

# New methods for timely estimates: nowcasting euro area GDP growth using quantile regression

GIAN LUIGI MAZZI, JAMES MITCHELL

2020 edition





**New methods for timely estimates:  
nowcasting euro area GDP growth  
using quantile regression**

**GIAN LUIGI MAZZI, JAMES MITCHELL | 2020 edition**

Manuscript completed in February 2020.

Printed by the Publications Office in Luxembourg.

The Commission is not liable for any consequence stemming from the reuse of this publication.

Luxembourg: Publications Office of the European Union, 2020

© European Union, 2020

Reuse is authorized provided the source is acknowledged.

The reuse policy of European Commission documents is regulated by Decision 2011/833/EU (OJ L 330, 14.12.2011, p. 39).

Copyright for the photographs: Cover: Anatolii Stoiko/Shutterstock

For any use or reproduction of photos or other material that is not under copyright of the European Union, permission must be sought directly from the copyright holders.

For more information, please consult: <https://ec.europa.eu/eurostat/about/policies/copyright>

The information and views set out in this publication are those of the authors and do not necessarily reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein.

Collection: Statistical working papers

Theme: General and regional statistics

ISBN 978-92-76-16865-2

ISSN 2315-0807

doi: 10.2785/26603

KS-TC-20-004-EN-N

## Abstract

This paper uses an application to explore the utility of quantile regression methods in producing (density) nowcasts. Our quantile regression modelling strategy is designed to reflect important nowcasting features, namely the use of mixed-frequency data, the ragged-edge and increasingly large numbers of indicators (*big data*). An unrestricted mixed data sampling strategy within a Bayesian quantile regression is used to accommodate a large mixed frequency dataset when nowcasting; we use a shrinkage prior to avoid parameter proliferation in what becomes a large dimensional quantile regression. In an application to euro area GDP growth, using over 400 mixed frequency indicators, we find that the quantile regression approach does not produce as accurate density nowcasts overall as the density combination approach of Mazzi et al. (2014) unless the indicators are orthogonalised and shrunk to a smaller number. The quantile regression approach overstates the uncertainties associated with GDP growth. This is reflected by high probability estimates of GDP at Risk (i.e. negative GDP growth) even during the period of strong GDP growth prior to the global financial crisis. However, when the nowcasts are formed early, at 30 or 15 days before the end of the quarter of interest, the quantile regression approach is better able to detect the ensuing recession than the density combination approach.

**Keywords:**

**Authors:** Gian Luigi Mazzi <sup>(1)</sup>, James Mitchell <sup>(2)</sup>

**Acknowledgement:** This work has been carried out under the supervision of the Eurostat project manager Matyas Meszaros. Special thanks goes to George Kapetanios and Massimiliano Marcellino for helpful comments.

---

<sup>(1)</sup> glmazzi@pt.lu

<sup>(2)</sup> james.mitchell@wbs.ac.uk

# Table of contents

<b>1.</b>	<b>Introduction</b> .....	<b>7</b>
<b>2.</b>	<b>Bayesian MIDAS quantile regression</b> .....	<b>9</b>
2.1	MIDAS Quantile Regression .....	9
2.2	Bayesian estimation.....	10
2.3	Lasso prior .....	11
2.4	Density nowcasts.....	12
<b>3.</b>	<b>Real-time data</b> .....	<b>14</b>
3.1	Indicator variables: aggregate and disaggregate.....	14
<b>4.</b>	<b>Reconsidering Mazzi et al. (2014): a comparison of QR against density nowcast combination methods</b> .....	<b>16</b>
4.1	Combination density nowcasts .....	16
4.2	Nowcasting component models.....	17
4.3	Combined density nowcasts .....	17
<b>5.</b>	<b>The trade-off between the timeliness and accuracy of nowcasts</b> .....	<b>18</b>
5.1	The information set as within-quarter data accrues .....	18
<b>6.</b>	<b>Results: nowcasting euro area GDP growth</b> .....	<b>20</b>
6.1	Evaluating the nowcast densities.....	21
6.1.1	GDP AT RISK: PROBABILITY OF A RECESSION .....	22
6.1.2	REDUCING THE SETS OF INDICATORS IN QR: DATA SHRINKAGE VIA PRINCIPAL COMPONENTS ANALYSIS (PCA) .....	23
<b>7.</b>	<b>Conclusion and future research</b> .....	<b>24</b>
<b>8.</b>	<b>Bibliography</b> .....	<b>25</b>

# Figures

Figure 1: Conditional Mean Nowcasts: from the Bayesian QR Approach .....21

Figure 2: GDP At Risk: probability of negative GDP growth from the recursive weight density combination approach of Mazzi et al. (2014) and the Bayesian QR approach .....22

## Tables

Table 1: Quarterly and monthly data used.....	15
Table 2: Average logarithmic score .....	21
Table 3: Average logarithmic score from QR with PCA used for data shrinkage .....	23



## Abbreviations

ALD	Asymmetric Laplace densities
AR model	Autoregressive model
EA	euro area
EA12	euro area, 12 member states
ESI	Economic Sentiment Indicator
GDP	Gross Domestic Product
IP	industrial production
LOP	Law of One price
MCMC	Markov chain Monte Carlo
MIDAS	Mi(xed) Da(ta) S(ampling)
PCA	Principal component analysis
PVAR	Panel Vector Autoregressive Models
QR	Quantile Regression
UMIDAS	Unrestricted MIDAS

## Country codes

EL	Greece
IE	Ireland

# 1

## Introduction

Official quarterly GDP data are published with a delay. In order to form a view about the current state of the economy, policymakers therefore use a wide range of more timely and higher frequency indicator data to form nowcasts. This paper uses an application to explore the utility of quantile regression methods in producing (density) nowcasts of GDP growth from these indicators. Thus it fills a gap in the literature; to our knowledge, there is no existing literature on nowcasting applications of quantile regression methods, certainly accommodating the mixed-frequency and ‘ragged-edge’ nature of the increasingly big datasets that typically characterise recent nowcasting applications. Because of their flexibility, in modelling the entire density, quantile regressions may prove a useful, and a relatively simple, method of producing reliable density nowcasts.

The literature on nowcasting GDP growth is large. This reflects the fact that statistical offices publish ‘official’ GDP data at a lag. For example, even Eurostat’s so-called Flash estimates of quarterly GDP growth, for the euro area (EA), are currently published 30 days after the end of the quarter. And they were for many years, from 2003–16, in fact published more slowly (45 days after the end of the quarter). However, ahead of these estimates, many informal measures and indicators of economic activity (often higher frequency) become available. But without a formal means of assessing the utility of these data, and relating them to official Eurostat GDP data, it is impossible to know how much weight to place on them when forming a view about the current state of the economy. An accurate, but timely, impression of the state of the economy is important for policymakers.

Various methods of computing GDP nowcasts have therefore been proposed, as Mazzi et al. (2014) review. But quantile regression methods have not been considered in this context. Our emphasis is producing density nowcasts. Publication of uncertainty estimates, alongside the central estimate, provides a means of indicating to the user the ‘quality’ of the nowcast, as measured by the confidence associated with the nowcast. The nowcasts from the quantile regressions are then compared with other nowcasting strategies, in particular the density combination approach of Mazzi et al. (2014). This approach has been found to be successful in previous applications. And like our proposed quantile regression method, the density combination approach is designed specifically to accommodate the mixed-frequency and ‘ragged-edge’ nature of the increasingly big datasets that typically characterise recent nowcasting applications.

The remainder of this paper is structured as follows. Section 2.1 motivates the use of Bayesian MIDAS quantile regression when nowcasting with mixed frequency datasets. Section 2 explains how an Unrestricted or UMIDAS approach makes sense in our application, given that this involves treating each quarterly indicator as three monthly indicators and letting the data help decide the weight on each indicator. Future work that generalises to consider the production of higher frequency nowcasts (e.g. daily nowcasts of GDP growth) may wish to impose restrictions on the parameter space as in the traditional MIDAS approach of midas rather than use this more ‘data expensive’ UMIDAS approach. Section 2.2 then sets out a Bayesian approach to estimate the UMIDAS model. A Bayesian approach is attractive when modelling and nowcasting in the face of a large number of

indicators — as shrinkage priors can be imposed to overcome the curse of dimensionality. Section 2.3 sets out the specific Lasso priors used. Lasso is a natural method to consider as it performs both variable selection and regularisation (i.e. it imposes simplicity on the model by penalising the inclusion of extra parameters, to avoid over-fitting in-sample). The aim is to improve prediction (out-of-sample) performance. Lasso is increasingly used in statistics and machine learning (albeit focus is typically on least squares methods and mean squared error loss evaluation rather than the full density as here). So it is natural to consider the utility of Lasso in our quantile regression application. Section 2.4 then considers how full posterior predictive density nowcasts can be produced from the Bayesian UMIDAS quantile regression. Importantly these nowcasts accommodate parameter estimation uncertainties, often ignored when producing nowcasts and forecasts from quantile regressions based on classical estimation methods.

Section 3 explains the real-time data <sup>(3)</sup> used in the application nowcasting euro area GDP growth. A distinction is made between soft and hard indicators, and aggregate and disaggregate indicators. Section 4 considers how the density nowcast combination approach of Mazzi et al. (2014) is used to benchmark and compare the performance of the Bayesian UMIDAS quantile regression approach. Summary details of the approach of Mazzi et al. (2014) are provided, with the reader referred to Mazzi et al. (2014) for full details. Section 5 then explains how and when nowcasts are produced at discrete intervals throughout the quarter, as within-quarter information on the monthly indicators arrives. The production of these nowcasts reflects the publication lags associated with particular indicators. Section 6 provides the empirical results, comparing the Bayesian UMIDAS quantile regression approach with the density nowcast combination approach of Mazzi et al. (2014). Section 7 concludes.

---

<sup>(3)</sup> Real-time data refers here to data for which data revisions matter or data where the timing of data releases is important, one way or another; this follows common usage of this phrase in a now large literature on real-time data analysis in economic statistics and applied macroeconomics e.g. for a review see Croushore2011

# 2

## Bayesian MIDAS quantile regression

Traditionally when seeking to model, explain and indeed nowcast official statistics, using regression methods, focus tends to be on the conditional mean of the variable of interest. Quantile regression is a generalisation and models the conditional  $\tau$ -quantile of the dependent variable — for example the first decile ( $\tau=0.1$ ) or the ninth decile ( $\tau=0.9$ ). Quantile regression describes the relationship at different points in the conditional distribution of the variable of interest.

Given increasing recognition - in an era of fan charts - of the importance of modelling and understanding the risks associated with the central estimate or nowcast/forecast, quantile regression is attractive in modelling and nowcasting the full distribution; see Aastveit et al. (2018) for a recent review of density forecasting.

### 2.1 MIDAS Quantile Regression

Nowcasting, in particular, is characterised by using within-period information to provide a more timely estimate of the current period for a variable that is published by the statistical office with a greater lag. Focus in this paper is nowcasting current quarter GDP growth using within-quarter (specifically monthly) known information/data on indicators such as industrial production, retail sales and qualitative business surveys. See Mazzi et al. (2014) and Foroni & Marcellino (2014) for applications (specifically to the euro area) and a helpful review.

We therefore follow Ghysels (2014) and consider a Mixed data sampling (or MIDAS) quantile regression:

$$Q_{y_t}(\tau|\Omega_t) = \beta_0(\tau) + \beta_1(\tau) \left[ B(L^{1/m}; \theta) x_t^m \right], \tau \in (0, 1) \quad (1)$$

where  $t$  ( $t=1, \dots, T$ ) is the quarter,  $y_t$  is quarterly GDP growth,  $B(L^{1/m}; \theta) = \sum_{k=1}^K b(k; \theta) L^{(k-1)/m}$ , where  $K$  is the order of the lag polynomial,  $L^{s/m} x_t^m = x_{t-s-m}^m$ ,  $x_t^m$  denotes the (in principle, vector of) indicators,

where  $m = 1, 2, 3$  denotes the month in the quarter. The  $B(L^{1/m}; \theta)$  function provides a means of parsimoniously modelling the monthly indicators. The estimated parameters will differ with the quantile,  $\tau$ . MIDAS models, based on use of distributed lags, have found increasing application in nowcasting and forecasting the conditional mean. This includes work by Ghysels et al. (2007) and Clements & Galvão (2008). MIDAS provides a simple means of running regressions that allows the regressand and regressors to be sampled at different frequencies. Lags of  $y_t$  can be included too, as in the quantile autoregression of Koenker & Xiao (2006).

Given that our frequency mismatch is small (quarterly to monthly), we do not employ distributed lag functions like  $B(L^{1/m}; \theta)$ . There is less need to use functions like this, that impose parsimony on the set of parameters to be estimated, when as in our application introducing the higher frequency indicators does not increase the parameter space massively (as it would, for example, if we considered daily indicators). Instead, following the suggestion of Foroni et al. (2015), we use the so-called Unrestricted or UMIDAS approach. This can be derived by aggregation of a general dynamic linear model in high frequency; it involves estimation of

$$Q_{y_t}(\tau|\Omega_t) = \beta_0(\tau) + \beta_1(\tau)x_t^1 + \beta_2(\tau)x_t^2 + \beta_3(\tau)x_t^3, \tau \in (0,1) \quad (2)$$

i.e. when nowcasting quarterly GDP using monthly indicators, the UMIDAS approach estimates the model at the quarterly frequency with month 1 of the monthly indicator data forming one variable,  $x_t^1$ , month 2 another,  $x_t^2$ , and month 3 the third,  $x_t^3$ . In effect in this UMIDAS model we therefore have three times as many parameters to estimate compared to a model specified at the quarterly frequency, that of course would not therefore exploit the within-quarter data as they accrue.

Lima & Menf (2018) argue that the MIDAS approach is not useful to address the parameter proliferation problem in quantile regression. This is because different quantiles maybe affected by different high-frequency predictors over time. This makes the data transformations in MIDAS too restrictive for quantile forecasting. Lima & Menf (2018) propose the use penalised quantile regressions, of the type seen in Belloni & Chernozhukov (2011), to overcome the parameter proliferation induced when with many higher-frequency indicators. However, we propose a Bayesian approach. This also offers a way to shrink the parameter space in the face of parameter proliferation and has advantages when density forecasting.

## 2.2 Bayesian estimation

Unlike classical estimation methods, Bayesian inference provides the entire posterior distribution of the parameter of interest. In addition, Bayesian methods naturally allow for parameter uncertainty to be taken into account when making nowcasting and forecasting. But until relatively recently, while popular with linear regression models, Bayesian methods had been little applied to quantile regression.

Koenker & Machado (1999) showed that likelihood-based inference using independently distributed asymmetric Laplace densities (ALD) is directly related to the traditional quantile regression minimisation problem (seen below in (4)). Yu & Moyeed (2001) showed how Bayesian inference proceeds by forming the likelihood function based on ALD.<sup>(4)</sup>

A random variable  $U$  is said to follow the ALD if its probability density is given by

$$f_{\tau}(u) = \sigma\tau(1-\tau)\exp\{-\sigma\rho_{\tau}(u)\} \quad (3)$$

where  $\sigma$  is the scale parameter and  $\rho_{\tau}(u)$  is defined as in the traditional loss function minimised in estimation of quantile regression (via classical methods, as in Koen & Bassett (1978) and Koenker (2005)):

$$\min_{(\beta_0, \beta_{1:3})} \sum_{t=1}^T \rho_{\tau}(y_{t+h} - \beta_0 - \beta_1 x_t^1 - \beta_2 x_t^2 - \beta_3 x_t^3) \quad (4)$$

where

$$\rho_{\tau}(u) = u(\tau - I(u < 0)) \quad (5)$$

$$= u(\tau I(u > 0) - (1 - \tau)I(u < 0)) \quad (6)$$

<sup>(4)</sup> Carriero et al. (2015) consider how Bayesian methods can be used to estimate (multivariate, VAR) UMIDAS models.

$$= \frac{|u| + (2\tau - 1)u}{2} \quad (7)$$

is the check loss function, and  $I(\cdot)$  denotes the indicator function.

When  $\tau=0.5$ ,  $f_\tau(u) = (1/4)\exp\{-|u|/2\}$  which is the probability density function of a standard symmetric Laplace distribution. For all other  $\tau$ ,  $f_\tau(u)$  is asymmetric.

So minimisation of (4) is equivalent to maximising a likelihood function under the ALD with  $\sigma=1$  :

$$L(\mathbf{y}|\boldsymbol{\beta}) = \tau^T (1-\tau)^T \exp \left\{ - \sum_{t=1}^T \rho_\tau(y_t - \beta_0 - \beta_1 x_t^1 - \beta_2 x_t^2 - \beta_3 x_t^3) \right\} \quad (8)$$

Priors can then be placed on the vector of  $\beta_s$ ,  $\beta$ , and Bayesian estimation can proceed. But standard conjugate prior distributions are not available for the quantile regression formulation. So analytical solutions are not available. However, Markov chain Monte Carlo (MCMC) methods can still be used to extract the posterior distributions of the unknown parameters. This, in fact, allows for the use of basically any prior distribution.

Yu & Moyeed (2001) suggest the use of improper uniform priors showing that they produce a proper joint posterior. Kozumi & Kobayashi provide an extended discussion developing a more efficient Gibbs sampling algorithm for fitting the quantile regression model based on the following location-scale mixture representation of the asymmetric Laplace distribution for  $u$  as seen in (3):

$$u = \theta z + \kappa \varepsilon \sqrt{z\sigma^{-1}} \quad (9)$$

where  $z$  is a standard exponential variable with mean  $\sigma^{-1}$ ,  $\varepsilon$  a standard normal variable (independent of  $z$ ) and

$$\theta = \frac{1-2\tau}{\tau(1-\tau)} \text{ and } \kappa^2 = \frac{2}{\tau(1-\tau)} \quad (10)$$

Use of this mixture representation, (9), simplifies estimation. This is as the quantile regression model, (2), can be rewritten as:

$$y_t = \beta_0(\tau) + \beta_1(\tau)x_t^1 + \beta_2(\tau)x_t^2 + \beta_3(\tau)x_t^3 + \theta z_t + \kappa \sqrt{z_t} \varepsilon_t \quad (11)$$

so that the joint **conditional** density of  $y_t$  is normal with mean

$$\beta_0(\tau) + \beta_1(\tau)x_t^1 + \beta_2(\tau)x_t^2 + \beta_3(\tau)x_t^3 + \theta z_t \quad \text{and variance } \kappa^2 z_t :$$

$$f(\mathbf{y}|\boldsymbol{\beta}, \mathbf{z}) \propto \left( \prod_{t=1}^T z_t^{-0.5} \right) \exp \left\{ - \sum_{t=1}^T \frac{y_t - \beta_0(\tau) - \beta_1 x_t^1 - \beta_2 x_t^2 - \beta_3 x_t^3 - \theta z_t}{2\kappa^2 z_t} \right\} \quad (12)$$

Given this specification of the likelihood prior distributions can then be defined and Bayesian estimation using a Gibbs sampler can proceed. Yu & Moyeed (2001) prove that all these posterior moments exist when the prior for  $\boldsymbol{\beta}$  is normal.

## 2.3 Lasso prior

To overcome the curse of dimensionality given that we wish to harness the information contained in many indicator variables when nowcasting, a Lasso QR is used. Lasso QR involves the following adapted minimisation:

$$\min_{(\beta_0, \beta_{1:3})} \sum_{t=1}^T \rho_{\tau}(y_t - \beta_0 - \beta_1 x_t^1 - \beta_2 x_t^2 - \beta_3 x_t^3) + \lambda \|\beta\|_1 \quad (13)$$

where  $\lambda$  is a nonnegative regularisation parameter. The new second term in (13) is a  $l_1$  penalty. As  $\lambda$  increases, the Lasso continuously shrinks QR coefficients towards zero.

Lasso, while originally developed for least squares, is a natural method to consider in QR too — as it performs both variable selection and regularisation. By forcing the absolute value of the coefficients  $\beta_j$  to be less than a fixed value it forces certain coefficients to zero. Thereby Lasso identifies a simpler model that does not include all the coefficients. The idea is similar to in ridge regression, where again the inclusion of extra coefficients is penalised. But in ridge regression while coefficients are shrunk, they are not set to zero as in Lasso.

Following Li et al. (2010), Lasso QR imposes a Laplace prior on the  $\beta_j$

$$p(\beta_j | \sigma, \lambda) = \left( \frac{\lambda}{2\sigma} \right) \exp(-\lambda |\beta_j| / \sigma)$$

which can be rewritten into a mixture of the following hierarchical priors integrating out  $\gamma_j$

$$|\beta_j| \left( \sigma, \gamma_j \right) \sim N(0, \sigma^2 \gamma_j) \quad (14)$$

$$\gamma_j | \sigma(\tau) \sim \exp(\lambda^2 / 2) \quad (15)$$

The Laplace distribution is sharply peaked at zero, compared to the Gaussian density, explaining how Lasso sets some coefficients to zero.

Following li et al. (2010), gamma priors are placed on  $\sigma^2$  and  $\left( \frac{\lambda}{\sigma} \right)^2$ , leading to a Bayesian hierarchical model. Posterior computation is then relatively simple: following li et al. (2010) one can sequentially sample from the posteriors of each unknown parameter conditional on all other parameters using a Gibbs sampling algorithm.

## 2.4 Density nowcasts

Having estimated the Bayesian Lasso quantile regression in-sample ( $t=1, \dots, T$ ), quantile nowcasts can be computed given  $x_{T+1}$ :

$$\hat{Q}_{y_{T+1}}(\tau | \Omega_{T+1}^j)^r = \hat{\beta}_0^r(\tau) + \hat{\beta}_1^r(\tau) x_{T+1}^1 + \hat{\beta}_2^r(\tau) x_{T+1}^2 + \hat{\beta}_3^r(\tau) x_{T+1}^3, \tau \in (0, 1) \quad (16)$$

where  $\hat{\beta}_k^r$  ( $k=0, \dots, 3$ ) denotes the  $r$ -th draw from the posterior parameter distribution, and  $\Omega_{T+1}^j$  denotes the  $j$ -th available information set or conditioning information. As explained in section 5.1 below, the information set,  $j$ , increases as within-quarter indicator data accrue.

Recall that the quarter  $T+1$  values of the indicator variables are published ahead of the quarter  $T+1$  values for  $y_t$  and can therefore be exploited when nowcasting. The nowcasts,  $\hat{Q}_{y_{T+1}}(\tau | \Omega_{T+1}^j)$ , can be evaluated when  $y_{T+1}$  is subsequently published.

Following Gaglianone & Lima (2012) and Korobolis (2017) we collect together  $r=1, \dots, R$  draws from the quantile forecast  $\hat{Q}_{y_{T+1}}(\tau | \Omega_{T+1}^j)$  across  $\tau \in [0.05, 0.01, \dots, 0.90, 0.95]$  and then construct the full posterior density nowcast — using a Gaussian kernel to smooth. Note Gaglianone & Lima suggest that quantile in the extreme tails, less than 0.05 and greater than 0.95, should not be modelled directly. This process delivers the density nowcast:

$$p^{QR}(y_{T+1} | \Omega_{iT+1}^j). \quad (17)$$

By publishing the whole density nowcast for GDP growth the statistics office ensures the user is free to extract from the density any feature of concern to them. This feature might be the conditional mean. But interest often focuses in tail events, which require an explicit statement about the uncertainty associated with the nowcast. Users of GDP growth forecasts may be concerned about the probability of (a one-quarter) recession i.e. of negative GDP growth; indeed the IMF now publish GDP at risk measures (see IMF2017). These probability event forecasts can readily be extracted from the density forecast; and we compute GDP at risk measures in the application below.

The trend towards forecasters publishing density forecasts is also explained by the obvious advantages they bring when communicating with the public. It reminds them that the statisticians/forecasters themselves expect the point forecasts to be 'wrong'. It also lets users assess the balance of risks associated with the nowcast.



# 3

## Real-time data

While GDP growth is published at a quarterly frequency, many indicator variables are available at a higher frequency. These comprise our dataset of indicators to help nowcast GDP growth. Following Giannone et al., we distinguish between quantitative (“hard”) and qualitative (“soft”) indicator variables, with the soft indicators typically published ahead of hard data. And we consider both EA (aggregate) and country-level (disaggregate) indicators. Examination of country-level indicator data might prove helpful if, following Hendry & Hubrich, these disaggregates contain information over and above that in aggregate indicators. Moreover, some countries publish their hard data more quickly than others, indeed more rapidly than Eurostat publishes the corresponding aggregate.

We exploit real-time (aggregate and disaggregate) data vintages. This means that our out-of-sample simulations are genuinely, rather than ‘pseudo’, real-time.

We focus on nowcasting the EA-12 GDP growth aggregate, given data availability constraints. In order to make our application realistic we need both a decent sample size and we need to use real-time data. Limited availability of country-level real-time data vintages, and lack of historical data for more recent euro area aggregates, means we restrict attention to the EA-12. In any case, the EA-12 comprises the bulk of economic activity in more recent EA aggregates, with which it also correlates highly in growth rates. So the results below are meaningful even when interest rests with the current aggregate.

### 3.1 Indicator variables: aggregate and disaggregate

We denote the soft and hard indicators,  $x_{soft,t}^m$  and  $x_{hard,t}^m$ , respectively, where  $m = 1, 2, 3$  denotes the month in quarter  $t$ . As soft indicators, we consider the Economic Sentiment Indicator (ESI) and the spread between short term and 10 year Euro interest rates (available from the ECB). The ESI, published by the European Commission, is a widely used composite indicator. It combines various information from qualitative business tendency surveys, including expectations questions, into a single cyclical confidence indicator. As hard indicators we consider real-time monthly industrial production (IP) data.

As well as considering these data at the euro area aggregate, EA-12, level, we examine them at the disaggregate (national) level for each of the twelve euro area countries. The EA-12 comprise Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal and Spain. Again real-time data (vintages) are used for these national data. We consider the three monthly releases of national IP, but use only the first release values of national GDP, rather than the three within quarter estimates produced by Eurostat for the EA; this reflects variability across the European countries in terms of their publication of within quarter GDP data. We supplement the national qualitative survey data published by the European Commission (except for Greece and Ireland where data are unavailable) with additional business survey data for Germany, from Ifo, on the business climate, situation and expectations, given it is the largest EA economy. Use of these disaggregate data considerably increases the set of indications available; and allows for the possibility that a specific data series from a given country may help explain the aggregate over and

above the aggregate information itself. Examination of hard country-level data can also prove helpful given that some countries, as discussed below, publish their hard data more quickly than others and indeed more quickly than Eurostat publishes the corresponding aggregate. These disaggregate data can therefore be exploited when nowcasting the aggregate; cf. Hendry & Hubrich.

The soft variables are published at the end of the month to which they refer. The hard variables for month  $m$  tend to be published, both for the euro area aggregate and most countries, around the middle of month  $m+2$  (i.e. at about  $t+45$  days). The Quarterly National Accounts, which include the GDP data, are also updated around the middle of each month.

As Table 1 summarises, we therefore use real-time data triangles for real GDP and industrial production, for the EA aggregate and the twelve countries, available from Eurostat's real-time (EuroIND) database. The qualitative survey data are not revised (in a significant manner at least). Models are estimated on data vintages back to 2001 with data back to at least 1991Q1. Seasonally adjusted data are used when available. It is important to use real-time data, namely data available at the time rather than the latest release, given data are revised.

**Table 1: Quarterly and monthly data used**

	Real-time	EA-12	Country-level	Seas. Adj.	Source	Sample dates back to:
<b>Quarterly GDP</b>	Yes	Yes	All 12 Countries	Yes	Eurostat	1991Q1
<b>Monthly Indicators</b>						
<b>Industrial Production</b>	Yes	Yes	All 12 Countries	Yes	Eurostat	1980M1
<b>ESI</b>	n/a	Yes	exc. EL and IE	Yes	EC	1985M1
<b>Business Climate</b>	n/a	No	Germany	Yes	Ifo	1991M1
<b>Interest rate spread</b>	n/a	Yes	No	No	ECB	1991M1

# 4

## Reconsidering Mazzi et al. (2014): a comparison of QR against density nowcast combination methods

In order to evaluate the relative performance of QR, i.e.  $p^{QR}(y_{T+1} | \Omega_{iT+1}^j)$ , when nowcasting we compare accuracy against the nowcast combination methodology of Mazzi et al. (2014). This approach has been previously applied successfully to nowcast EA GDP growth.

Mazzi et al. (2014) construct density nowcasts for EA GDP growth by taking combinations across a large number of competing ‘component’ models using the same set of indicator variables discussed above in Section 3.1. In their approach, model uncertainty, in particular uncertainty about what indicator variables should be used, is explicitly accommodated. The approach is also based on the belief that the candidate component models might all be incorrectly specified; but some might work reasonably well at some points in time. The component models might also differ in how they adapt to structural changes, including recessions. But consideration of many components allows the modeller to explore a wide range of uncertainties. The resulting combination reflects the model uncertainty by taking a weighted average across many (simple) component models, with the component models distinguished by what indicator variables they consider. The post-data weights on the components can be time-varying and reflect the relative fit of the individual model forecast densities. The combination becomes very flexible as the number of component models rises; and aims to approximate an unknown but likely complex (non-linear and non-Gaussian) data-generating-process.

### 4.1 Combination density nowcasts

Specifically, Mazzi et al. (2014) use the linear opinion pool to combine density nowcasts. We summarise their approach below; for further details of the model space and the number of components,  $N$ , to be considered see Mazzi et al. (2014). This involves producing density nowcasts from a large set of component models, which differ in terms of the indicators variables, and transformations thereof, they consider. Mazzi et al. (2014) use simple regression based component models to link GDP growth and the indicator(s) — as now explained.

## 4.2 Nowcasting component models

Mazzi et al. (2014) also take a UMIDAS approach and estimate, in principle when the full quarter's data are available, three component models for each indicator. These involve relating quarterly GDP growth,  $y_t$  to  $x_{k,t}^m$  ( $m=1,2,3; t=1,\dots,T$ ).  $x_{k,t}^m$  denotes the  $k$ -th indicator variable drawn from the information set  $\Omega_t^j$  where  $j$  ( $j=1,\dots,4$ ) denotes the first, second, third and fourth nowcast formed at  $t-30$ ,  $t-15$ ,  $t+0$  and  $t+15$  days, respectively. Each successive nowcast exploits an ever larger information set. This reflects the fact that with the passage of time more and more (aggregate and disaggregate) indicator data become available.

Specifically,  $N_j$  component models (one model for each element in  $\Omega_t^j$ ) are separately estimated of the form:

$$y_t = \beta_0 + \beta_1 x_{k,t}^m + e_t; (m=1,2,3), \quad (18)$$

where  $e_t$  is assumed to be normally distributed.

## 4.3 Combined density nowcasts

Given these  $i = 1, \dots, N_j$  component models, the combination densities for GDP growth are given by the linear opinion pool:

$$p^{lop}(y_{T+1} | \Omega_{T+1}^j) = \sum_{i=1}^{N_j} w_{i,T+1,j} g(y_{T+1} | \Omega_{i,T+1}^j), \quad (19)$$

where  $N_j$  ( $j=1,\dots,4$ ) where  $N_{j+1} > N_j$ ;  $g(y_{T+1} | \Omega_{i,T+1}^j)$ ; are the nowcast forecast densities of  $y_{T+1}$  from component model  $i$ ,  $i = 1, \dots, N_j$ , each conditional on one element (indicator/transformation) from the pooled (across  $l$ ) information set  $\Omega_{T+1}^j$ . These densities are obtained having estimated (18). The non-negative weights,  $w_{i,T+1,j}$ , in this finite mixture sum to unity. Furthermore, the weights may change with each recursion in the evaluation period.

# 5

## The trade-off between the timeliness and accuracy of nowcasts

Using both the Bayesian Lasso quantile regression and the Mazzi et al. (2014) density forecast combination approach we produce nowcasts of quarterly GDP growth for the EA-12 to four timescales: t-30, t-15, t+0, t+15.  $t$  denotes quarter, so that t-30, for example, means that a nowcast for quarter  $t$  is produced 30 days before the end of the quarter for which we want a quarterly GDP estimate. In contrast, the nowcast produced at t+15 is produced 15 days after the end of the quarter of interest.

At all four timescales we know the value of GDP in the previous quarter. But this (t-1) estimate may be measured by the first (Flash), second or third release from Eurostat. Eurostat's Flash GDP estimate for the current quarter is currently released at about t+30 days, but until 2016 was released at about t+45 days.

### 5.1 The information set as within-quarter data accrues

The information set available to both QR and the combination approach — at t-30, t-15, t+0 and t+15 days — accumulates as follows, where:

1.  $j=1$ . t-30: 30 days before the end of the quarter.

$$\Omega_t^1 = \left( \{x_{soft,t}^m\}_{m=1}^2, \{x_{hard,t-l}\}_{l=1}^{p_1}, \{y_{t-l}\}_{l=1}^{p_2} \right) \quad (20)$$

$N_l$  = no of elements of  $\Omega_t^1$ .  $p_1$  and  $p_2$  denote the number of lags of the quarterly variables  $x_{k,t}$  ( $k=hard$ ) and  $y_t$ ; and  $x_{soft,t}^m$  means the  $m=1$  (first) and  $m=2$  (second) month's soft data for quarter  $t$  are used. We do not consider lagged quarterly values of the soft data; this seems reasonable, since when the nowcasts are produced we always have in our information set at least two months of within-quarter information on the soft indicators. But, to accommodate dynamics, we do consider previous quarter information about the hard indicators and GDP growth itself, given that these variables are published at a greater lag. In particular, for previous quarter GDP growth,  $y_{t-1}$ , we consider all three EA national accounts (vintage) estimates ending with ( $T-1$ ) information; plus we consider lagged values of the GDP vintage containing the first release of GDP for data up to quarter  $T$  (since this is available at about t+45 days, it has been known for about 15 days at  $j=1$ ). Simultaneous consideration of multiple vintages means that, implicitly without modelling, the revisions process to GDP is accommodated; our density combination exercise does not simply use only the most recent vintage. In sum, (20) means  $\Omega_t^1$  includes two

months of within-quarter soft data, as well as previous quarter hard indicator data and lagged GDP data.  $\Omega_t^1$  is then related to  $y_t$ , as measured by the first release of GDP growth, via (18).

2.  $j=2$ .  $t-15$ : 15 days before the end of the quarter.

$$\Omega_t^2 = \left( \{x_{soft,t}^m\}_{m=1}^2, x_{hard,t}^1, \{x_{hard,t-l}\}_{l=1}^{p1}, \{y_{t-l}\}_{l=1}^{p2} \right) \quad (21)$$

$N_2 = no$  of elements of  $\Omega_t^2$ . This means  $\Omega_t^2$  now includes the first month of within-quarter hard data, as well as  $\Omega_t^1$ .  $\Omega_t^2$  is related to  $y_t$ , as measured by the second release of GDP growth data.

3.  $j=3$ .  $t+0$ : 0 days after the end of the quarter.

$$\Omega_t^3 = \left( \{x_{soft,t}^m\}_{m=1}^3, x_{hard,t}^1, \{x_{hard,t-l}\}_{l=1}^{p1}, \{y_{t-l}\}_{l=1}^{p2} \right) \quad (22)$$

$N_3 = no$  of elements of  $\Omega_t^3$ .  $\Omega_t^3$  now includes the final month of within-quarter soft data, as well as  $\Omega_t^2$ . In practice, given that we use the same GDP release as at  $j=2$  to measure  $y_t$ , to avoid duplicating component forecasts we do not re-consider indicators already used at  $j=2$ . This means the only new component forecast at  $j=3$  involves regressing  $x_{soft,t}^3$  on  $y_t$ .

4.  $j=4$ .  $t+15$ : 15 days after the end of the quarter.

$$\Omega_t^4 = \left( \{x_{soft,t}^m\}_{m=1}^3, \{x_{hard,t-l}^m\}_{m=1}^2, \{x_{hard,t-l}\}_{l=1}^{p1}, \{y_{t-l}\}_{l=1}^{p2} \right) \quad (23)$$

$N_4 = no$  of elements of  $\Omega_t^4$ .  $\Omega_t^4$  now includes the second month of within-quarter hard data, as well as  $\Omega_t^3$ .  $\Omega_t^4$  is related to  $y_t$ , as measured by the third release of GDP growth data.

Given these assumptions, and the availability of the aggregate and disaggregate data, as in Mazzi et al. (2014), the number of indicators considered (i.e. included in the quantile regression or considered in the density combination approach) at each of the four time horizons is:  $N_1=214$ ,  $N_2=293$ ,  $N_3=351$ ,  $N_4=430$ .

In each case ( $j=1, \dots, 4$ ), in the density combination approach each element (i.e. indicator) from  $\Omega_t^j$  is related to  $y_t$  via (18) for  $t=1, \dots, T$ . We then use this model, and its estimated coefficients from the sample  $t=1, \dots, T$ , and the quarter  $T+1$  values of the indicator variables,  $\Omega_{T+1}$ , to nowcast.

In the proposed QR approach, all  $N_j$  indicators are considered simultaneously with Lasso used, in effect, for variable selection.

# 6

## Results: nowcasting euro area GDP growth

We compare the accuracy of density nowcasts of euro area GDP growth from the two approaches (QR and density forecast combination) at the four horizons ( $j=1,\dots,4$ ) in recursive out-of-sample experiments using real-time data. The evaluation period is fixed to match Mazzi et al. (2014): 2003Q2–2010Q4 — to ensure replicability and comparison. We note that the evaluation sample starts in 2003Q2 as this is when Eurostat published its first Flash estimate for GDP growth. Models are estimated on data vintages back to 2001, leaving a little over two years as a training sample, with data used in (in-sample) estimation dating back to 1991Q1.

To be clear, our recursive simulation strategy designed to mimic real-time use of the two approaches is as follows. The first recursive estimation (for both the QR and density forecast combination approaches) involves taking data as available to the researcher at the  $j=1,\dots,4$  points relating to  $t=2001Q1$ . So at  $j=1$  ( $t-30$  days) means data are available up to the end of February 2001;  $j=2$  ( $t-15$  days) means data are available up to mid March 2001;  $j=3$  ( $t=0$  days) means data are available up to the end of March 2001 and  $j=4$  ( $t+15$  days) means data are available up to mid April 2001. The  $j=1,\dots,4$  models, based on the different information sets available, are then estimated on data from 1991Q1 to 2000Q4 with the 2001Q1 values for the known (monthly) indicators then used to produce the nowcast for GDP growth in 2001Q1. The second recursive estimation repeats this first step, but adds an additional quarter's worth of data to the information set: so the models are now estimated on data from 1991Q1 to 2001Q1 and used to nowcast 2001Q2. This process is repeated until data from 1991Q1 to 2010Q3 are used to estimate the four models then used to nowcast 2010Q4.

The nowcasts are evaluated by defining the 'outturn',  $y_\tau$ , as the first (Flash) GDP growth estimate from Eurostat. The exercise could be repeated for different definitions of the outturn, say the second or third Quarterly National Account release. But as our primary interest is in accelerating delivery of national accounts data, the first estimate is the natural benchmark.

We focus on evaluation using the logarithmic scores,  $\log S$ , of the density nowcasts from QR and the combination approach, i.e.  $p^{QR}(y_{T+1}|\Omega_{T+1}^j)$  and  $p^{lop}(y_{T+1}|\Omega_{T+1}^j)$ , against the subsequent GDP growth outturn  $y_{T+1}$ . So the average  $\log S$  over the evaluation period running from  $t = \underline{\tau}$  to  $t = \bar{\tau}$  are:

$$\log S_j^{QR} = \frac{1}{\bar{\tau} - \underline{\tau}} \sum_{t=\underline{\tau}}^{\bar{\tau}-1} \ln p^{QR}(y_{t+1}|\Omega_t^j) \quad (24)$$

$$\log S_j^{lop} = \frac{1}{\tau - \tau} \sum_{t=\tau}^{t=\tau} \ln p^{lop}(y_\tau | \Omega_\tau^j) \quad (25)$$

for  $j=1, \dots, 4$ . For benchmarking purposes, we also compare against an autoregressive (AR) density. This is an AR(1), with one lag, assuming Gaussianity that is recursively estimated over the out-of-sample window. The logarithmic scoring rule is intuitively appealing as it gives a high score to a density nowcast that provides a high probability to the value  $y_{T+1}$  that materialises. We note that alternative scoring rules to the logarithmic score might also be used, when for example interest lies in a specific region of the density; for a review of different classes of scoring rule see Gneiting & Raftery (2007). In section (6.1.1) below, we focus on left tail events as characterised by the probability of negative GDP growth.

## 6.1 Evaluating the nowcast densities

Table 2 presents, for  $j=1, \dots, 4$ , the average logarithmic scores of  $p^{QR}(y_\tau | \Omega_\tau^j)$  and  $p^{lop}(y_\tau | \Omega_\tau^j)$  over the evaluation period.

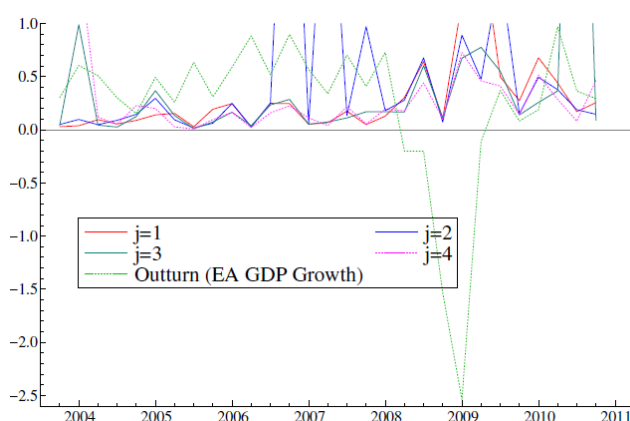
**Table 2: Average logarithmic score**

	$\log S_j^{QR}$	$\log S_j^{lop}$	$\log S_j^{AR}$
$t-30:j=1$	-1.48	-0.85	-0.84
$t-15:j=2$	-1.44	-0.80	-0.87
$t+0:j=3$	-1.37	-0.79	-0.87
$t+15:j=4$	-1.41	-0.50	-0.84

We clearly see from Table 2 that the QR densities are much less accurate than those from both the combination method (LOP) and the AR benchmark. While we see for QR a slight improvement in accuracy as time passes, the LOP dominates and clearly benefits from the arrival of additional within quarter information. While it is only as accurate as the AR benchmark at  $j=1$ , by 15 days after the end of the quarter ( $j=4$ ) it is much more accurate. QR remains much less accurate.

The poor performance of QR — we emphasise this is *on average* over our evaluation sample — raises the question, why? To help diagnose, we first plot the conditional mean nowcasts in Figure 1. Figure 1 shows that the QR approach only picked up the recession with a one quarter lag. But otherwise the point nowcasts track the outturn (EA GDP growth, in %) quite well. We also note that the density nowcasts from QR are highly non-Gaussian and often exhibit multi-modalities. Secondly, we consider now the implied probability of a recession from both  $p^{QR}(y_\tau | \Omega_\tau^j)$  and  $p^{lop}(y_\tau | \Omega_\tau^j)$

**Figure 1: Conditional Mean Nowcasts: from the Bayesian QR Approach**





### 6.1.1 GDP AT RISK: PROBABILITY OF A RECESSION

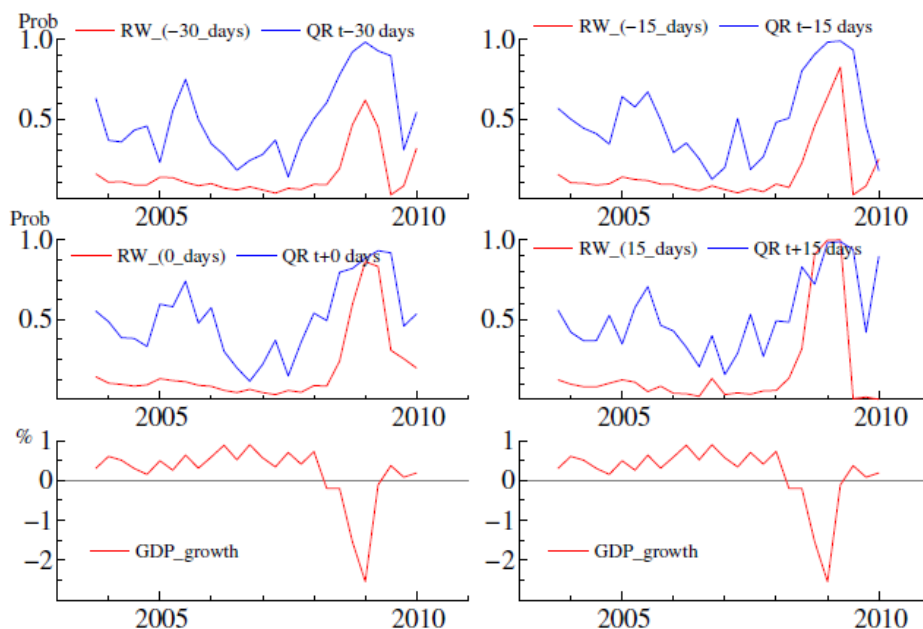
To evaluate further the accuracy of these density nowcasts we evaluate not the entire density, as in Table 2, but the probability forecast of an event of specific interest - namely *GDP at Risk*, i.e. the probability of negative GDP growth.

We rely on graphical evaluation of these probability event forecasts, rather than formal statistical tests given the relatively small sample sizes; e.g., see Clements (2004). This exploratory analysis illustrates our main findings. Specifically, Figure 2 extracts from the QR and combined density nowcasts the implied probability of a (one quarter) recession. In the bottom right panel of the figure we present the outturn, as measured by Eurostat's first (Flash) estimate, for quarterly GDP growth.

Figure 2 shows that the combination, LOP, gave about a 10 % chance to a recession from 2003 until 2008, i.e. before the crisis. But QR gave a much more volatile and higher chance, around 40 % to 50 %. QR was clearly (with the advantage of hindsight) overstating the probability of negative growth. And it also helps us understand the poor performance of QR in Table 2. The density nowcasts formed from QR are too wide (as well as too volatile in terms of the time-variation in skewness and kurtosis): hence, the high probability that growth is less than 0 %. In contrast, the density combination method produces sharper density nowcasts.

However, more encouragingly, QR is able to detect the recession in real-time. The probabilities of negative GDP growth rise to 1 in the depths of the crisis. In fact, these probabilities are higher at  $j=1$  (so  $t-30$  days) and at  $j=2$  so ( $t=15$  days) than from LOP. So this is more positive news for QR. It indicates that while not so well-calibrated overall (hence the worse averaged logarithmic score statistics in Table 2), QR can still offer value-added for regions of the density of specific interest at specific points in time. This said, QR was overstating the chance of negative growth prior to the crisis.

**Figure 2: GDP At Risk: probability of negative GDP growth from the recursive weight density combination approach of Mazzi et al. (2014) and the Bayesian QR approach**



### 6.1.2 REDUCING THE SETS OF INDICATORS IN QR: DATA SHRINKAGE VIA PRINCIPAL COMPONENTS ANALYSIS (PCA)

These mixed results, and evidence that QR is overstating uncertainty, suggest the need to impose additional parameter or data parsimony in estimation. One option is to explore the use of different priors, that allow for different types of dependencies between the indicators. Here we take up an alternative but pragmatic strategy, also designed to acknowledge that the Lasso prior may not be working well for our sample of indicators given that there is considerable dependence between them. Specifically, from the  $N_j > 200$  indicators recursively throughout the evaluation sample we select the first 30 principal components (based on the size of the associated eigenvalue) and use these in the Bayesian QR. A large number of principal components is chosen deliberately, to ensure they explain almost all of the variation in the underlying dataset. So all 30 principal components are considered simultaneously, with Lasso used to select those that best explain GDP growth in the QR. The advantage of this strategy is that by construction it orthogonalises the indicators, rendering them more appropriate for Lasso given Lasso does not accommodate cross-variable dependencies.

We denote the new PCA based QR density nowcast,  $QR,PCA$ . To isolate the effects of non-Gaussian features on these QR densities, we also fit a Gaussian density to the  $r = 1, \dots, R$  draws from the quantile forecast  $\hat{Q}_{y_{T+1}}(\tau | \Omega_{T+1}^j)$  across  $\tau \in [0.05, 0.01, \dots, 0.90, 0.95]$ . We denote this density  $QR,PCA,N$ .

**Table 3: Average logarithmic score from QR with PCA used for data shrinkage**

	$\text{Log } S_j^{QR,PCA}$	$\text{Log } S_j^{QR,PCA,N}$	$\text{Log } S_j^{QR}$	$\text{Log } S_j^{lop}$	$\text{Log } S_j^{AR}$
T-30:j= 1	-0.83	-0.61	-1.48	-0.85	-0.84
T-15:j= 2	-0.85	-0.60	-1.44	-0.80	-0.87
t+0:j= 3	-0.97	-0.80	-1.37	-0.79	-0.87
t+15:j= 4	-0.90	-0.84	-1.41	-0.50	-0.84

Table 3 considers the relative performance of these two new QR-based density nowcasts comparing them with the density nowcasts considered in Table 2. Table 3 shows that at  $j = 1$  and  $j = 2$  the new PCA-based QR nowcasts outperform clearly QR using all  $N_j > 200$  indicators: the log scores statistics rise from -1.4 to around -0.8. This improvement also now matches *lop* from the density combination approach of Mazzi et al. (2014). Interestingly, when Gaussianity is imposed on the QR densities further improvements in nowcast accuracy are seen. The log scores rise further to around -0.6.  $QR,PCA,N$  is now offering more accurate density nowcasts than *lop* from the density combination approach of Mazzi et al. (2014).

However, the QR approach does not gain in accuracy as  $j$  increases. This contrasts the density combination approach where clear gains are seen at  $j = 4$ .

# 7

## Conclusion and future research

Official quarterly GDP data are published with a delay. In order to form a view about the current state of the economy, policymakers therefore use a wide range of more timely and higher frequency indicator data to form nowcasts.

We propose a UMIDAS strategy within a Bayesian quantile regression to accommodate a large mixed frequency dataset when nowcasting. We use a shrinkage prior (the double-exponential/Laplace prior), that leads to Lasso, to avoid parameter proliferation in what becomes a large dimensional quantile regression. The modelling strategy is designed to reflect important nowcasting features, namely the use of mixed-frequency data, the ragged-edge and increasingly large numbers of indicators (*big data*).

In a real-time application to euro area GDP growth, using over 400 mixed frequency indicators, we find that the quantile regression approach does not produce as accurate density nowcasts overall as the density combination approach of Mazzi et al. (2014) unless the indicators are orthogonalised and shrunk to a smaller number prior to QR estimation. The quantile regression approach overstates the uncertainties associated with GDP growth. This is reflected by high probability estimates of *GDP at Risk* (i.e. negative GDP growth) even during the period of strong GDP growth prior to the global financial crisis. However, when the nowcasts are formed t-30 or t-15 days before the end of the quarter of interest, the quantile regression approach is better able to detect the ensuing recession than the density combination approach. But later in the quarter the density combination approach dominates. Further gains in accuracy for the quantile regression nowcasts formed at t-30 and t-15 days are also found if Gaussianity is imposed on the QR densities. Imposing symmetry may appear to be a high price to pay - to smooth out the volatilities observed over time in the QR densities - but it does improve density nowcast accuracy. It renders the QR nowcast densities more accurate than those from the density combination approach of Mazzi et al. (2014) when the nowcasts are formed 'early' at t-30 or t-15 days.

Future work should consider the use of alternative priors, within the proposed Bayesian QR framework, to assess whether these help improve the performance of QR when producing density nowcasts. Lasso is in effect a variable selection method. Priors that allow for *sets* (or groups) of variables to be selected may improve the accuracy of QR relative to the flexible density combination approach of Mazzi et al. (2014). These methods generalise Lasso by allowing for different types of dependencies between the indicators. Also given that QR, with autoregressive components, can be interpreted as the reduced form consistent with a range of models that allow for temporal instabilities and structure change (cf. Cai et al. (2000)) comparisons should be made with time-varying parameter models. While the density combination approach does allow for temporal changes in the importance of different indicators, it would be interesting to relate this time-variation to QR. One could then better understand if and when QR is expected to perform competitively.

# 8

## Bibliography

- Aastveit, K. A., Mitchell, J., Ravazzolo, F. and van Dijk, H. (2018): *The evolution of forecast density combinations in economics*. Tinbergen Institute Discussion Papers, pp. 18–069/III. Tinbergen Institute.
- Belloni, A. and Chernozhukov, V. (2011): *L-1 — penalized quantile regression in high-dimensional sparse models*. The Annals of Statistics, 39(1), pp. 82–130.
- Cai, Z., Fan, J. and Yao, Q. (2000): *Functional-coefficient regression models for nonlinear time series*. Journal of the American Statistical Association, 95(451), pp. 941–956.
- Carriero, A., Clark, T. E. and Marcellino, M. (2015): *Realtime nowcasting with a Bayesian mixed frequency model with stochastic volatility*. Journal of the Royal Statistical Society Series, A 178(4), pp. 837–862.
- Clements, M. P. (2004): *Evaluating the Bank of England density forecasts of inflation*. Economic Journal, 114, pp. 844–866.
- Clements, M. P. and Galvão, A. (2008): *Macroeconomic forecasting with mixed-frequency data*. Journal of Business and Economic Statistics, 26, pp. 546–554.
- Froni, C. and Marcellino, M. (2014): *A comparison of mixed frequency approaches for nowcasting euro area macroeconomic aggregates*. International Journal of Forecasting, 30(3), pp. 554–568.
- Froni, C., Marcellino, M. & Schumacher, C. (2015): *Unrestricted mixed data sampling (MIDAS): MIDAS regressions with unrestricted lag polynomials*. Journal of the Royal Statistical Society Series, A 178(1), pp. 57–82.
- Gaglianone, W. P. and Lima, L. R. (2012): *Constructing Density Forecasts from Quantile Regressions*. Journal of Money, Credit and Banking, 44(8), pp. 1589–1607.
- Ghysels, E. (2014): *Conditional Skewness with Quantile Regression Models: SoFIE Presidential Address and a Tribute to Hal White*. Journal of Financial Econometrics, 12(4), pp. 620–644.
- Ghysels, E., Sinko, A. and Valkanov, R. (2007): *Midas regressions: Further results and new directions*. Econometric Reviews, 26(1), pp. 53–90.
- Giannone, D., Reichlin, L. & Small, D. (2008): *Nowcasting: The real time informational content of macroeconomic data releases*. Journal of Monetary Economics, 55, pp. 665–76.
- Gneiting, T. and Raftery, A. E. (2007): *Strictly proper scoring rules, prediction, and estimation*. Journal of the American Statistical Association, 102, pp. 359–378.
- Hendry, D. F. and Hubrich, K. (2011): *Combining disaggregate forecasts or combining disaggregate information to forecast an aggregate*. Journal of Business and Economic Statistics, 29, pp. 216–227.

- IMF (2017): *Financial conditions and growth at risk*. Global financial stability report, October, chapter 3, International Monetary Fund.
- Koenker, R. (2005): *Quantile Regression*. Cambridge University Press.
- Koenker, R. and Machado, J. A. (1999), *Gmm inference when the number of moment conditions is large*. *Journal of Econometrics*, 93(2), pp. 327–344.
- Koenker, R. W. and Bassett, Gilbert, J. (1978): *Regression Quantiles*. *Econometrica*, pp. 46, 33–50.
- Koenker, R. and Xiao, Z. (2006): *Quantile autoregression*. *Journal of the American Statistical Association*, 101, pp. 980–990.
- Korobilis, D. (2017): *Quantile regression forecasts of inflation under model uncertainty*. *International Journal of Forecasting*, 33, pp. 11–20.
- Kozumi, H. & Kobayashi, G. (2011): *Gibbs sampling methods for bayesian quantile regression*. *Journal of Statistical Computation and Simulation*, 81(11), pp. 1565–1578.
- Li, Q., Xi, R. & Lin, N. (2010): *Bayesian regularized quantile regression*. *Bayesian Anal*, 5(3), pp. 533–556.
- Lima, L. R. & Menf, F. (2018): *Quantile forecasting with mixed-frequency data*. *International Journal of Forecasting*. Forthcoming.
- Mazzi, G. L., Mitchell, J. & Montana, G. (2014): *Density nowcasts and model combination: Nowcasting euro area gdp growth over the 2008–09 recession*. *Oxford Bulletin of Economics and Statistics*, 76(2), pp. 233–256.
- Yu, K. & Moyeed, R. A. (2001): *Bayesian quantile regression*. *Statistics & Probability Letters*, 54(4), pp. 437 – 447.

## **Getting in touch with the EU**

### **In person**

All over the European Union there are hundreds of Europe Direct Information Centres. You can find the address of the centre nearest you at: <https://europa.eu/contact>

### **On the phone or by e-mail**

Europe Direct is a service that answers your questions about the European Union. You can contact this service

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696 or
- by electronic mail via: <https://europa.eu/contact>

## **Finding information about the EU**

### **Online**

Information about the European Union in all the official languages of the EU is available on the Europa website at: <https://europa.eu>

### **EU Publications**

You can download or order free and priced EU publications from EU Bookshop at: <http://bookshop.europa.eu>. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see <https://europa.eu/contact>)

### **EU law and related documents**

For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex at: <https://eur-lex.europa.eu>

### **Open data from the EU**

The EU Open Data Portal (<http://data.europa.eu/euodp/en/data>) provides access to datasets from the EU. Data can be downloaded and reused for free, both for commercial and non-commercial purposes.

# New methods for timely estimates: nowcasting euro area GDP growth using quantile regression

In this paper we are discussing the possibility of using quantile regression models to provide nowcasting and short-term forecasts for macroeconomic variables. The method is then applied to produce density nowcasting of euro area and GDP comparing them with the results obtained by using a combination density approach.

---

**For more information**

<https://ec.europa.eu/eurostat/>