Erasmus University Rotterdam
Erasmus School of Economics
Bachelor Thesis Double Degree Bsc2 in Econometrics and Economics

GDP Growth Nowcasting Using Mixed Frequency Gaussian Vector Autoregressive Processes and Outlier Adjusted Stochastic Volatility

Stefan van Diepen (573588)

Supervisor:	Aishameriane Schmidt
Second Assessor:	dr. Wendun Wang
Date:	1st July 2024

Abstract

Accurate GDP nowcasts are important to improve decision making. This paper investigates whether existing mixed frequency nowcasting methods can be improved by using a Gaussian process as underlying model. Furthermore, this paper researches whether using an outlier adjusted stochastic volatility (SVO) specification is useful to further improve nowcasts in this setting. We find that a Gaussian process does not lead to improved point nowcasts of GDP growth. Moreover, adding an SVO covariance matrix further deteriorates the performance of the point nowcasts. A Gaussian process does improve the full nowcasted GDP distribution when facing samples including economic turmoil. Using an SVO specification further improves the nowcasted distribution of GDP growth in crisis times. However, as both models lead to worse distributional nowcasts when excluding Covid from our sample, the Gaussian process, both with and without outlier adjusted stochastic volatility, fails to improve existing models.

1 Introduction

Every quarter, estimates of GDP growth are released all over the world. However, these figures are published with a delay. For example, the GDP growth figures of the first quarter of 2024, which ended on the 31^{st} of March, are only published far into the second quarter of 2024. For reference, this GDP growth figure was released on the 7th of June for the EU, and 30th of May for the US. Hence, looking at actual growth figures leads to a large delay in decision making based on GDP growth data. It is thus important to obtain accurate estimates of GDP growth at an earlier time.

Therefore, this paper introduces two GDP growth nowcasting models. I construct a mixed frequency Gaussian vector autoregressive process (MF-GVARP) model, both with and without an outlier adjusted stochastic volatility (SVO) covariance matrix. The nowcasting performance of these two models is compared against the performance of the mixed frequency Bayesian vector autoregressive trees model and mixed frequency vector autoregression model as used by Huber, Koop, Onorante, Pfarrhofer, and Schreiner (2023). This paper uses the models to make nowcasts of GDP growth for four European countries for two samples: one ranging until Q4-2019, and one sample including Covid-19 observations ranging until Q2-2020.

GDP nowcasting is a widely debated topic in academic literature. Schorfheide and Song (2015) construct a mixed frequency vector autoregression (MF-VAR) model, which uses data from monthly and quarterly frequencies to make nowcasts for the low-frequency (quarterly) variables. When producing nowcasts for several macro-economic variables, including GDP, they find that including monthly data from within the quarter that is nowcasted improves nowcasting performance.

Furthermore, Huber et al. (2023) take inspiration from the MF-VAR model introduced by Schorfheide and Song (2015) and find that using a mixed frequency Bayesian additive vector autoregressive trees (MF-BAVART) model improves nowcasting performance when compared to an often used MF-VAR model. This improvement is especially large when looking at the volatile COVID-19 period. In the MF-BAVART model, Huber et al. (2023) use the model developed by Chipman, George, and McCulloch (2010) in a mixed frequency framework, and argue that the improvement this model introduces can partially be attributed to trees being useful in handling

outliers in volatile periods. However, the motivation of the use of Bayesian additive regression trees (BART) over other models in the mixed frequency framework is not immediately clear. Also, Huber et al. (2023) assume a constant covariance matrix, even though the volatility of GDP growth is not constant over time. Therefore, this paper proposes to use a different model in the mixed frequency framework, and suggests to use a different, time-varying, covariance matrix, resulting in the MF-GVARP and MF-GVARP-SVO models. The following two paragraphs will outline the motivation for the specification of these two models.

First of all, Hauzenberger, Marcellino, Pfarrhofer, and Stelzer (2024) combine a mixed data sampling (MIDAS) model with a Gaussian process prior. They also specify two alternative priors: a restricted linear version of this MIDAS model and a Bayesian additive regression trees (BART) prior. The BART specification they use is similar to the BART model used by Huber et al. (2023). They find that using a Gaussian process in the MIDAS framework outperforms other models when nowcasting GDP, noteably also the BART model considered. This outperformance is present in the whole sample they consider, as well as in the subsamples excluding Covid-19 observations, and only the Covid-19 period. Even though the Gaussian process will be less suited for handling outliers, this gives rise to the idea of replacing the BART specification used in the nowcasting framework of Huber et al. (2023) by the Gaussian process model used in the MIDAS specification of Hauzenberger et al. (2024).

Secondly, Carriero, Clark, Marcellino, and Mertens (2022) find that adding an outlier adjusted stochastic volatility specification to a Bayesian vector autoregression (BVAR) model, improves the forecasting performance for many macroeconomic variables. This is mainly attributed to the better ability to handle large outliers in the data. Furthermore, the stochastic volatility in the covariance matrix will allow for time varying volatility in the model. The research by Carriero et al. (2022) handles forecasting, and different macroeconomic data than GDP growth. However, due to the relatively close relation between nowcasting and forecasting, and the wide variety of macroeconomic variables for which their model worked, the specification outlined by Carriero et al. (2022) is worth pursuing in a GDP nowcasting setting. Therefore, this paper will use the covariance matrix of Carriero et al. (2022) in combination with the Gaussian process prior to enhance the ability of the model to handle outliers.

When considering the societal relevance of accurate GDP growth nowcasts, one should consider that GDP growth figures affect many aspects of everyday life. For example, higher GDP growth typically coincides with a lot of job opportunities and increasing salaries, whereas negative GDP growth oftentimes coincides with more economic uncertainty and higher unemployment (Tumanoska, 2020). Therefore, GDP growth figures affect most working people. Moreover, business managers and entrepreneurs adjust business plans based on economic outlooks. As higher or lower GDP growth directly impacts economic activity and thus sales, GDP growth figures affect investment decisions by firms (Farooq, Ahmed, & Khan, 2021). Similarly, finance professionals are affected by the state of the economy, as they need to decide which loans to hand out or which companies to invest in. These decisions are heavily influenced by GDP growth figures. Lastly, in order to smooth the business cycle, politicians need to take policy measures to ensure that the economy does not overheat in times of economic prosperity, while also ensuring a soft landing when crises emerge. It is thus important for policy makers to have an accurate GDP

growth figure, as this will enable them to take well substantiated policy measures.

This paper examines whether using a Gaussian process prior distribution in the framework of Huber et al. (2023) leads to improved GDP nowcasts. Secondly, we add an outlier adjusted stochastic volatility covariance matrix to test if this improves the performance of a mixed frequency model with a Gaussian process prior. Both the point nowcasts, as well as the entire nowcasted density are compared by using the root mean squared error, cumulative log predictive score, and the transformed probability integral transform. I find that using a Gaussian process in the mixed frequency framework results in worse point nowcasts when compared to the MF-BAVART and MF-VAR models. However, if we include the Covid-19 observations in the data, the MF-GVARP model outperforms the MF-BAVART and MF-VAR models based on the entire distribution of the nowcasted GDP growth. Furthermore, the MF-GVARP-SVO model produces point nowcasts that perform substantially worse than the MF-GVARP model for both samples under consideration in this paper. Therefore, adding an outlier adjusted stochastic volatility covariance matrix does not improve the performance of the model.

The rest of this paper is structured as follows. Section 2 describes the data used in the research. Section 3 introduces the models used to nowcast GDP growth, as wel as the estimation methods used to estimate these models. This section also outlines several evaluation metrics that are used to compare the nowcasting performance of the models. Section 4 describes and discusses the result of the research, Section 5 discusses the economic interpretation of the results, after which Section 6 concludes the paper.

2 Data

This section discusses the data used in this paper. The chosen dataset is exactly the same as the data used by Huber et al. (2023). This dataset consists of data on 4 countries: France, Germany, Italy, and Spain. The set contains quarterly GDP growth data, as well as data on 5 monthly macroeconomic variables, spanning a period from April/Q2 2005 until June/Q2 2020. The data is log-differenced if the variable is expressed to be a growth rate. More precisely, the following data is used in the research. GDP growth is obtained from the OECD real time database. The growth rate of the economic sentiment indicator (ESI) is provided by the European Commission. Industrial production growth (IP) is obtained from the OECD real time database. The growth rate of the purchasing managers' index (PMI) is obtained from Markit. The growth rate of monthly car registrations (CAR) is obtained from the European Automobile Manufacturers Association, and the 1-year ahead bond yields (I) are obtained from Macrobond. GDP growth and bond yield data are expressed in percentage terms, whereas the other variables are expressed in decimal values. The range of values these variables take on over time differs. Therefore, to ensure the models do not over- or under-estimate the effect of these variables, all data is standardized in all models.

Table 1 shows the summary statistics for all variables under consideration for all countries. The table shows that the different variables take on differing ranges of values. For example, the total range of values of the GDP growth in Germany is 12.843, whereas this is only 0.328 for the ESI. For some differences, this is explained by the difference in scaling, (such as bond yields against PMI). However, for other variables, this difference arises due to a different underlying

economic dynamic. For example, because new cars are typically a luxury good, the registrations of these vehicles differ more with economic conditions than the industrial production. Furthermore, besides different ranges and variances, the other characteristics of the variables also differ substantially. Note that some variables have a positive mean (and median), such as bond yields, whereas others have a negative mean (but not median), such as the change in car registrations. This difference is due to economic conditions and market trends which affect the variables differently. For example, bond yields have a positive mean because bond yields are never negative in normal economic circumstances, whereas GDP growth has a negative average for most countries in the sample, as the COVID-19 observations push the mean growth downwards (note that the median values are positive, as the economy typically grows steadily over time, with periods of crisis alternating the periods of growth). The standard deviation of the variables also differs a lot, with variables with larger ranges having higher standard deviations.

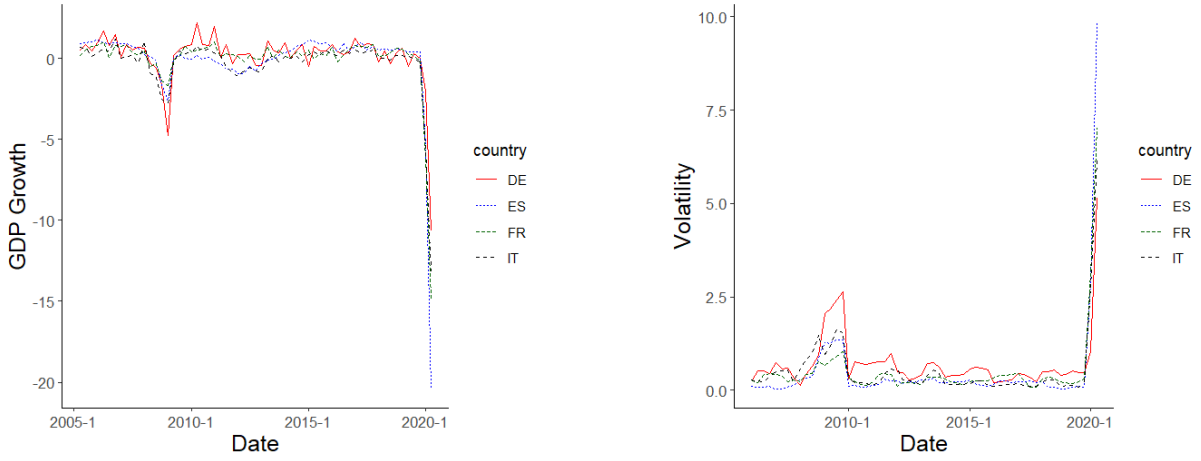
Lastly, the skewness and kurtosis that characterise the variables are very different. The bond yields have a positive skewness, but small kurtosis for all countries, indicating that the distribution does not have fat tails, but is skewed to the right. This makes sense for bond yields as they are typically not below zero. The other variables have a negative skewness in most countries, and a large positive kurtosis. This indicates that most variables under consideration have a distribution with a high peak and fat tails that is skewed to the left. This can be explained by the presence of 2 crises in our sample, the Covid-19 period, and the financial crisis of 2008. These periods cause the overall data to have a few large negative values. However, there are no large positive values, as large positive growth rates for macroeconomic data occur only rarely, and if they do occur, they typically occur in countries with exceptionally high growth for a period of time, and not just a few observations (such as post world war 2 Germany). These high growth observations would therefore not be outlying values in the same way as the crisis observations are, and are in any case not present in the sample. The high peak and fat tails can be explained by the underlying characteristic of the macroeconomic variables. Typically, the macroeconomic environment is relatively stable, leading to a distribution with a high peak. However, when instability arises, for instance during crisis times, a few large negative values occur, leading to a fatter left tail (high kurtosis and negative skewness). Following these crises, typically a few observations of positive growth occur, but these are observations are smaller than the crisis shrinkage. This leads to a fatter right tail as well (but this is not enough to balance the skewness).

Country	Variable	Mean	St. Deviation	Skewness	Kurtosis	Minimum	Maximum	Percentile_25	Median	Percentile_75
Germany										
	Industrial Production	0.000	0.025	-4.326	40.220	-0.226	0.092	-0.008	0.002	0.010
	Economic Sentiment Index	-0.001	0.027	-4.061	37.604	-0.244	0.084	-0.008	0.001	0.011
	Car Registrations	-0.002	0.090	-1.874	15.369	-0.496	0.347	-0.018	0.006	0.025
	Purchasing Mangers Index	0.000	0.039	-1.420	22.579	-0.285	0.185	-0.015	0.002	0.015
	Bond Yields	1.381	1.647	0.951	2.728	-0.356	5.393	-0.057	0.604	2.158
	GDP	0.155	1.706	-4.627	28.473	-10.638	2.205	0.074	0.482	0.827
Spain										
	Industrial Production	-0.002	0.027	-5.069	57.352	-0.258	0.141	-0.005	0.001	0.005
	Economic Sentiment Index	-0.001	0.032	-4.393	43.823	-0.304	0.104	-0.011	0.001	0.014
	Car Registrations	-0.004	0.253	-0.400	51.629	-2.057	2.065	-0.034	0.004	0.035
	Purchasing Managers Index	0.000	0.039	-1.420	22.579	-0.285	0.185	-0.015	0.002	0.015
	Bond Yields	1.381	1.647	0.951	2.728	-0.356	5.393	-0.057	0.604	2.158
	GDP	-0.148	2.827	-6.309	45.163	-20.437	1.141	-0.055	0.419	0.855
France										
	Industrial Production	-0.001	0.030	-2.717	37.240	-0.235	0.182	-0.011	0.000	0.010
	Economic Sentiment Index	-0.002	0.038	-5.224	53.354	-0.377	0.135	-0.013	0.002	0.013
	Car Registrations	0.000	0.181	1.167	49.486	-1.276	1.579	-0.027	0.002	0.029
	Purchasing Managers Index	0.000	0.039	-1.420	22.579	-0.285	0.185	-0.015	0.002	0.015
	Bond Yields	1.381	1.647	0.951	2.728	-0.356	5.393	-0.057	0.604	2.158
	GDP	-0.073	2.147	-5.861	39.262	-14.868	1.031	0.038	0.247	0.651
Italy										
	Industrial Production	-0.002	0.043	-0.329	49.853	-0.335	0.351	-0.011	0.002	0.009
	Economic Sentiment Index	0.000	0.031	-1.199	11.636	-0.191	0.122	-0.012	0.000	0.016
	Car Registrations	-0.002	0.311	3.475	70.994	-1.902	3.140	-0.028	-0.002	0.028
	Purchasing Managers Index	0.000	0.039	-1.420	22.579	-0.285	0.185	-0.015	0.002	0.015
	Bond Yields	1.381	1.647	0.951	2.728	-0.356	5.393	-0.057	0.604	2.158
	GDP	-0.309	1.949	-5.196	33.277	-13.191	1.167	-0.229	0.142	0.379

Table 1: Summary Statistics

The table shows the summary statistics for all variables under consideration for all countries in the sample. The sample ranges from April/Q2 2005 until June Q2/2020. GDP growth and bond yields are in percentages, the other variables are expressed in decimal values.

The final part of this section describes characteristics of GDP growth that are used in this paper to motivate the choice of the models. Figures 1a and 1b show the GDP growth and GDP growth volatility of the countries in the sample. The GDP growth volatility displayed is the 4-quarter rolling volatility of GDP growth. The figures show that GDP growth and the GDP growth volatility display a positive correlation among the countries. This is to be expected, as the countries are all part of the same economic and political regions, and are therefore subject to similar macroeconomic shocks. For example, because of the interconnectedness of the countries, we expect the German economy to react in a similar way to an oil price shock as the French economy. Also, the countries are expected to experience similar shocks at the same moment in time. Furthermore, Figure 1b clearly shows periods with relatively stable and low volatility, as well as periods with relatively high volatility. GDP growth volatility is therefore not stable over time, but fluctuates substantially. The periods of low and stable volatility are noticeably periods with relatively stable macroeconomic conditions, namely the 2010s and, to a lesser extent, the period before the financial crisis. The two high volatility periods are the periods surrounding the financial crisis of 2008 and the Covid-19 crisis in 2020. Comparing the graph with GDP growth volatility to the graph with GDP growth, the figures clearly show that the periods of high volatility are caused by large negative spikes in GDP growth, whereas low volatility periods coincide with small positive but stable GDP growth numbers. These high volatility periods can therefore be described by crisis periods, and are notably caused by a few crisis observations.



(a) GDP Growth

(b) GDP Growth Volatility

Figure 1: GDP Growth and 12 Month Rolling Volatility of GDP Growth for Germany (DE), Spain (ES), France (FR) and Italy (IT) Using Data From Q2-2005 Until Q2-2020.

3 Methodology

This section explains the methods that are used to nowcast GDP growth. The time of observation of a variable is denoted as t for $t \in (1, .., T)$. The latent states of the variable observed at quarterly frequency (GDP growth) are denoted as $y_{q,t}$, with q denoting the quarter, and t being the month under consideration. The variables observed at monthly frequency are denoted as $y_{w,t}$, and the entire vector of variables at time t is denoted as $y_t = (y'_{q,t}, y'_{w,t})'$. Note that the variable under consideration is denoted by m, $\forall m \in (1, .., M)$, for which p = 5 lags are added in

the models. $X_t = (y'_{t-1}, ..., y'_{t-p})'$ is a Q(=Mp)-dimensional vector containing all the variables under consideration, as well as their lags at time T. Lastly, $X = (X_1, .., X_T)'$ is the TxQ matrix containing all the data including the lagged values, and $Y = (y_1, .., y_T)'$ is the TxM matrix containing all data on the variables.

The framework used for GDP nowcasting in this paper closely follows the framework as outlined by Huber et al. (2023). First of all, monthly and quarterly GDP growth rates have the relation indicated in Equation 1, where the right hand side of the equation is scaled s.t. the scales of the monthly and quarterly growth rates are comparable (Huber et al., 2023). Therefore, this equation will be used to convert monthly estimates of GDP growth to quarterly estimates that can be compared to the actual released GDP growth figures.

$$y_{Q,t} = \frac{1}{9}y_{q,t} + \frac{2}{9}y_{q,t-1} + \frac{1}{3}y_{q,t-2} + \frac{2}{9}y_{q,t-3} + \frac{1}{9}y_{q,t-4} \tag{1}$$

In this paper, all models are vector autoregressive models. We assume that our vector of time series data, y_t, follows the process as outlined in Equation 2, where $F(X_t) = (f_1(X_t), ..., f_M(X_t))'$ is a vector of functions that map the covariates to each of the dependent variables, and Σ_t is the covariance matrix of the system.

$$y_t = F(X_t) + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \Sigma_t) \tag{2}$$

This paper proposes 3 specifications of the function $f_m(X_t)$, and two ways to define the covariance matrix Σ_t. First of all, Section 3.1 will implement Bayesian additive regression trees, to obtain the MF-BAVART model as introduced by Huber et al. (2023). Secondly, Section 3.2 will introduce a Gaussian process which takes inspiration from Hauzenberger et al. (2024), to get a mixed frequency Gaussian vector autoregressive process. This section will also introduce a second version of the covariance matrix, one with an outlier adjusted stochastic volatility specification. Finally, Section 3.3 outlines a linear model to obtain a simpler mixed frequency vector autoregression model (MF-VAR), as used more frequently in macro-economic nowcasting. The other subsections will discuss the estimation of the models, as well as evaluation metrics used to evaluate the nowcasting performance of the models.

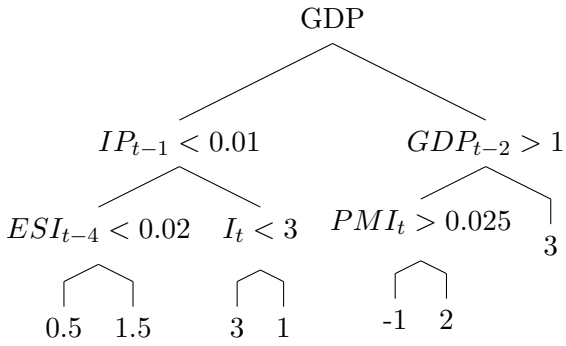
3.1 Mixed Frequency Bayesian Additive Vector Autoregressive Trees

In this section, the mixed-frequency Bayesian additive vector autoregressive trees (MF-BAVART) model is explained. This model closely follows the MF-BAVART model as introduced by Huber et al. (2023). First of all, the covariance matrix in Equation 2 is constant in this case, $\Sigma_t = \Sigma$. Furthermore, every $f_m(X_t)$ is assumed to be a BART model as specified in Equations 3 and 4.

$$f_m(X_t) \approx \sum_{s=1}^{S} g_{ms}(X_t|\tau_{ms}, \mu_{ms}), \tag{3}$$

$$g(X|\tau, \mu) = \sum_{r=1}^{b} \mu_r \mathbb{1}(X \in \mathcal{A}_r) \tag{4}$$

Here, S denotes the number of trees in the model, which is set to 250 in this paper. μ_{ms} denotes the terminal nodes of tree ms, and τ_{ms} denotes the tree structure of tree ms. $\mathcal{I}$ is the indicator function, and $\mathcal{A}_r$ indicates the r^{th} partition set of the tree. When considering a simple example of a single tree $g(X|\tau, \mu)$, this might look as follows.



In a tree model, predictions are made by classifying a dependent variable along the tree, and assigning the appropriate value of the partition set to the observation. In this example, GDP growth is classified according to the variables described in the tree. Therefore, the actual values of the variables will decide which branch of the tree is chosen at every step, and which terminal node is chosen as prediction of GDP growth. In this example, the tree has the following terminal nodes $[-1, 0.5, 1, 1.5, 2, 3]$. In an actual BART model, many such trees are obtained, which are added to acquire the desired model. Using a BART model has several advantages when modeling data that has nonlinear properties, such as samples that include crisis data. The BART model enables the system to handle outliers, as the outliers can be absorbed by either different branches of a tree, or different trees altogether (Huber et al., 2023). However, because of the great flexibility the BART model offers, issues regarding overfitting on the training data might arise. This can arise as creating separate branches or trees for every observation might create very low training errors, but leads to poor predicting performance. Therefore, regularization priors need to be imposed on the structure of the trees, in order to reduce overfitting, while at the same time attempting to keep the flexibility offered by the model.

Equations 5 and 6 are priors imposed on the tree structure to simplify the rest of the prior specification. These priors impose the tree structure (τ_{ms}, μ_{ms}), and the terminal node parameters to be independent of each other.

$$p((\tau_{m1}, \mu_{m1}), .., (\tau_{ms}, \mu_{ms})) = \prod_s p(\mu_{ms}|\tau_{ms})p(\tau_{ms}), \tag{5}$$

$$p(\mu_{ms}|\tau_{ms}) = \prod_i \mu_{i,ms}|\tau_{ms}) \tag{6}$$

Because Equation 5 and 6 impose independence between the different tree structure and the terminal nodes, all that is left is specifying separate priors for the the tree generating process, the terminal nodes, and the covariance matrix. Therefore, what follows in this section is a description of these priors. Firstly, similarly to Huber et al. (2023), formula 7 specifies the prior for the tree generating process, namely the probability that a node is not a terminal node. This prior therefore limits the degree of overfitting and keeps the trees simple. The prior is specified as follows:

$$\alpha(1+n)^{-\beta}, \quad \alpha \in (0,1), \quad \beta > 0, \tag{7}$$

where α and β are scalar parameters that specify the penalty used to control the complexity of the tree structures. Note that lower values of α, and higher values for β lead to less complex trees. Similarly to Huber et al. (2023), this paper chooses the values $\alpha = 0.95$, and $\beta = 2$. These values are chosen as they are shown to perform well in a wide variety of estimation problems. Furthermore, the threshold, that is used when determining which branch along the tree is taken when making predictions should be specified. This threshold is chosen to take a uniform distribution over the possible values of the independent variable under consideration. Similarly, the variable chosen in the splitting rule is chosen randomly from the columns of X.

The prior for the terminal nodes is chosen similarly to Huber et al. (2023), and is specified as follows

$$\mu_{i,ms}|\tau_{ms} \sim \mathcal{N}(0, \sigma^2_{\mu m}), \tag{8}$$

$$\sigma_{\mu m} = \frac{1}{2\gamma\sqrt{S}}, \tag{9}$$

with $\mu_{i,ms}$ indicating the i^{th} terminal node of tree ms, γ being a positive constant set equal to 2, and S the number of trees. This prior specification leads to a wider prior variance when the range of values of the variable under consideration is larger. This should help the model return wider predictive distributions when facing outliers, as outliers will cause a larger prior variance, and thus a less restrictive prior. Furthermore, an increase in the number of trees or the value for γ decreases the prior variance, and thus leads to a more restrictive prior. It is desirable for an increasing number of trees to result in a more restrictive prior, as increasing the number of trees increases the flexibility of the model. To limit the degree of overfitting, the prior variance should therefore be more restrictive in these instances. Note that the value of γ implies that the range the dependent variable takes on is a 95% confidence interval for the conditional mean of the prior distribution (Huber et al., 2023).

Lastly, the prior that handles the covariance matrix will be discussed. First of all, the variance-covariance matrix Σ is decomposed to enable the specification of separate priors for the variances and covariances. This decomposition is defined as $\Sigma = QHQ'$. Q is a lower triangular matrix of covariances and H is a diagonal matrix of variances, $\text{diag}(\sigma^2_1, ..\sigma^2_M)$. The variances are sampled from an inverse Chi-square distribution, to restrict the flexibility of the model. The Chi-square distribution used is the same as specified by Huber et al. (2023). Furthermore, a horsehoe prior is imposed on the matrix Q. Let q_{mz} be the (m, z) element of matrix Q, with z<m. A horseshoe prior is imposed on every element q_{mz}, thus imposing the horseshoe prior on every element below the diagonal.

$$q_{mz}|\tau_{mz}, \lambda \sim \mathcal{N}(0, \tau^2_{mz}\lambda^2), \quad \tau_{mz}, \lambda \sim C^+(0,1) \tag{10}$$

Equation 10 specifies the horseshoe prior, as introduced in Carvalho, Polson, and Scott (2010),

and $C^+(0,1)$ is used to denote a half-cauchy distribution. λ is a scalar that serves as a common factor to regularize the covariances, whereas $\tau_{m,z}$ also regularizes the covariances, but can differ across elements.

The horseshoe prior is a shrinkage approach, which functions as a way to reduce the dimensionality of a large estimation problem. This prior has two main characteristics. The first important characteristic is the ability to not shrink strong signals too heavily, which is made possible due to the half-cauchy distribution. At the same time, the horseshoe prior heavily shrinks the coefficients that are close to zero. These two properties enable the prior to reduce the dimensionality of large estimation problems. Note that another shrinkage method, such as a discrete mixture method or other shrinkage priors, could have been chosen. However, due to the desirable properties of the horseshoe prior and the possible issues that can arise with discrete mixture methods, which include the size of the set of solutions, the horseshoe prior is chosen in this paper (Carvalho et al., 2010).

3.2 Mixed Frequency Gaussian Vector Autoregressive Process

Huber et al. (2023) use a BART model as underlying model for the mixed frequency nowcasting process. The motivation of this BART model mostly relates to its ability to handle outliers, and therefore its suitability for data samples including crisis observations. However, Hauzenberger et al. (2024) introduce a framework for MIDAS models that incorporates a prior on the process. Specifically, BART and Gaussian process (GP) priors are tested against a linear model with a horseshoe prior imposed on the coefficients of the model. In their research, the GP prior outperforms the BART prior in all samples, including the sample of Covid-19 data, with a more pronounced outperformance in the sample preceding COVID-19.

This result inspires this paper to use a Gaussian process instead of a BART model in the mixed frequency framework. Therefore, this section outlines a Gaussian process used to create a mixed frequency Gaussian vector autoregressive process (MF-GVARP) model, taking strong inspiration from Huber et al. (2023), as well as Hauzenberger et al. (2024). Secondly, this section outlines an outlier adjusted stochastic volatility (SVO) covariance matrix that is used in the MF-GVARP model to create the MF-GVARP-SVO model.

The MFGVARP model assumes every model $f_m(X_t)$ in Equation 2 to be a Gaussian process. A Gaussian process is a non-parametric regression method that assumes the target variable to be represented by a draw from a multivariate normal (Gaussian) distribution. The entire process can be specified by the mean and covariance function of the multivariate normal distribution. Therefore, every $f_m(X_t)$ can be defined as in Equation 11 (Ebden, 2015).

$$f_m(X_t) \sim \mathcal{GP}(0, K(X_t, X_t'), \tag{11}$$

This paper chooses the mean function to be zero, and the covariance function K to be a radial basis kernel function. Using the full data matrices X and Y (as is done in Equation 4), Equation 11 takes the form of Equation 12.

$$f_m \sim \mathcal{N}(0, K(X, X')) \tag{12}$$

The kernel function determines the covariance of the observations, and thus imposes the similarity between the data at time t_1 and t_2, where similarly to Hauzenberger et al. (2024), the (t_1, t_2) element of $k(X, X')$ is given by $k(x_{t1}, x_{t2})$. This paper uses a radial basis function, which is given by $k(x_{t1}, x_{t2}) = exp(-\phi||x_{t1} - x_{t2}||^2)$. Caputo, Sim, Furesjö, and Smola (2002) find that setting ϕ equal to any number between the 10% and 90% quantiles of $||X_t - X_t'||$ results in optimal predictive performance of the model. Therefore, ϕ is chosen to be the mean value of these 10% and 90% quantiles (Karatzoglou, Smola, Hornik, & Zeileis, 2004). The covariance matrix used in the MF-GVARP model is the same as used for the MF-BAVART model, and is described in Section 3.1.

After specifying the MF-GVARP model, there is a final point that stays unaddressed. The main argument Huber et al. (2023) suggest in favour of their BART specification, is its ability to handle outliers. This stays unaddressed in the current GP specification, as argued by Li, Li, and Shao (2021) among others. Furthermore, as we have seen in section 2, the volatility of GDP growth is not actually constant over time. Therefore, we introduce a covariance matrix that is time-varying and improves the model's ability to handle outliers. Carriero et al. (2022) introduce a covariance matrix with a stochastic volatility specification and an outlier correction, which they find to insulate forecasts from outlying values. This specification also ensures that the covariance matrix is time-varying. The outlier adjusted stochastic volatility (SVO) covariance matrix will be used in this paper as a covariance matrix for the MF-GVARP model, creating the MF-GVARP-SVO model. From the specifications considered in the paper by Carriero et al. (2022), the SVO specification has been chosen, as this is argued to be especially adequate for data containing rare events that double volatility, which is present in the GDP data (see Figures 1a and 1b). Therefore, the covariance matrix of Equation 2 will be specified as: $\Sigma_t = QB_t\Omega_t B_t' Q'$. Here, Q determines the covariance, similarly to Q in section 3.1, with Q being a lower triangular matrix, with all diagonal elements equal to 1, and the off-diagonal nonzero elements of Q following a horseshoe prior as specified in Equation 10. Contrary to Section 3.1, Ω_t is a time-varying diagonal matrix, with every diagonal element following a stochastic volatility specification as specified by Huber et al. (2023) and Kastner (2016). Therefore, the m^{th} diagonal element corresponding to the m^{th} variable follows Equations 13 - 18.

$$\sigma^2_{m,t}\text{-sv} = \sigma^2_{m,t} + exp(\eta_t), \tag{13}$$

$$\eta_t = \mu + \phi(\eta_{t-1} - \mu) + \sigma^2_{\eta t}, \tag{14}$$

$$\eta_0 = \mathcal{N}(\mu, \frac{\sigma^2_\eta}{1 - \phi^2}), \tag{15}$$

$$\mu \sim \mathcal{N}(0, 10), \tag{16}$$

$$\phi \sim \beta(5, 1.5), \tag{17}$$

$$\sigma^2_\eta \sim \gamma(0.5, 10) \tag{18}$$

In this case, the variances in Ω_t are time-varying. This time-varying property of the covariance matrix is motivated in section 2, where the volatility of GDP growth is found to fluctuate over time. The specifics of the stochastic volatility process are obtained from Kastner (2016). Here, the volatility is assumed to consist of two components, a component that is similar to the variance component of the MF-BAVART model in section 3.1, and an unobserved time-varying variance component. The logarithm of this unobserved component is assumed to follow a first-order autoregressive process over time. Priors are set on each of the parameters of the autoregressive process as specified in Equations 15 - 18.

Lastly, B_t is a diagonal matrix, which is introduced to improve the model performance in the presence of outliers. The element of B corresponding to the m^{th} variable at time t is defined as

$$b_{m,t} = \begin{cases} 1, & \text{with probability } 1 - p_m \\ U(2, 20), & \text{with probability } p_m \end{cases} \tag{19}$$

with p_m being sampled from a $\beta(1, 1)$ prior. This matrix increases the variance of the m^{th} variable with probability p_m. This specification should help insulate the Gaussian process from placing too much weight on extreme observations, and thus increase nowcasting performance, especially in samples where outlying observations distort model performance. The $\beta(1, 1)$ prior is chosen as a generalisation of the β prior used by Carriero et al. (2022). However, as this $\beta(1, 1)$ distribution is essentially a uniform distribution on the range [0,1], the prior is not that restrictive. Therefore, the model could give nowcasted distributions that are too wide with this specification, in which case a different prior could be more suitable. This paper leaves this topic up to further research.

3.3 Mixed Frequency Vector Autoregression

As a third model, this section introduces a Mixed Frequency Vector Autoregression (MV-VAR) model, similar to the model used by Huber et al. (2023), that draws inspiration from Schorfheide and Song (2015). This model is created by imposing every $F_m(X_t)$ in Equation 4 to be a linear model, such that

$$F_m(X_t) = EX_T. \tag{20}$$

The horseshoe prior from Equation 10 is imposed on $vec(E)$ as a regularization. The priors imposed on the covariance matrix Σ are the same as used in 3.1. Therefore, the only difference between the MF-VAR and MF-BAVART model is the specification of $f_m(X_t)$, which is linear in the MF-VAR case, and nonparametric in the MF-BAVART model.

3.4 Estimation

All three models outlined in sections 3.1, 3.2 and 3.3 will be estimated using the framework described in Huber et al. (2023). The estimation procedure consists of several steps, that are very similar for the models under consideration in this paper. The main difference will be in the steps concerning the underlying model (BART, GP or a linear model), and the covari-

ance parameters (with or without outlier adjusted stochastic volatility specification). Firstly we describe the overall estimation framework, as outlined by Huber et al. (2023), as well as the posterior used to draw from the horseshoe prior. The estimation is done using a Markov Chain Monte Carlo (MCMC) algorithm, specifically a Metropolis within gibbs sampler. The overall estimation framework consists of an MCMC algorithm that samples from the conditional posterior distributions of the models and updates all variables accordingly in each iteration of the process. In every iteration, the following steps are executed.

First of all the model (BART, GP or a linear model) is simulated, as explained in Sections 3.4.2, 3.4.3 and 3.3. As a second step, the covariance matrix is constructed. Next, the latent states obtained from the estimated models are simulated using forward filtering backwards sampling (FFBS). If we want to make predictions, we finally draw from the predictive distribution. At every iteration, all the relevant parameters are updated as described in the previous subsections. The estimation of the models and the method used to obtain the covariance matrices are explained in the following sections. Section 3.4.1 describes the estimation of the horseshoe prior, as well as the latent states and the predictive distribution. Section 3.4.2 describes the estimation of the Gaussian process defined in Section 3.2, and Section 3.4.3 describes the estimation of the MF-BAVART model as described in Section 3.1. Note that in this section, we denote v_t the variable under consideration at time t, with $v = (v_1, ..v_T)$.

3.4.1 Latent States and Horse Shoe Prior

This section outlines common steps of the estimation processes, namely the estimation of the horseshoe prior, the latent state drawing, and the generation of forecasts. Draws for the horse shoe prior on the covariance matrix and the E matrix in Equation 20 are obtained as follows. Let Q denote the matrix on which the horse shoe prior is imposed. q_m is sampled from a multivariate Gaussian posterior, that is specified as in Huber et al. (2023).

$$q_m|\Xi \sim \mathcal{N}(u_m, \Omega_m), \tag{21}$$

$$\Omega_m = (Z_m' Z_m + V_m^{-1})^{-1}, \tag{22}$$

$$u_m = \Omega_m Z_m' \tilde{Y_m}. \tag{23}$$

Here, $\tilde{Y_m} = Y_m - f_m(X)$, and the auxiliary parameters are obtained as follows.

$$\tau_{mi}^2|\Xi \sim \mathcal{G}^{-1}(1, \frac{1}{w_{mi}} + \frac{q_{mi}^2}{2\lambda^2}, \forall i \in (1, .., j-1), \tag{24}$$

$$\lambda^2|\Xi \sim \mathcal{G}^{-1}(\frac{M(M-1)+2}{4}, \frac{1}{\zeta} + \frac{1}{2}\sum_i \sum_m (\frac{q_{mi}^2}{\tau_{mi}^2})), \tag{25}$$

$$w_{mi}|\Xi \sim \mathcal{G}^{-1}(1, 1 + \tau_{mi}^{-2}), \tag{26}$$

$$\zeta|\Xi \sim \mathcal{G}^{-1}(1, 1 + \frac{1}{\lambda^2}). \tag{27}$$

Because of the highly nonlinear nature of the nonparametric MF-BAVART and MF-GVARP models, using simple FFBS methods is not appropriate to draw the latent states of quarterly GDP growth. Therefore, similar to Huber et al. (2023), a linear approximation is used. To

produce this linear approximation, the matrix X is projected on F, $\tilde{O} = X^{+}F$, where + is used to denote the Moore-Penrose inverse. $\tilde{O}$ is used to create a linear approximation of the model, using

$$y_t = \tilde{O}' X_t + \epsilon_t \tag{28}$$

This linear approximation of y_t has Gaussian shocks. Therefore, standard FFBS algorithms can be used to approximate the latent states of GDP growth.

Lastly, forecasts are made similarly to Huber et al. (2023), by directly sampling from the predictive distribution

$$y_{t+1}|\Xi, Y = N(f(X_{t+1}), \Sigma). \tag{29}$$

Draws from Equation 29 are draws of the distribution of latent monthly GDP growth percentages. These values are then transformed into quarterly GDP growth nowcasts using Equation 1. The quarterly predictions are stored in every iteration of the model. These quarterly predictions are then used as a distribution for the nowcast of GDP for the quarter under consideration.

3.4.2 Gaussian Process Estimation

A Gaussian process specifies the data to follow a multivariate normal distribution with a specified mean and covariance function. Therefore, a closed form solution for the posterior distribution can be derived. The model is subsequently fitted with maximum likelihood estimation to maximize the probability of the posterior likelihood function. [1] If we assume that the target variable is observed with some noise, then the observed value is denoted as $s(X_t)$. If in addition we assume this noise to be Gaussian with mean zero, the posterior distribution is specified as follows.

$$p(Y|s) = \prod_t (p(Y_t - s(X_t)) \frac{1}{\sqrt{(2\pi)^T det(K)}} exp(\frac{1}{2} s' K^{-1} s) \tag{30}$$

Here, K is the kernel matrix as defined in section 3.2. Substituting $s = Kl$, with l being a linear combination of the matrix K, results in the likelihood of Equation 31.

$$ln(p(v|Y) = -\frac{1}{2\sigma^2}||Y - Kv||^2 - \frac{1}{2} v' K v + c \tag{31}$$

Maximizing this likelihood w.r.t l results in the linear combination $l = (K + \sigma^2 \mathcal{I})^{-1} v$, which can be used to predict and sample from the model, via $v_t = K(X_t, X_t')l$. Note that every iteration of the implemented MCMC algorithm samples from the estimated Gaussian process, as explained in Section 3.4. (Karatzoglou et al., 2004).

The final part of this section will explain the estimation of the outlier adjusted stochastic volatility covariance matrix of the MF-GVARP-SVO model. In order to specify this estimation process, it is important to note that two parts change relative to the specification used in the MF-BAVART model by Huber et al. (2023). Namely, the volatility of the variables, as defined in H, becomes time-varying using a stochastic volatility specification. Furthermore, a

[1] The Gaussian process will be fitted with an the kernlab R package

diagonal matrix B_t is added to improve the models ability to work with outliers in the data. The B_t matrix is obtained relatively simply, by sampling from the distributions as specified in Section 3.2. The stochastic volatily estimation is more involved, because a closed form posterior distribution can not be derived. Therefore, a Metropolis-Hastings algorithm is used to simulate the posterior.[2] The idea behind this technique is to draw from the prior distributions of μ, ϕ, and $\sigma_{\eta t}^2$, as defined in Equation 13, to simulate the posterior distribution. The prior distribution, as well as the distribution of the dependent and independent variables can be combined to obtain the posterior distribution of the latent log-variances η_t. This combination is made with the idea of Bayesian statistics in mind, namely that the posterior distribution of the parameter vector ξ can be denoted as $p(\xi|x,v) = \frac{p(x,y|\xi)}{p(x,y)} \propto p(x,v|\xi)p(\xi)$. The distribution $p(x,y)$ can not be easily calculated. However, the algorithm used ensures that this probability cancels out in the estimation process, and thus only the distribution $p(x,v|\xi)p(\xi)$ specifies the posterior distribution. The distribution of the latent states of volatility is estimated in every iteration and used to estimate these latent states, and obtain estimates of the volatilities in H.

3.4.3 MF-BAVART Estimation

A Bayesian Additive Regression Tree (BART) is estimated similarly to Huber et al. (2023), using the framework outlined by Chipman et al. (2010). The estimation of the BART uses two steps, the estimation of the trees, and the estimation of the terminal nodes. Firstly, using Equation 32, the tree structures can be sampled independently from the terminal nodes.[3]

$$p(\tau_{mz}|R_{mz}, q_m, Z_m, \sigma_m) \propto p(\tau_{mz}) \int p(R_{mz}|\mu_{mz}, \tau_{mz}, q_m, Z_m, \sigma_m) p(\mu_{mz}|\tau_{mz}, q_m, \sigma_m) d\mu_{mz} \quad (32)$$

Here R_{mz} is is a vector of residual values that is dependent on previously fitted trees, $R_{mz} = Y_m - \sum_{w \neq d} g_{mw}(X|\tau_{mw}, \mu_{mw}) - Z_m q_m$. Similarly to Huber et al. (2023), $Z_m = (Z_{1m}, .., Z_{mT})$ is a vector containing the shocks of the previously estimated trees. σ_m and q_m are the prior values drawn from the horseshoe prior and variance matrix as specified in Section 3.4.1.

The tree structures are estimated by drawing from Equation 32, which is done with a Metropolis-Hastings algorithm. In this algorithm, a candidate tree is drawn from the posterior distribution in Equation 32, and is accepted with a certain probability. After estimating the tree structures using Equation 32, the terminal nodes are drawn from a multivariate normal distribution, derived from Equations 8 and 9.

3.5 Evaluation Metrics

To compare the performance of the different models, several performance metrics are used. This paper uses the root mean squared error to evaluate point nowcasts, and the cumulative log predictive score and probability integral transform to evaluate the entire predictive density of the nowcasts provided by the models. The following subsection highlights the performance metrics used. In this section, $y_{\hat{Q},c}$ denotes the GDP growth nowcast of quarter Q made at time

[2]Note that the R package stochvol is used to implement the stochastic volatility specification, and obtain the latent states of the variance process.

[3]The dbarts package in R is used to estimate the BART model.

c, $\forall c \in (1, .., C)$, where C is the total amount of nowcasts made. Similarly, $y_{Q,c}$ is the actual value of GDP growth corresponding to the nowcast.

3.5.1 Root Mean Squared Error

The root mean squared error (RMSE) will be used to compare the performance of the point nowcasts of the different models. The RMSE is calculated as follows:

$$RMSE = \sqrt{\sum_{c=1}^{C}(y_{Q,c} - \hat{y}_{Q,c})^2}, \tag{33}$$

where C is the amount of nowcasts in the test sample. The RMSE is taken instead of a mean squared error to obtain a metric that has the same scale as the nowcasts, which eases interpretability. Note that lower values of the RMSE are preferred.

3.5.2 Cumulative Log Predictive Score

Cumulative log predictive scores (LPS) are used to compare the entire nowcasted distribution of GDP growth of different models. The LPS are computed as

$$LPS = \sum_{c=1}^{C} log(\phi(x_c)), \tag{34}$$

where ϕ denotes a normal distribution with a mean and standard deviation equal to the mean and standard deviation of the vector of nowcasted GDP growths at time c. Here, the vector of nowcasted GDP growth is meant to be the list of nowcasts at time c, obtained by storing the nowcasts made at every iteration of the MCMC sampler, as described in section 3.4. Larger values for the LPS are preferred, as smaller values indicate worse nowcasting performance.

3.5.3 Transformed Probability Integral Transform

To compare the absolute nowcasting performance of the models, transformed probability integral transforms (PIT) will be used. Similarly to Huber et al. (2023) and Clark (2011), these values are obtained as follows. First, the probability integral transform ($PIT_{standard}$) is obtained using Equation 35.

$$PIT_{standard} = \frac{\#(\text{nowcasts } < \text{actual GDP Growth})}{\#\text{nowcasts}} \tag{35}$$

Here, the nowcasts are the list of nowcasts obtained from storing the nowcasts made in every iteration of the MCMC sampler. Therefore, Equation 35 uses an approximation of the empirical CDF of the nowcast made at time t, to approximate the value of the empirical CDF evaluated at the actual (realised) value of GDP growth in the corresponding quarter.

To get a metric that can be used to evaluate the absolute performance of the models, the standard PITs are transformed using Equation 36, where Φ^{-1} is the inverse CDF of a standard normal distribution.

$$PIT = \Phi^{-1}(PIT) \tag{36}$$

These PITs should follow a standard normal distribution (Huber et al., 2023). Therefore, the mean (μ), variance (σ^2) and a first order autoregressive coefficient ($AR(1)$) (as calculated by an AR(1) process) will be presented. Perfectly calibrated models will have PITs with a variance of 1, and a mean and $AR(1)$ coefficient of 0.

4 Results

This section discusses the results of the models introduced in Section 3. All models are estimated with 1000 draws of the MCMC algorithm, using the first 500 draws as burn-in. The models are used to make GDP nowcasts from Q1-2011 until Q2-2020, using an expanding window approach. This indicates that data is added to the training set as it is made available as time passes. The set of training data therefore grows in every time step of the estimation process. Furthermore, data is assumed to be known directly after the end of the corresponding month or quarter, resulting in monthly variables being added to the training set at the start of the next month, and GDP growth of the previous quarter at the start of next quarter. Therefore, models using data until April 2011 will have GDP growth of Q1 2011 included in the data, and observations on all monthly variables up to and including April 2011.

The nowcasts are made as follows. To make a nowcast of the first quarter of GDP growth in January 2011, the latent state of GDP growth in January is estimated, and the latent states of GDP growth for February and March are forecasted, after which the GDP nowcast is made using Equation 1. This is done similarly in February, where only the latent state of GDP growth in March is forecasted using data until February 2011, and in March, when the latent states of GDP growth are estimated, but no forecasts are made. Therefore, 3 estimates of GDP growth are obtained in each quarter. One made in every month of the quarter, denoted as M/Q 1, M/Q 2 or M/Q 3. All models give density estimates of the quarterly GDP growth. If point nowcasts are discussed, they are created by taking the median of the density estimate.

4.1 Model Performance

Table 2 shows the RMSE and LPS for the different models for every country. These metrics have been calculated three times for every country, once for every month of the quarter. This enables the evaluation of the models in every month of the quarter separately, namely for the first, second, and third month respectively. Such a comparison is useful as it might be worth using different models depending on the month of the quarter if the distinction in model performance is substantial. The table shows these metrics for two periods. The left side of the table has a nowcasting period running from Q1-2011 until Q4-2019. The right side of the table also includes the Covid periods of the sample, the first and second quarter of 2020. The separation of the data in two samples enables the comparison of the nowcasting performance of the models in periods with and without outlying crisis observations.

4.1.1 Pre-Covid

Table 2 clearly shows that RSMEs decrease for all models when we move further into the quarter. This is as expected as new information is released, which is information of the same quarter as the GDP growth nowcast made. Therefore, this additional data should hold some information about GDP growth in the quarter. For example, if the industrial production growth is very high in the first month of the quarter, this would indicate high economic activity in the first quarter. The GDP growth would therefore increase in this quarter, which would be reflected in the GDP growth nowcast.

Furthermore, when comparing the models, the table shows mixed results on the performance of MF-BAVART against MF-VAR in the sample until Q4-2019. For Germany and Spain, the RMSE of the MF-BAVART model is consistently lower than the MF-VAR equivalent. However, for France and Italy, this conclusion is switched, with the MF-VAR outperforming the MF-BAVART model. When including the the MF-GVARP model in the comparison, Table 2 shows that the RMSEs of this model underperform the MF-BAVART model in all instances. For the MF-VAR model this is less clear, with underperformance in most instances, but some periods with lower MSPEs as well. Therefore, based on the RMSEs in the sample until 2019, the Gaussian process does not seem to improve point nowcasts over the usage of BART. This underperformance can be explained by the flexibility of Gaussian processes, which might lead the model to fit too closely to the training data, which results in poorer out of sample nowcasting performance.

When looking at the MF-GVARP-SVO model, Table 2 shows that the model has higher RMSEs than the standard MF-GVARP model without a SVO component. Remember that the covariance matrix in this model is defined such that an observation is given an increased variance with probability p_m. Therefore, this specification is build upon the assumption that there are at least some outliers present in the data. However, looking at the data in the sample until Q4-2019, there are not that many outlying observations present, especially in the nowcasting period. Therefore, imposing this prior on the covariance matrix might increase the variance imposed on certain variables too much, which can lead to worse nowcasting results.

Looking at the LPS values displayed in the left hand side of Table 2, consistent with Huber et al. (2023), the MF-BAVART model seems to outperform the MF-VAR slightly. Overall, the MF-BAVART model has higher LPS values, and the differences between the performance of the models are larger if it is the better model of the two, apart from for France. Therefore, the density nowcasts of the MF-BAVART model slightly outperform the MF-VAR model.

Moreover, using the MF-GVARP model results in higher LPS values for all instances compared to both the MF-BAVART and MF-VAR models. Using this specification therefore seems to result in poorer overall distributional nowcasts. This can be attributed to a too high variance imposed in the model, which results in a posterior distribution which is too wide, assigning a low density to the actual realised value of GDP growth. This gives rise to the idea that perhaps a different prior on the variance σ_m^2 could be useful to reduce both the RMSE and increase the LPS in samples with stable macro-economic conditions. Table 2 also shows that using the MF-GVARP-SVO model does not results in consistently lower LPS than using the MF-GVARP model, with roughly half of the values lower, and half higher. In this case, using the SVO co-

variance matrix results in nowcasts that are wider than required. In economically stable times, such as the nowcasting period under consideration here, this will result in too low probability densities being assigned to the actual value of GDP growth, which results in a lower LPS.

4.1.2 Full Sample

Looking at the right side of Table 2, we see that including the pandemic observations clearly leads to worse overall GDP growth nowcasts. This can be seen by the increase in RMSE and decrease in LPS in Table 2. This worse performance is due to the last few COVID observations, as these are the only differing data points between the left and right parts of the table. Section 2 showed that these Covid-19 observations are clearly outlying datapoints, with much lower GDP growth numbers compared to the rest of the sample. Therefore, nowcasting models seem to have issues with accurately predicting these datapoints.

Furthermore, when comparing the MF-BAVART and MF-VAR models, the conclusion of the comparison between the two models based on the RMSE does not change much. There is still no model that clearly outperforms the other based on this metric. There are some instances for which the MF-BAVART model performs better, but this is not always the case.

Including the MF-GVARP model in the comparison, table 2 shows that the performance of the point nowcasts of the MF-GVARP model display relative improvement compared to the point nowcasts of the MF-VAR and MF-BAVART models when including the pandemic period. This model outperforms the MF-BAVART model in roughly 40% of the time, and the MF-VAR model in 50% of the time. Therefore, the flexibility of the MF-GVARP model seems to lead to relatively improved nowcasts when including crisis observations. However, clear outperformance of the model is not evident.

Interestingly, the MF-GVARP-SVO model still underperforms the MF-GVARP model in this sample, indicated by higher RMSEs in most cases. This is especially evident by the RMSEs for the first month of the quarter, caused by GDP nowcasts that are off by a large amount in the first month of the second quarter of 2020. The model predicts GDP growth values that are much too high for this period, due to the outlier specification of the covariance matrix.

When looking at the LPS, the worsening of the model performance is also clear for all models when including the COVID-19 period. Again in this case, the MF-BAVART model outperforms the MF-VAR model slightly, with the difference between the two becoming more pronounced. The model therefore seems to have better density forecasts in the presence of large volatility and outliers compared to the MF-VAR model. This is consistent with the conclusion from Huber et al. (2023), and can be explained as follows. Because of the underlying BART model in the mixed frequency framework, there are two ways this model handles outliers. The model can either create a new tree with very few observations, or it can create a new node in a more complicated tree. Both methods will absorb these outlying observations. Furthermore, as these nodes of outlying observations have few observations, they add little information to the likelihood, which is reflected in the wide variance of the distributional GDP nowcast (Huber et al., 2023).

Furthermore, Table 2 clearly shows that the relative performance of the MF-GVARP model compared to the MF-BAVART and MF-VAR models has improved enormously. The LPS values

of the MF-GVARP model have increased only relatively little when including these pandemic observations, resulting in substantial outperformance using the nowcasts until Q2-2020. This can be explained by a wider distribution of GDP nowcasts, as will be illustrated in section 4.2. These wider distributions will assign a higher likelihood to the outlying observations of the GDP growth data points influenced by COVID. Therefore, the corresponding LPS will be higher.

Looking at the MF-GVARP-SVO model, contary to the RMSEs, the LPS actually show a clear outperformance of the MF-GVARP-SVO model compared to the other models. The MF-GVARP-SVO model has higher LPS values than the MF-GVARP model for all observations. Therefore, the covariance specification seems to correctly shield against outlying values when looking at the entire distribution of the GDP nowcast. This can be attributed to the prior imposed on the covariance matrix, which can also be used to explain the worse performance in the sample until Q4-2019. In the full sample, until Q2-2020, there are outlying observations added to the data (the Covid period). The covariance specification of the MF-GVARP-SVO model predicts these outliers with a probability p_m. Therefore, larger variances are imposed with probability p_m. These larger variances are actually appropriate in this case, which leads to relatively more accurate nowcasted distributions of GDP growth.

Concluding, based on the LPS, the MF-GVARP and MF-GVARP-SVO models seem to perform better relative to the MF-BAVART and MF-VAR models in more uncertain times, e.g. times of economic turmoil such as the COVID-19 period. Also, similarly to Huber et al. (2023), this paper finds that the MF-BAVART model performs better than the MF-VAR model in uncertain times. These findings can be explained by a quickly widening nowcasted GDP growth distribution returned by the MF-GVARP and MF-GVARP-SVO models, and to a lesser extent the MF-BAVART model, when the uncertainty of the estimate increases. For the MF-GVARP models this is mostly due to the great flexibility of Gaussian processes. A Gaussian process is a very flexible model, with the ability to fit to any nonlinear relationship in the data. This results in very flexible models, with the ability to closely fit the data fed to the model. However, this has the drawback of overfitting on the training data. Moreover, in economically stable times, the MF-GVARP and MF-GVARP-SVO models cannot improve the MF-BAVART and MF-VAR models, potentially caused by predicting too wide distributions. Lastly, adding a SVO specification to the covariance matrix of the model results in worse point nowcasts. However, it does result in a better overall distribution of the nowcasts as measured by the LPS.

	Until 2019-Q4								Until Q1-2020							
	MF-BAVART		MF-VAR		MF-GVARP		MF-GVARP-SVO		MF-BAVART		MF-VAR		MF-GVARP		MF-GVARP-SVO	
	RMSE	LPS	RMSE	LPS	RMSE	LPS	RMSE	LPS	RMSE	LPS	RMSE	LPS	RMSE	LPS	RMSE	LPS
Germany																
M/Q 1	0.539	-33.951	0.625	-40.881	0.782***	-63.911***	0.929***	-48.282	1.334	-59.837	2.526	-327.643	2.365	-77.192	3.277	-56.629
M/Q 2	0.521	-33.133	0.569	-42.728	0.676***	-72.382***	0.699*	-37.557	1.769	-121.416	2.186	-329.132	1.967	-110.920	1.118	-46.845
M/Q 3	0.475	-28.267	0.552*	-37.309	0.568*	-56.468**	0.518	-43.663	1.708	-109.601	2.136	-236.424	1.513	-81.077	1.390	-58.350
Spain																
M/Q 1	0.286	-12.197	0.366**	-23.110	0.311	-37.030	1.113**	-76.235***	2.237	-493.212	3.340	-555.403	3.171	-90.198	18.038	-85.621
M/Q 2	0.269	-10.248	0.342**	-17.655	0.345**	-27.506*	0.839***	-65.007***	3.146	-523.810	2.543	-616.076	3.780	-301.738	2.410	-72.181
M/Q 3	0.248	-5.142	0.337**	-16.206**	0.314	-28.536**	0.498***	-26.451**	2.865	-294.976	2.314	-631.698	2.828	-218.511	3.052	-64.031
Italy																
M/Q 1	0.373	-10.524	0.304	-11.550	0.510*	-29.275**	1.029***	-58.256***	1.557	-221.829	1.093	-133.399	2.362	-384.997	11.201	-98.766
M/Q 2	0.367	-10.986	0.305	-10.024	0.396	-15.636	0.608**	-30.523***	2.832	-249.809	1.128	-159.560	2.210	-107.062	1.937	-103.079
M/Q 3	0.354	-10.858	0.331	-9.754	0.393	-26.437*	0.449	-21.448	1.552	-43.906	1.424	-129.908	1.927	-60.340	7.861	-59.208
France																
M/Q 1	0.331	-18.678	0.312	-9.565	0.387*	-30.078	0.515**	-29.074	1.374	-371.846	1.413	-219.100	2.921	-160.071	21.596	-126.885
M/Q 2	0.310	-15.757	0.295	-7.541*	0.392**	-46.342**	0.426**	-31.042*	2.410	-481.872	2.502	-347.897	2.216	-292.631	2.016	-124.258
M/Q 3	0.302	-14.147	0.283	-5.226*	0.368**	-31.518**	0.317	-17.321	2.467	-95.400	1.841	-294.039	2.092	-303.810	6.829	-123.959

Table 2: RMSE and LPS of the Different Models for Every Country and Month of Quarter

The difference in model performance is tested for every model and evaluation metric with a Diebold-Mariano test with as base the MF-BAVART model. Significance is denoted as * for $p < 0.1$, ** for $p < 0.05$ and *** for $p < 0.01$.

4.2 Posterior Interval Width

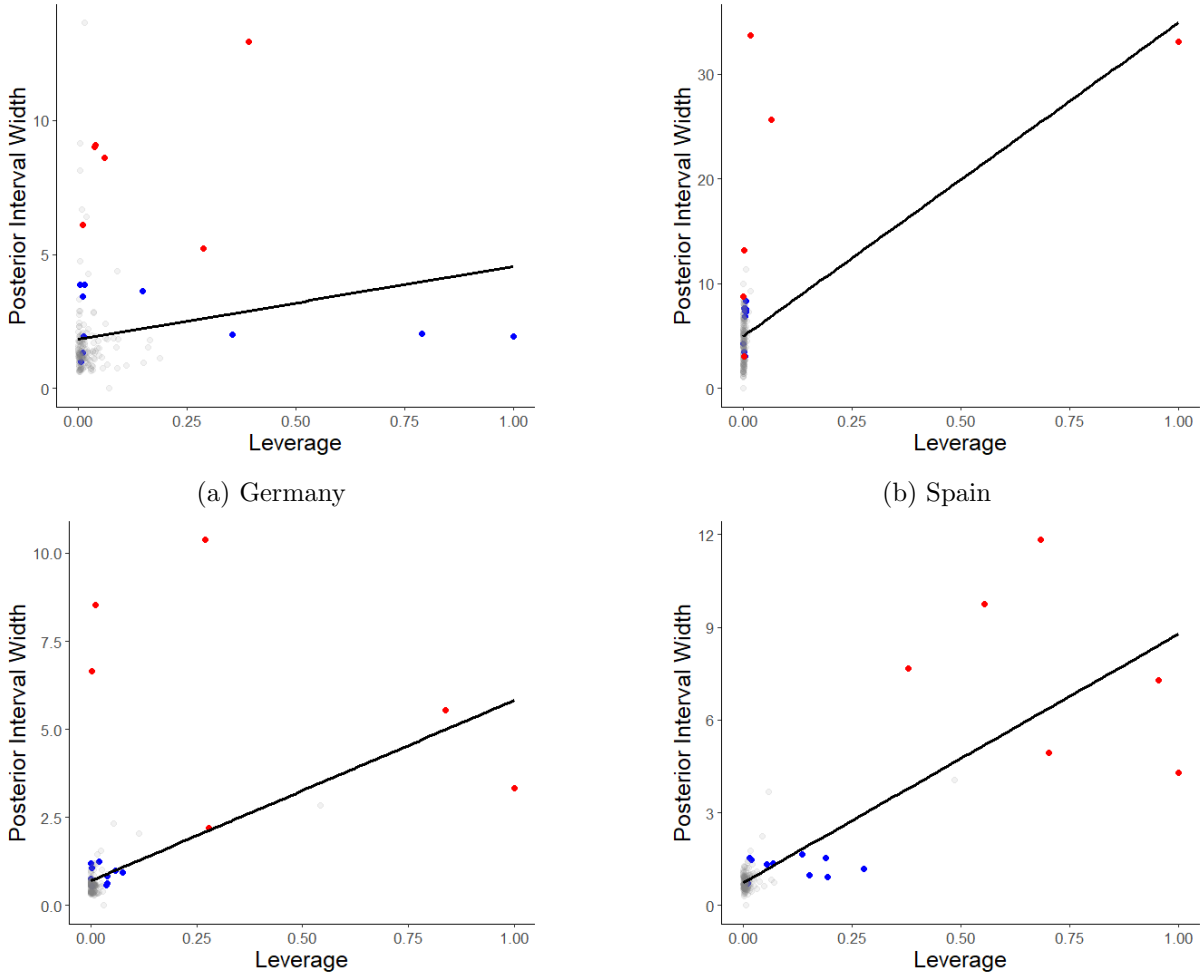
This section provides some intuition on why the MF-BAVART model outperforms the MF-VAR model in pandemic times, and why the GVARP models outperform both the MF-BAVART and MF-VAR models based on the LPS in the same period. The widening of the distribution of the GDP nowcasts when uncertainty increases is illustrated using Figure 2 and 3. Figure 2 shows scatter plots of the posterior interval width against the leverage for the nowcasts generated by the MF-BAVART model. Similarly to Huber et al. (2023), the posterior interval width is calculated as the difference between the 95^{th} and 5^{th} percentile of the nowcasts produced in the MCMC algorithm. The leverage is calculated as $diag(X(X'X)^{-1}X')$, where $X = (x_1, .., x_T)$, with x_t denoting he posterior mean of the latent GDP growth at time t. This leverage is scaled to be between zero and one using Equation 37.

$$y_{scaled} = \frac{y - min(y)}{max(y) - min(y)} \tag{37}$$

The means for the latent GDP growth are not the same as the forecasts used in the previous sections. Instead, X is taken to be the mean of the latent state estimates from the estimation of the model using all available data (from Q2-2005 until Q1-2020).

Figure 2 shows a positive correlation between the leverage vector and the posterior interval width for all countries. This indicates that the model generates a wider posterior distribution, indicated by a wider set of possible values, if it encounters outlying observations, as indicated by a higher leverage. This is especially visible for the pandemic observations (the red dots), and to a lesser extent the financial crisis observations (blue dots). This indicates that part of the improved performance of the MF-BAVART model is attributable to the ability of the model to increase the variance of the posterior distribution in times of greater uncertainty. This characteristic of the model is also reflected in the outperformance compared to the MF-VAR model based on the LPS in table 2, and can be explained by the flexibility of the BART model when facing outliers, as explained in section 4.1.

This characteristic of the MF-BAVART model increases its performance when nowcasting GDP growth due to the nature of GDP growth processes. Figure 1a and 1b in Section 2 have shown that GDP is characterised by longer periods of relatively stable and low growth, alternated by periods of recession, in which there are large negative GDP changes. Therefore, a GDP nowcasting model should have the ability to shield the other nowcasts from these large negative GDP growth spikes, while at the same time predicting these spikes as they occur. Correctly widening the distribution when uncertainty increases (crises observations) ensures that these negative GDP spikes are assigned a higher likelihood in the nowcasted GDP growth distribution, which results in outperformance based on the LPS.



(a) Germany

(b) Spain

(c) France

(d) Italy

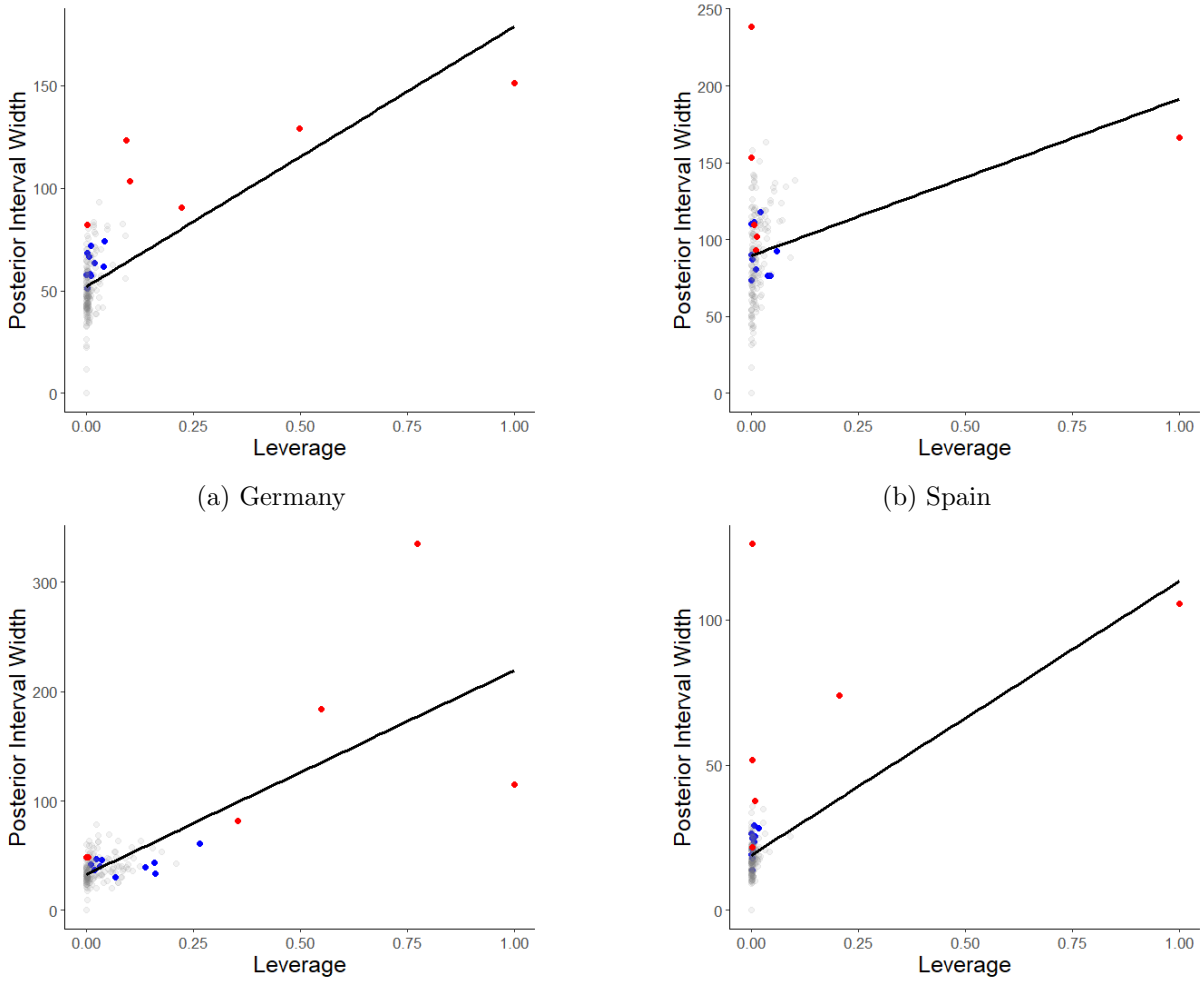
Figure 2: Scatterplot of Posterior Interval Width Against the Leverage Vector for the Different Countries for the MF-BAVART Model

The blue points indicate datapoints of the financial crisis, and the red points indicate data from the Covid-19 period.

Furthermore, Figure 3 shows the same scatterplot of posterior interval width against leverage for the MF-GVARP-SVO model. The figure displays the same positive correlation, indicating that the model predicts a wider distribution GDP growth when outlying observations are encountered. Furthermore, the figures show that the posterior interval widths for the MF-GVARP-SVO model are much larger (8x-30x as large) than for the MF-BAVART model, especially when looking at the pandemic observations. This indicates that the MF-GVARP-SVO model correctly widens the distribution by a larger amount when it encounters outlying observations. This results in the higher LPS values in Table 2, as the rare observations get higher assigned likelihoods.

However, even for those leverages close to zero, the posterior interval widths are much higher for the MF-GVARP-SVO model than for the MF-BAVART model. The MF-GVARP-SVO model thus generates wider distributional nowcasts even when no outliers are encountered. This is not desired behaviour, as we want the model to nowcast a distribution that accurately discribes the distribution of the GDP growth. Therefore, a tight distribution is desired in times where

GDP growth can be nowcasted with greater certainty. This property of generating 'too' wide distributional nowcasts can result in median values that are further off, and LPS values that are lower (as the LPS formula implies higher LPS values for distributions that assign more weight to the actual realised value). These unnecessarily wide nowcasts may therefore be a reason for the poorer performance of the models in the period excluding the COVID observations. These wide nowcasts are a property that might be caused by a prior on the covariance that is not restrictive enough.



(a) Germany

(b) Spain

(c) France

(d) Italy

Figure 3: Scatterplot of Posterior Interval Width Against the Leverage Vector for the Different Countries for the MF-GVARP-SVO Model

The blue points indicate datapoints of the financial crisis, and the red points indicate data from the Covid-19 period.

4.3 Model Calibration

Next, the PIT, as introduced in Section 3.5.3, is used to discuss the calibration of the models. Similarly to Table 2, Table 3 shows the PIT for all models for 2 samples, namely the GDP nowcasts until Q4-2019, and the nowcasts until Q2-2020. The table shows that the MF-VAR and MF-BAVART models perform similarly based on these statistics. However, based on the

means, the MF-BAVART model is calibrated worse in the sample until Q4-2019, with means deviating more from 0 in roughly 67% of the cases. Looking at the σ^2 coefficients, mixed results are presented. Depending on the country, either of them might be calibrated very well (coefficients close to 1), or not that well at all (coefficients deviate up to roughly 1 from 1). The AR(1) coefficients also display mixed results. For France they are smaller for the MF-BAVART model, but for Italy they favour the MF-VAR model.

Including the pandemic period shows a more pronounced difference between the two models. More specifically, the calibration of the MF-VAR model gets relatively slightly better based on the AR(1) parameters, outperforming the MF-BAVART model roughly 67% of the time. Looking at the σ^2, again the MF-VAR is better calibrated, having a variance closer to 1 in roughly 67% of the cases. Based on the μ there is no clear difference between the two models, with both models having values closer to 0 in roughly 50% of the cases. Therefore, it is hard to draw conclusions on the relative performance of the models based on this table. Even though the MF-VAR model seems slightly better calibrated, this is not consistent across all periods and metrics.

What can be said is that the calibration of the models does not seem to change that much when including COVID-19 observations, especially when comparing the change in calibration to the change in RMSE and LPS. Furthermore, both the models are relatively well calibrated, with the means in the table being close to zero and the AR(1) parameters relatively close to zero as well in most of the cases. However, the variance of the transformed PIT is too high in most instances, sometimes even being two times larger than expected.

Looking at the MF-GVARP model for the first sample period, Table 3 shows that this model is calibrated slightly worse than the MF-BAVART model. The MF-GVARP model performs slightly worse based on the means, with values deviating more from 0 in a little more than half of the instances. Furthermore, the σ^2 of the model is too high, with values deviating substantially more from 1 than for the MF-BAVART and MF-VAR models in all cases. The AR(1) coefficients show mixed results. Here the coefficients are closer to zero than for the other two models under consideration in roughly half of the cases. Therefore, the model seems calibrated slightly worse than the MF-BAVART and MF-VAR model, which is especially clear when considering the variance of the PIT.

When looking at the MF-GVARP-SVO model, we see it performs relatively similar to the MF-GVARP model when looking at the mean values of the PIT, with values closer to zero in roughly half of the cases. The variances of the transformed PIT are closer to one for the MF-GVARP-SVO model than for the MF-GVARP in all cases. Based on the AR(1) coefficients, the MF-GVARP-SVO model performs worse, with values further from 0 than the MF-GVARP model in most cases. Overall, the MF-GVARP-SVO model performs relatively similar compared to the MF-GVARP model, indicating that adding the SVO specification does not improve model calibration in this sample period. However, it is clear that the variance of the transformed PIT is closer to one than for the MF-GVARP model.

Including the Covid observations improves the calibration of the MF-GVARP models slightly when looking at the means. Now, the MF-GVARP model performs relatively similar to the MF-BAVART and MF-VAR based on this statistic. However, in this period, the σ^2 are still too large

in all instances. The AR(1) coefficients give mixed results, with values that are higher than the corresponding values of the MF-BAVART and MF-VAR models in roughly half of the instances. The MF-GVARP-SVO model improves in calibration relative to the MF-GVARP model when using the full sample. The mean values of the PIT are closer to zero in slightly more than 50% of the cases, and the variances are closer to one in all but 1 instance. However, the AR(1) coefficients still indicate worse calibration for the MF-GVARP-SVO model, with values farther from 0 in most instances.

Therefore, based on these findings it is relatively difficult to see if the MF-GVARP model is better calibrated than the MF-GVARP and MF-BAVART models. What is evident however, is that the variance coefficient of the GVARP model is too high. This can be caused by the wide predictive distribution of the model. This wide predictive distribution, will lead to quite a few predicted values that are large positive or negative. This will result in too high values for σ^2 in the probability integral transform.

Furthermore, looking at the relative calibration of the MF-GVARP-SVO model compared to the MF-GVARP model, the table shows that the calibrations of the two models do not differ much. However, including the Covid period seems to improve the relative calibration of the MF-GVARP-SVO model, especially when looking at the means. This indicates that the outlier adjusted stochastic volatility specification does lead to better calibration in the presence of outliers. However, the SVO covariance matrix does not lead to a consistently better calibration throughout all measures and observations when including the Covid-19 period. Moreover, we see that adjusting the covariance matrix to have a SVO specification ensures that the variances of the PIT improve in value.

4.4 Economic Interpretation

Accurate GDP nowcasts are useful for a wide variety of people, from business owners trying to decide whether to invest in expanding their business, to politicians deciding what policies to implement. It is therefore important to put the results of Section 4 into economic perspective.

This paper shows that using a Gaussian process in a mixed frequency framework does not improve the MF-BAVART model of Huber et al. (2023). The point nowcasts provided by the MF-GVARP model underperform the existing MF-BAVART model when measured by RMSE. Also, the model cannot be used to improve the nowcasted density of GDP growth in economically stable times, as it produces too wide estimates of the distribution of GDP growth in these periods.

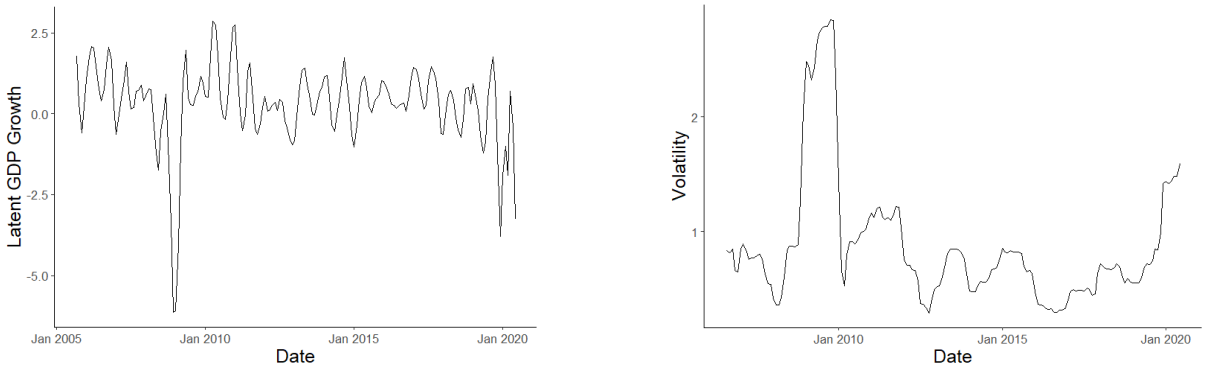
One exception occurs when it is clear the economy is heading into unstable times, as was the case at the outbreak of Covid-19, when businesses had to shut down and the economy grinded to a halt. In these instances, the MF-GVARP model can be used to improve the nowcasted densities of GDP growth by correctly predicting wider GDP growth distributions. However, it is difficult to know when a crisis will occur beforehand. Therefore, it is not clear when to use this model. Also, because policy makers and other individuals typically also want to get the best point estimate of GDP growth, the MF-GVARP model as defined in this paper cannot be used by policy makers and other individuals to improve upon existing GDP growth models.

	Until Q4-2019												Until Q1-2020											
	MF-BAVART			MF-VAR			MF-GVARP			MF-GVARP-SVO			MF-BAVART			MF-VAR			MF-GVARP			MF-GVARP-SVO		
M/Q	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Germany																								
μ	-0.012	0.060	0.094	0.111	0.031	0.024	0.063	0.333	0.213	0.034	0.180	0.282	-0.163	-0.094	-0.062	-0.047	-0.123	-0.129	-0.092	0.164	0.071	-0.062	0.075	0.131
σ^2	1.934	1.956	1.692	1.890	1.995	1.602	2.663	2.719	2.496	1.033	1.096	1.958	2.250	2.292	2.053	2.245	2.320	1.947	2.962	3.100	2.743	1.152	1.301	2.272
AR(1)	-0.556	-0.508	-0.456	-0.513	-0.454	-0.438	-0.374	-0.522	-0.456	0.105	-0.162	-0.073	-0.388	-0.333	-0.260	-0.345	-0.276	-0.267	-0.230	-0.373	-0.356	0.192	-0.171	0.022
Spain																								
μ	-0.346	-0.300	-0.122	-0.391	-0.291	-0.418	0.104	-0.417	-0.259	0.022	0.098	0.229	-0.468	-0.435	-0.267	-0.522	-0.428	-0.547	-0.053	-0.547	-0.391	-0.084	0.041	0.072
σ^2	1.889	1.902	1.709	2.018	1.908	1.830	2.667	2.912	2.372	0.543	0.617	1.355	2.067	2.139	2.005	2.226	2.148	2.041	2.978	3.064	2.566	0.735	0.655	1.741
AR(1)	-0.237	-0.285	-0.241	-0.208	-0.303	-0.231	0.109	-0.152	-0.255	0.359	0.690	0.342	-0.148	-0.143	-0.067	-0.098	-0.161	-0.121	0.179	-0.053	-0.167	0.371	0.680	0.441
France																								
μ	-0.110	-0.061	0.087	-0.078	-0.131	-0.077	0.103	-0.080	-0.102	0.172	0.281	0.086	-0.211	-0.209	-0.069	-0.202	-0.275	-0.224	-0.032	-0.198	-0.248	0.011	0.141	-0.070
σ^2	1.948	2.076	2.152	0.922	0.955	1.019	2.480	2.737	2.828	1.264	2.134	2.127	2.071	2.370	2.486	1.170	1.290	1.366	2.693	2.862	3.069	1.672	2.392	2.462
AR(1)	-0.155	-0.067	-0.083	-0.181	-0.233	-0.287	-0.305	-0.153	-0.177	0.395	0.344	-0.383	-0.028	0.099	0.078	0.095	0.091	0.009	-0.169	-0.058	-0.064	0.399	0.376	-0.153
Italy																								
μ	-0.209	-0.183	-0.116	-0.087	-0.102	-0.080	0.198	0.053	-0.030	-0.185	-0.167	-0.102	-0.349	-0.324	-0.267	-0.182	-0.123	-0.227	0.036	-0.084	-0.149	-0.327	-0.271	-0.097
σ^2	0.899	0.989	1.062	0.422	0.513	0.712	2.128	1.505	1.753	1.996	1.367	1.647	1.215	1.308	1.426	0.619	0.799	1.074	2.498	1.781	1.937	2.260	1.524	2.006
AR(1)	0.026	0.183	0.165	-0.032	0.037	0.158	0.167	-0.037	-0.091	0.485	0.325	0.064	0.216	0.329	0.294	0.144	-0.083	0.300	0.271	0.103	0.014	0.488	0.274	-0.016

Table 3: Summary Statistics of Transformed Probability Integral Transform for the Different Models

Additionally, this paper finds that adding an outlier adjusted stochastic volatility covariance matrix to the MF-GVARP model does not improve the GPD growth nowcasts. Even though the distribution is widened in times of crisis, the performance of the point nowcasts worsens substantially. Also, the predicted distribution of GDP growth is too wide when in stable macroeconomic periods. Therefore, adding this SVO specification does not improve GDP growth predictions. The policy implications of this finding are clear. As using an outlier adjusted stochastic volatility covariance matrix worsens the nowcasting performance, this specification of the covariance matrix should be avoided when nowcasting GDP growth.

To understand why the SVO covariance matrix cannot improve the GDP growth nowcasts, Figure 4 shows the estimates of the latent monthly GDP growths for Germany as estimated by the MF-BAVART model (the best performing model in this paper). These latent states are estimated by the model that nowcasts GDP growth in Q2-2020 using all available data, resembling a nowcast made in the last month of the quarter. The figure shows that the monthly latent states of GDP growth are relatively stable. Also, even though the volatility of the latent states of GDP growth varies, most of this variation is related to crises, with spikes occuring during the financial crisis and the Covid-19 period. Outside of these observations, GDP growth volatility does not seem to vary that much over time. This is to be expected, as GDP growth itself is relatively stable. This is especially noticeable when compared to other macroeconomic variables such as inflation, for which stochastic volatility has often been used successfully in models, such as by Stock and Watson (2007). Figures 5a and 5b in appendix A show that monthly inflation of Germany does indeed display more variation, as does the volatility of inflation. Therefore, the stability of GDP growth ensures that the covariance matrix of the mixed frequency framework does not need a time varying outlier adjusted stochastic volatility component. Thus, the SVO specification is not suitable to improve GPD growth nowcasts.



(a) GDP Growth

(b) GDP Growth Volatility

Figure 4: Latent GDP Growth of Germany

The figure shows the latent monthly GDP growth of Germany estimated by the MF-BAVART model, ranging from September 2005 until June 2020.

5 Conclusion

This paper investigates whether the performance of a mixed frequency GDP nowcasting model can be improved by replacing the Bayesian additive regression trees model in a mixed frequency

framework by a Gaussian process. Additionally the paper adds an outlier adjusted stochastic volatility covariance matrix to check if this can improve nowcasting performance, especially in crisis times. This paper researches this by using a data from Q2-2005 until Q2-2020. The dataset contains GDP growth values, as well as several macroeconomic variables available on four countries: Germany, Spain, Italy, and France. We nowcasts GDP growth using a mixed frequency vector autoregressive (MF-VAR) model and a mixed frequency Bayesian additive vector autoregressive trees (MF-BAVART) model, and introduces two new models. First of all, the paper introduces a mixed frequency Gaussian vector autoregressive process (MF-GVARP) model, and adds an outlier adjusted stochastic volatility covariance matrix to this specification, creating the mixed frequency Gaussian vector autoregressive process with outlier adjusted stochastic volatility (MF-GVARP-SVO).

We conclude that the MF-BAVART model outperforms the simpler MF-VAR model, especially when comparing them based on the entire distribution as measured by the cumulative log-predictive score (LPS). This outperformance seems to stem from the ability of the MF-BAVART model to widen the distribution in more uncertain and volatile economic times. Furthermore, the MF-GVARP model underperforms the MF-BAVART and MF-VAR models when excluding 2020 (the COVID-19 observations) from the sample. This underperformance is evident both for point estimates, as well as the entire predicted distribution. However, when including the COVID-19 data in the sample, the relative nowcasting performance of the MF-GVARP model improved. Even though the MF-GVARP model still underperformed the MF-VAR and MF-BAVART models based on point nowcasts, the MF-GVARP improves the entire distributional nowcasts of GDP growth based on the LPS. This improvement largely arises due to the ability of widening the distribution even further when in times of economic turmoil. Also, adding an outlier adjusted stochastic volatility covariance matrix to the MF-GVARP model did not result in improved GDP nowcasts. This result seems to be attributable to the stability of GDP growth over time, ensuring an outlier adjusted stochastic volatility specification is not necessary in the mixed frequency framework.

The policy implications of the results in this paper are not immediately clear. Using an outlier adjusted stochastic volatility covariance matrix cannot improve the performance of the models when nowcasting GDP, and should therefore not be used by policy makers. However, the MF-GVARP model performs relatively well when forecasting the distribution of GDP growth when including the Covid-19 observations. This model can thus be used to improve the nowcasted distribution, but only when in crisis times. Because it is difficult to know in advance when such a time will occur, it is in practice not possible to use this model to improve GDP nowcasts. Therefore, it will in practice be difficult to use either models to improve GDP nowcasts for policy makers.

Finally, there is still room for additional research in this field. The underperformance of the MF-GVARP model could be caused by a too flexible model, partially caused by a variance prior that is not restrictive enough. Therefore, it might be worth to research whether imposing more restrictive priors on the variance can improve the nowcasting performance in stable economic situations. Additionally, the methods used in this paper use a forward filtering backwards sampling (FFBS) algorithm to estimate the latent states of the variables that are observed at a

quarterly frequency. In order to use this FFBS algorithm, the model is linearized first. It might be more appropriate to use different algorithms such as particle filters to estimate the latent states. As the MF-GVARP model did not improve the performance of point nowcasts, this paper would also suggest to research the appropriateness of different models in the mixed frequency framework used in this paper. Lastly, it is worth exploring how the models used in this paper perform compared to other often used nowcasting techniques such as MIDAS or dynamic factor models.

References

Caputo, B., Sim, K., Furesjö, F., & Smola, A. (2002, 01). *Appearance-based object recognition using svms: Which kernel should i use.*

Carriero, A., Clark, T. E., Marcellino, M., & Mertens, E. (2022). Addressing COVID-19 Outliers in BVARs with Stochastic Volatility. *The Review of Economics and Statistics*, 1-38.

Carvalho, C. M., Polson, N. G., & Scott, J. G. (2010). The horseshoe estimator for sparse signals. *Biometrika*, *97*(2), 465–480.

Chipman, H. A., George, E. I., & McCulloch, R. E. (2010). Bart: Bayesian additive regression trees. *The Annals of Applied Statistics*, *4*(1), 266–298.

Clark, T. E. (2011). Real-time density forecasts from bayesian vector autoregressions with stochastic volatility. *Journal of Business & Economic Statistics*, *29*(3), 327–341.

Ebden, M. (2015). *Gaussian processes: A quick introduction.*

Farooq, U., Ahmed, J., & Khan, S. (2021, January). Do the macroeconomic factors influence the firm's investment decisions? A generalized method of moments (GMM) approach. *International Journal of Finance & Economics*, *26*(1), 790-801.

Hauzenberger, N., Marcellino, M., Pfarrhofer, M., & Stelzer, A. (2024). *Nowcasting with mixed frequency data using gaussian processes.*

Huber, F., Koop, G., Onorante, L., Pfarrhofer, M., & Schreiner, J. (2023). Nowcasting in a pandemic using non-parametric mixed frequency vars. *Journal of Econometrics*, *232*(1), 52-69.

Karatzoglou, A., Smola, A., Hornik, K., & Zeileis, A. (2004). kernlab - an s4 package for kernel methods in r. *Journal of Statistical Software*, *11*(9), 1–20.

Kastner, G. (2016). Dealing with stochastic volatility in time series using the r package stochvol. *Journal of Statistical Software*, *69*(5), 1-30.

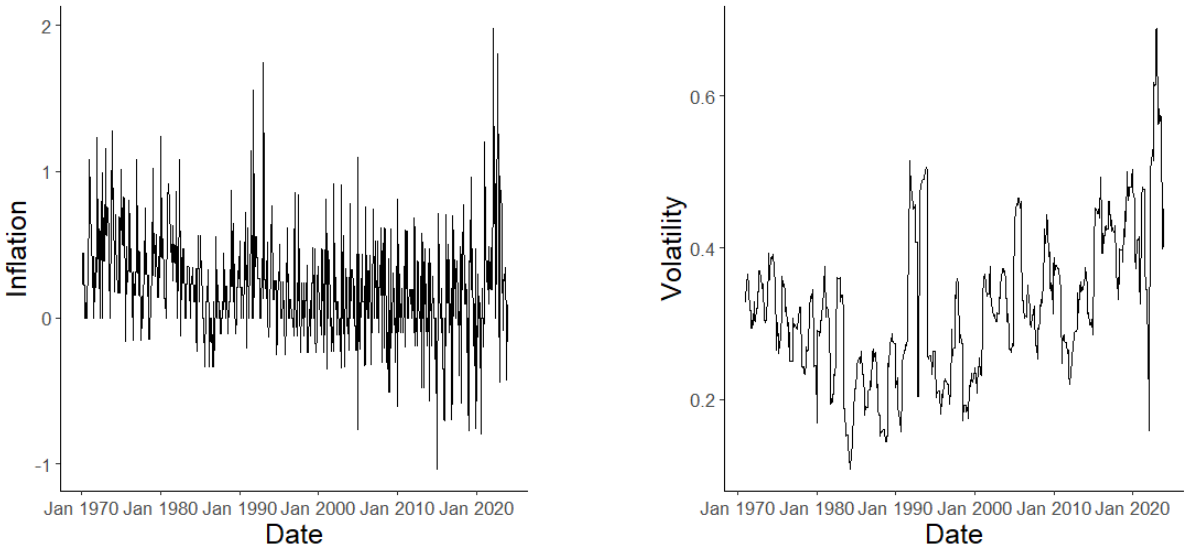
Li, Z.-Z., Li, L., & Shao, Z. (2021). Robust gaussian process regression based on iterative trimming. *Astronomy and Computing*, *36*, 100483.

Schorfheide, F., & Song, D. (2015). Real-time forecasting with a mixed-frequency var. *Journal of Business & Economic Statistics*, *33*(3), 366–380.

Stock, J. H., & Watson, M. W. (2007). Why Has U.S. Inflation Become Harder to Forecast? *Journal of Money, Credit and Banking*, *39*(s1), 3-33.

Tumanoska, D. (2020, 03). The relationship between economic growth and unemployment rates: Validation of okun's law in panel context. *Research in Applied Economics*, *12*, 33.

A Inflation



(a) Inflation

(b) Rolling 12 Month Volatility of Inflation

Figure 5: Inflation and Rolling 12 Month Volatility of Inflation in Germany From February 1970 Until December 2023

Figures 5a and 5b display the inflation, as well as the rolling 12 month volatility of inflation, in Germany from February 1970 until December 2023. The inflation data is obtained from the World Bank. The figures clearly show that inflation varies substantially from month to month. Also, the volatility of inflation varies considerably in short periods of time. Therefore, when comparing these results to Figure 4, GDP seems more stable over time.

B Code

This section explains the code that is used to generate the nowcasts of the models used in the paper, as well as all the figures and tables presented. The attached zipfile contains all the code and the data used to generate the results in this paper. The folder contains a README.md file that explains the code. Note that once the working directory is set to this folder, with a directory ending in a "/, all files should run. (as an example, for my computer this would be done with: setwd(C:/Users/vandi/OneDrive/Documenten/Academic Year 2023-2024/Thesis/Code_Stefan_van_Diepen_573588/")). The file contains several subfolders, namely Code, Data, Display, and Figures, that are explained in the remainder of this appendix.

The "Code" subfolder contains the code that is used to create nowcasts using the models of the paper. This subfolder contains a read_data.R file, which contains 2 functions, one that reads the input data as provided by Huber et al. (2023), and one that reads the nowcasts made by the models. Note that this function is required as the nowcasts produced by the models of this paper are stored in excel files to ensure the nowcasts do not have to be generated every time a table is altered slightly, which is necessary as the models take a long time to run.

Furthermore, the code subfolder contains a utils.R file which contains several fuctions that

are used to make the nowcasts for the different models. These functions mainly ensure the proper data is selected for every nowcast and process the nowcasted distributions of GDP growth, such that the tables and figures can easily be generated from the stored nowcasts.

The Code folder also contains two subfolders, a folder "Models", which contains the code that is used to run the models and a folder "Nowcast", that is used to actually generate the nowcasts of the different models. The "Models" subfolder contains the file mfbavart_func_var.R, which contains the function that is used to run all models (MF-BAVART, MF-VAR, MF-GVARP, and MF-GVARP-SVO), with the model being selected depending on the parameters "model" (indicating which model is used) and sv (indicating whether the SVO specification of the covariance matrix is used). Secondly, this subfolder contains the file aux_func.R, which contains auxiliary functions that are used by the function in the mfbavart_func_var.R file to estimate the models. The aux_func.R file is provided by the github of Michael Pfarrhofer (repository mf-bavart), and the mfbavart_func_var.R file is the "mfbavart_func.R" file from the same github repository, but changed such that it accomodates the estimation of MF-VAR and MF-GVARP models as well. Note that this github repository contains the code for the paper "Nowcasting in a Pandemic using Non-Parametric Mixed Frequency VARs" by Huber et al. (2023). Furthermore, the mfbavart_func_var.R file draws inspiration from the mfbavart.R file that is obtained from an old forked repository of the repository of Michael Pfarrhofer, to obtain nowcasts using the MF-VAR model.

The "Nowcast" subfolder that is contained in the "Code" subfolder contains 4 files that are used to obtain the nowcasts from the models described in the paper. Every file in this folder calls the function in mfbavart_func_var.R via the helper functions defined in utils.R, with the proper parameters. For the MF-BAVART model, the model is called with "model = MF_BAVART", for the MF_VAR model, the model is called with "model = MF_VAR". For the MF_GVARP model, the model is called with "model = MF_GVARP" and "sv = False", and the MF-GVARP-SVO model is called with "model = "MF_GVARP" and "sv = True".

The "Data" subfolder contains 2 folders. One subfolder contains input data, as provided by Hubert et. al and described in Section 2, as well as inflation data used for the figure in the appendix as provided by the world bank. The second subfolder contains the nowcasts generated by the models in the paper, as obtained by the files in the "Models/Nowcast" subfolder.

The "Display" subfolder contains the R files that are used to create the figures and tables that are presented in the paper. Table_RMSE_LPS.R contains the code that creates table 2, presenting the LPS and RMSE evaluation metrics for all models and countries. Figure_PIW_Leverage.R contains the code that creates the scatterplot of Figures 2 and 3 and PIT_table.R contains the code that creates the Table 3, presenting the information on the transformed probability integral transforms. Economic_Figure.R contains the code that is used to make Figures 4, 4a, 5a, and 5b used in the economic interpretation sections well as the inflation graph in the appendix. This includes both the plot of the latent states of GDP growth, as well as the figure of inflation in Germany. The file Summary_Stats.R contains the code to create Table 1 and Figures 1a and 1b, used to describe the data in section 2. The utils_display.R file contains all auxiliary functions that are used by the files mentioned above to make the tables and graphs presented in the paper.

Lastly, the "Figures" subfolder contains all the figures that are created for the paper in the

R files of the "Display" folder. This includes some older versions of figures that are not used in the paper. The versions of the figures used in the paper have a name ending in "_used".