Acknowledgements: The research project was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the 4th Call for Action "Science and Society" - Emblematic Action – "Interventions to address the economic and social effects of the COVID-19 pandemic"

Modelling GDP growth: the Economic Impact of COVID-19 Pandemic Using Univariate Regression and Dynamic Panel Models

Ioannis Vrontos Department of Statistics Athens Univeristy of Economics and Business Joint work with J. Galakis, E. Panopoulou, and S. Vrontos

University of Crete, Septemb[er](#page-0-0) [20](#page-1-0)[24](#page-0-0)

- • Introduction and aim of the analysis
- Modelling GDP growth
	- Standard linear regression models
	- Quantile regression models
	- Panel data models
- **•** Estimation, model selection and comparison
- **•** Data, empirical analysis and findings
- **Conclusions and future research**

つくい

- The primary goal of this study is to effectively measure the impact of a severe exogenous shock, such as the COVID-19 pandemic on aggregate economic activity in Greece and other five Euro Area economies, namely Germany, France, Italy, Spain and Belgium.
- The class of linear and quantile predictive regression models is proposed for the analysis of real GDP growth, and a Bayesian approach for model selection is developed.
- Dynamic Panel models are also implemented for modeling the economic impact of COVID-19 pandemic. The proposed class of panel models allows different economic dependent variables to be affected by common observed factors or covariates, as well as by different cross-sectional specific factors, and can account for cross-section correlations/covariances of the unobserved error terms.

 $4.11 \times 4.41 \times 4.71 \times$

• The standard regression model can be used to forecast the dependent variables based on a set of predictor variables through the following equation:

$$
y_{t+1} = \alpha + \sum_{k=1}^{K} \beta_k x_{k,t} + \epsilon_{t+1},
$$
 (1)

∽≏ດ

where y_{t+1} is the dependent variable, i.e. real GDP growth at time $t+1$, $x_{k,t}$ are the predictor variables, $k = 1, \ldots, K$ at time t, and ϵ_{t+1} is the innovation process assumed to be independent and identically distributed with mean zero and variance $\sigma^2.$

• Regression-type model specifications [\(1\)](#page-3-0) suggests that the conditional mean of the predictive distribution of real GDP growth y_{t+1} given a set of K predictors $x_{1,t}, \ldots, x_{K,t}$ is equal to

$$
E(y_{t+1}|x_{1,t},\ldots,x_{K,t}) = \alpha + \sum_{k=1}^{K} \beta_k x_{k,t}
$$
 (2)

マーティ ミトマ ミト

へのへ

Estimation using: Least Squares (LS) method, Maximum Likelihood (ML) approach.

Quantile Regression model Specification-Motivation

- The primary motivation for using quantile regression models to predict real GDP growth stems from the dynamic nature of the economic time series that are asymmetrically impacted by different economic, financial, geopolitical and other random events and shocks, such as market crises and economic downturns, the COVID-19 pandemic, or the recent war in Ukraine.
- Due to country specific or to world-wide events and episodes, economic series may exhibit non-linearities, fat tails, excess kurtosis and deviations from normality. In the presence of such characteristics, the conditional mean approach may not capture the effects of different predictors to the entire distribution of the series under consideration and may provide estimates that are not robust.
- In addition, the quantile regression approach estimates the potential differential effect of a set of predictors on various quantiles in the conditional distribution, and provides a natural generalization of the standard conditional mean approach. $\left\{ \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \end{array} \right\}$, $\left\{ \begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right\}$

We also consider the predictive quantile regression model of the form:

$$
y_{t+1} = \alpha^{(\tau)} + \sum_{k=1}^{K} \beta_k^{(\tau)} x_{k,t} + \epsilon_{t+1},
$$
\n(3)

∽≏ເ

where $\tau \in (0,1)$, y_{t+1} is real GDP growth at time $t+1$, $x_{k,t}$ is the value of predictor k at time $t, \, k=1,\ldots,K, \, \alpha^{(\tau)}$ and $\beta^{(\tau)}_k$ $\mathbf{k}^{\left(7\right) }$ are the regression parameters, i.e. the intercept and the betas, associated with the τ th quantile, and ϵ_{t+1} is an unknown error term. The errors ϵ_{t+1} are assumed independent from an error distribution $g_{\tau}(\epsilon)$ with τ -quantile equal to 0, i.e. $\int_{-\infty}^{0} g_{\tau}(\epsilon) d\epsilon = \tau$.

• Quantile regression-type models [\(3\)](#page-6-0) suggest that the τ th conditional quantile of y_{t+1} given $x_{1,t}, \ldots, x_{K,t}$ is

$$
Q_{\tau}(y_{t+1}|x_{1,t},\ldots,x_{K,t}) = \alpha^{(\tau)} + \sum_{k=1}^{K} \beta_k^{(\tau)} x_{k,t}
$$
 (4)

へのへ

where the intercept and the regression coefficients depend on τ .

Bayesian approach to model comparison

- Bayesian approach to model comparison deals with the uncertainty regarding the set of predictors that should enter the linear and quantile regression models.
- **•** This is a probabilistic approach to inference that is based on the estimation of the posterior probabilities of different predictive regression models.
- **•** Estimation of the posterior model probabilities is achieved by designing a Markov chain Monte Carlo (MCMC) algorithm that visits (jumps between) a variety of regression model specifications.
- Inference can be based on the most probable model or a subset of most probable models, weighted by their (normalised) posterior model probabilities. The latter approach takes into account model uncertainty (Kass and Raftery, 1995).

 $4.49 \times 4.72 \times$

• The posterior probability of model γ given the dependent variable y and the set of predictor variables x can be computed by:

$$
p(\gamma|\mathbf{y}, \mathbf{x}) = \frac{p(\mathbf{y}|\gamma, \mathbf{x})p(\gamma)}{\sum_{\delta} p(\mathbf{y}|\delta, \mathbf{x})p(\delta)}
$$
(5)

マター・エヌ マラン

∽≏ເ

where

 $p(\mathbf{y}|\gamma,\mathbf{x})=\int p\left(\mathbf{y}|\gamma,\mathbf{x},\theta_{\gamma}\right) p\left(\theta_{\gamma}|\gamma,\mathbf{x}\right) d\theta_{\gamma}$ is the marginal likelihood of model γ $p(\mathbf{y}|\gamma, \mathbf{x}, \theta_{\gamma})$ is the likelihood given model γ $p(\theta_{\gamma}|\gamma, \mathbf{x})$ is the prior density of θ_{γ} under model γ $p(\gamma)$ is the prior probability for model γ $\gamma = (\gamma_1, \ldots, \gamma_K)'$, $\gamma_k = 1$ if the k th predictor is included in the model, and $\gamma_k = 0$, otherwise.

Prior Specification

- The predictive regression model specifications are identified by a parameter vector $\theta = (\gamma, \theta_{\gamma})$.
- It is convenient to use

$$
p(\gamma, \theta_{\gamma}) = p(\theta_{\gamma}|\gamma) p(\gamma)
$$

and specify $p(\gamma)$ and $p(\theta_{\gamma}|\gamma)$, separately.

• The prior distribution that is adopted for γ is of the form

$$
p(\gamma) = \prod_{k=1}^{K} \pi_k^{\gamma_k} (1 - \pi_k)^{1 - \gamma_k}
$$

• Under this prior, each predictor x_k , $k = 1, ..., K$ enters the predictability model independently with probability $p (\gamma_k = 1) = \pi_k$. Assigning $\pi_k = \pi = 0.5$ for all k, yields the uniform prior, $p\left(\gamma\right) =1/2^{K}$, that is often used as a representation of ignorance, i.e. it implies that the analyst is agnostic about the predictor variables that will enter the model. イロト イ押 トイモト イモト

Prior Specification

- With regard to the prior distribution for the parameters $\theta_\gamma=\big(\theta_R^\gamma,\sigma^2\big)'$ of the predictive linear regression model γ an independent conjugate prior distribution is assumed, that is, a multivariate normal $N\left(\mu , c\sigma^2 {\bf{V}} \right)$ for the vector $\theta_R = (\alpha,\beta_1,\ldots,\beta_k)'$, and an inverted Gamma $IG(d/2,\nu/2)$ prior for σ^2 .
- Choose $\mu = 0$, which reflects prior ignorance about the location of the means of the regression coefficients, $c=T$, and ${\bf V} = ({\bf F}'{\bf F})^{-1}$, where \bf{F} is the corresponding design matrix that replicates the covariance structure of the data and yields the g-prior of Zellner (1986).
- \bullet The hyperparameters d and ν are chosen in such way that the prior mean $E\left(\sigma^2\right)=\frac{v}{d-2},\,d>2$ equals the maximum likelihood estimate of σ^2 , i.e. $\hat{\sigma}^2$, and the prior variance $Var(\sigma^2) = \frac{2}{d-4} \left(\frac{\nu}{d-2}\right)^2$ equals to $100\widehat{\sigma}^2$.

K 御 ▶ K 唐 ▶ K 唐 ▶

Calculation of the marginal likelihood

- For the standard linear regression models, with parameter vector $\theta_\gamma=\left(\theta_R',\sigma^2\right)'=\left(\alpha,\beta_1,\ldots,\beta_k,\sigma^2\right)',$ and assuming a normal inverse gamma prior, the marginal likelihood $p(\mathbf{y}|\gamma, \mathbf{x})$ can be evaluated analytically, since the model parameters are integrated out. See, for example, Zellner (1971), O'Hagan and Forster (2004).
- **•** For the predictive quantile regression models, analytic evaluation of the marginal likelihood is not possible. In our analysis, the marginal likelihood is estimated using the Bayesian Information Criterion (BIC) approximation that is given by

$$
\ln \widehat{p}(\mathbf{y}|\gamma, \mathbf{x}) = \ln p\left(\mathbf{y}|\gamma, \mathbf{x}, \widehat{\theta_{\gamma}}\right) - \frac{\dim(\theta_{\gamma})}{2}\ln(T-1),
$$

where $\widehat{\theta}_{\gamma}$ denotes the ML estimate of θ_{γ} , $dim(\theta_{\gamma})$ is the dimension of θ_{γ} , and T is the sample size.

The BIC approximation is efficient, quite intuitive, less accurate, and can be used without introducing a prior density for the regression parameters θ_{γ} in the underlying quantile re[gre](#page-11-0)[ssi](#page-13-0)[on](#page-11-0) [m](#page-12-0)[od](#page-0-0)[el.](#page-57-0)

- **•** Identification of the most important predictive variables for inclusion in a model is difficult, especially in problems, where the number of potential variables is large and, as a consequence, the number of possible models can be vast.
- In such cases, it is computationally prohibitive to compute the posterior probability of each possible model.
- To address the problem, a Markov chain Monte Carlo (MCMC) algorithm, which efficiently searches over such high dimensional model spaces to detect the models with the highest posterior model probabilities, is proposed.
- This algorithm enables the analysis of high dimensionality datasets with respect to the number of predictor variables.
- The proposed algorithm enables us to implement Bayesian Model Averaging and to account, therefore, for model uncertainty.

イロメ マ桐 メメモ レマモメ

 Ω

- The Metropolis-Hastings type algorithm simulates a Markov chain sequence of models $\gamma^{(1)},\,\gamma^{(2)},\,\gamma^{(3)},... ,$ that under certain regularity conditions (see, for example, Smith and Roberts, 1993, and Tierney, 1994) converges to the equilibrium distribution $p(\gamma|\mathbf{y})$.
- **•** The Metropolis-Hastings algorithm is constructed as follows: Starting with an initial model $\gamma^{(0)}$, iteratively simulate the transitions from the current model $\gamma^{(i)}$, at the i-th iteration, to model $\gamma^{(i+1)}$ at the next iteration by using the two steps:
	- simulate a candidate model $\gamma^{'}$ from a proposal distribution $q\left(\gamma^{(i)}, \gamma^{'}\right)$
	- set $\gamma^{(i+1)}=\gamma^{'}$ with probability

$$
a\left(\gamma^{(i)},\gamma^{'}\right)=\min\left\{\frac{q\left(\gamma^{'},\gamma^{(i)}\right)}{q\left(\gamma^{(i)},\gamma^{'}\right)}\frac{p\left(\mathbf{y}|\gamma^{'}\right)}{p\left(\mathbf{y}|\gamma^{(i)}\right)}\frac{p\left(\gamma^{'}\right)}{p\left(\gamma^{(i)}\right)},1\right\}
$$

メロメ メタメ メミメ メミメ

へのへ

otherwise, set $\gamma^{(i+1)} = \gamma^{(i)}$.

- Transition kernels $q\left(\gamma^{(i)},\gamma'\right)$ used to generate candidate models γ' from $\gamma^{(i)}$ by randomly choosing among the following steps:
	- Birth: Randomly select a predictor variable from those possible (i.e. those not present in the current model) and add it in the subset to create a new proposed model with one additional variable
	- Death: Randomly select a predictor from those present in the current model and delete it from the subset to create a new proposed model with one less variable
	- Change: Randomly select a predictor variable from those present in the current model and change it with a new one from the remaining variables
- These steps permit the algorithm to move efficiently through models of the same (through Change step), or different (through Birth and Death steps) dimensionality, i.e. number of predictors, in [or](#page-14-0)der to generate a sample from the posterior [dis](#page-16-0)[tr](#page-14-0)[ib](#page-15-0)[u](#page-16-0)[tio](#page-0-0)[n o](#page-57-0)[f](#page-0-0) γ .

- The proposed Bayesian stochastic search algorithm can be used for the identification of the most important predictors for real GDP growth in the quantile regression models.
- This approach is based on, and extends, the algorithms of Vrontos, Vrontos, and Giamouridis (2008), Giannikis and Vrontos (2011), and Vrontos (2012), who used Bayesian model selection techniques to identify important risk factors and predictor variables in univariate and multivariate regression models with GARCH-type models.
- The proposed method is also based on, and extends, the approach of Dellaportas and Vrontos (2007) and Galakis, Vrontos and Xidonas (2022), who used Bayesian model selection techniques to identify the most probable tree topologies (non-linear thresholds) for tree-structured multivariate GARCH models, and tree-structured quantile regression models, respectively, assuming a fixed (known) number of regressors in the respective model specifications.
- As a result, the proposed predictive quantile regression model and Bayesian approach to inference allows for a[uto](#page-15-0)[ma](#page-17-0)[t](#page-15-0)[ic](#page-16-0) [m](#page-17-0)[od](#page-0-0)[el](#page-57-0) [se](#page-0-0)[lec](#page-57-0)[tio](#page-0-0)[n](#page-57-0) Ω [Modelling GDP growth](#page-0-0)
- Dynamic panel data models have been extensively used in financial and economics studies. The specific class of models possesses interesting characteristics, due to its dynamic structure that allows for lagged dependent variables in the right-hand side of the model equation, as well as the panel structure that allows the modeling of both cross-sectional and time series dynamics and relations.
- The following panel data models are considered:

$$
y_{i,t} = \delta_i + \sum_{p=1}^{P} \alpha_{p,i} f_{p,t} + \sum_{q=1}^{Q} \gamma_{q,i} f_{q,i,t} + \epsilon_{i,t}
$$
 (6)

$$
y_{i,t} = \delta_i + \sum_{p=1}^{P} \alpha_{p,i} f_{p,t} + \sum_{q=1}^{Q} \gamma_{q,i} f_{q,i,t} + \epsilon_{i,t}, \text{ where } \epsilon_{i,t} = \phi \epsilon_{i,t-1} + u_{i,t},
$$
\n(7)

The above model has the following nonlinear representation:

$$
y_{i,t} = \delta_i + \phi(y_{i,t-1} - \delta_i) + \sum_{p=1}^P \alpha_{p,i}(f_{p,t} - \phi f_{p,t-1}) + \sum_{q=1}^Q \gamma_{q,i}(f_{q,i,t} - \phi f_{q,i,t-1}) + u_{i,t}
$$

or

$$
y_{i,t} = \phi y_{i,t-1} + \delta_i (1-\phi) + \sum_{p=1}^P \alpha_{p,i} (f_{p,t} - \phi f_{p,t-1}) + \sum_{q=1}^Q \gamma_{q,i} (f_{q,i,t} - \phi f_{q,i,t-1}) + u_{i,t}
$$

or

$$
y_{i,t}^* = \delta_i (1 - \phi) + \sum_{p=1}^P \alpha_{p,i} f_{p,t}^* + \sum_{q=1}^Q \gamma_{q,i} f_{q,i,t}^* + u_{i,t}
$$

by denoting $y_{i,t}^* = y_{i,t} - \phi y_{i,t-1}$, $f_{p,t}^* = f_{p,t} - \phi f_{p,t-1}$, and $f_{q,i,t}^* = f_{q,i,t} - \phi f_{q,i,t-1}.$

 $4.171 \pm$

 $2Q$

Panel Data Models

- A useful parametrization of the specific dynamic panel model can be derived as follows: Let $y^*_t = \left(y^*_{1,t}, \ldots, y^*_{N,t} \right)'$ be a $N \times 1$ vector of observations for the economic dependent variables (conditional on ϕ) on N cross-section units at time t, $t = 2, \ldots, T$. Suppose also that there are $1 + P + NQ$ available covariates that can be used to model each of the N cross-sectional dependent variables, and denote by $z_t = (z_{1,t}, z_{2,t}, \ldots, z_{1+P+NQ,t})' =$ $(1-\phi, f_{1,t}^*, \ldots, f_{P,t}^*, f_{1,1,t}^*, \ldots, f_{Q,1,t}^*, f_{1,2,t}^*, \ldots, f_{Q,2,t}^*, \ldots, f_{1,N,t}^*, \ldots, f_{Q,N}^*$ the $(1 + P + NQ) \times 1$ vector of covariates at time t, $t = 2, \ldots, T$.
- Then, y_t^* can be modeled as follows:

$$
y_t^* = (I_N \otimes z_t') S\beta + u_t, u_t \sim N_N(0, \Sigma)
$$

or

$$
y_t^* = x_t'\beta + u_t, \ u_t \sim N_N(0, \Sigma)
$$

The dynamic panel model can be written in a matrix form of a Seemingly Unrelated Regression (SUR) model as follows:

$$
Y^* = X\beta + U, \ U \sim N_{N(T-1)}(0, I_{T-1} \otimes \Sigma)
$$

where $Y^{*}=\left(y_{2}^{*'},\ldots,y_{T}^{*'}\right)'$ is the $N(T-1)\times 1$ vector of realizations of the dependent variables, $X = (x_2, \ldots, x_T)^{\prime}$ is the $N(T - 1) \times N(1 + P + Q)$ matrix with covariates, and $U = (u_2', \ldots, u_T')'$ is the $N(T-1) \times 1$ vector of innovations.

∽≏ເ

- The proposed dynamic panel model is general, in the sense that it allows common covariates and specific cross-sectional unit factors to affect the dependent variables
- It allows for cross-sectional dependence of the innovation process, and in this way, it may capture unobserved random effects (e.g pandemic effects), and provides a natural link/connection to the system of dependent variables.

Under the assumption of the multivariate normal error distribution, the likelihood function for the dynamic panel model can be written as:

 $L(Y^*|X,\beta,\Sigma,\phi) =$

$$
= (2\pi)^{-(T-1)N/2} |\Sigma|^{-(T-1)/2} \exp \left\{ -\frac{1}{2} \sum_{t=2}^{T} \left[(y_t^* - x_t'\beta)' \Sigma^{-1} (y_t^* - x_t'\beta) \right] \right\}
$$

$$
= (2\pi)^{-(T-1)N/2} |\Sigma|^{-(T-1)/2} \exp \left\{ -\frac{1}{2} tr \left[\Sigma^{-1} \sum_{t=2}^{T} (y_t^* - x_t'\beta) (y_t^* - x_t'\beta)' \right] \right\}
$$

$$
= (2\pi)^{-(T-1)N/2} |\Sigma|^{-(T-1)/2} \exp \left\{ -\frac{1}{2} (Y^* - X\beta)' (I_{T-1} \otimes \Sigma^{-1}) (Y^* - X\beta) \right\}.
$$

 $2Q$

オター・エヌ オライ

Bayesian Inference - Prior Specification

- The Bayesian approach to inference requires specifying prior distributions on the model parameters $\theta = (\beta, \Sigma, \phi)'$.
- The prior specification considered is the following. Independent prior distributions on β , Σ , and ϕ , i.e. $\pi(\beta, \Sigma, \phi) = \pi(\beta)\pi(\Sigma)\pi(\phi)$ are assumed.
- \bullet A multivariate normal prior distribution for the parameter vector β is assumed:

$$
\pi(\beta) = (2\pi)^{-\frac{N(1+P+Q)}{2}} |A_0|^{-\frac{1}{2}} \exp \left\{-\frac{1}{2} (\beta - B_0)' A_0^{-1} (\beta - B_0) \right\}.
$$

• For Σ an inverted Wishart prior distribution is adopted:

$$
\pi(\Sigma) = \left[2^{vN/2} \pi^{N(N-1)/4} \prod_{i=1}^{N} \Gamma\left(\frac{v+1-i}{2}\right) |Q|^{-v/2} \right]^{-1} |\Sigma|^{-(v+N+1)/2} \exp\left[-\frac{1}{2} tr\left(\Sigma^{-1} Q \right) \right]
$$

へのへ

An appropriate prior distribution on the autoregressive coefficient with support in the stationary region, e.g. a Uniform prior $\phi \sim U(-1,1)$ is considered. $4.11 \times 4.41 \times 4.71 \times$

Bayesian Inference - MCMC Sampling Scheme

- The MCMC sampling scheme is constructed by iteratively and successively sampling Σ , β , and ϕ from their full conditional posterior distributions via two Gibbs and one Metropolis-Hastings step.
- The conditional posterior distributions of the model parameters are the following:

$$
\pi(\Sigma | Y, X, \beta, \phi) \equiv IW\left((T-1) + v, \sum_{t=2}^{T} (y_t^* - x_t'\beta) (y_t^* - x_t'\beta)' + Q\right),
$$

$$
\pi (\beta | Y, X, \Sigma, \phi) \equiv N_{N(1+p+q)} (\beta_1, A_1),
$$

$$
A_1 = (A_0^{-1} + X' (I_{(T-1)} \otimes \Sigma^{-1}) X)^{-1}, \beta_1 = (\beta'_0 A_0^{-1} + Y^{*'} (I_{(T-1)} \otimes \Sigma^{-1}) X) A_1
$$

$$
\pi(\phi \mid Y, X, \Sigma, \beta) \propto \exp \left\{-\frac{1}{2} \sum_{t=2}^{T} \left(y_t^* - x_t^{\prime} \beta\right)^{\prime} \Sigma^{-1} \left(y_t^* - x_t^{\prime} \beta\right)\right\} \pi(\phi).
$$

Modeling and forecasting GDP growth

- The research focuses on measuring the impact of the COVID-19 pandemic primarily on the economic activity of Greece, as well as five Euro-zone countries for comparison reasons; Germany, France, Italy, Spain and Belgium, i.e. include economies of different size and level of significance in the analysis.
- A country's economic activity is represented by the quarterly change of its underlying Real GDP (RGDP) growth. RGDP growth is measured by the difference of the natural logarithm of each country's chain linked index, for the period between the first quarter of 2001 and the third quarter of 2021.
- As it is well documented in the literature that there is sizeable autocorrelation in the underlying RGDP growth series, lagged values of the individual RGDP series are employed in the models $(RGDP_{t-1}, RGDP_{t-2}, RGDP_{t-3}, RGDP_{t-4})$ to address the high autocorrelation issue, as well as to extract their information content for projected estimates.

K ロト K 御 ト K ヨ ト K ヨ

- The first set of explanatory and/or predictor variables used in the regression framework are perceived to be common for all countries, as they are supposed to capture global effects, given the universal nature of the COVID-19 pandemic. The set of common factors included three variables, the oil price, an uncertainty index (World uncertainty index), and a COVID-19 related factor (World Pandemic Index).
- The rationale is to employ factors that reflect the economic effects of a significant exogenous impact, such as the outbreak of the COVID-19 pandemic (Chudik et al., 2021).
- Even though, the pandemic had asymmetric effects on different continents-regions of the world, a global measure could provide significant insights regarding the scale and spread of the pandemic and its likely impact on the world economy.

K ロ ト K 倒 ト K ミ ト

- Moving on to the country specific factors, the forecasting variables are representative of categories related to output and productivity, the labour market, the housing market, orders and inventories, money and credit, interest rates, prices, the financial markets, as well as business and consumer confidence surveys. Other financial market and economic activity related indicators, such as the change in passenger car registrations and real productivity growth are included in the analysis.
- The data is obtained from several economic and financial databases and cover the period between the first quarter of 2001 and the third quarter of 2021 (83 quarterly observations). More specifically, the Federal Reserve Bank of St. Louis' FRED database, the Eurostat database, the OECD Statistics database, and the European Central Bank's database were the primary data sources.

Table 1: Set of Predictor Variables - Overview Code Predictor Variables Transformation

The table reports detailed information about the set of predictor variables and the corresponding transformation used in the analysis; Δln denotes first differences of logarithms, and Δln denotes first differences.

メロメ メタメ メミメ メミメ

 $2Q$

Linear Regression Model: Posterior model probabilities

1-4: lagged real GDP Growth $(RGDP_{t-1}, RGDP_{t-2}, RGDP_{t-3}, RGDP_{t-4}),$ 5: WTI oil price (OIL), 6: World Uncertainty Index (WUI), 7: World Pandemic Uncertainty Index (WPI), 8: real productivity growth (RPROD), 9: car registrations (CREG), 10: rate of unemployment (UNEM), 11: growth in the Consumer Price Index (CPI), 12: growth in the Producer Price Index (PPI), 13: Construction Volume Index of Production (CONPROD), 14: long-run interest rates (LONGR), 15: stock index return (STOCK), 16: [OE](#page-28-0)C[D L](#page-30-0)[ea](#page-28-0)[din](#page-29-0)[g I](#page-30-0)[ndic](#page-0-0)[ator](#page-57-0) [\(L](#page-0-0)[EAD](#page-57-0)[\).](#page-0-0)
 $\Box \rightarrow \Box \rightarrow \Box \rightarrow \Box \rightarrow \Box$

 Ω

1-4: lagged real GDP Growth ($RGDP_{t-1}$, $RGDP_{t-2}$, $RGDP_{t-3}$, $RGDP_{t-4}$), 5: WTI oil price (OIL), 6: World Uncertainty Index (WUI), 7: World Pandemic Uncertainty Index (WPI), 8: real productivity growth (RPROD), 9: car registrations (CREG), 10: rate of unemployment (UNEM), 11: growth in the Consumer Price Index (CPI), 12: growth in the Producer Price Index (PPI), 13: Construction Volume Index of Production (CONPROD), 14: long-run interest rates (LONGR), 15: stock index return (STOCK), 16: OECD Leading Indicator (LEAD).

K ロ ▶ K 御 ▶ K 君 ▶ K 君 ▶

重

 $2Q$

1-4: lagged real GDP Growth (RGDP_{t-1}, RGDP_{t-2}, RGDP_{t-3}, RGDP_{t-4}), 5: WTI oil price (OIL), 6: World Uncertainty Index (WUI), 7: World Pandemic Uncertainty Index (WPI), 8: real productivity growth (RPROD), 9: car registrations (CREG), 10: rate of unemployment (UNEM), 11: growth in the Consumer Price Index (CPI), 12: growth in the Producer Price Index (PPI), 13: Construction Volume Index of Production (CONPROD), 14: long-run interest rates (LONGR), 15: stock index return (STOCK), 16: OECD Leading Indicator (LEAD).

イロメ イ御 トイヨ トイヨメー

 $2Q$

4: lagged real GDP Growth (RGDP_{t−1}, RGDP_{t−2}, RGDP_{t−3}, RGDP_{t−4}), 5: WTI oil price (OIL), 6: World Uncertainty Index (WUI), 7: World Pandemic Uncertainty Index (WPI), 8: real productivity growth (RPROD), 9: car registrations (CREG), 10: rate of unemployment (UNEM), 11: growth in the Consumer Price Index (CPI), 12: growth in the Producer Price Index (PPI), 13: Construction Volume Index of Production (CONPROD), 14: long-run interest rates (LONGR), 15: stock index return (STOCK), 16: OECD Leading Indicator (LEAD).

メロメ メ団 メメミメメミメ

4: lagged real GDP Growth (RGDP_{t−1}, RGDP_{t−3}, RGDP_{t−3}, RGDP_{t−4}), 5: WTI oil price (OIL), 6: World Uncertainty Index (WUI), 7: World Pandemic Uncertainty Index (WPI), 8: real productivity growth (RPROD), 9: car registrations (CREG), 10: rate of unemployment (UNEM), 11: growth in the Consumer Price Index (CPI), 12: growth in the Producer Price Index (PPI), 13: Construction Volume Index of Production (CONPROD), 14: long-run interest rates (LONGR), 15: stock index return (STOCK), 16: OECD Leading Indicator (LEAD).

メロメ メ団 メメミメメミメ

1-4: lagged real GDP Growth (RGDP_{t−1}, RGDP_{t−2}, RGDP_{t−3}, RGDP_{t−4}), 5: WTI oil price (OIL), 6: World Uncertainty Index (WUI), 7: World Pandemic Uncertainty Index (WPI), 8: real productivity growth (RPROD), 9: car registrations (CREG), 10: rate of unemployment (UNEM), 11: growth in the Consumer Price Index (CPI), 12: growth in the Producer Price Index (PPI), 13: Construction Volume Index of Production (CONPROD), 14: long-run interest rates (LONGR), 15: stock index return (STOCK), 16: OECD Leading Indicator (LEAD).

メロト メタト メミト メミト

1-4: lagged real GDP Growth (RGDP_{t−1}, RGDP_{t−2}, RGDP_{t−3}, RGDP_{t−4}), 5: WTI oil price (OIL), 6: World Uncertainty Index (WUI), 7: World Pandemic Uncertainty Index (WPI), 8: real productivity growth (RPROD), 9: car registrations (CREG), 10: rate of unemployment (UNEM), 11: growth in the Consumer Price Index (CPI), 12: growth in the Producer Price Index (PPI), 13: Construction Volume Index of Production (CONPROD), 14: long-run interest rates (LONGR), 15: stock index return (STOCK), 16: OECD Leading Indicator (LEAD).

メロト メタト メミト メミト

1-4: lagged real GDP Growth (RGDP_{t−1}, RGDP_{t−2}, RGDP_{t−3}, RGDP_{t−4}), 5: WTI oil price (OIL), 6: World Uncertainty Index (WUI), 7: World Pandemic Uncertainty Index (WPI), 8: real productivity growth (RPROD), 9: car registrations (CREG), 10: rate of unemployment (UNEM), 11: growth in the Consumer Price Index (CPI), 12: growth in the Producer Price Index (PPI), 13: Construction Volume Index of Production (CONPROD), 14: long-run interest rates (LONGR), 15: stock index return (STOCK), 16: OECD Leading Indicator (LEAD).

メロト メタト メミト メミト

Ioannis Vrontos, AUEB [Modelling GDP growth](#page-0-0)

 $2Q$

Ioannis Vrontos, AUEB [Modelling GDP growth](#page-0-0)

 $2Q$

 299

The table reports posterior means and standard deviations (in parentheses) of the parameters of the common predictor variables in the Dynamic Panel data model $y_{i,t}=\delta_i+\sum_{p=1}^P\alpha_{p,i}f_{p,t}+\epsilon_{i,t}$, where $\epsilon_{i,t}=\phi\epsilon_{i,t-1}+u_{i,t}$ or $Y^*=X\beta + U,\; U\sim N_{N(T-1)}\left(0,I_{T-1}\otimes\Sigma\right)$ of the real GDP growth for the analysed countries based on the sample period from 2001:Q2 to 2021:Q3. The common predictor variables are: the WTI oil price (OIL), the World Uncertainty Index (WUI), and the World Pandemic Uncertainty Index (WPI).

メロト メタト メミト メミト

Panel model: common covariates with cross-sectional dependence - Estimates of the covariance matrix Σ

The table reports posterior means and standard deviations (in parentheses) of the elements of the covariance matrix Σ of the residuals of the Dynamic Panel data model $y_{i,t}=\delta_i+\sum_{p=1}^P\alpha_{p,i}f_{p,t}+\epsilon_{i,t},$ where $\epsilon_{i,t}=\phi\epsilon_{i,t-1}+u_{i,t}$ or $Y^*=X\beta + U,\; U\sim N_{N(T-1)}\;(0,I_{T-1}\otimes\Sigma)$ of the real GDP growth of the analysed countries based on the sample period from 2001:Q2 to 2021:Q3.

イロト イ押 トイラト

Figure 1: Convergence diagrams of the posterior sample of the parameters of the Dynamic Panel model with common covariates and autoregressive dynamics, $y_{i,t} = \delta_i + sum_{p=1}^P \alpha_{p,i} f_{p,t} + \epsilon_{i,t},$ where $\epsilon_{i,t} = \phi \epsilon_{i,t-1} + u_{i,t}$ of the real GDP growth for the analysed countries based on the sample period from 2001:Q2 to 2021:Q3.

 $2Q$

メイヨメ

Figure 2: Histograms of the posterior sample of the parameters of the Dynamic Panel model with common covariates and autoregressive dynamics, $y_{i,t} = \delta_i + \sum_{p=1}^{P} \alpha_{p,i} f_{p,t} + \epsilon_{i,t}$, where $\epsilon_{i,t} = \phi \epsilon_{i,t-1} + u_{i,t}$ of the real GDP growth for the analysed countries based on the sample period from 2001:Q2 to 2021:Q3.

∢ ⊔ ∤ ਤਜ਼ਾ ⊁

つくい

Figure 3: Convergence diagram and histogram of the posterior sample of the common ϕ autoregressive parameter of the Dynamic Panel model with common covariates and autoregressive dynamics,

 $y_{i,t}=\delta_i+\sum_{p=1}^P\alpha_{p,i}f_{p,t}+\epsilon_{i,t},\,$ where $\epsilon_{i,t}=\phi\epsilon_{i,t-1}+u_{i,t}$ of the real GDP growth for the analysed countries based on the sample period from 2001:Q2 to 2021:Q3. (a) convergence diagram of the common ϕ autoregressive parameter, (b) histogram of the common ϕ autoregressive parameter.

Figure 4: Convergence diagrams and histograms of the posterior sample of variances and covariances of the Dynamic Panel model with common covariates and autoregressive dynamics,

 $y_{i,t}=\delta_i+\sum_{p=1}^P\alpha_{p,i}f_{p,t}+\epsilon_{i,t},\,$ where $\epsilon_{i,t}=\phi\epsilon_{i,t-1}+u_{i,t}$ of the real GDP growth for the analysed countries based on the sample period from 2001:Q2 to 2021:Q3.

Ioannis Vrontos, AUEB [Modelling GDP growth](#page-0-0)

 Ω

化重氮化医氮

Panel model: common and country specific covariates with cross-sectional dependence of the errors

 $2Q$

Panel model: common and country specific covariates with cross-sectional dependence of the errors (cont'ed)

 Ω

The table reports estimates of the correlation coefficients and of the elements (variances and covariances) of the

 $y_{i,t} = \delta_i + \sum_{p=1}^P \alpha_{p,i} f_{p,t} + \sum_{q=1}^Q \gamma_{q,i} f_{q,i,t} + u_{i,t}$ or the equivalent Seemingly Unrelated Regressions (SUR) representation $Y = X\beta + U$, $U \sim N_{NT} (0, I_T \otimes \Sigma)$ with common and country specific predictor variables, and cross-sectional dependence in the error processes of the real GDP growth for the analysed countries based on the sample period from 2001:Q2 to 2021:Q3.

イロメ マ桐 メメモ レマモメ

- **•** Our analysis confirms that the outbreak of the pandemic had a profound negative effect on the economies under study, and reveals that different predictor variables are able to explain different quantiles of the underlying real GDP growth distribution for the six Euro Area countries.
- The impact of the COVID-19 pandemic was not symmetric across the Euro Area economies under study.
- The quantile regression models reveal that there are several predictors that affect the lower or/and the upper quantiles of each country's real GDP growth distribution, and, thus, uncover interesting relations of real GDP growth with the predictor variables, suggesting that this modelling approach improves the ability to adequately explain real GDP series compared to the standard conditional mean approach.

 $4.11 \times 4.49 \times 4.72 \times$

∽≏ເ

Empirical Findings

- The proposed dynamic panel data model with common and specific covariates that allows for cross-sectional dependence of the error processes is introduced to investigate the impact of several macroeconomic and financial covariates, and that of the COVID-19 pandemic in particular, on the real economic activity of six Euro Area countries.
- The findings of the underlying analysis reveal that the exogenous shock caused by the pandemic has a significant negative effect on the real GDP growth of all the countries under study. The factors that proxy the state and spread of the pandemic show up in all estimated models.
- (Future work) Model comparison/selection for Dynamic panel data model with common and country specific predictor variables, and cross-sectional dependence in the error processes
- (Future work) Develop tree-structured linear and quantile regression models to account for possible non-lineariti[es](#page-49-0)

 Ω

Future Work: Tree-structured Regression models

- **•** Tree-structured regression models can be constructed based on Denison, Mallick and Smith (1998)
- \bullet A tree T has a root node whose descendant nodes can be divided into internal or split nodes, denoted by s and terminal nodes, denoted by tn .
- The proposed model is based on a binary tree, that can be uniquely defined by the splitting nodes, the predictor variables $X_{j,t}$, $j = 1, ..., N$, $t = 1, ..., T$ on which the nodes are split, and the rules with which these splits are made. These variables can be defined as $s_i, s_i^{var}, s_i^{rule}$ $(i = 1, ..., s_{\text{max}})$, respectively.
- **In our analysis, we are interested in exploring the dynamics of** financial returns under negative and/or positive returns of the explanatory factors under consideration. We consider, therefore, that the splitting rules are zero, i.e. $s_i^{rule}=0$.

 $(1 - 1)$ $(1 - 1)$ $(1 - 1)$ $(1 - 1)$ $(1 - 1)$ $(1 - 1)$ $(1 - 1)$ $(1 - 1)$

つくい

Tree-structured Regression models

- To uniquely define the split nodes, we follow the labeling of Denison, Mallick and Smith (1998). The root node that is always present in the model, is the first split node and its position is labelled as position 1, so that $s_1 = 1$. Any descendant (child) splitting node's position s_i is uniquely defined given its parent position s_i^{parent} by setting $s_i = 2s_i^{parent}$, if the node is the left child, and $s_i = 2s_i^{parent} + 1$, if the node is the right child.
- The positions of the terminal nodes are defined similarly and are fully determined by the tree structure given by the split nodes. Note that the total number of terminal nodes is $k = s_{\text{max}} + 1$, where s_{max} is the maximum number of internal nodes, while the maximum number of terminal nodes is 2^N .

 4 \overline{m} \rightarrow 4 \overline{m} \rightarrow

- The number of terminal nodes determines the complexity of the underlying tree structure. A change in the number of terminal nodes k leads to a change in the topology of the tree and the dimension of the model.
- **•** Inference is carried out assuming that the true model is unknown. The parameter vector of a tree T with k terminal nodes (and $k-1$ internal nodes) is denoted by $\theta_T^{(k)} = \left(s_1, s_1^{var}, ..., s_{k-1}, s_{k-1}^{var} \right)$, considering $s_i^{rule} = 0$ for $i = 1, ..., k - 1$. The tree structure is determined uniquely, therefore, by the number of terminal nodes k and the parameter vector $\theta_T^{(k)}$ $T^{(k)}$ of tree T (with k terminal nodes).

∽≏ເ

Tree-structured Regression models

- \bullet We will denote a specific model-tree T with k terminal nodes and tree parameters $\theta^{(k)}_T$ $T^{(k)}$ by $T_{k,\theta}$.
- Bayesian inference about k and $\theta_T^{(k)}$ $T^{(\kappa)}_{T}$ is based on the posterior distribution $p\left(k,\theta_T^{(k)}|R,X\right)$ that can be written as

$$
p\left(k,\theta_T^{(k)}|R,X\right)=p\left(k|R,X\right)p\left(\theta_T^{(k)}|k,R,X\right),
$$

where R , X are the observed asset returns and common predictors, respectively. We generate a sample from the posterior distibution of $\left(k,\theta^{(k)}_T\right)$ using a Markov chain Monte Carlo algorithm that visits (jumps between) models-trees of different dimensionality.

The vector of returns can be modeled using a tree-structured regression model of the form

$$
R_{i,t+1} = \sum_{j=1}^{k} \left[a_j +_{i=1}^{N} \beta_{ij} X_{i,t} + \sigma_j z_t \right] I \left(\mathbf{X}_t \in P_j \right) \tag{8}
$$

Under the assumption of a normal distribution for the error term, the parameter vector of the tree-structured linear regression model [\(8\)](#page-55-0) can be estimated by maximizing the log-likelihood:

$$
\ln p(R|\theta, X) = -\frac{T-1}{2} \ln (2\pi) - \frac{1}{2} \sum_{t=1}^{T-1} \ln \left[\sum_{j=1}^{k} \sigma_j^2 I(\mathbf{X}_t \in P_j) \right]
$$

$$
-\frac{1}{2} \sum_{t=1}^{T-1} \frac{\left[(R_{i,t+1}) - \sum_{j=1}^{k} \left(a_j + \sum_{i=1}^{N} \beta_{ij} X_{i,t} \right) I(\mathbf{X}_t \in P_j) \right]^2}{\sum_{j=1}^{k} \sigma_j^2 I(\mathbf{X}_t \in P_j)}
$$

Tree-structured quantile Regression models

The tree-structured quantile predictive regression specification that models the τ conditional quantile of $R_{i,t+1}$ given $X_{1t},..., X_{Nt}$ is the following

$$
Q_{R_{i,t+1}}(\tau|X_{jt}) = \sum_{j=1}^{k} \left[\alpha_j^{(\tau)} + \sum_{i=1}^{N} \beta_{ij}^{(\tau)} X_{i,t} \right] I(\mathbf{X}_t \in P_j)
$$
(9)

Under the assumption of an asymmetric Laplace distribution for the error term, the parameter vector of the tree-structured quantile regression model [\(9\)](#page-56-0) can be estimated by maximizing the log-likelihood given by

$$
\ln p^{(\tau)}\left(R|\theta^{(\tau)},X\right) = (T-1)\left[\ln\left(\tau(1-\tau)\right)\right] - \sum_{t=1}^{T-1} \ln\left[\sum_{j=1}^{k} \sigma_j^{(\tau)} I\left(\mathbf{X}_t \in P_j\right)\right]
$$

$$
-\sum_{t=1}^{T-1} \left\{\sum_{j=1}^{k} \frac{1}{\sigma_j^{(\tau)}} \rho_{\tau}\left[(R_{i,t+1}) - \left(\alpha_j^{(\tau)} + \sum_{i=1}^{N} \beta_{ij}^{(\tau)} X_{i,t}\right)\right] I\left(\mathbf{X}_t \in P_j\right)\right\}
$$

つくい

THANK YOU

Ioannis Vrontos, AUEB [Modelling GDP growth](#page-0-0)

メロトメ 御 メメ ミトメ きょうき

 299