to the release of the GDP figure for the previous quarter. Overall, the results improve by almost 50% when comparing the accuracy of the first predictions to that of the last nowcast made for all the models. The relative order of the performance of selected models across the nowcasting horizons is stable with the exception of the XGBoost regressor. The accuracy of XGboost relative to the other models fluctuates significantly, as is apparent from it being the best overall model for one horizon but also having the worst performance out of all the models in the third month of the nowcasted quarter.

Table 5.1 shows the average RMSE for all the selected models, both absolutely and relative to the RMSE of the benchmark model.

Table 5.1: Average RMSE of selected models

	RMSE (absolute)	**RMSE (relative to RW)**
RW	1,37	1,00
DFM	1,06	0,78
MIDAS	1,29	0,94
Factor MIDAS	1,22	0,89
XGBoost	1,24	0,91

Note: The figures refer to RMSE averaged over all nowcasting horizons in the out of sample nowcasting exercise starting in the first quarter of 2012 and ending with last quarter of 2019.

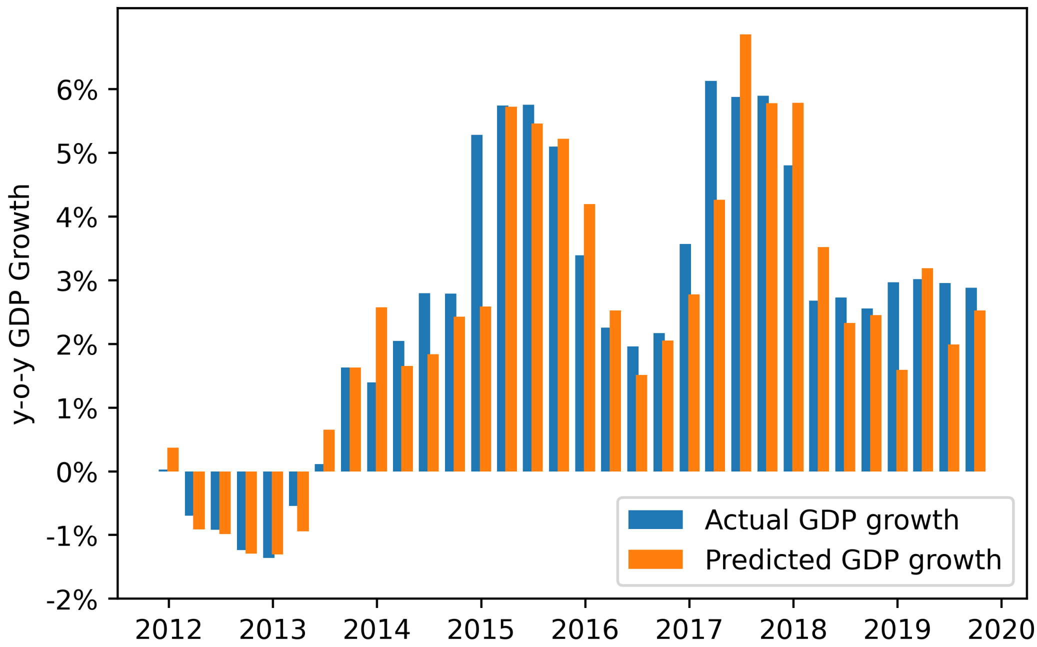
All 4 of the selected models perform on average better in the out-of-sample nowcasting exercise than the naive benchmark. Overall, the best model is the DFM, which reaches RMSE lower by 22% than the benchmark model. It is followed by the factor augmented MIDAS and the XGBoost algorithm, whose RMSE is able to deduct 11% and 9% from the RW model's RMSE respectively. The predictions of the MIDAS model, which are constructed as a weighted average of the predictions of a set of bivariate MIDAS regressions, have on average prediction error 6% lower than the benchmark as measured by the RMSE. These results provide further evidence of the very good performance of the DFM in GDP nowcasting tasks. Compared to the benchmark model, its nowcasting error is in line with the existing literature, where it has been shown that generally, DFM outperforms benchmark models by 15% – 30%.

Figure 5.5 shows the GDP as predicted by the best performing individual model, the DFM, along with the actual GDP growth taken from the final data vintage from August 2021. The displayed predictions are made on the 30th day of the last month of every quarter. We can see that the model nowcasts Czech

GDP YoY growth well across the entire 8–year period, as the predictions never divert noticeably from the actual GDP growth even during spikes and dips.

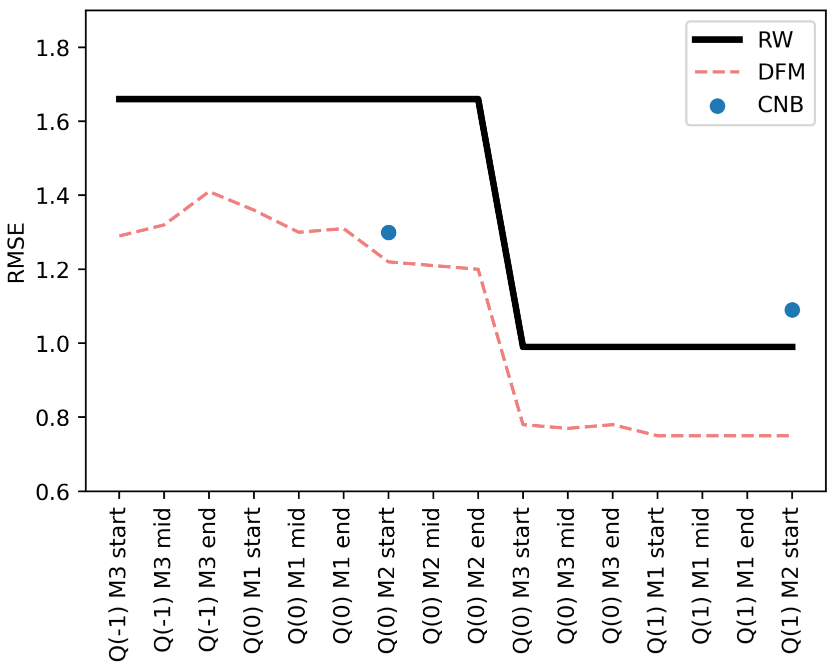
One interesting observation that can be made is that the predictions of GDP growth seem to be, in most cases, a slight delay of the actual GDP growth. This can be attributed to the lag of release of many variables, especially the hard macroeconomic data, which is usually released around 40 days after the end of the reference month. As was shown, the hard series are the most important GDP growth predictor overall, and therefore, the timeliness of the information is an inherent issue of the nowcasting models.

Figure 5.5: GDP growth predicted by the DFM



Now that the nowcasting performance of the individual models has been discussed and those with the best accuracy have been identified, it is possible to compare them with the official forecasts of the CNB. CNB announces their GDP forecast typically at the beginning of the 2nd month of each quarter, although the exact dates differ from quarter to quarter. For the quarters that are part of our out-of-sample nowcasting exercise, the CNB announced the forecast as early as on the last day of the first month of the reference quarter and as late as on the 7th day of the second month of the reference quarter. For simplicity, I assume that the CNB nowcast is released simultaneously with the 7th nowcasting round in each quarter, which is done on the 10th day of the second month in every quarter. In each quarter, CNB announces predictions for the last quarter, current quarter, and then for the following quarters up to 5 quarters ahead. For comparison with the predictions of our models, we are only interested in those CNB nowcasts that fall within the same nowcasting horizons examined in this analysis. Therefore, there are two relevant CNB nowcast in each quarter: the one during the reference quarter and the one made at the beginning of the 2nd month of the next quarter, which is also the last nowcasting round in our analysis. Figure 5.6 shows the RMSE of the selected best-performing model (the DFM) along with the average RMSE of official CNB nowcasts, evaluated using the GDP figures from the last available vintage. However, it is important to note that while the CNB nowcasts were made in a purely real-time setting, I use pseudo-real-time data vintages that account for the lags of individual data releases but do not account for subsequent revision that some of the variables are subject to. This could potentially favor the performance of the models used in this thesis compared to the GDP predictions of the CNB.

Figure 5.6: RMSE of the DFM model compared with CNB nowcast



The RMSE of the CNB nowcast made at the beginning of the second month of the reference quarter is slightly higher than the RMSE of the best model used in this analysis, the DFM. The CNB nowcast outperforms comfortably the benchmark model for this nowcasting horizon. However, the accuracy of the CNB official nowcast does not improve significantly in the backcast, which is made at the beginning of the second month of the following quarter. Here, the CNB nowcast is not only beaten by the DFM model's backcast, but it also fails to outperform the benchmark model. This result provides strong support to the usefulness of purely statistical models, and the DFM in particular, in comparison to the official CNB forecast adjusted with expert judgment. Another essential thing to consider is that CNB only releases its forecasts once each quarter. Therefore, at any point, CNB's best prediction is its last available forecast, which can be up to 3 months ago for any moment of reference. On the contrary, the models used in this thesis are not dependent on expert judgment and can be updated automatically for any release of relevant data. As shown, the performance of the models generally improves with additional data released before, during, and after the nowcasted quarter.

It is important to emphasize that all the results presented in this section are reached with the out-of-sample predictions that exclude the period affected by the COVID–19 pandemic. As argued earlier, the pandemic period was a completely unprecedented shock to the global economy and could cause bias in comparing the performance of the models in GDP nowcasting. Because of that, I opted to exclude the covid period from the baseline analysis. However, all the nowcasts were also produced for the quarters affected by the pandemic and presented in the subsection 5.5. The results indicate that the performance of all the nowcasting models diminished substantially during the covid-effected period. The CNB judgment-adjusted official nowcasts in this period are superior to every model estimated.

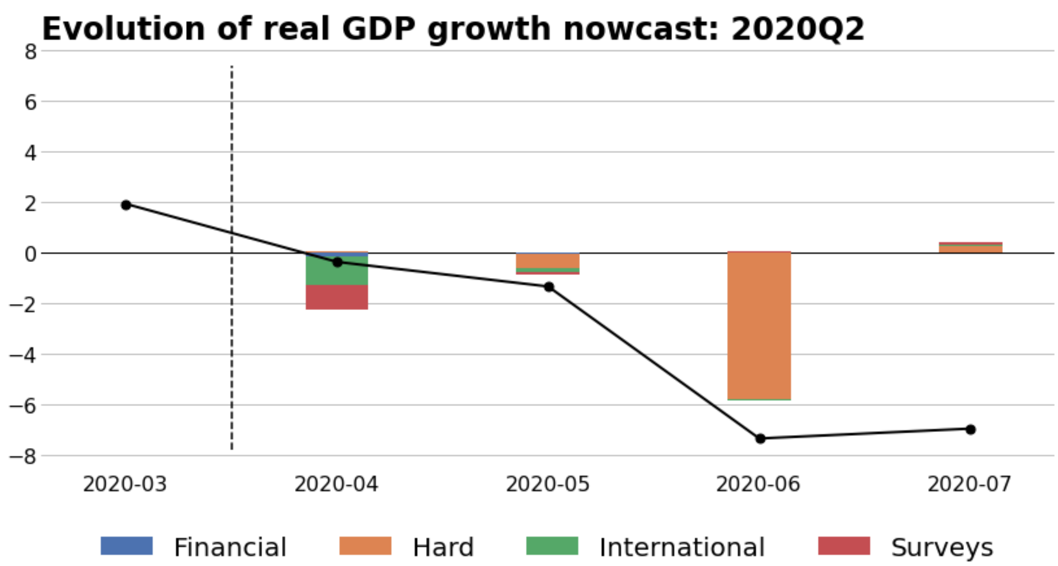
5.3 Impact of new data releases

As nowcasts are updated after new data is released throughout the quarter, it is of interest to examine how different data publications impact the predictions and the accuracy of a specific model. We can, in theory, observe how the RMSE changes for any data release of any model, but this does not link the change in nowcast with a specific variable, only with the group of variables newly released for the given vintage. As was shown on the different models's nowcasting performance, the most significant improvement of predictive accuracy comes with the release of GDP for the previous quarter in the first half of the 2nd month of the current quarter. For the other vintages, corresponding to new data releases, the nowcasting improvement has been generally modest.

However, the DFM allows for a more detailed decomposition of the source of changes of every new nowcast. As the DFM creates forecasts for all the variables, it will enable to attribute a change of any nowcast to the difference between the actual value of a specific variable and its expected value before its release. Banbura & Modugno (2013) develop a methodology to quantify the impact of changes in the information set on the nowcasts revision for each variable, as presented in chapter 4.

First, figure 5.7 shows how the *news* arriving throughout the quarter effect the DFM's nowcast of the GDP growth for the second quarter of 2020. This quarter was chosen because it is the first quarter in which the COVID-19 pandemic heavily impacted GDP growth, and it is a good example of how new data releases can affect the model's predictions. The effect of news is summarized for the four groups of variables available in the dataset.

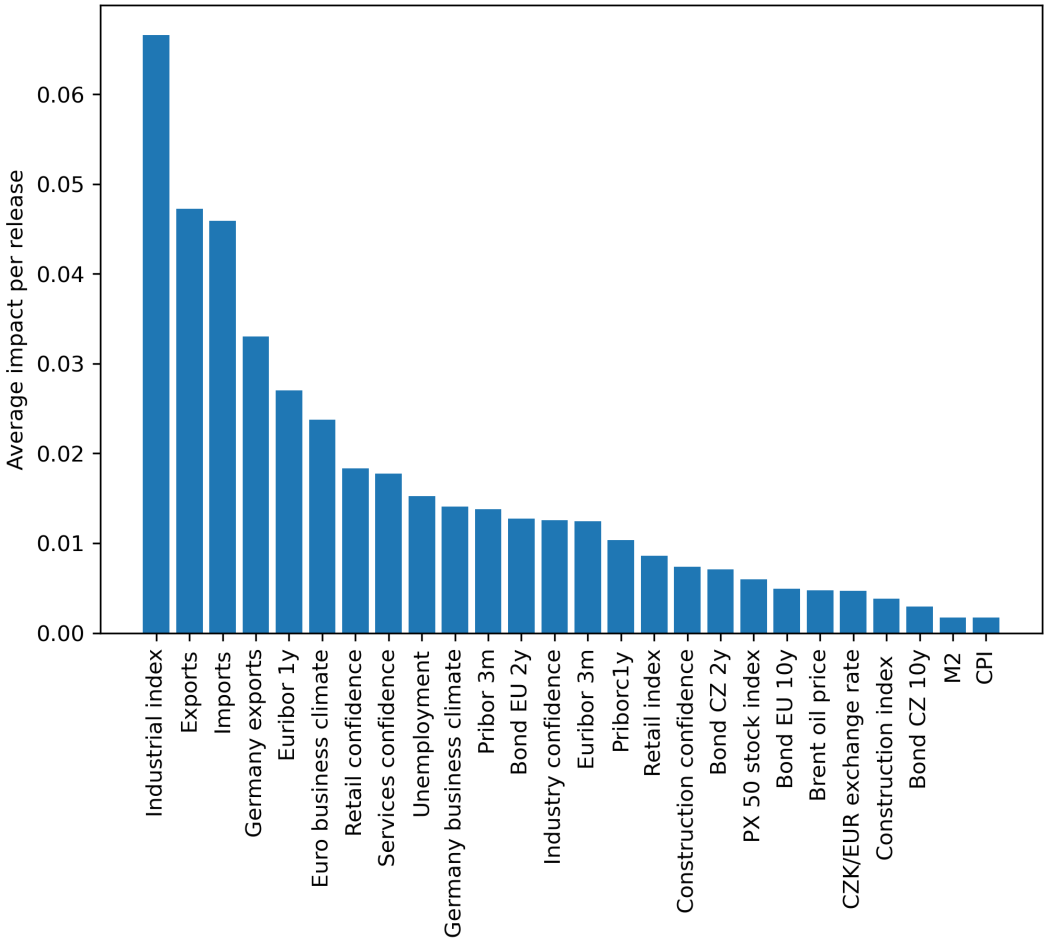
Figure 5.7: The impact of new data releases on 2020Q2 GDP nowcast



The figure does not show the impact of every new data release, but the monthly frequency is sufficient to display all the changes occurring, as each variable is only released once a month. The baseline nowcast, conducted in March of 2020, still paints a relatively rosy picture of the economy as the GDP is predicted to grow YoY. The new data released in April causes the prediction to be adjusted downwards due to global variables and newly released confidence indicators, which were lower than their previous forecast. The same trend continues in May; however, here, the change can be attributed to the publication of the hard series. It is important to remember that the hard variables are almost exclusively released with a lag of about 40 days, so it takes longer for them to signal turns in the economy. The most considerable downgrade of the GDP growth nowcast occurs in June and is almost solely due to the release of the GDP growth for the previous quarter. As was described previously, the release of GDP in the prior quarter was the single most significant factor in increasing nowcast accuracy. Finally, the nowcast did not change much in July, still predicting the GDP to decrease by about 6.5% YoY. The actual GDP growth in the second quarter of 2020 was –10.9%, so even the last nowcast underpredicted the magnitude of the downturn significantly, but the effect of new data releases is noticeable.

Figure 5.8 reports the average impact of new data releases for all the variables from every nowcasting round. The mean is taken from the absolute value of the effects so that only the size of the impact matters, not its direction. It is apparent that the hard variables were the source of the majority of the updates to GDP nowcasts, with the Industrial index, Export, and Imports having the most significant average impact. The hard variables are followed by the exogenous variables and several confidence indicators, respectively. The domestic financial variables did not significantly impact the nowcast adjustments. These results align with the in-sample analysis of the importance of individual variables. However, it is important to emphasize that the average impact of new data releases does not correlate entirely with the predictive power of a given variable, as it may also be the effect of frequent unanticipated changes for this series.

Figure 5.8: The average impact of variables on nowcast adjustments



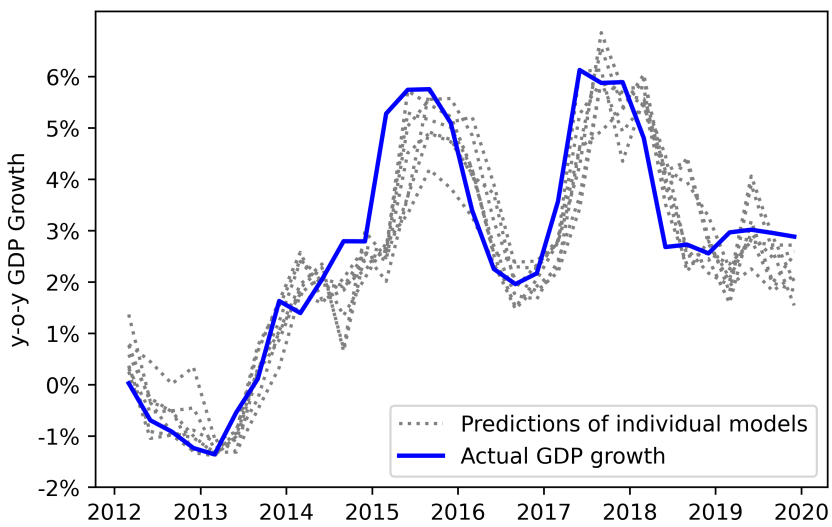
5.4 Results of model averaging

A similar strategy to Bjornland *et al.* (2012) is adopted for nowcast averaging. The basic idea consists in a weighted average of point nowcasts of individual models, with the weights being proportional to the inverse of each model's RMSE. The averaging process is done separately for each nowcasting horizon.

The potential benefit of the nowcast combination can be threefold. First, it can help mitigate uncertainty connected to the predictions of individual models. For the nowcasting models used here, the Diebold Mariano test finds the differences between the nowcasting models as statistically insignificant for almost all cases. The exception is the DFM, whose nowcast performance is superior to XGBoost at a 95% significance level. This provides a clear indication that there remains a significant degree of uncertainty in selecting a specific model as the most appropriate one to nowcast Czech GDP growth. The second source of benefit from model averaging is relevant in cases where particular models predict well in most periods, but few of the predictions have severe errors. The weighted averaging process can help eliminate these outliers and stabilize the predictive performance over the entire sample period. Finally, in some instances, model averaging can actually result in better predictions than even the best individual model. This is possible when some of the individual models overpredict the target variable while others predict lower values than the actual observation of the target variable. The combination can potentially average these errors out.

Figure 5.9 provides some insight into the possible benefits of model averaging for the GDP nowcasting done in this thesis. It plots the GDP growth predicted by seven individual models at the end of the reference quarter along with the actual GDP growth. Even though, in most cases, all models either overestimate or underestimate GDP growth, there are a few instances in which some models predict higher than actual while others underpredict GDP growth. This is apparent especially in the year 2016. During these periods, the weighted averaging process can be especially effective. Another interesting observation is that the relative precision of the individual nowcasting model changes at many points – some models predict GDP growth better than others for the specific quarter but are beaten in other quarters. This is also promising in regards to the potential gains of nowcast averaging.

Figure 5.9: GDP predicted by the selected models



Two alternatives are considered for the model averaging itself: the weighted average of the seven best individual models and the simple average of these models. I opt not to include more models in the averaging process for two reasons. Firstly, models from the same class that only have different lag structure or different lag polynomial, in the case of MIDAS, appear to result in nearly identical predictions in most cases. Therefore, their inclusion in the weighted average construction would not provide additional information. Furthermore, Bjornland *et al.* (2012) find that the forecast average of the eight best models yielded better predictions than those reached by a weighted average of more models. The seven models used for the averaging process include DFM with one global factor; DFM with a global, real, and nominal factor; MIDAS model with the exponential Almon lag polynomial; MIDAS-DFM joint model; XGBoost; Random Forest and the Support Vector Regression.

The weighted average of point nowcasts is slightly superior to the simple mean. The figure 5.10 plots the RMSE of the weighted average of the nowcasts, the RMSE of the best individual model (the DFM), the benchmark model as well as the RMSE of the official CNB nowcasts for all nowcasting horizons. The RMSE of the model combination is close to the RMSE of the DFM for all the horizons, but the ensemble model does not beat the DFM at any nowcasting horizon. Before the GDP for the previous quarter is released at the beginning of the second month in the nowcasted quarter, the accuracy of model combination is nearly identical to that of the best performing individual model. From that point on, the DFM beats it by a small margin. The ensemble model outperforms the benchmark model for all nowcasting horizons and beats the CNB GDP nowcast at both horizons for which they are announced. Overall, the average RMSE is nearly identical to that of the best individual model, the DFM, with the latter having the average RMSE lower by 0.9%. The performance of the weighted average of 7 nowcasting models is very good, considering it predicts the YoY GDP growth on par with the DFM and beats all the other individual models. The ensemble model's usefulness consists of the mitigation of uncertainty, which is inherently present to the choice of an appropriate model to nowcast GDP growth.

Figure 5.10: The RMSE of the ensemble model

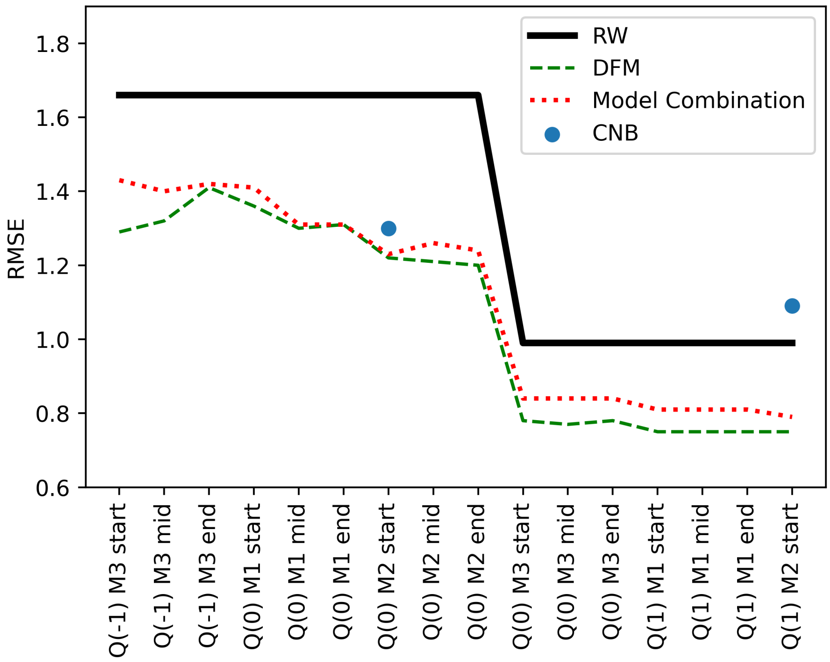
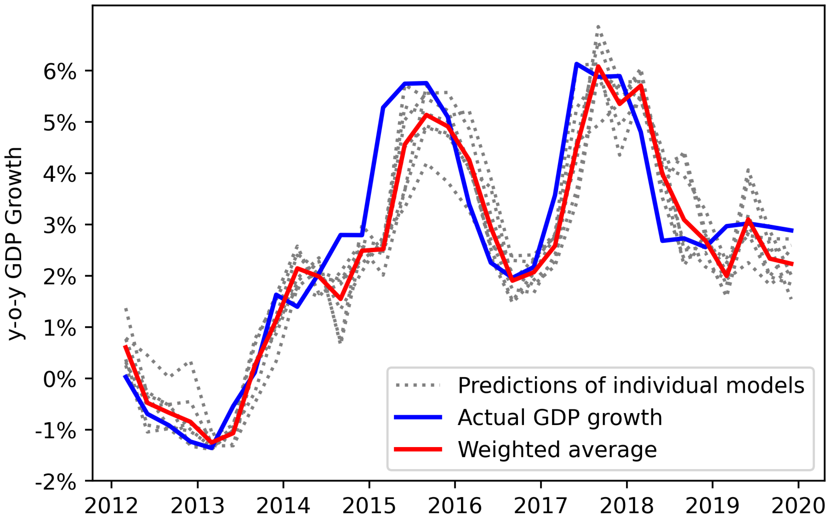


Figure 5.11 shows the YoY GDP growth predicted by the weighted average of individual models compared to the predictions of the individual models made at the end of each quarter. The predictions are also compared to the actual GDP growth. It is apparent that the weighted average rarely provides the best prediction. However, its performance is very stable – only in 6 of the 32 nowcasted quarters does its error surpass 1%. The ensemble model predicts GDP growth better than any individual model in 4 quarters, 2 of them in 2016. In these two quarters, 4 of the individual models underpredict the YoY GDP growth while the remaining three predicted it to be higher than it actually was. This perfectly illustrates the situations in which model averaging can outperform even the best individual model.

Figure 5.11: GDP predicted by the ensemble model



5.5 The impact of COVID-19 on nowcasting models

For all the models presented up to this point, the baseline nowcasting exercise excluded the period affected by the COVID-19 pandemic. As argued earlier, this period was non-standard in recent history to such a degree that its inclusion in the relatively modest-sized sample could cause bias in the comparison of the performance of individual models. However, it is important to address the performance of all the models also in this turbulent period and observe how purely statistical nowcasting methods were able to deal with these unprecedented economic conditions.

Therefore, this section replicates the entire analysis conducted on the baseline period (2012Q1-2019Q4) with the inclusion of covid-affected quarters (2012Q1-2021Q2). The main takeaways are the following:

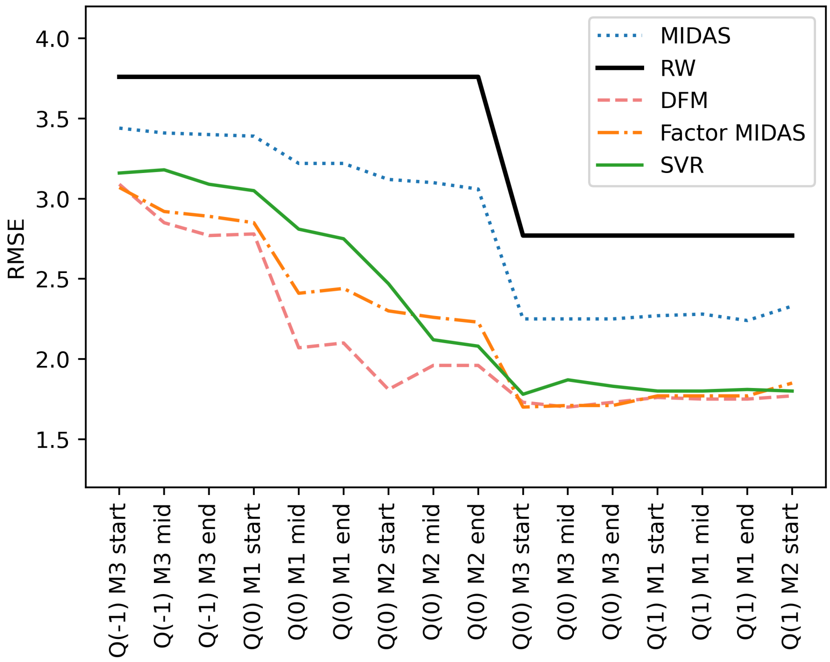
- The performance of all the models decreases substantially during the global pandemic
- All the models beat the benchmark model by an even larger margin
- The performance of the models relative to each other remains very similar
- The impact of new data releases during the nowcasted quarter on nowcast accuracy is greater during the pandemic

- The official CNB nowcasts outperform the model-based nowcasts

As was done in the baseline analysis, first, I select the best specification for each model class. In some cases, the best-performing specifications are the same as in the baseline analysis – DFM (real, nominal) and the Factor MIDAS specification with 1 factor modeled with one lag. For the MIDAS regression, the specification with beta lag polynomial slightly edges the specification with the exponential Almon lag polynomial. In the case of ML algorithms, the SVR outperforms the two tree-based algorithms by a small margin. Lastly, the RW remains the best performing model from the three considered benchmarks.

Figure 5.12 presents the RMSE of the selected models across all nowcasting horizons. This time all the models perform better than the benchmark for every horizon. The DFM performs the best for longer-term predictions. For nowcasts starting with the last month of the nowcasted quarter, the DFM, Factor MIDAS, and the SVR provide almost identical accuracy. The MIDAS regression lacks noticeably behind these three models.

Figure 5.12: RMSE of the selected models



One interesting observation is that the predictive accuracy improves substantially with new data releases throughout the nowcasted quarter. This was not the case to such a degree for the baseline analysis, where the only substantial gain in accuracy occurred through the release of the GDP for the previous quarter. The performance of the models does not improve in the month after the nowcasted quarter.

The Diebold – Mariano test is conducted for every pair of estimated models, but no statistically significant differences were found at 95% significance level. This is not necessarily surprising considering the relatively short sample and supports pursuing a model combination approach as model uncertainty remains an issue.

Table 5.2 reports the average RMSE of all the models both in absolute terms and relative to the RW benchmark. The RMSE of all the models is significantly greater than in the baseline analysis, which excluded the period affected by the global pandemic. For illustration, the RMSE of the DFM almost doubled. The results show that the DFM remains the best performing individual model, but the accuracy of the Factor MIDAS model and the SVR is not far behind. The models are able to improve upon the benchmark RW model by a more significant margin than in the baseline analysis – DFM outperforms it by about 37%.

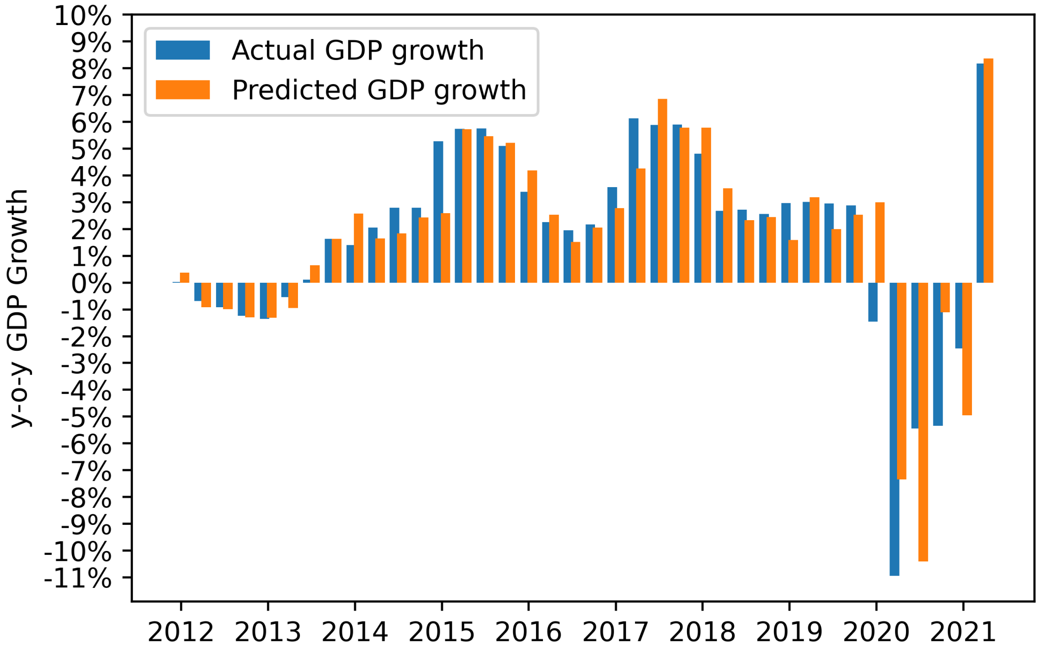
Table 5.2: Average RMSE of selected models (including the pandemic)

	RMSE (absolute)	**RMSE (relative to RW)**
RW	3,33	1,00
DFM	2,10	0,63
MIDAS	2,83	0,85
Factor MIDAS	2,23	0,67
SVR	2,34	0,70

Note: The figures refer to RMSE averaged over all nowcasting horizons in the out of sample nowcasting exercise starting in the first quarter of 2012 and ending with second quarter of 2021.

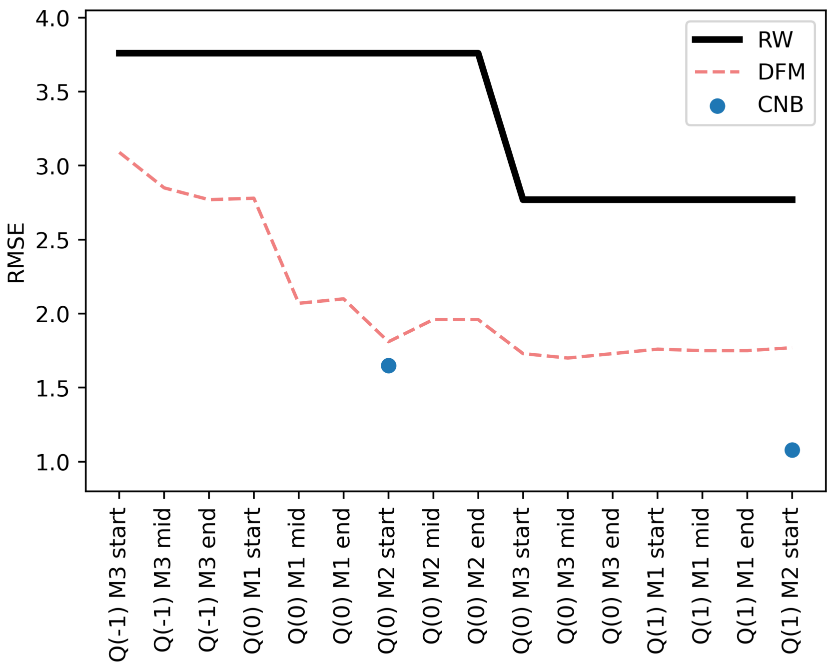
Figure 5.13 plots the GDP YoY growth as predicted by the DFM at the end of each quarter. The predictions clearly start to deteriorate with the first quarter of 2020 as the COVID-19 pandemic begins to take its toll on the economy of the Czech Republic. From the first quarter of 2020 to the first quarter of 2021, the average error of the DFM is 4.1%. The predictive accuracy seemingly gets back on track in the second quarter of 2021; therefore, it is of interest to observe how the DFM will nowcast Czech GDP growth in the following quarters.

Figure 5.13: GDP growth predicted by the DFM



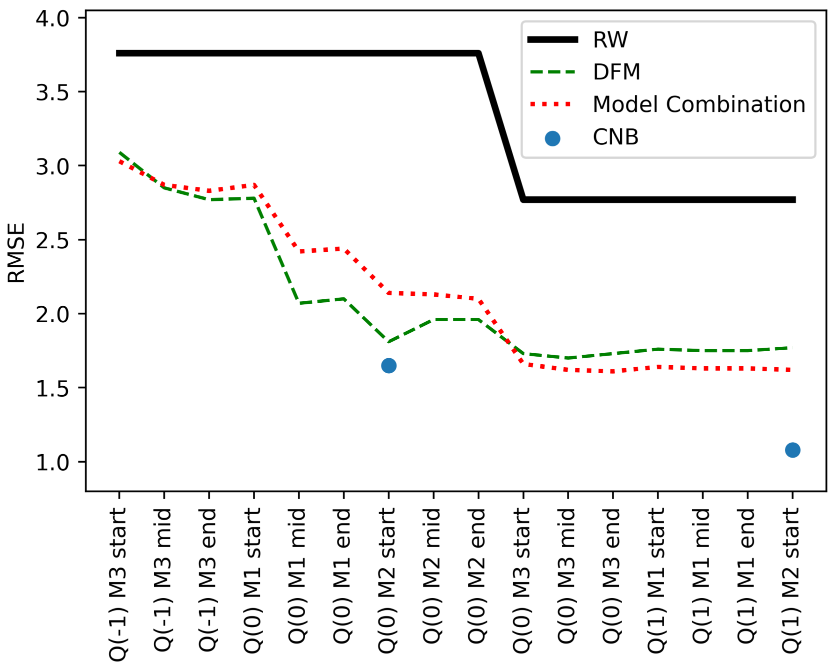
As was shown on the average RMSE of the selected models, the prediction error increased significantly during the global pandemic. Figure 5.14 shows how the accuracy of the DFM compares with the official GDP nowcast released by the CNB when the COVID-effected period is included. The CNB predictions outperform the DFM both for the nowcast during the reference quarter and the backcast conducted at the beginning of the second month of the following quarter. For the predictions made during the nowcasted quarter, the DFM is not lacking much to the accuracy of CNB nowcasts, with the latter having the RMSE lower by 9%. However, in the backcast, the official predictions of the central bank are clearly superior, beating the DFM's results by 39%. These results indicate that expert judgment adds substantial value to the nowcasts during the very non-standard economic situation caused by the pandemic, especially for short-term predictions.

Figure 5.14: RMSE of the DFM and the CNB official forecast



Finally, model averaging was also conducted with the inclusion of the additional six quarters that were affected by the pandemic. Again, the seven best-performing individual models were included in the weighted-averaging process. The figure 5.15 plots the RMSE of the model combination for every nowcasting horizon, along with the RMSE of the DFM, the CNB, and the benchmark model. The ensemble model clearly outperforms the benchmark also with the inclusion of the period affected by the pandemic and for seven shortest horizons also slightly beats the DFM. For longer-term predictions, the DFM is superior to the ensemble model. The average RMSE of the DFM is lower by about 2% than the average RMSE of the combined nowcasting model. Apart from the DFM, the ensemble model outperforms all the other individual models.

Figure 5.15: RMSE of the model combination



Chapter 6

Conclusion

This thesis compares the performance of several methods in nowcasting YoY GDP growth in the Czech Republic in a pseudo-real time setting. The models used include those commonly employed in GDP nowcasting literature, such as the DFM or MIDAS, but also ML algorithms that have been proposed only recently in this field. The performance of all models is evaluated in an out-of-sample nowcasting exercise starting from 2012. Additionally, an approach to nowcast averaging is adopted to examine the potential gains of model combination. Finally, the effect of new data releases on the accuracy of the nowcasts is examined and discussed.

The results suggest that the DFM outperforms other models in the pseudo-real-time nowcasting exercise. Furthermore, the remaining models provided better GDP predictions than a benchmark RW model. The differences between nowcasts of the individual models and the naive benchmark were statistically significant even with limited sample size. Nowcast averaging also yielded promising results. The predictions of the ensemble model, constructed as a weighted average of 7 individual nowcasting models, were on par with the best individual model. The nowcasting performance of estimated models was also compared with the official CNB nowcast of GDP growth. In the baseline model, which excludes the period impacted by the global pandemic starting in 2020, the DFM provides comparable accuracy to the CNB nowcasts at the start of the quarters and beats it in the nowcasts and backcasts conducted in the last month of the quarter and the first month of the following quarter.

These results provide strong evidence of the good predictive performance of the DFM in nowcasting Czech GDP growth. The ability of this nowcasting method to leverage the latest available data is a highly useful tool for policymakers. I intend to implement this nowcasting framework on a website that would display the current predictions of the model for GDP growth in the Czech Republic, updated daily. Another contribution of this thesis is the comparison of the performance of a wide range of methods, some of which have been used only scarcely in the GDP nowcasting literature and never in the case of the Czech Republic. A comprehensive comparison of the performance of the methods is conducted. The result shows that the factor augmented MIDAS and several ML algorithms perform well in nowcasting Czech GDP growth. Further evidence of the relevance of the estimated nowcasting models is gained from the comparison of their predictions with the official CNB nowcasts of the GDP, as several models appear to be superior for short-term predictions. Finally, using a weighted average of several nowcasting models proved not only to address the issue of model uncertainty but also to perform very well in terms of the accuracy of the predictions.

The out-of-sample nowcasting exercise was conducted for the period 2012-2019 that ends before the start of the COVID-19 pandemic as well as for the period 2012-2021. The relative performance of the individual nowcasting models did not change significantly during the global pandemic, with DFM still providing the most accurate nowcasts. All the estimated models beat the benchmark model's predictions by a more considerable margin when the period affected by the pandemic is included in the analysis. However, the accuracy of the predictions of all models diminishes significantly during the pandemic both in absolute terms as well as comparatively with the official CNB judgment-adjusted nowcasts.

Future research in GDP nowcasting for the Czech Republic or other countries can focus on the effect of different phases of the economic cycle on the performance of individual methods. The global pandemic highlighted the sharp decline of nowcasting accuracy of most models in times of increased volatility and rapidly changing economic conditions. There is already growing literature experimenting with alternative high-frequency indicators, such as electricity consumption or google searches, to predict development in economic activity in a more timely manner than the standard methods relying primarily on monthly series. Future research could merge these two approaches and create a combined model that would dynamically assign weights to individual models based on the economic cycle's current phase.

Bibliography

Aastveit, K. M., K. R. Gerdrup, A. S. Jore, & L. A. Thorsrud (2011): "Nowcasting GDP in real-time: A density combination approach." *Working Paper Series 11*, Norges Bank.

Adam, T., O. Michalek, & A. Michl (2021): "The Rushin Index: A Weekly Indicator of Czech Economic Activity." *Working Paper Series* .

Arnostova, K., D. Havrlant, L. Ruzicka, & P. Toth (2011): "Short-Term Forecasting of Czech Quarterly GDP Using Monthly Indicators." *Czech Journal of Economics and Finance (Finance a uver)* **61(6)**: pp. 566–583.

Banbura, M., D. Giannone, M. Modugno, & L. Reichlin (2013): "Now-Casting and the Real-Time Data Flow." *Working Paper Series* **(1564)**: pp. 566–583.

Banbura, M., D. Giannone, & L. Reichlin (2010): "Nowcasting." **1275**.

Banbura, M. & M. Modugno (2013): "Maximum likelihood estimation of factor models on data sets with arbitrary pattern of missing data." *Journal of Applied Econometrics* .

Barhoumi, K., O. Darne, & L. Ferrara (2010): "Are disaggregate data useful for factor analysis in forecasting French GDP?" *Journal of Forecasting* **29(1-2)**: pp. 132–144.

Bates, J. M. & C. W. J. Granger (1969): "The combination of forecasts." *Operational Research Quarterly* **20(4)**: pp. 451–468.

Benda, V. & L. Ruzicka (2007): "Short-term Forecasting Methods Based on the LEI Approach: The Case of the Czech Republic." *Research and Policy Notes* .

Bernanke, B. & J. Boivin (2003): "Monetary policy in a data-rich environment." *Journal of Monetary Economics* **50(3)**: pp. 525–546.

Bjornland, H. C., K. Gerdrup, A. Jore, C. Smith, & L. A. Thorsrud (2012): "Does Forecast Combination Improve Norges Bank Inflation Forecasts?" *Oxford Bulletin of Economics and Statistics* **74(2)**: pp. 163–179.

Bragoli, D., L. Metelli, & M. Modugno (2014): "The Importance of Updating: Evidence from a Brazilian Nowcasting Model." *FEDS Working Paper* **94**.

Chakraborty, C. & A. Joseph (2017): "Machine learning at central banks." *Working Papers* **(674)**.

Chernis, T. & R. Sekkel (2017): "A dynamic factor model for nowcasting Canadian GDP growth." *Empirical Economics* **53(1)**: pp. 217–234.

Chikamatsu, K., N. Hirakata, Y. Kido, & K. Otaka (2018): "Nowcasting Japanese GDPs." *Working Paper Series* **(18)**.

Chin, D. M. & P. J. Miller (1996): "Using monthly data to improve quarterly model forecasts." *Quarterly Review* **20(1)**: pp. 16–33.

Clements, M. & A. B. Galvao (2008): "Macroeconomic Forecasting With Mixed-Frequency Data." *Journal of Business & Economic Statistics* **26**: pp. 546–554.

Doz, C., D. Giannone, & L. Reichlin (2006): "A Quasi-Maximum Likelihood Approach for Large, Approximate Dynamic Factor Models." *Review of Economics and Statistics* **94**.

Franta, M., D. Havrlant, & M. Rusnak (2016): "Forecasting Czech GDP Using Mixed-Frequency Data Models." *Journal of Business Cycle Research* **12(2)**: pp. 165–185.

Ghysels, E., P. Santa-Clara, & R. Valkanov (2004): "The MIDAS Touch: Mixed Data Sampling Regression Models." *CIRANO Working Papers* .

Giannone, D., L. Reichlin, & S. Simonelli (2009): "Nowcasting Euro Area Economic Activity in Real-Time: The Role of Confidence Indicators." *National Institute Economic Review* **210**: pp. 90–97.

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

Grui, A. & R. Lysenko (2017): "Nowcasting Ukraine's GDP Using a Factor-Augmented VAR (FAVAR) Model." *Visnyk of the National Bank of Ukraine* **242**: pp. 5–13.

Guagliano, C. & C. Mantovani (2014): "Real-Time Evaluation of GDP in Some Eurozone Countries." *SSRN Electronic Journal* **10**.

Gul, S. & A. Kazdal (2021): "Nowcasting and Short-term Forecasting Turkish GDP: Factor-MIDAS Approach." *Working Papers* **(2111)**.

Havranek, T., R. Horvath, & J. Mateju (2012): "Monetary transmission and the financial sector in the Czech Republic." *Economic Change and Restructuring* **45**: pp. 135–155.

Higgins, P. (2014): "GDPNow: A Model for GDP Nowcasting." *Working Paper 7*, Federal Reserve Bank of Atlanta.

Horvath, R. (2013): "Do Confidence Indicators Help Predict Economic Activity? The Case of the Czech Republic." *Czech Journal of Economics and Finance (Finance a uver)* **62**: pp. 398–412.

Jonsson, K. (2020): "Machine Learning and Nowcasts of Swedish GDP." *Journal of Business Cycle Research* **16(2)**: pp. 123–134.

Klein, L. R. & E. Sojo (1989): "Combinations of High and Low Frequency Data in Macroeconometric Models." *Economics in Theory and Practice: An Eclectic Approach* pp. 3–16.

Makridakis, S., E. Spiliotis, & V. Assimakopoulos (2018): "Statistical and Machine Learning forecasting methods: Concerns and ways forward." *PLOS ONE* **13(3)**: pp. 1–26.

Marcellino, M. & C. Schumacher (2008): "Factor-MIDAS for now- and forecasting with ragged-edge data: A model comparison for German GDP." *CEPR Discussion Papers* **(6708)**.

Mariano, R. S. & Y. Murasawa (2003): "A new coincident index of business cycles based on monthly and quarterly series." *Journal of Applied Econometrics* **18(4)**: pp. 427–443.

Parigi, G. & G. Schlitzer (1995): "Quarterly forecasts of the Italian business cycle by means of monthly economic indicators." *Journal of Forecasting* **14(2)**: pp. 117–141.

Proietti, T. & A. Giovannelli (2021): "Nowcasting monthly GDP with big data: A model averaging approach." *Journal of the Royal Statistical Society* **184(2)**: pp. 683–706.

Richardson, A., T. van Florenstein, & T. M., Vehbi (2019): " Nowcasting New Zealand GDP using machine learning algorithms." .

Rusnak, M. (2013): "Nowcasting Czech GDP in Real Time." *CNB Working Paper Series* .

Schumacher, C. & J. Breitung (2008): "Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data." *International Journal of Forecasting* **24(3)**: pp. 386–398.

Soybilgen, B. & E. Yazgan (2021): "Nowcasting US GDP Using Tree-Based Ensemble Models and Dynamic Factors." *Computational Economics* **57(1)**: pp. 387–417.

Timmermann, A. (2006): "Forecast Combinations." *Handbook of Economic Forecasting* **1**: pp. 135–196.

Trehan, B. (1989): "Forecasting growth in current quarter real GNP." *Economic Review* **2**: pp. 39–52.

Appendix A

Data

Table A.1: List of variables

Group	Variable	Revisions	Lag	Source
Hard	GDP	Y	68 to 71	CZSO
Hard	CPI	N	8 to 11	CZSO
Hard	Unemployment	Y	8 to 11	CZSO
Hard	Industrial Production Index	Y	37 to 45	CZSO
Hard	Imports	Y	38 to 45	CZSO
Hard	Exports	Y	38 to 45	CZSO
Hard	Construction Index	Y	38 to 45	CZSO
Hard	Retail Index	Y	38 to 45	CZSO
Hard	Import Price Index	N	38 to 45	Trading Economics
Hard	Export Price Index	N	38 to 45	Trading Economics
Financial	PRIBOR 3M	N	0	CNB
Financial	PRIBOR 1Y	N	0	CNB
Financial	Credit	N	30 to 31	Trading Economics
Financial	CZ BOND 2Y	N	0	CNB
Financial	CZ BOND 10Y	N	0	CNB
Financial	CZK/EUR Exchange Rate	N	0	Investing.com
Financial	PX-50 Stock Index	N	0	Investing.com
Financial	M2	Y	30 to 31	CZSO
Surveys	Construction Confidence Indicator	N	-7 to -2	CZSO
Surveys	Retail Confidence Indicator	N	-7 to -2	CZSO
Surveys	Trade Confidence Indicator	N	-7 to -2	CZSO
Surveys	Services Confidence Indicator	N	-7 to -2	CZSO
International	Brent Oil Price	N	0	eia.gov
International	EU Bond 2Y	N	0	CNB
International	EU Bond 10Y	N	0	CNB
International	Euribor 3M	N	0	ECB
International	Euribor 1Y	N	0	ECB
International	Euro Area Business Climate	N	-4 to -1	EC
International	Germany Business Climate	N	-10 to -4	IFO
International	Germany Exports	Y	40	WorldBank

Notes:Lag shows the typical delay in publication of the figure, negative numbers mean that data was released before the before the end of the reference month or quarter. CZSO denotes the Czech Statistical Office, CNB denotes the ARAD database of the CNB, ECB means the European Central Bank's Statistical Data Warehouse, EC stands for European comission and IFO denotes the IFO institue. GDP is collected at quarterly frequency while the rest of the variables is at monthly frequency.

Figure A.1: Variables

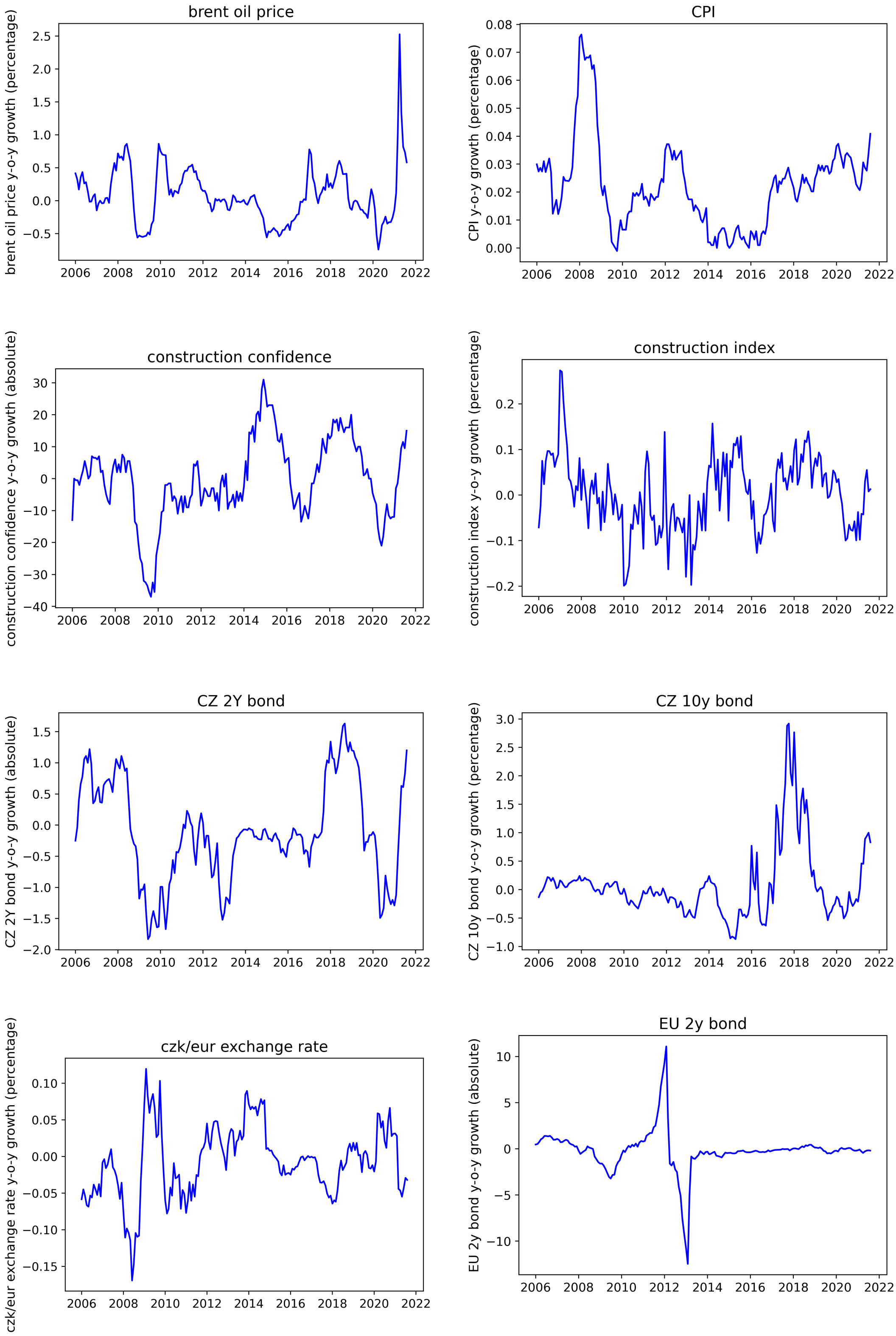


Figure A.2: Variables continuation

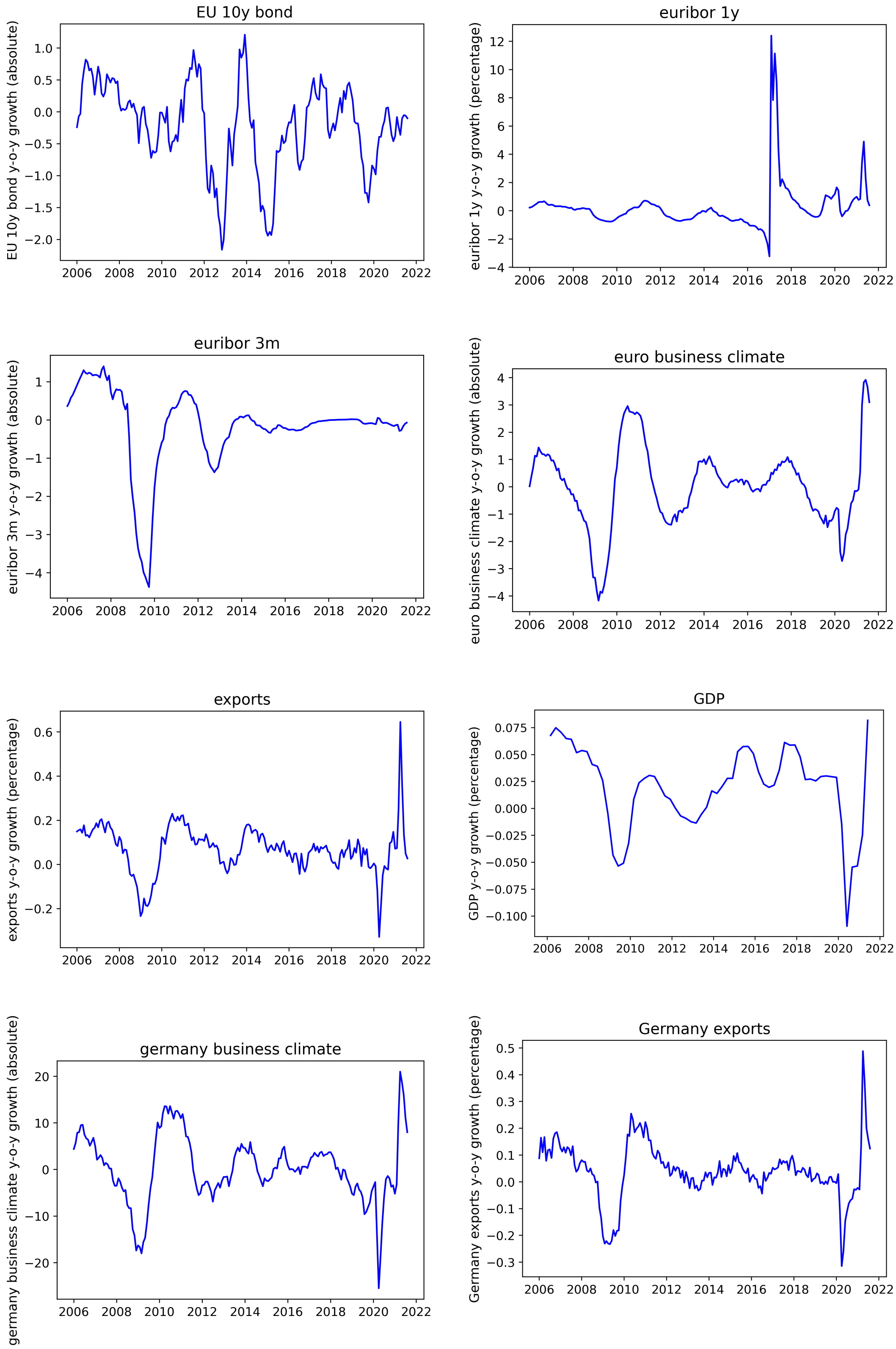


Figure A.3: Variables continuation

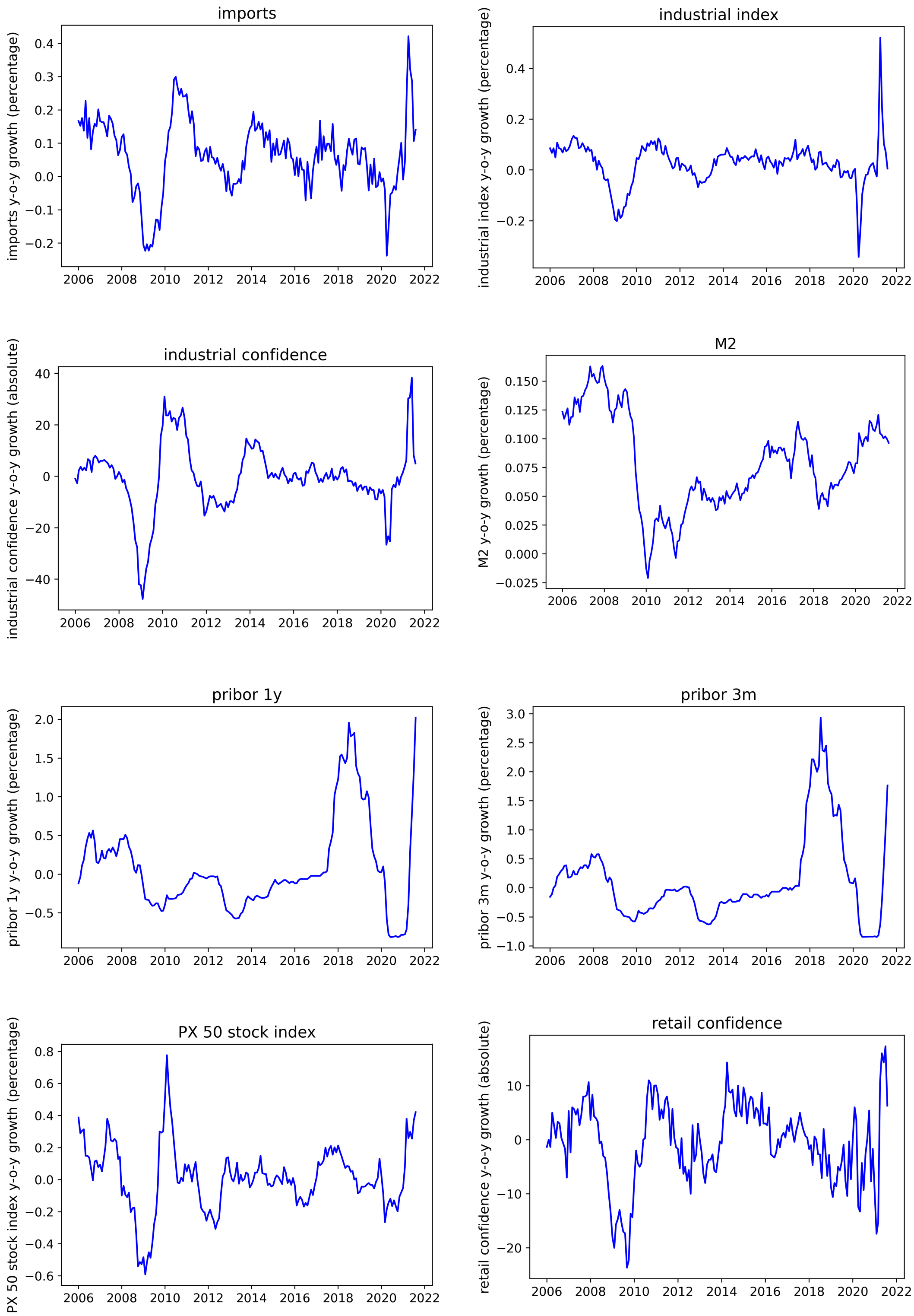
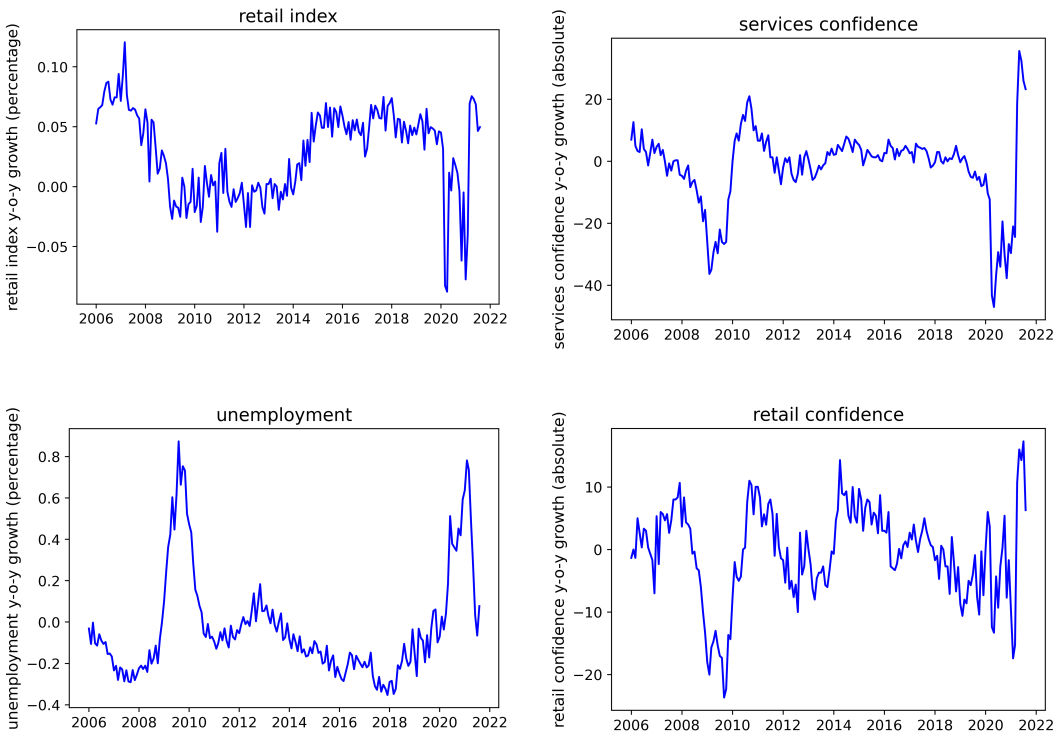


Figure A.4: Variables continuation



Appendix B

Further Results

Table B.1: The RMSE of estimated models (excluding covid period)

Horizon	DFM	Factor MI-DAS	MIDAS	XGBoost	RW	Combination	CNB
Q(-1) M3 start	1,29	1,55	1,58	1,59	1,66	1,43	
Q(-1) M3 mid	1,32	1,53	1,57	1,5	1,66	1,4	
Q(-1) M3 end	1,41	1,54	1,56	1,49	1,66	1,42	
Q(0) M1 start	1,36	1,5	1,54	1,5	1,66	1,41	
Q(0) M1 mid	1,3	1,47	1,51	1,37	1,66	1,31	
Q(0) M1 end	1,31	1,47	1,6	1,36	1,66	1,31	
Q(0) M2 start	1,22	1,45	1,58	1,18	1,66	1,23	1,3
Q(0) M2 mid	1,21	1,43	1,57	1,43	1,66	1,26	
Q(0)M2 end	1,2	1,41	1,5	1,39	1,66	1,24	
Q(0) M3 start	0,78	0,93	0,95	1,03	0,99	0,84	
Q(0) M3 mid	0,77	0,92	0,95	1,14	0,99	0,84	
Q(0) M3 end	0,78	0,94	0,95	1,13	0,99	0,84	
Q(1) M1 start	0,75	0,92	0,94	0,97	0,99	0,81	
Q(1) M1 mid	0,75	0,92	0,93	0,97	0,99	0,81	
Q(1) M1 end	0,75	0,92	0,93	0,96	0,99	0,81	
Q(1) M2 start	0,75	0,92	0,93	0,9	0,99	0,79	1,09

Note: The models used for calculation of the RMSE are the best performing specifications for each model class in out of sample nowcasts from 1st quarter of 2012 til last quarter of 2019. Q (0) is the quarter before the nowcasted quarter, Q (0) referers to the nowcasted quarter and Q(1) is the following quarter. M1,M2,M3 are the first, second and third months of the quarter respectively. Start, mid and end refer to 10th, 24th and 30th days of the relevant month respectively.

Table B.2: The RMSE of estimated models (including covid period)

Horizon	DFM	Factor MIDAS	MIDAS	SVR	RW	Combination	CNB
Q(-1) M3 start	3,09	3,07	3,44	3,16	3,76	3,03	
Q(-1) M3 mid	2,85	2,92	3,41	3,18	3,76	2,87	
Q(-1) M3 end	2,77	2,89	3,4	3,09	3,76	2,83	
Q(0) M1 start	2,78	2,85	3,39	3,05	3,76	2,87	
Q(0) M1 mid	2,07	2,41	3,22	2,81	3,76	2,42	
Q(0) M1 end	2,1	2,44	3,22	2,75	3,76	2,44	
Q(0) M2 start	1,81	2,3	3,12	2,47	3,76	2,14	1,65
Q(0) M2 mid	1,96	2,26	3,1	2,12	3,76	2,13	
Q(0)M2 end	1,96	2,23	3,06	2,08	3,76	2,1	
Q(0) M3 start	1,73	1,7	2,25	1,78	2,77	1,66	
Q(0) M3 mid	1,7	1,71	2,25	1,87	2,77	1,62	
Q(0) M3 end	1,73	1,71	2,25	1,83	2,77	1,61	
Q(1) M1 start	1,76	1,77	2,27	1,8	2,77	1,64	
Q(1) M1 mid	1,75	1,77	2,28	1,8	2,77	1,63	
Q(1) M1 end	1,75	1,77	2,24	1,81	2,77	1,63	
Q(1) M2 start	1,77	1,85	2,33	1,8	2,77	1,62	1,08

Note: The models used for calculation of the RMSE are the best performing specifications for each model class in out of sample nowcasts from 1st quarter of 2012 till the second quarter of 2021. Q (0) is the quarter before the nowcasted quarter, Q (0) referers to the nowcasted quarter and Q(1) is the following quarter. M1,M2,M3 are the first, second and third months of the quarter respectively. Start, mid and end refer to 10th, 24th and 30th days of the relevant month respectively.

The following figures plot the errors of the individual nowcasting models compared to the errors of the official CNB nowcasts. The notation *Q(0) M2 start* refers to the nowcasts made on the 10th day of the second month of the reference quarter, while *Q(1) M2 start* means the nowcasts were conducted on the 10th day of the second month of the following quarter. These two dates correspond to the release of the CNB GDP forecasts.

Figure B.1: Nowcasting error of the DFM and CNB (Q(0) M2 start)

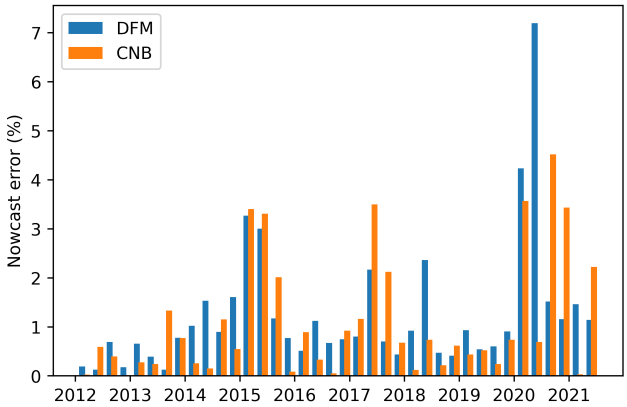


Figure B.2: Nowcasting error of MIDAS and CNB (Q(0) M2 start)

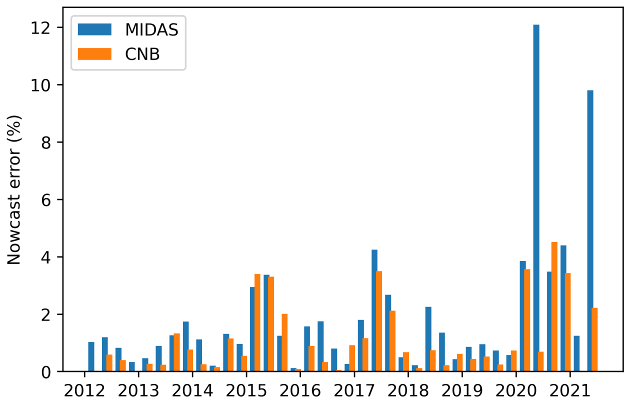


Figure B.3: Nowcasting error of Factor MIDAS and CNB (Q(0) M2 start)

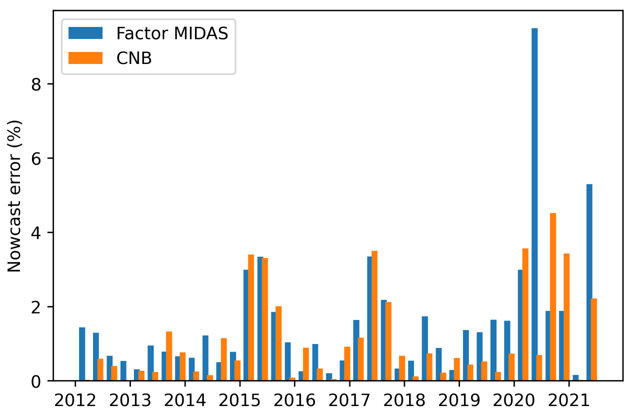


Figure B.4: Nowcasting error of the SVR and CNB (Q(0) M2 start)

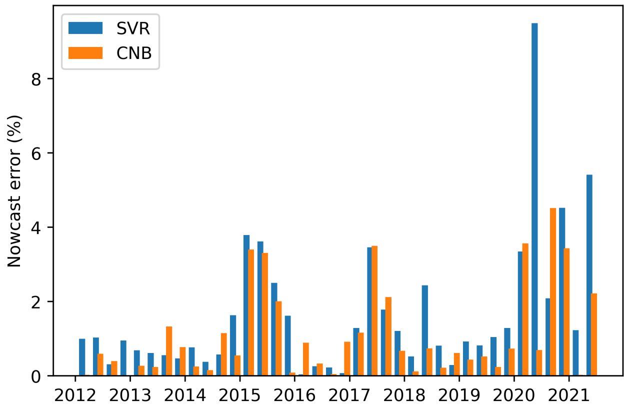


Figure B.5: Nowcasting error of XGBoost and CNB (Q(0) M2 start)

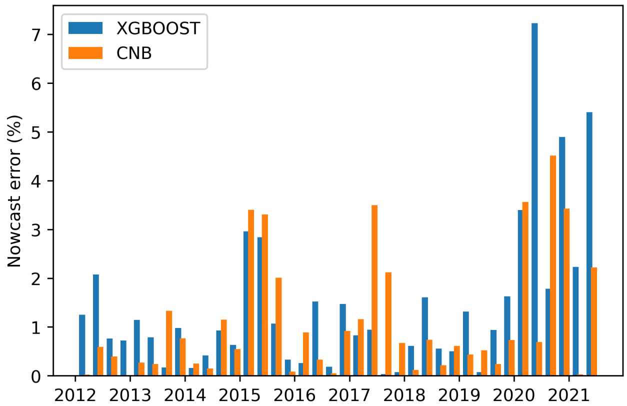


Figure B.6: Nowcasting error of the RF and CNB (Q(0) M2 start)

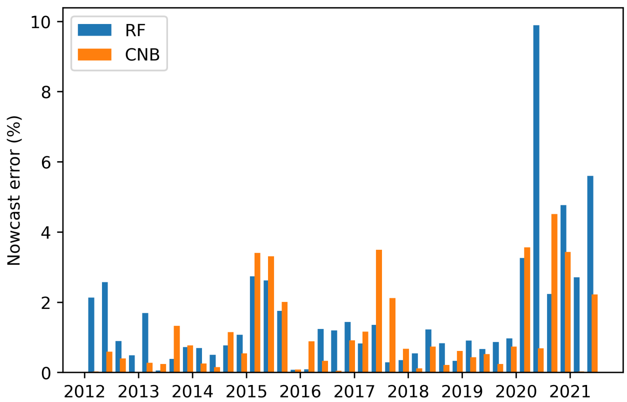


Figure B.7: Nowcasting error of the ensemle model and CNB (Q(0) M2 start)

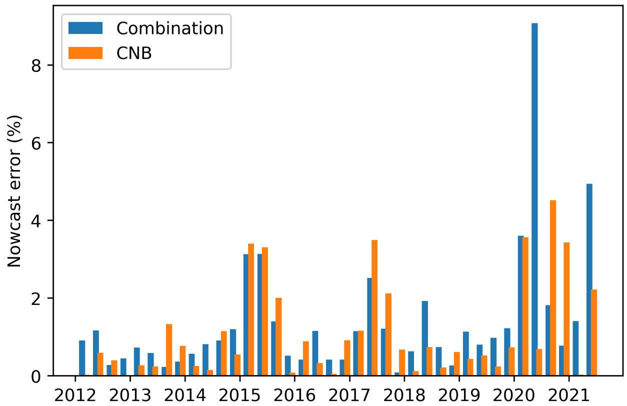


Figure B.8: Nowcasting error of RW and CNB (Q(0) M2 start)

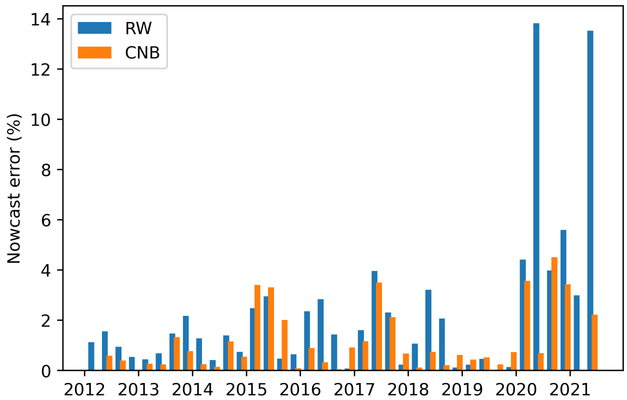


Figure B.9: Nowcasting error of the DFM and CNB (Q(1) M2 start)

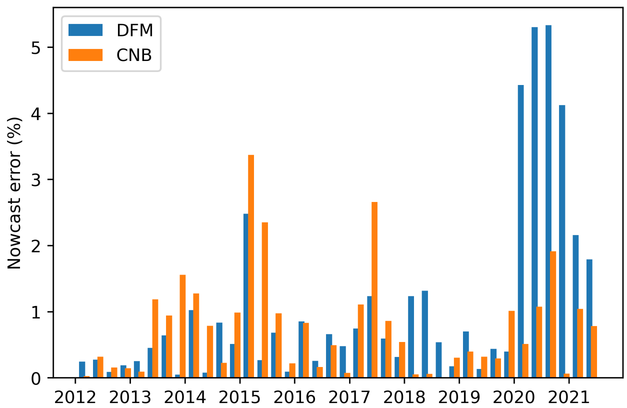


Figure B.10: Nowcasting error of MIDAS and CNB (Q(1) M2 start)

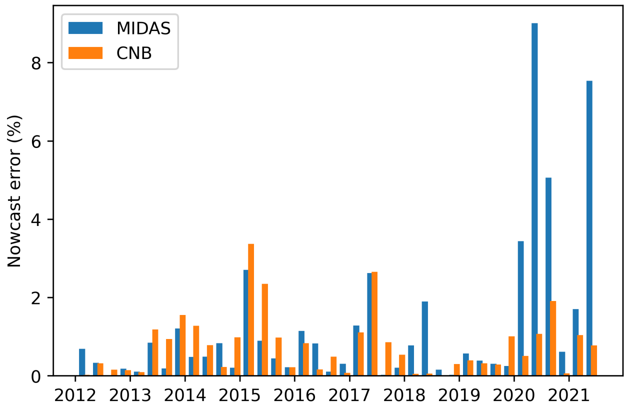


Figure B.11: Nowcasting error of Factor MIDAS and CNB (Q(1) M2 start)

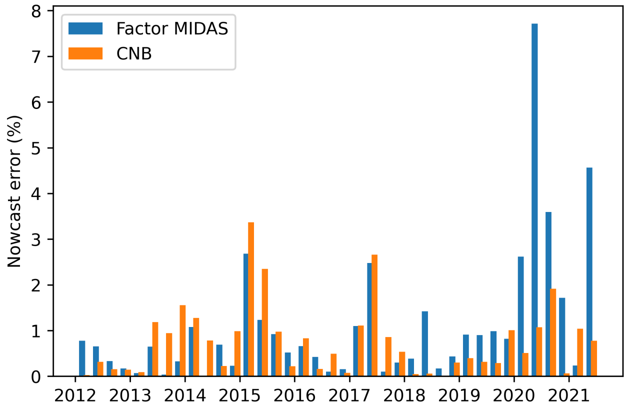


Figure B.12: Nowcasting error of the SVR and CNB (Q(1) M2 start)

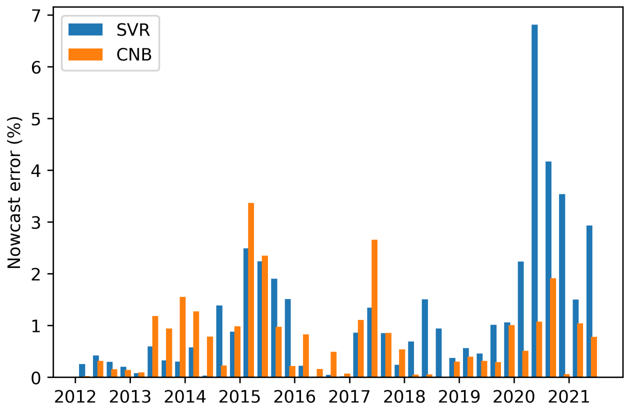


Figure B.13: Nowcasting error of XGBoost and CNB (Q(1) M2 start)

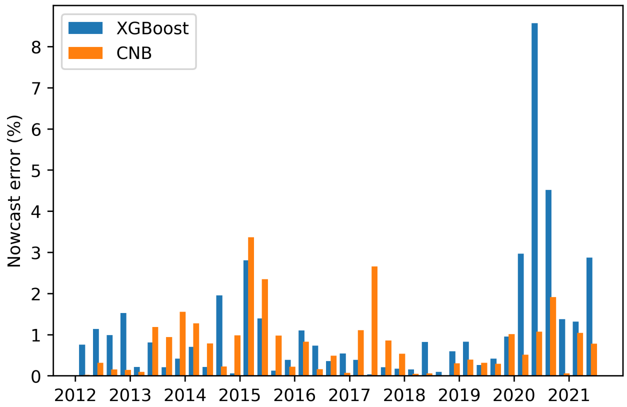


Figure B.14: Nowcasting error of the RF and CNB (Q(1) M2 start)

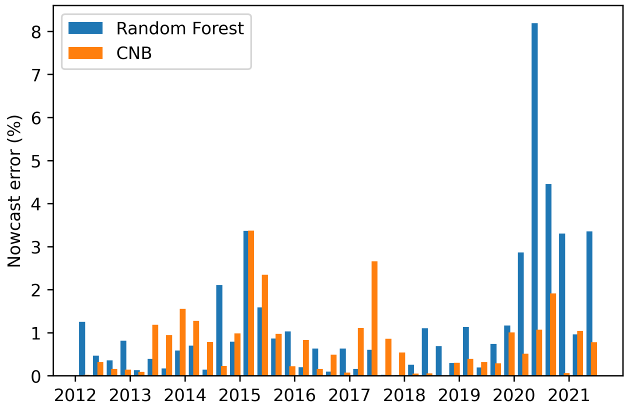


Figure B.15: Nowcasting error of the ensemle model and CNB (Q(1) M2 start)

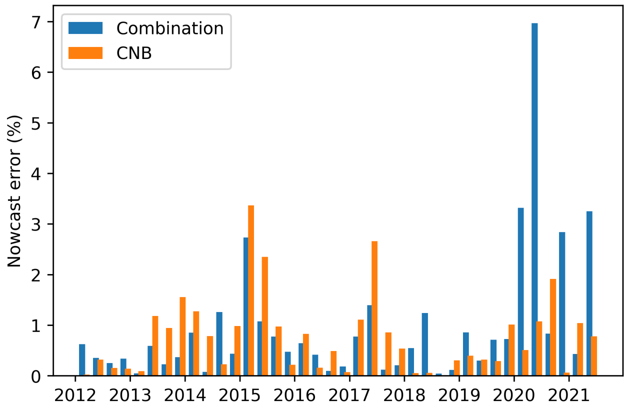


Figure B.16: Nowcasting error of RW and CNB (Q(1) M2 start)