Annals of Operations Research
https://doi.org/10.1007/s10479-024-06378-7

ORIGINAL RESEARCH

Bayesian dynamic quantile model averaging

Mauro Bernardi[1] · **Roberto Casarin**[2] · **Bertrand B. Maillet**[2,3,4] · **Lea Petrella**[5]

Received: 10 June 2024 / Accepted: 28 October 2024

Abstract
This article introduces a novel dynamic framework to Bayesian model averaging for time-varying parameter quantile regressions. By employing sequential Markov chain Monte Carlo, we combine empirical estimates derived from dynamically chosen quantile regressions, thereby facilitating a comprehensive understanding of the quantile model instabilities. The effectiveness of our methodology is initially validated through the examination of simulated datasets and, subsequently, by two applications to the US inflation rates and to the US real estate market. Our empirical findings suggest that a more intricate and nuanced analysis is needed when examining different sub-period regimes, since the determinants of inflation and real estate prices are clearly shown to be time-varying. In conclusion, we suggest that our proposed approach could offer valuable insights to aid decision making in a rapidly changing environment.

Keywords Bayesian model averaging · Dynamic model averaging · Markov chain Monte Carlo · Quantile regression · Finance · Forecasting · Inflation

1 Introduction

The study of extremes, risks and quantiles of phenomena is a broad area of interest in Operations Research, since large movements of variables are of great importance in many fields related to natural disasters, namely geophysics (e.g. earth-quakes prediction Pisarenko and Lyubushin, 1997), hydrology and environmental studies (such as floods, tsunami, severe

✉ Roberto Casarin
r.casarin@unive.it

Mauro Bernardi
mauro.bernardi@unipd.it

Bertrand B. Maillet
bertrand.maillet@univ-reunion.fr

Lea Petrella
lea.petrella@uniroma1.it

1 University of Padova, Padua, Italy

2 Ca' Foscari University of Venice, Venice, Italy

3 University of La Reunion, Saint-Denis, France

4 Variances, Paris, France

5 Sapienza University of Rome, Rome, Italy

Published online: 21 November 2024

droughts; e.g. Lin et al., 2018), bio-medecine (Huang et al., 2017) and, economics and finance (see, e.g. Amédée-Manesme & Barthélémy, 2018; Benkraiem & Zopounidis, 2021; Bellini et al., 2021), specifically when forecasting is the ultimate objective (see, e.g. Chalamandaris & Vlachogiannakis, 2018; Tian et al., 2020; Candila et al., 2023).

Among the mobilized techniques related to extremes, quantile regressions (denoted QR, hereafter) provide a simple way to model the conditional quantiles of a response variable with respect to some covariates in order to have a more satisfactory representation of its conditional distribution than one can have with traditional linear regression. Since the seminal works by Koenker and Basset (1978) and Koenker (2005), QR has become popular in economics and finance (Behr, 2010; Chun et al., 2012; Rockafellar and Royset, 2013; Candila et al., 2023), especially in recent years (Lin et al., 2018; Ben Ameur et al., 2022; Tsionas, 2020; Ben Bouheni et al., 2022; Tian et al., 2020), as a simple, robust and distribution free modelling tool. More specifically, a QR approach is mostly appropriate not only when the underlying model is non-linear, or the innovations are non-Gaussian, but also when modelling the distribution tails is the primary interest of the analysis. See Koenker (2005), Lum and Gelfand (2012) and Huang et al. (2017) for a review of the recent advances in QR modeling and inference. However, two greatest challenges in predicting extreme events are the robustness of the forecast accuracy in a changing environment and the correct representation of the statistical features of the variable under studies, as explicated below.

One of the most crucial issues in QR analyses is, indeed, related to model specification. The literature primarily focuses on shrinkage penalization approaches for simultaneous handling of both the estimation and model selection. For instance, Wu and Liu (2009) examine the properties of variable selection based on the least absolute shrinkage and selection operator (LASSO) and smoothly clipped absolute deviation (SCAD) penalties for static models, while Wang et al. (2012) investigate SCAD penalization for high-dimensional QR models. Noh et al. (2012) advocate a model selection method for time-varying QR, based on coefficient basis expansion and penalization. Lee et al. (2014) proposed instead a simpler alternative, based on Bayesian information criterion. A concern in the QR estimation process emerges from the fact that the objective function is not differentiable with respect to the parameters, which makes the derivation of the asymptotics of QR estimators almost unfeasible. Thus, we followed a Bayesian model averaging (BMA) approach to overcome these difficulties (see, e.g. Ji et al., 2012; Mamatzakis & Tsionas, 2021).

From a computational point of view, Bayesian model selection requires the estimation and the post-processing of the 2^M models, where $M \in \mathbb{N}$ being the number of regressors. Badly, it is almost impossible to be run as soon as M is even only moderately large. For this main reason, a computationally lighter version of the stochastic search variable selection (SSVS) by George and McCulloch (1993), has been proposed for QR models in Meligkotsidou et al. (2009), Chen et al. (2013) and Alhamzawi and Yu (2012). In this paper, we propose an alternative approach, extending the Dynamic Model Averaging (DMA) by Raftery et al. (2010) to QR models. Recently, DMA for linear models has gained significant attention in econometrics (Koop and Korobilis, 2012, 2013; Koop and Tole, 2013). To the best of our knowledge, it remains the only computationally feasible algorithm when handling a large number of regressors. Several extensions of the DMA technique have already been proposed. McCormick et al. (2012) introduced a DMA algorithm for binary outcome regression models. Belmonte and Koop (2014) proposed an alternative model-based mechanism for the time-varying model switching. However, an extension to more general models, such as QR models, is still absent in the literature, primarily due to the computational cost associated with increased model complexity. In this article, we demonstrate how this cost can be significantly reduced by implementing the Sequential MCMC (SMCMC) algorithm recently proposed by Guhaniyogi

et al. (2018), which is an efficient and self-tuning approach for simulating both parameters and latent variables.

In summary, our novel Bayesian Dynamic Quantile Model Averaging (BDQMA) combines three complementary overlays as an attempt for mitigating the specification and estimation errors inherent to our operational research problem: (1) a time-varying parameter QR setting (Koenker, 2005); (2) a BMA framework (Steel, 2020); and (3) a DMA approach (Raftery et al., 2010) with the use of SMCMC (Guhaniyogi et al., 2018). The contribution to the literature is manifolds. First, we extend the model selection procedure for QR models by (Kim, 2007) to time-varying parameters QR models. Second, we extend the classical DMA for linear models by Raftery et al. (2010) to conditionally linear models. Our article makes a significant contribution also to the model combination literature Stone (1961); Hall and Mitchell (2007); Billio et al. (2012); Fawcett et al. (2015); Gneiting and Ranjan (2013); Bassetti et al. (2018); Casarin et al. (2013); Busetti (2017); Billio et al. (2013); Aastveit et al. (2014, 2018); Casarin et al. (2023); McAlinn and West (2019).

We investigate the effectiveness of the time-varying parameter quantile regression model, supported by a sequential estimation procedure, in extracting quantile signals from synthetic datasets designed to replicate real-world features such as abrupt and smooth changes in the relationship between the dependent variables and the covariates. This validation highlights the model's robustness and applicability to dynamic financial and macroeconomic time series, where capturing extreme quantiles and handling time-varying volatility are particularly important. We illustrate further the potential of our methodology with an application to two real-world datasets. Financial and macroeconomic time series are widely documented in the literature as exhibiting distinctive features, such as heavy tails in the conditional distribution and time-varying volatility (e.g. Engle, 1982; Bollerslev, 1986). Addressing these characteristics is particularly important when the study focuses on extreme quantiles (e.g. Taylor, 2020; 2022; Bonaccolto et al., 2022; Bellini et al., 2022; Candila et al., 2023). Accordingly, our first application examines inflation (following Koop and Korobilis, 2012), while the second focuses on the U.S. real estate market (following Risse & Kern, 2016; Amédée-Manesme and Barthélémy, 2018).

The rest of the article is structured as follows. Section 2 introduces our BDQMA approach. Section 3 first underscores the effectiveness of our methodology using synthetic data. Section 4 and 5 provide two empirical applications, whilst Sect. 6 encapsulates our concluding remarks.

2 Bayesian dynamic quantile model averating

The building blocks of our BDQMA approach are: (i) time-varying parameter quantile regression (QR) modelling within a Bayesian framework; (ii) dynamic model averaging (DMA) within a Bayesian framework, and (iii) sequential Markov chain Monte Carlo (SMCMC) for sequential estimation of model risk and prediction of extreme events. Figure 1 provides a graphical illustration of the three blocks. In the following subsections, we provide a detailed explanation of each building block and how they contribute to the overall methodology.

2.1 Time-varying quantile regressions

Regarding the first block, the dynamic QR model used in this paper extends the one in Bernardi et al. (2015) to the case all the parameters evolve randomly over time. Let

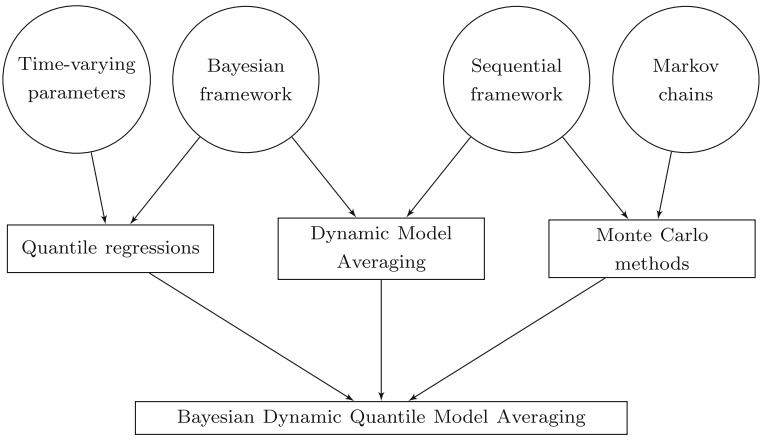


Fig. 1 The three main modeling and inference blocks (boxes) of our Bayesian Dynamic Quantile Model Averaging (BDQMA) framework together with the main features of the blocks (circles)

$\mathbf{x}_t = \left(1, x_{2,t}, \ldots, x_{M,t}\right)^\top, t = 1, 2, \ldots, T$ be the collection of M exogenous variables. The dependent variable, y_t, is a linear function of $\mathbf{x}_t$, and, as in Harvey (1989), the parameter vector $\boldsymbol{\beta}_t$ follows a multivariate random walk:

$$y_t = \mathbf{x}_t^\top \boldsymbol{\beta}_t + \xi_t \tag{1}$$

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\zeta}_t, \tag{2}$$

$$\boldsymbol{\beta}_0 \sim \mathsf{MVN}_M\left(\boldsymbol{\beta}_{0|0}, \mathbf{P}_{0|0}\right), \tag{3}$$

where $\xi_t \sim \mathsf{AL}(\tau, 0, \sigma), t = 1, 2, \ldots, T$ are i.i.d. variables with centred Asymmetric Laplace (AL) distribution, with $\tau \in (0, 1)$ the quantile level, $\sigma \in \mathbb{R}^+$ the scale parameter and $\boldsymbol{\beta}_0 \in \mathbb{R}^M$ the initial state of the evolution of $\boldsymbol{\beta}_t$ defined by the transition Eq. (2). We assume that $\boldsymbol{\beta}_0 \in \mathbb{R}^M$ has null mean and diffuse variance-covariance matrix $\mathbf{P}_{0|0} = \kappa \mathbf{I}_M$, with $\kappa \to +\infty$ and that the error terms $\boldsymbol{\zeta}_t \sim \mathsf{MVN}_M(\mathbf{0}, \boldsymbol{\Omega})$, are i.i.d. and independent of the measurement equation errors $\xi_s, \forall t, s = 1, 2, \ldots, T$. Hereafter, $\mathsf{MVN}_M(\cdot, \cdot)$ denotes the M-dimensional Gaussian random variable, while $\mathsf{N}(\cdot, \cdot)$ is exclusively reserved for representing univariate Gaussian random variables.

The linear state space model introduced in Eq. (1–3) for representing time-varying conditional quantiles deviates from Gaussianity due to the assumptions imposed on the measurement innovation terms. Consequently, traditional optimal filtering techniques, such as those relying on Kalman filter recursions, cannot be applied (see, Durbin and Koopman, 2012). Nonetheless, drawing upon the insights of Kozumi and Kobayashi (2011) and Bernardi et al. (2015), it is possible to leverage the representation of the centered Asymmetric Laplace probability density function, $f(\xi_t|\tau, \sigma)$, as a location-scale continuous Normal mixture. This approach yields:

$$f(\xi_t|\tau, \sigma) = \int_0^\infty \frac{1}{\sqrt{2\pi\delta\sigma\omega}} \exp\left\{-\frac{1}{2\delta\sigma\omega}(\xi_t - \lambda\omega)^2\right\} \frac{1}{\sigma} \exp\left\{-\frac{1}{\sigma}\omega\right\} d\omega, \tag{4}$$

for $t = 1, 2, \ldots, T$. Setting $\lambda = \frac{1-2\tau}{\tau(1-\tau)}$ and $\delta = \frac{2}{\tau(1-\tau)}$ ensures that the τ-level quantile of the measurement error ξ_t becomes zero. Consequently, it becomes evident that our non-Gaussian state space model admits the following conditionally Gaussian and linear state

space representation:

$$y_t = \mathbf{x}_t^\top \boldsymbol{\beta}_t + \lambda \omega_t + \varepsilon_t, \quad \varepsilon_t \overset{\text{iid}}{\sim} \mathsf{N}\left(0, \delta\sigma\omega_t\right), \tag{5}$$

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\zeta}_t, \qquad \boldsymbol{\zeta}_t \overset{\text{iid}}{\sim} \mathsf{MVN}_M\left(0, \boldsymbol{\Omega}\right), \tag{6}$$

$$\boldsymbol{\beta}_0 \sim \mathsf{MVN}_M\left(\boldsymbol{\beta}_{1|0}, \mathbf{P}_{1|0}\right), \tag{7}$$

where ω_t, $t = 1, 2, \ldots, T$, are i.i.d. with an exponential distribution of parameter σ^{-1}, that is $\omega_t \sim \mathsf{Exp}\left(\sigma^{-1}\right)$. We further assume the following prior distributions for the parameters σ and $\boldsymbol{\Omega}$:

$$\sigma \sim \mathsf{IG}\left(a_0, b_0\right), \qquad \boldsymbol{\Omega} \sim \mathsf{IW}\left(c_0, \mathbf{C}_0\right), \tag{8}$$

which are Inverse Gamma and Inverse Wishart distributions, with probability density functions $p\left(\sigma|a_0, b_0\right) \propto \sigma^{-(a_0+1)} \exp\left\{-\frac{b_0}{\sigma}\right\}$ and $p\left(\boldsymbol{\Omega}|c_0, \boldsymbol{\Omega}_0\right) \propto |\mathbf{C}_0|^{c_0} |\boldsymbol{\Omega}|^{-\left(c_0+\frac{M+1}{2}\right)} \exp\left\{-\text{trace}\left(\mathbf{C}_0\boldsymbol{\Omega}^{-1}\right)\right\}$, respectively, where $(a_0, b_0, c_0, \boldsymbol{\Omega}_0)$ are given hyperparameters.

2.2 Bayesian dynamic model averaging

For the second building block of our method, we propose a dynamic procedure for selecting and combining dynamic quantile regression (QR) models. The challenge of determining which regressors to include in the model can be addressed from two perspectives. One approach is to use a variable dimension model (see, e.g. Marin and Robert, 2007). However, in this study, we take a different route, opting for a model selection approach that requires estimating all possible sub-models. Drawing inspiration from the seminal work of Raftery et al. (2010), we implement a recursive model selection process over time to continually evaluate model adequacy. Expanding on the framework introduced by Raftery et al. (2010), we introduce a time-varying model index L_t, which takes values in $\{1, 2, \ldots, K\}$, representing the selected model at each time point t. We assume it has a Markov dynamics with transition matrix

$$\mathsf{P}(L_t = l|L_{t-1} = j) = \pi_{lj}, \qquad l, j \in \{1, 2, \ldots, K\}. \tag{9}$$

Let $\mathbf{x}_{kt} = (1, x_{i_1 t}, \ldots, x_{i_{m^{(k)}} t})^\top$ be the set of $(m^{(k)} + 1)$ variables in the model k, with $i_j \in \{2, 3, \ldots, M\}$, $j = 1, \ldots, m^{(k)}$, $i_j \neq i_l \; \forall l, j$ and $m^{(k)} \in \{0, 1, \ldots, M-1\}$. Then, we posit that each model can be expressed in the form of Eq. (1)-(3), where the dynamic regression parameter vector $\boldsymbol{\beta}_t^{(k)}$ and the covariate vector $\mathbf{x}_t^{(k)}$ are unique to each model $k = 1, 2, \ldots, M$. This assumption implies that at each time t and for each model $k = 1, 2, \ldots, K$, the system variables represented by the dynamic QR parameters $\boldsymbol{\beta}_t$ and the model indicator variable L_t, i.e. $\left\{\boldsymbol{\beta}_t, L_t\right\}$, move to a new state according to the transition matrix specified in Eq. (9) and the transition kernel defined by the dynamics of latent states in Eq. (2). The updated states comprise the vector of predicted values at time $t+1$ for each model-specific quantile regression parameter vector $\boldsymbol{\beta}_{t+1}^{(k)}$ and the updated time-varying probability $\pi_{t+1}^{(k)}$ associated with that model.

As anticipated in the Introduction, the dynamic quantile regression model specified in Eq. (10) exhibits non-Gaussian characteristics, rendering the iterative application of the linear Kalman filter and smoother impractical for obtaining updated estimates of the latent dynamics over time. To address this limitation, we once again leverage the representation of the Asymmetric Laplace error term in the measurement equation (Eq. (1)) as a location-scale

mixture of Normals. This representation hinges on the introduction of an additional latent factor $\omega_t^{(k)}$, unique to each model k, following an i.i.d. Exponential distribution with shape parameter $\sigma^{(k)}$, i.e. $\omega_t^{(k)} \sim \mathsf{Exp}(1/\sigma^{(k)})$, i.i.d. $\forall t$.

For the model selection purposes, we derive from Eq. (5–6) the following:

$$\begin{aligned} y_t &= \sum_{k=1}^{K} \mathbb{1}_{\{k\}}(L_t)\left(\mathbf{x}_t^{(k)\top}\boldsymbol{\beta}_t^{(k)} + \lambda\omega_t^{(k)} + \sqrt{\delta\sigma^{(k)}\omega_t^{(k)}}\varepsilon_t^{(k)}\right) \\ \boldsymbol{\beta}_t^{(k)} &= \boldsymbol{\beta}_{t-1}^{(k)} + \boldsymbol{\zeta}_t^{(k)}, \end{aligned} \tag{10}$$

where $\varepsilon_t^{(k)} \overset{\text{i.i.d.}}{\sim} \mathsf{N}(0,1)$ and $\boldsymbol{\zeta}_t^{(k)} \overset{\text{i.i.d.}}{\sim} \mathsf{MVN}_{k+1}(\mathbf{0}, \boldsymbol{\Omega}^{(k)})$. The parameters λ, δ, and τ are defined as above. The description of the model is completed by the prior specification for the initial state given in Eq. (3).

2.3 Sequential Markov chain Monte Carlo

In the third block, we adopt a Bayesian approach to inference for QR models, and introduce auxiliary variables to represent the original model in Eq. (1)-(3) as a conditionally Gaussian and linear state space model. Let $\mathbf{z}_{1:t} = (\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_t)$ denote the sequence of vectors $\mathbf{z}_u$ up to time t. Then, the complete-data likelihood of the unobservable components $\boldsymbol{\beta}_{1:T}$ and $\boldsymbol{\omega}_{t:T}$ and all parameters $\boldsymbol{\gamma} = (\sigma, \boldsymbol{\Omega})$ is given by:

$$\begin{aligned} &\mathcal{L}\left(\boldsymbol{\gamma}, \boldsymbol{\omega}_{1:T}, \boldsymbol{\beta}_{1:T} | \mathbf{y}_{1:T}, \mathbf{x}_{1:T}\right) \propto f\left(\mathbf{y}_{1:T} | \mathbf{x}_{1:T}, \boldsymbol{\beta}_{1:T}, \boldsymbol{\omega}_{1:T}, \boldsymbol{\gamma}\right) f\left(\boldsymbol{\beta}_{1:T}, \boldsymbol{\omega}_{1:T} | \boldsymbol{\gamma}\right) \pi(\boldsymbol{\gamma}) \\ &\propto \prod_{t=1}^{T} f\left(y_t | \boldsymbol{\beta}_t, \omega_t, \sigma, \mathbf{x}_t\right) f\left(\omega_t | \sigma\right) f\left(\boldsymbol{\beta}_1\right) \prod_{t=2}^{T} f\left(\boldsymbol{\beta}_t | \boldsymbol{\beta}_{t-1}, \boldsymbol{\Omega}\right) \\ &\propto \prod_{t=1}^{T} \left(\sigma \times \omega_t\right)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma\delta}\sum_{t=1}^{T}\frac{\left(y_t - \lambda\omega_t - \mathbf{x}_t^{\top}\boldsymbol{\beta}_t\right)^2}{\omega_t}\right\} \\ &\quad \times \exp\left\{-\frac{\sum_{t=1}^{T}\omega_t}{\delta}\right\} \exp\left\{-\frac{\boldsymbol{\beta}_1^{\top}\boldsymbol{\beta}_1}{2\kappa}\right\} \\ &\quad \times |\boldsymbol{\Omega}|^{-\frac{T-1}{2}} \exp\left\{-\frac{1}{2}\sum_{t=2}^{T}\left(\boldsymbol{\beta}_t - \boldsymbol{\beta}_{t-1}\right)^{\top}\boldsymbol{\Omega}^{-1}\left(\boldsymbol{\beta}_t - \boldsymbol{\beta}_{t-1}\right)\right\}. \end{aligned}$$

The Gaussian scale mixture representation poses challenges for sequential inference of latent variables and parameters. In this model, the DMA approach introduced by Raftery et al. (2010) may become computationally burdensome, as exact estimation of latent variables and parameters across different models requires posterior approximation. Standard Markov chain Monte Carlo (MCMC) schemes for posterior approximation are not feasible in such cases. To address this issue, we propose combining a sequential MCMC procedure (SMCMC) (Yang and Dunson, 2013) with the DMA approach to create a feasible sequential DMA procedure for latent variable models. Additionally, we introduce enhancements to reduce computational time. Firstly, we demonstrate how the forgetting factor, as introduced by Raftery et al. (2010) for the state covariance, can expedite computations without compromising the validity of the SMCMC procedure. Secondly, we illustrate that computing costs, which typically increase linearly over time, can be stabilized using fixed-lag backward sampling within the SMCMC transition kernel. Subsequent subsections present the SMCMC for the single-model sce-

nario, its extension to the multi-model scenario, and strategies for mitigating computational complexity.

2.3.1 Sequential model estimation

Let $\boldsymbol{\theta}_t = \left(\sigma, \boldsymbol{\Omega}, \boldsymbol{\omega}_{1:t}, \boldsymbol{\beta}_{1:t}\right)$, $t \in \mathbb{N}$, denote the time series of augmented parameter vectors, taking values in measurable spaces $\left(\mathbb{R}^{d_t}, \mathcal{B}\left(\mathbb{R}^{d_t}\right)\right)$, where the dimension non-decreasingly grows as $d_t = d_{t-1} + d$, for $t \geq 1$. We assume that the augmented parameter vector can be decomposed as $\boldsymbol{\theta}_t = \left(\boldsymbol{\theta}_{t-1}, \boldsymbol{\eta}_t\right)$, where $\boldsymbol{\eta}_t = \left(\omega_t, \boldsymbol{\beta}_t\right)$ represents the latent variable vector of dimension d associated with the observation y_t.

The proposed quantile regression model with time-varying parameters possesses a prior distribution $\pi_t\left(\boldsymbol{\theta}_t\right) = \pi\left(\boldsymbol{\gamma}\right)\pi\left(\boldsymbol{\omega}_{1:t}, \boldsymbol{\beta}_{1:t}\right)$, which satisfies the compatibility condition:

$$\pi_{t+1}\left(\boldsymbol{\theta}_t\right) = \int_{\mathbb{R}^d} \pi_t\left(\boldsymbol{\theta}_t, \boldsymbol{\eta}_{t+1}\right) d\boldsymbol{\eta}_{t+1}, \tag{11}$$

enabling a simplification of notation. Furthermore, we denote the posterior distribution at time t, with respect to the Lebesgue measure of $\boldsymbol{\theta}_t$, as $\pi_t(\boldsymbol{\theta}_t) = \pi\left(\boldsymbol{\theta}_t|\mathbf{y}_{1:t}\right) \propto \pi\left(\boldsymbol{\gamma}\right)\mathcal{L}\left(\boldsymbol{\gamma}, \boldsymbol{\omega}_{1:t}, \boldsymbol{\beta}_{1:t}|\mathbf{y}_{1:t}, \mathbf{x}_{1:t}\right)$. Henceforth, in the notation, we omit the conditioning on the dependent variables $\mathbf{y}_{1:t}$ and the covariates $\mathbf{x}_{1:t}$ from the posterior.

In the SMCMC algorithm, L parallel inhomogeneous Markov chains are employed to generate samples $\boldsymbol{\theta}_t^{(l,j)}$ with $j = 1, \ldots, m_t$, $l = 1, \ldots, L$ and $t = 1, 2, \ldots, T$ from the sequence of posterior distributions π_t, $t = 1, 2, \ldots, T$. Each Markov chain within the population is defined by a sequence of transition kernels $\mathcal{K}_t\left(\boldsymbol{\theta}, A\right)$, $t \in \mathbb{N}$, which are operators from $\left(\mathbb{R}^{d_{t-1}}, \mathcal{B}\left(\mathbb{R}^{d_{t-1}}\right)\right)$ to $\left(\mathbb{R}^{d_t}, \mathcal{B}\left(\mathbb{R}^{d_t}\right)\right)$, such that $\mathcal{K}_t\left(\boldsymbol{\theta}, \cdot\right)$ is a probability measure for all $\boldsymbol{\theta} \in \mathbb{R}^{d_{t-1}}$, and $\mathcal{K}_t\left(\cdot, A\right)$ is measurable for all $A \in \mathcal{B}\left(\mathbb{R}^{d_t}\right)$.

The kernel $\mathcal{K}_t\left(\boldsymbol{\theta}, A\right)$ has stationary distribution π_t and results from the composition of a jumping kernel, $\mathcal{J}_t$, and a transition kernel, $\mathcal{T}_t$, expressed as:

$$\mathcal{K}_t(\boldsymbol{\theta}, A) = \mathcal{J}_t \circ \mathcal{T}_t^{m_t}\left(\boldsymbol{\theta}, A\right) = \int_{\mathbb{R}^{d_t}} \mathcal{J}_t\left(\boldsymbol{\theta}, d\boldsymbol{\theta}'\right)\mathcal{T}_t^{m_t}\left(\boldsymbol{\theta}', A\right),$$

where the fixed dimension transition is defined as:

$$\mathcal{T}_t^{m_t}\left(\boldsymbol{\theta}, A\right) = \mathcal{T}_t \circ \mathcal{T}_t^{m_t-1}\left(\boldsymbol{\theta}, A\right) = \int_{\mathbb{R}^{d_t}} \mathcal{T}_t(\boldsymbol{\theta}, d\boldsymbol{\theta}')\mathcal{T}_t^{m_t-1}\left(\boldsymbol{\theta}', A\right), \tag{12}$$

with $m_t \in \mathbb{N}$, and $\mathcal{T}^0$ is the identity kernel. We assume that the jumping kernel satisfies

$$\mathcal{J}_{t+1}\left(\boldsymbol{\theta}_t, \boldsymbol{\theta}_{t+1}\right) = \mathcal{J}_{t+1}\left(\boldsymbol{\theta}_t, \boldsymbol{\eta}_t\right)\delta_{\boldsymbol{\theta}_t}(\boldsymbol{\theta}_{t+1}), \tag{13}$$

where $\mathcal{J}_{t+1}\left(\boldsymbol{\theta}_t, \boldsymbol{\eta}_{t+1}\right) = \mathcal{J}_{t+1}\left(\boldsymbol{\theta}_t, \left(\boldsymbol{\theta}_t, \boldsymbol{\eta}_{t+1}\right)\right)$. This condition ensures that the error propagation through the jumping kernel can be effectively managed across the iterations of SMCMC. This result, presented in Theorem 1, is a straightforward extension of the findings established in the L_1-norm by Guhaniyogi et al. (2018) and related works. Let us introduce the v-norm, denoted as $||\mu(\cdot)||_v = \sup|f| \leq v|\mu(f)|$, where $v : \mathbb{R}^n \to [1, \infty]$ and μ represents a signed measure (see Meyn and Tweedie, 1993, Ch.14).

Theorem 1 *For any probability density function $p(\cdot)$, and for $\boldsymbol{\theta}_{t-1} \in \mathbb{R}^{d_{t-1}}$ the next inequality is met:*

$$||\pi_t - \mathcal{J}_t \circ p||_v \leq \sup_{\boldsymbol{\theta}_{t-1} \in \mathbb{R}^{d_t - 1}} ||\pi_t(\cdot|\boldsymbol{\theta}_{t-1}) - \mathcal{J}_t(\boldsymbol{\theta}_{t-1}, \cdot)||_{\widetilde{v}} + ||\pi_t - p||_v, \tag{14}$$

where $\widetilde{v} = \int\limits_{\mathbb{R}^{d_{t-1}}} v\left(\left(\boldsymbol{\theta}_{t-1}, \boldsymbol{\eta}_t\right)\right) d\boldsymbol{\theta}_{t-1}$.

Proof See Appendix A. □

The following findings outlined in Guhaniyogi et al. (2018) demonstrate the convergence of the chain with transition kernel $\mathcal{J}_t \circ \mathcal{T}^{m_t}$ towards the target distribution, under highly general conditions.

Theorem 2 *Let us assume the following conditions hold:*

(i) *(Geometric ergodicity) For each t, there exists a function $V_t : \mathbb{R}^{d_t} \to [1, \infty)$, $C > 0$ and $\rho_t \in (0, 1)$ such that:*
   (a) $\int_{\mathbb{R}^{d_t}} V_t(\boldsymbol{\theta}_t)^2 \pi_t(\boldsymbol{\theta}_t) d\boldsymbol{\theta}_t \leq C$;
   (b) $\int_{\mathbb{R}^{d_t}} V_t(\boldsymbol{\theta}_t)^2 \pi_t(\boldsymbol{\eta}_t|\boldsymbol{\theta}_t) d\boldsymbol{\eta}_t = V_{t-1}(\boldsymbol{\theta}_{t-1})$, *where* $\boldsymbol{\theta}_t = \left(\boldsymbol{\theta}_{t-1}, \boldsymbol{\eta}_t\right)$;
   (c) *for all* $\boldsymbol{\theta}_t \in \mathbb{R}^{d_t}$, $||\mathcal{T}_t(\boldsymbol{\theta}_t, \cdot) - \pi_t(\cdot)||_{v_t} \leq V_t(\boldsymbol{\theta}_t)\rho_t$.

(ii) *(Stationary convergence) The stationary distribution π_t of $\mathcal{T}_t$ satisfies*

$$\alpha_t = 2\sqrt{C} d_H(\pi_t, \pi_{t-1}) \to 0,$$

   where π_t is the marginal posterior of $\boldsymbol{\theta}_{t-1}$ at time t.

(iii) *(Jumping consistency) For a sequence of $\lambda_t \to 0$ the following holds:*

$$\sup_{\boldsymbol{\theta}_{t-1} \in \mathbb{R}^{d_t - 1}} ||\mathcal{J}_t(\boldsymbol{\theta}_{t-1}, \cdot) - \pi_t(\cdot|\boldsymbol{\theta}_{t-1})||_{\widetilde{v}_t},$$

   where $\widetilde{v}_t = \int\limits_{\mathbb{R}^d} v((\boldsymbol{\theta}, \boldsymbol{\eta})) d\boldsymbol{\theta}_{t-1}$.

Let $\varepsilon_t = \rho_t^{m_t}$. Then for any initial distribution π_0,

$$||\mathcal{K}_t \circ \cdots \circ \mathcal{K}_1 \circ \pi_0 - \pi_t||_{v_t} \leq \sum_{s=1}^{t} \left(\prod_{u=s}^{t} \varepsilon_u\right)(\alpha_s + \lambda_s).$$

Proof See Appendix A. □

Regarding the assumptions of geometric ergodicity and jumping consistency, several works have investigated convergence rates of Gibbs samplers for Bayesian models. For instance, Román and Hobert (2012) establish geometric ergodicity for a range of improper priors in Bayesian linear models. Additionally, studies by Jones and Hobert (2004) and Papaspiliopoulos and Roberts (2008) are relevant in this context. In the present paper, as we employ a Gibbs sampler for a conditionally linear and Gaussian model with proper conditionally conjugate prior distributions, these assumptions are met in accordance with Román and Hobert (2012).

To employ SMCMC, one must specify the transition kernel $\mathcal{T}_{t+1}$ and the jumping kernel $\mathcal{J}_{t+1}$ at the iteration $t+1$. The transition kernel $\mathcal{T}_t$ at the iteration t enables each parallel chain to explore the sample space of dimension d_t and generate samples $\boldsymbol{\theta}_t^{(l,j)}$ from the posterior distribution π_t, given the previously generated samples. The collapsed Gibbs sampler

Algorithm 1: Collapsed Gibbs sampler.

Data: $\mathbf{y}_{1:T}, \mathbf{x}_{1:T}$
for $t = 1, 2, \ldots, T$, *and* $l = 1, 2, \ldots, L$ **do**
  draw $\sigma^{(l)}$ from $f(\sigma|\boldsymbol{\beta}_{1:t}^{(l)}, \mathbf{y}_{1:t}, \mathbf{x}_{1:t})$ marginalizing the Gibbs with respect to the $\boldsymbol{\omega}_{1:t}$;
  draw $\omega_s^{(l)}$ from $f(\omega_s|\sigma^{(l)}, \boldsymbol{\Omega}^{(l)}, \boldsymbol{\beta}_{1:t}^{(l)}, \mathbf{y}_{1:t}, \mathbf{x}_{1:t})$, $s = 1, 2, \ldots, t$;
  draw $\boldsymbol{\beta}_{1:t}^{(l)}$ from $f(\boldsymbol{\beta}_{1:t}|\sigma^{(l)}, \boldsymbol{\Omega}^{(l)}, \boldsymbol{\omega}_{1:t}^{(l)}, \mathbf{y}_{1:t}, \mathbf{x}_{1:t})$ with a multi-move proposal distribution;
  draw $\boldsymbol{\Omega}^{(l)}$ from $f(\boldsymbol{\Omega}|\boldsymbol{\beta}_{1:t}^{(l)})$;
end

algorithm is detailed in Algorithm 1. Within this algorithm, we opt for a transition kernel akin to a blocked (or multi-move) kernel.

The sampling strategy for the full conditional distributions is elaborated in Appendix B.1. Leveraging the stochastic representation of the AL distribution in Eq. (4), we implement a partially collapsed SMCMC strategy based on Gibbs-type updating with data augmentation (Liu, 1994; Van Dyk and Park, 2008). The core concept of the complete collapsed Gibbs-type simulation scheme is to avoid simulations from the full conditional distributions of blocks of model parameters whenever possible by analytically marginalizing them out. As demonstrated by Park and Van Dyk (2009), this approach offers several advantages over systematic sampling, such as reduced computational time and enhanced convergence rate of the sampler. In our scenario, we can only partially integrate out the latent variables $\omega_{1:s}^{(l)}$, for $s = 1, 2, \ldots, t$, from the full conditional of the scale parameters $\sigma^{(l)}$ in the preceding step 3, for $l = 1, 2, \ldots, L$, thereby reducing the algorithm to a partially collapsed Gibbs sampler type move. Concerning the convergence of partially collapsed Gibbs moves, it is notable that updating the parameters in the specified order ensures that the posterior distribution at each time point $t = 1, 2, \ldots, T$ corresponds to the stationary distribution of the generated Markov chain. In fact, combining steps 1 and 2 produces sample draws from $\pi(\sigma^{(l)}, \boldsymbol{\omega}_{1:t}^{(l)}|\boldsymbol{\beta}_{1:t}^{(l)}, \mathbf{y}_{1:t}, \mathbf{x}_{1:t})$, for $t = 1, 2, \ldots, T$ and $l = 1, 2, \ldots, L$, i.e. the conditional posterior distribution. The partially collapsed Gibbs sampler thus represents a blocked version of the ordinary Gibbs sampler (Van Dyk and Park, 2008; Park and Van Dyk, 2009).

To initialize the Gibbs sampling, we begin by generating a random draw from the joint prior distribution of the parameters defined in Eq. (8). Conditionally on this draw, we simulate the initial values of the augmentation variables $\omega_1^{(l)}$, for $l = 1, 2, \ldots, L$. Subsequently, at iteration $t + 1$, the jumping kernel $\mathcal{J}_{t+1}$ facilitates the transition of chains from a space of dimension d_t to one of dimension d_{t+1}. We designate the jumping kernel of the l-th parallel chain as the transition kernel of a Gibbs sampler, characterized by the following full conditional distributions:

1. draw $\omega_{t+1}^{(l)}$ from $f(\omega_{t+1}|\sigma^{(l)}, \boldsymbol{\beta}_{t+1}^{(l)}, y_{t+1}, \mathbf{x}_{t+1})$;
2. draw $\boldsymbol{\beta}_{t+1}^{(l)}$ from $f(\boldsymbol{\beta}_{t+1}|\sigma^{(l)}, \boldsymbol{\Omega}^{(l)}, \boldsymbol{\beta}_t^{(l)}, \omega_{t+1}^{(l)}, y_{t+1}, \mathbf{x}_{t+1})$.

The specifics of the sampling strategy for the full conditional distributions are detailed in Appendix B.2. Thanks to the location-scale mixture representation of the AL distribution, all the full conditional distributions involved in the transition and jumping kernels have a known closed-form representation. This feature is especially advantageous when T is large or when multiple models are estimated simultaneously, as in the current scenario, because the availability of conditional sufficient statistics that can be tracked helps mitigate the increase in storage and computational burden over time.

2.3.2 Monitoring convergence

The number of iterations of the Sequential MCMC sampler at each time m_t is determined based on the cross-chain correlation (see Guhaniyogi et al., 2018). Specifically, we set the number of iterations at time t, denoted as m_t, to be the smallest integer s such that $r_t(s) \leq 1-\epsilon$, where $r_t(s)$ represents the rate function associated with the transition kernel $\mathcal{T}_t$ and ϵ is a predefined threshold level. An upper bound for the rate function can be estimated by the lag-s chain autocorrelation, which is calculated online using information provided by the output of all the parallel chains. This estimation is performed as follows:

$$\widehat{r}_t(s) = \max_{j=1,2,\ldots,p} \frac{\sum_{l=1}^{L}(\theta_j^{(s+1,t,l)} - \bar{\theta}_j^{(s+1,t)})(\theta_j^{(1,t,l)} - \bar{\theta}_j^{(1,t)})}{\left(\sum_{l=1}^{L}(\theta_j^{(s+1,t,l)} - \bar{\theta}_j^{(s+1,t)})^2\right)^{\frac{1}{2}}\left(\sum_{l=1}^{L}(\theta_j^{(1,t,l)} - \bar{\theta}_j^{(1,t)})^2\right)^{\frac{1}{2}}}, \quad (15)$$

where $\theta_j^{(s,t,l)}$, is the j-th element of the vector $\boldsymbol{\theta}^{(s,t,l)}$ staking the fixed parameters $\boldsymbol{\gamma}^{(l)}$ and the latent states generated up to time t, i.e. $(\boldsymbol{\omega}_{1:t}^{(l)}, \boldsymbol{\beta}_{1:t}^{(l)})$, by the l-th chain at the the s-th iteration. Moreover, $\bar{\theta}_j^{(s,t)} = L^{-1}\sum_{l=1}^{L}\theta_j^{(s,t,l)}$ is the average of the draws of the s-th iteration over the L parallel chains.

2.3.3 Sequential model selection

Consider scenarios where multiple models are evaluated and estimated simultaneously, especially when there is uncertainty about the correct specification of the data-generating process. We proceed under the assumption that the entire array of models can be represented as in Eq. (10), where the indicator variable $L_t \in \{1, 2, \ldots, K\}$ has a Markovian transition kernel, selecting the operative model at each time step $t = 1, 2, \ldots, T$. Furthermore, it is assumed that the divergence among models primarily arises from the presence or absence of specific explanatory variables within the predictor vector $\mathbf{x}_t$. Assuming there are $M-1$ potential explanatory variables, with the intercept being a constant across all models, the total number of models becomes $K = 2^{M-1}$. Even with a moderate number of variables, such as $M = 10$, the count of models can increases substantially, reaching, for instance, 1, 024, thereby making the formulation of the transition matrix a formidable task. Following the approach outlined by Raftery et al. (2010), we adopt a forgetting factor methodology, as described in the following.

Let $\pi_{t+1|t}^{(k)} = \mathsf{P}(L_{t+1} = k|\mathbf{y}_{1:t}, \mathbf{x}_{t+1}^{(k)})$ and $\pi_{t+1|t+1}^{(k)} = \mathsf{P}(L_{t+1} = k|\mathbf{y}_{1:t+1}, \mathbf{x}_{t+1}^{(k)})$, where again $\mathbf{y}_{1:t} = (y_1, y_2, \ldots, y_t)$. The approximated Hamilton filter recursion involves the following two steps. At each time point t, the first step predicts and updates the model probabilities as follows:

$$\pi_{t+1|t}^{(k,l)} = \frac{(\pi_{t|t}^{(k,l)})^{\alpha} + \upsilon}{\sum_{j=1}^{K}(\pi_{t|t}^{(j,l)})^{\alpha} + \upsilon}, \qquad \forall l = 1, 2, \ldots, L, \quad (16)$$

$$\pi_{t+1|t+1}^{(k)} \propto \sum_{l=1}^{L} \pi_{t+1|t}^{(k,l)} f(y_{t+1}|\mathbf{y}_{1:t}, \mathbf{x}_{t+1}^{(k)}, L_t = k, \omega_{t+1}^{(k,l)}). \quad (17)$$

In Eq. (16), the exponent α serves as a forgetting factor, typically set close to one, while the constant $\upsilon > 0$ is a small number introduced to mitigate rounding errors due to machine precision, with υ set to $\upsilon = 0.001/K$, following the methodology outlined in Raftery et al. (2010).

To circumvent the necessity of explicitly specifying and computing the entire transition matrix of the Markov chain L_t, which comprises $K(K-1)$ entries, we introduce the forgetting factor α. This factor allocates prior weight to information preceding time t, with a higher α parameter implying reduced weight assigned to subsequent observations at time $t+1$. Essentially, higher values of α indicate less prior confidence in the informational significance of new observations during the model probability update process.

As argued by Raftery et al. (2010), although this approach lacks an explicit specification of the transition matrix, it does not pose a problem for inference in model probabilities as long as the data provides sufficient information regarding which models better capture the data dynamics. Regarding the updating of model probabilities in Eq. (17), it is obtained as the Monte Carlo average of the approximated predictive distributions $f(y_{t+1}|\mathbf{y}_{1:t}, \mathbf{x}_{t+1}^{(k)}, L_t = k, \omega_{t+1}^{(k,l)})$ over the entire population of independent Markov chains, effectively marginalizing out the simulated latent factor ω_{t+1}. The predictive distribution of the observable at time $t+1$ is obtained as a by-product of the Kalman filtering updating equations for the approximated Gaussian model:

$$f(y_{t+1}|\mathbf{y}_{1:t}, \mathbf{x}_{t+1}^{(k)}, L_t = k, \omega_{t+1}^{(k,l)}) \approx \mathsf{N}(y_{t+1}|\widehat{y}_{t+1|t}^{(k,l)}, V_{t+1}^{(k,l)}), \tag{18}$$

where $\widehat{y}_{t+1|t}^{(k,l)} = \lambda\omega_{t+1}^{(k,l)} + \mathbf{x}_{t+1}^{(k)\top}\boldsymbol{\beta}_{t+1|t}^{(k,l)}$ is the predictive mean of y_{t+1} for the l-th chain and model $k \in \{1, 2, \ldots, K\}$ and the variance-covariance matrix equal to $V_{t+1}^{(k,l)} = \left(\delta\sigma^{(k,l)}\omega^{(k,l)} + \mathbf{x}_{t+1}^{(k)\top}\mathbf{P}_{t+1|t}^{(k,l)}\mathbf{x}_{t+1}^{(k)}\right)^{-1}$, from the Kalman filter recursion given in Appendix B.1.

The updated model probabilities are normalised as follows:

$$\widetilde{\pi}_{t+1|t+1}^{(k,l)} = \frac{\pi_{t+1|t+1}^{(j,l)}}{\sum_{j=1}^{K}\pi_{t+1|t+1}^{(j,l)}}, \qquad \forall l = 1, 2, \ldots, L, \tag{19}$$

to get a proper probability measure.

The second step involves updating the latent factors $\boldsymbol{\beta}_{t+1}^{(l,k)}$ and $\omega_{t+1}^{(l,k)}$, for $l = 1, 2, \ldots, L$, and this can be easily accomplished by running the SMCMC algorithm defined in the previous section sequentially for each model $k = 1, 2, \ldots, K$. This process is then iterated as a new sample observation becomes available. The sequence is initialized at $t = 0$, with the initial model probabilities $\pi_{0|0}^{(k,l)} = \frac{1}{K}$ set uniformly for all the models $k = 1, 2, \ldots, K$, and for all the parallel chains of the SMCMC sampler $l = 1, 2, \ldots, L$. Static model parameters $\left(\sigma^{(k,l)}, \boldsymbol{\Omega}^{(k,l)}\right)$ are initialized by drawing from their prior distributions defined in Eq. (8).

2.3.4 Dynamic quantile averaging

Forecasting with SMCMC methods offers a streamlined approach to integrating parameter uncertainty. The following steps delineate the process of generating one-step-ahead quantile forecasts from all competing models at each time point t, and subsequently merging them using updated model weights. These procedures are executed at the conclusion of each SMCMC iteration within the sampling phase, leveraging the complete parameter set of the model denoted as $\boldsymbol{\theta}_t^{(k)}$. At every time instance during the estimation period, an approximation of the predictive mean for the τ-level quantile function $q_\tau(\mathbf{x}, \boldsymbol{\beta}) = \mathbf{x}^\top\boldsymbol{\beta}$, can be derived for the k-th model, as follows:

$$\widehat{q}_{\tau,t+1}^{(k)}(\mathbf{x}_{t+1}^{(k)}, \boldsymbol{\beta}_{t+1}^{(k)}) = \mathsf{E}\left[q_\tau(\mathbf{x}_{t+1}^{(k)}, \boldsymbol{\beta}_{t+1}^{(k)})|\mathbf{y}_{1:t}, \mathbf{x}_{t+1}^{(k)}, \boldsymbol{\beta}_{1:t}^{(k)}\right]$$

$$= \mathbf{x}_{t+1}^{(k)\top} \mathsf{E}\Big[\boldsymbol{\beta}_{t+1}^{(k)} | \mathbf{y}_{1:t}, \mathbf{x}_{t+1}^{(k)}, \boldsymbol{\beta}_{1:t}^{(k)}\Big]. \tag{20}$$

Here, $\widehat{\boldsymbol{\beta}}_{t+1}^{(k)} = \mathsf{E}\Big[\boldsymbol{\beta}_{t+1}^{(k)} | \mathbf{y}_{1:t}, \mathbf{x}_{t+1}^{(k)}, \boldsymbol{\beta}_{1:t}^{(k)}\Big]$ represents the mean of the marginal posterior distribution of the regression parameters for model k, i.e. $\boldsymbol{\beta}_{t+1}^{(k)}$. This can be obtained by numerically marginalizing out the nuisance parameters $(\sigma, \boldsymbol{\Omega})^{(k)}$ and the latent factors $\omega_{1:t+1}^{(k)}$ using the population of chains:

$$\begin{aligned}\widehat{\boldsymbol{\beta}}_{t+1}^{(k)} &= \mathsf{E}\Big(\boldsymbol{\beta}_{t+1}^{(k)} | \mathbf{y}_{1:t}, \mathbf{x}_{t+1}^{(k)}, \boldsymbol{\beta}_{1:t}^{(k)}\Big) \\ &\approx \frac{1}{L}\sum_{l=1}^{L} \mathsf{E}\Big(\boldsymbol{\beta}_{t+1}^{(k)} | \mathbf{y}_{1:t}, \mathbf{x}_{t+1}^{(k)}, \boldsymbol{\beta}_{1:t}^{(k,l)}, \boldsymbol{\omega}_{1:t}^{(k,l)}, \sigma^{(k,l)}, \boldsymbol{\Omega}^{(k,l)}\Big).\end{aligned} \tag{21}$$

In the right term, the expectation is analytically provided by the Kalman filter predictive Eq. (S.6), averaged over the entire set of parameters $(\sigma^{(k,l)}, \boldsymbol{\Omega}^{(k,l)}, \boldsymbol{\omega}_{1:t}^{(k,l)})_{l=1}^{L}$ generated by the parallel chains, as follows:

$$\mathsf{E}\left(\boldsymbol{\beta}_{t+1}^{(k)} | \mathbf{y}_{1:t}, \mathbf{x}_{t+1}^{(k)}, \boldsymbol{\beta}_{1:t}^{(k)}\right) \approx \frac{1}{L}\sum_{l=1}^{L}\left(\boldsymbol{\beta}_{t-1|t-2}^{(k,l)} + \mathbf{P}_{t-1|t-2}^{(k,l)} \mathbf{x}_{t-1} V_{t-2}^{(k,l)} \nu_{t-2}^{(k,j)}\right), \tag{22}$$

where all elements of the expression are given in Appendix B.1. The resulting Rao-Blackwellized estimate of the quantile function is more efficient than a simple average over the samples from the full conditional distribution of $\boldsymbol{\beta}_{t+1}^{(k)}$, as argued by (Robert & Casella, 2004, pp. 130–134).

The model-averaged one-step-ahead τ-level quantile prediction is then formed by combining previous forecasts using the predicted model probabilities $\pi_{t+1|t}^{(k)}$, for each competing model $k = 1, 2, \ldots, K$:

$$\widehat{q}_{\tau,t+1}^{\mathsf{DMA}}(\mathbf{x}_{t+1}^{(k)}, \boldsymbol{\beta}_{t+1}^{(k)}) = \sum_{k=1}^{K} \pi_{t+1|t}^{(k)} \widehat{q}_{\tau,t+1}^{(k)}(\mathbf{x}_{t+1}^{(k)}, \boldsymbol{\beta}_{t+1}^{(k)}) \tag{23}$$

$$= \sum_{k=1}^{K} \pi_{t+1|t}^{(k)} \mathbf{x}_{t+1}^{(k)} \widehat{\boldsymbol{\beta}}_{t+1}^{(k)}, \tag{24}$$

where the predicting model probabilities $\pi_{t+1|t}^{(k)}$, for $k = 1, 2, \ldots, K$ have been obtained by averaging single model predictive probabilities defined in Eq. (16) over the L parallel chains:

$$\pi_{t+1|t}^{(k)} = \frac{1}{L}\sum_{l=1}^{L} \pi_{t+1|t}^{(k,l)}, \qquad \forall k = 1, 2, \ldots, K, \tag{25}$$

and Eq. (24) can be obtained by substituting for the definition of the predicted model specific τ-level quantile in Eq. (20) into Eq. (23). The multi-model predictions of the τ-level quantile function of the response variable y_t at time t are obtained through a weighted average of the model-specific τ-level quantile predictions $\widehat{q}_{\tau,t}^{(k)}(\mathbf{x}_t^{(k)}, \boldsymbol{\beta}_t^{(k)})$. These weights correspond to the posterior predictive model probabilities for sample t, $\pi_{t+1|t}^{(k)}$, derived as in Eq. (25). Furthermore, since model predictive probabilities in Eq. (25) are obtained by averaging $\pi_{t+1|t}^{(k,l)}$ over the L independent parallel chains, we can provide approximated confidence intervals

for the estimated probabilities by calculating their Monte Carlo variance as follows:

$$\text{var}\left(\pi_{t+1|t}^{(k)}\right) = \frac{1}{L-1}\sum_{l=1}^{L}\left(\pi_{t|t-1}^{(k,l)} - \pi_{t|t-1}^{(k)}\right)^2. \tag{26}$$

2.3.5 Reducing the computational complexity

Fixed-lag backward sampling is an effective approach to alleviate computational demands, particularly when dealing with a large number of observations. This involves replacing the fixed-interval backward smoothing algorithm outlined in Appendix B.3 with a fixed-lag smoother. The fixed-lag smoothing procedure updates the dynamic latent states within a specified lag $h \leq t$ surrounding the current observation y_t, while keeping the remaining states from 1 to $h-1$ unchanged. At each iteration, the resulting sampling procedure only simulates the quantile regression coefficient $\boldsymbol{\beta}_{t-h|t}$ at time t, for a given lag $h > 0$. For further details, refer to Anderson and Moore (1979) and Simon (2006).

Remark 1 The fixed-lag smoother can be viewed as a form of local-dynamic quantile regression, which extends the non-parametric approach introduced by Yu and Jones (1998).

3 BDQMA efficiency on simulated series

We now focus on simulated examples specifically designed to assess the effectiveness of the SMCMC algorithm and the BDQMA approach in accurately recovering the true parameter values of the underlying Data Generating Process (DGP). These simulations are designed to reflect real-world challenges, such as abrupt and smooth changes in the relationships between dependent variables and covariates. These simulations provide crucial insights into the robustness and adaptability of the methods in handling complex dynamics, further validating their performance in practical scenarios.

Throughout all simulated scenarios, the sample length remains fixed at $T = 200$ observations, aligning with the sample size of the real data examples presented in subsequent sections. The covariates are generated from a uniform distribution on $\left(-\frac{T}{2}, \frac{T}{2}\right)$, denoted as $x_{i,t} \sim \mathsf{U}\left(-\frac{T}{2}, \frac{T}{2}\right)$, i.i.d. for $i = 1, 2, \ldots, M$ and $t = 1, 2, \ldots, T$. The innovation term follows a Gaussian distribution, i.e. $\varepsilon_t \sim \mathsf{N}\left(0, \nu_t^2\right)$, independently for $t = 1, 2, \ldots, T$, with heteroskedastic variance. Furthermore, the true DGP is specified as a time-varying parameters regression model:

$$y_t = \mathbf{x}_t^\top \boldsymbol{\beta}_t^* + \varepsilon_t, \qquad \forall t = 1, 2, \ldots, T, \tag{27}$$

where $M = 2$, $\mathbf{x}_t = \left(1, x_{1,t}, x_{2,t}\right)$ and the quantile regression parameters dynamics $\boldsymbol{\beta}_t^* = \left(\beta_{1,t}^*, \beta_{2,t}^*, \beta_{3,t}^*\right)^\top$ are defined as follows.

3.1 Smooth change in quantiles

We assume the true quantile parameters exhibit the following dynamics: the constant term $\beta_{1,t}^*$ remains fixed at $\beta_1^* = -2.5$, the coefficient $\beta_{2,t}^*$ undergoes a change in slope at a specific time point $t = 100$, and the third parameter $\beta_{3,t}^*$ follows a smooth, sinusoidal transition

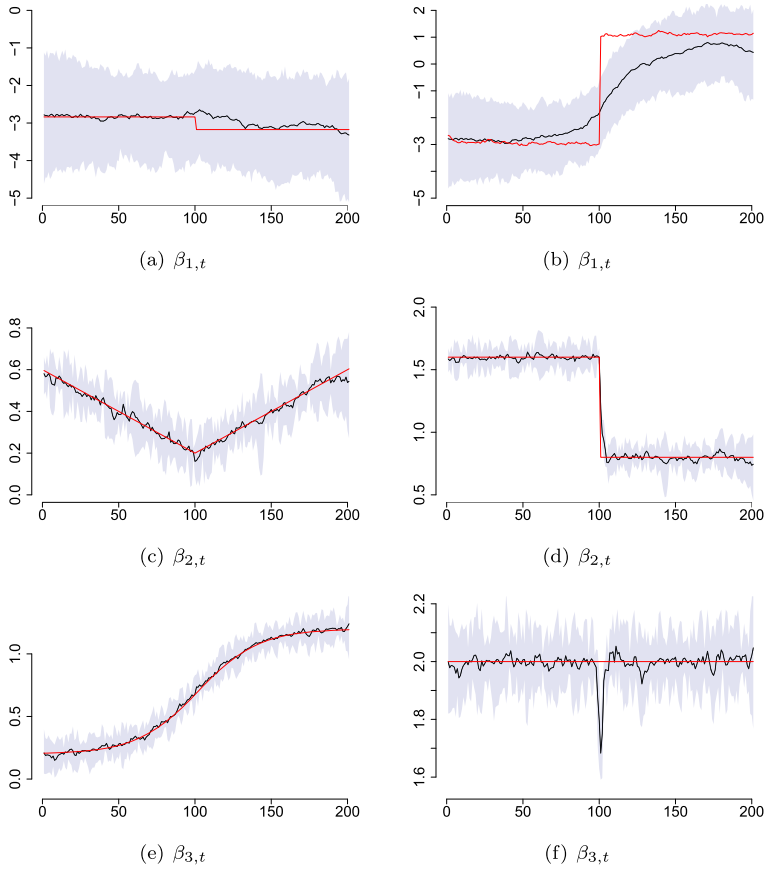


Fig. 2 Posterior mean of the regression parameters $\boldsymbol{\beta}_t = (\beta_{1,t}, \beta_{2,t}, \beta_{3,t})$, $t = 1, 2, \ldots, T$, for the simulated data in Example 3.1 *(left panel)* and 3.2 *(right panel)*, with quantile level $\tau = 0.25$ and $N = 100$ parallel chains. In each plot: true parameters (red), posterior medians (dark) and 95% HPD regions (gray areas)

between two distinct levels:

$$\beta^*_{2,t} = \left(0.6 - \frac{0.4t}{100}\right)\mathbb{1}_{(-\infty,100]}(t) + \left(\frac{0.4t}{100} - 0.2\right)\mathbb{1}_{(100,\infty)}(t)$$

$$\beta^*_{3,t} = a + b\left(1 + \exp\left\{\frac{-c\,(2t - T - 2)}{T}\right\}\right)^{-1}$$

$$v_t^2 = \mathbb{1}_{(-\infty,100]}(t) + 0.25\mathbb{1}_{(100,\infty)}(t),$$

for $t = 1, \ldots, T$, with $a = 0.2$, $b = 1$ and $c = 5$.

3.2 Abrupt change in quantiles

We assume an abrupt change at time $t = 100$ in the constant term $\beta^*_{1,t}$ and the coefficient $\beta^*_{2,t}$, and a GARCH(1,1) dynamics for the conditional volatility of the innovation term v_t:

$$\beta^*_{1,t} = -2\mathbb{1}_{(-\infty,100]}(t) + 2\mathbb{1}_{(100,\infty)}(t)$$

$$\beta^*_{2,t} = 1.6\mathbb{1}_{(-\infty,100]}(t) + 0.8\mathbb{1}_{(100,\infty)}(t)$$
$$\nu_t^2 = a + b\nu_{t-1}^2 + c\varepsilon_{t-1}^2,$$

for $t = 1, \ldots, 200$, and $\beta^*_{3,t} = 2, \forall t = 1, 2, \ldots, T$, with $a = 0.05$, $b = 0.9$ and $c = 0.05$ and $\varepsilon_0 \sim \mathsf{N}(0, \frac{a}{1-b})$. It is worth noting that the relationship between the coefficients of the DGP and the coefficients $\boldsymbol{\beta}_t$ of the quantile regression is as follows: $\beta_{1,t} = \beta^*_{1,t} + \nu_t \Phi^{-1}(\tau)$, $\beta_{2,t} = \beta^*_{2,t}$, and $\beta_{3,t} = \beta^*_{3,t}$, where τ represents the quantile level and $\Phi(\cdot)$ denotes the cumulative density function of a standard Normal distribution. The prior hyperparameter settings are specified as follows: $a_0 = b_0 = 10^{-3}$, $c_0 = M + 1$, and $C_0 = 0.01\mathbf{I}_3$, where $\mathbf{I}_3$ denotes the 3-dimensional identity matrix.

In Fig. 2, we present the sequential estimates for the two simulated examples. The figure illustrates how the quantile regression model effectively captures both abrupt shifts and smooth transitions in the intercept $\beta_{1,t}$ (first row) in presence of both homoskedastic and heteroskedastic observation noise (Panel 2a, b, respectively). Moreover, the second and third rows demonstrate the ability of our BDQMA procedure to accurately detect both abrupt changes (Panel 2c, d) and smooth changes (Panel 2e) in the relationship between the dependent variable and the covariates.

4 Predicting US inflation when causality changes

Inflation has been a central concern for economists, as expectations around it play a critical role in shaping the decisions of economic agents and significantly impact both the economic and social development of nations (Phillips, 1958; Stock and Watson, 1999; Kumar and Wesselbaum, 2024). Our empirical exercise serves two primary objectives: (1) to explore the relevance of covariates beyond the unemployment rate and lagged inflation in predicting current inflation at various quantile levels; and (2) to assess whether the predictors for "high" and "low" inflation rates differ or remain consistent, as reflected by dynamically time-varying inclusion probabilities, $\pi^{(k)}_{t+1|t}$. In this study, we use a slightly modified version of the dataset presented in Koop and Korobilis (2012). Detailed descriptions of the dataset and initial analyses can be found in Appendixes C and D of the supplementary materials.

4.1 A generalized phillips curve

We employ our BDQMA approach to select the best subset of predictors for forecasting US inflation within a generalized Phillips curve framework. Our BDQMA approach is well-suited to address these issues due to its ability to consider dynamically evolving linear relationships between covariates and quantiles of the explained variable. Additionally, it can capture and reproduce structural breaks commonly observed in the evolution of economic variables such as inflation (see also Primiceri, 2005; Koop & Onorante, 2012; Stock and Watson, 2007). Furthermore, focusing on quantiles of the predicted variable aids in identifying periods characterized by different economic conditions, particularly those featuring low or high inflation levels. Building on the conditionally Gaussian representation of the BDQMA model, and drawing inspiration from the seminal works on inflation by Engle (1982) and Bollerslev (1986), we incorporate a conditional heteroskedastic volatility error term. This modification greatly enhances the model's flexibility in capturing asymmetric conditional distributions.

Regarding the model specification, we extend the autoregressive model of order p with exogenous covariates (ARX(p)) developed by Stock and Watson (1999), and previously

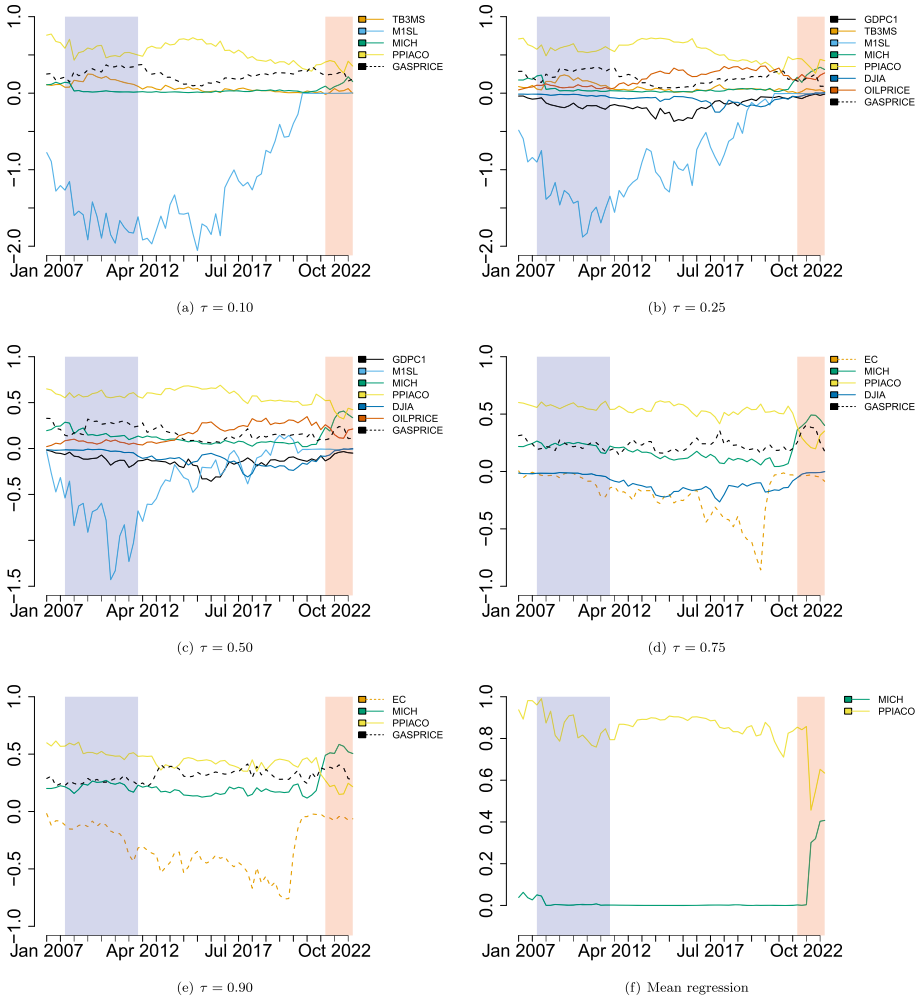


Fig. 3 US inflation data. Sequential estimates of the regression parameters by DMA (for the mean regression **f**) and BDQMA (for the quantile regression **a**–**e**) for the CPIAUCSL (Consumer Price Index for All Urban Consumers: All Items in U.S. City Average). For each quantile level τ the corresponding figure only reports those parameters having inclusion probability larger or equal to 0.7 for at least one quarter. See Table S.11 for a summary of the relevant covariate. The shaded areas identify two significant periods: the US Great Financial Crisis from 2007-Q4 to 2011-Q4 *(blue)*, and the Russian-Ukraine crisis *(red)*

considered by Koop and Korobilis (2012), within our BDQMA framework. The quantile model is therefore formulated as follows:

$$q_\tau(\mathbf{x}_t, \boldsymbol{\beta}) = \mathbf{z}_{t-1}^\top \boldsymbol{\psi} + \sum_{j=1}^{p} \phi_j y_{t-j}, \tag{28}$$

where $\boldsymbol{\beta} = (\boldsymbol{\psi}^\top, \phi_1, \ldots, \phi_p)^\top$ is the parameter vector, y_t represents inflation, defined as the percentage logarithmic change of the price index P_t, that is: $y_t = 100 \log(P_t / P_{t-1})$. The covariate set $\mathbf{x}_t$ includes lagged predictors $\mathbf{z}_{t-1}$, and lagged inflation, $(y_{t-1}, \ldots, y_{t-p})$.

Drawing from standard literature on quantile regression analysis, we concentrate on five quantile levels: $\tau = 0.10$ and $\tau = 0.25$ for lower quantile analyses, $\tau = 0.5$ for median analysis, and $\tau = 0.75$ and $\tau = 0.90$ for upper quantile analyses. These levels correspond to five distinct inflation thresholds at each time step. Due to space constraints, we present in Figs. 3, 4 the time-varying parameters and inclusion probabilities derived from our BDQMA model for all the quantile levels $\tau = (0.10, 0.25, 0.5, 0.75, 0.90)$, focusing solely on the response variable CPIAUCSL. For comparison, we also include Figs. 3f and 4f, which display the results for the Gaussian version of the BDQMA model. This Gaussian counterpart follows the Dynamic Model Averaging (DMA) approach introduced by Raftery et al. (2010), while incorporating the findings of Koop and Korobilis (2013) and Koop and Onorante (2012), adapted to the updated dataset. Similar results have been provided for the response variable CPILFESL are provided in Appendix D.4 of the supplementary materials (Fig. S.10 and S.11). To maintain readability, we include in the figures only those variables relevant for predicting the different conditional quantiles, i.e. those variables with inclusion probabilities exceeding 0.7. For comparison, we present a summary of the coefficient values obtained from a static mean and quantile regressions over the entire sample in Tab. S.4–S.5 and Tab. S.6–S.10 of the Appendix D of the supplementary materials accompanying the paper.

The median analysis provides a robust version of the time-varying regression analysis pioneered by Koop and Korobilis (2013) and Koop and Onorante (2012). Consistent with their findings, our results (see Figs. 3, 4, $\tau = 0.5$) show that coefficients of the regressors and their probabilities of inclusion in the model change over time. Moreover, the level of inflation persistence, indicated by the sum of the autoregressive coefficients (not reported), is approximately 0.5, consistent with the findings of Koop and Korobilis (2013). The primary disparity between mean regression and its robust related median version, extends beyond the quantity of pertinent regressors to elucidate inflation; it also encompasses their efficacy. Indeed, beyond the two regressors shared with the mean regression, i.e. the changes in the NAPM commodities price index (PPIACO) and expectations on future inflation (MICH), the significance of the remaining variables fluctuates over time in response to prevailing economic conditions. For instance, the prominence of oil and gas prices in explaining heightened inflation levels during the recent two years, catalyzed by the Russian-Ukraine crisis, contrasts with their lesser relevance during the Great Financial Crisis. See Appendix D for a detailed description of the results.

The analysis of lower and upper quantiles reveals significant variations in the coefficients of the regressors, which bear important policy implications. Specifically, in the static quantile regression, we observe that the autoregressive terms (ϕ_1, ϕ_2) (not reported) increase as we move towards higher quantile levels. For instance, at low inflation levels (e.g. $\tau = 0.10$, see Tab. S.4), the persistence is approximately 0.24, whereas for high inflation levels (e.g. $\tau = 0.90$, see Tab. S.4), we find evidence of higher persistence, reaching around 0.61. The dynamic regression reveals a similar pattern toward the end of the period (not reported), indicating greater inflation persistence at higher quantile levels. Notably, during the two inflation peaks in 1975 and 1981, the parameter ϕ_2 follows a predominantly negative trajectory, suggesting a reduction in inflation persistence.

The imperative of adopting a time-varying parameters analysis becomes abundantly clear upon scrutiny of Fig. 4. Consider, for instance, if we were to exclusively rely on the findings of a static regression model fitted across the entire sample, meticulously detailed in Tab. S.4. Such an approach might lead us to erroneously conclude that six variables out of the total 18 (EC, HOUST, USPRIV, GS10, T10Y3MM, M1SL, and PMI) hold no relevance across any inflation level at any given point in time (refer to Tab. S.4). While this assertion largely holds true for most scenarios, even under the BDQMA framework, which consistently excludes

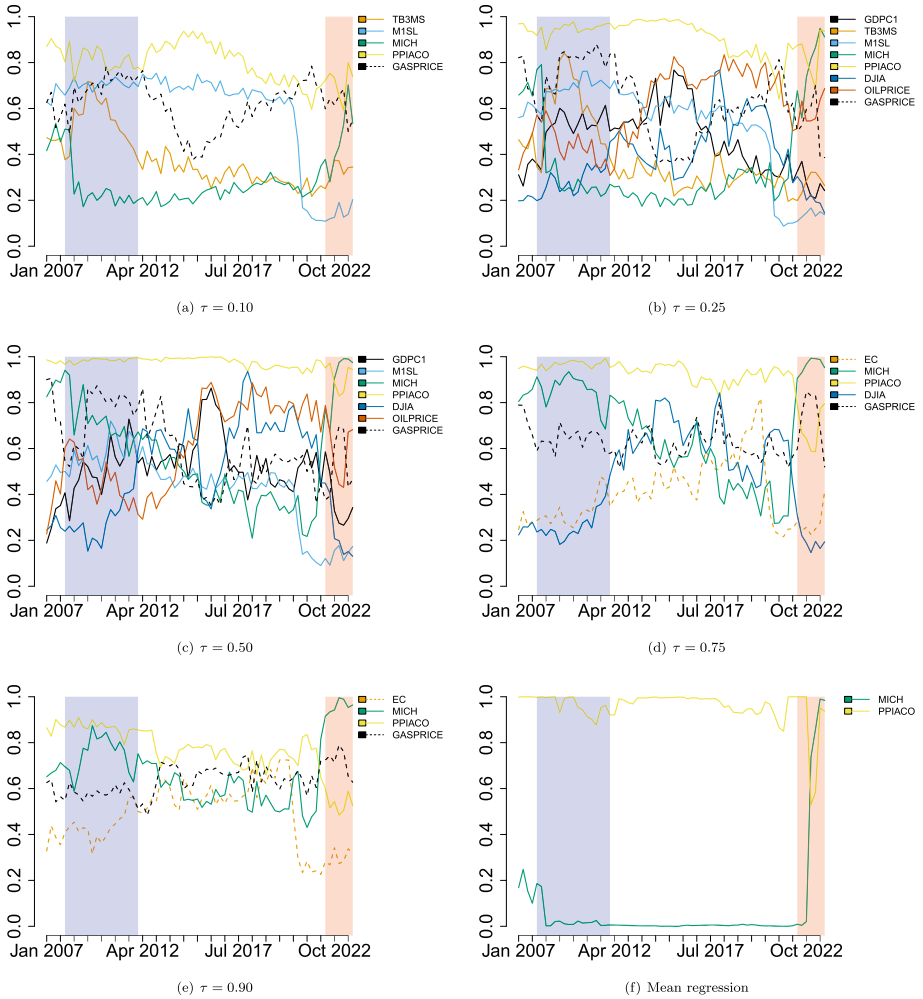


Fig. 4 US inflation dataset. Sequential update of the predicted inclusion probabilities $\pi_{t|t-1}$ by DMA (for the mean regression (**f**)) and BDQMA (for the quantile regression **a**–**e** for the CPIAUCSL (Consumer Price Index for All Urban Consumers: All Items in U.S. City Average). For each quantile level τ the corresponding figure only reports those parameters having inclusion probability larger or equal to 0.7 for at least one quarter. See Table S.11 for a summary of the relevant covariate. The shaded areas identify two significant periods: the US Great Financial Crisis from 2007-Q4 to 2011-Q4 *(blue)*, and the Russian-Ukraine crisis *(red)*

HOUST, USPRIV, GS10, T10Y3MM, and PMI across all quantile levels and time spans (as gleaned from the inclusion probabilities summary statistics in Tab. S.11), the same cannot be said for EC and M1SL. Moreover, it becomes evident that money supply (M1SL) and real personal consumption expenditures (EC) emerge as highly informative for quantiles below the median and the median itself, respectively, and for the quantile exceeding the median. In contrast, variables such as UNRATE, PRFI, T10YFMM, and NAPMSDI, though included in the static regression for at least one quantile, are consistently deemed irrelevant by BDQMA. While the explanatory power of inflation by the NAPM vendor deliveries index remains contentious in the literature, economic theory suggests that the remaining variables—

encompassing investment, the spread of long-term government bonds, the federal funds rate, and the unemployment rate—are pivotal determinants of inflation rates and are typically integrated into the Phillips curve. To shed light on this discrepancy, it is worth noting that for the static quantile regression, UNRATE and PRFI are relevant solely for the lower and two lower quantiles, respectively. However, the interest rate on US short-term federal funds (TB3MS) is included in both the static quantile regression and BDQMA for quantiles ranging from 0.1 to 0.5, thereby tempering the impact of other variables associated with federal funds interest rates. This nuanced analysis helps to unravel the complexity surrounding these variables. While some exhibit sporadic relevance across different regression techniques and quantiles, others, like the federal funds rate, consistently play a moderating role, highlighting the intricate interplay of economic factors influencing inflation dynamics. To conclude our exploration of the relevance of the proposed methodology in understanding the determinants of inflation dynamics, let us juxtapose the results obtained from static and dynamic quantiles for three pivotal variables: the growth rate of real GDP (GDPC1) and energy prices (OILPRICE and GASPRICE). or a detailed discussion of the static quantile regression results, please refer to Appendix D.2. These covariates, as dictated by economic theory, are overwhelmingly relevant and are certainly included in both static and dynamic quantile estimates, albeit for varying quantile levels. However, BDQMA outperforms its static counterpart by a considerable margin in terms of efficacy and accuracy of results. BDQMA not only offers insights into the presence of a relationship between the covariate and the response variable but also illuminates the period during which this relationship is most pronounced. For instance, in the case of energy prices, BDQMA reveals that GDP is influential during periods of low inflation levels, particularly for quantiles $\tau = 0.25$ and $\tau = 0.5$, with the relationship being strongest in the middle of the sample period. Conversely, for energy prices, the association is more robust during the latter part of the sampling period (see Fig. 4).

Regarding the sign and magnitude of the impact of the most relevant regressors, as measured by the estimated coefficients in Fig. 3, the dynamic regression largely corroborates the findings of the static regression for variables such as MICH, PPIACO, and energy prices. However, notable deviations are observed, particularly in the case of TB3MS, where the different impact is clearly evident when examining the results in Tab. S.4. Furthermore, a crucial variable like money supply (M1SL), which was not included in the static regression, displays a consistent negative sign across all periods and quantile levels equal to or below the median. Notably, its impact exhibits an intriguing U shape, underscoring the pronounced negative effect during the Great Financial Crisis. The dynamic regression results also reveals a larger number of significant covariates for most or part of the period. More specifically, some covariates, such as M1SL and TB3SL for instance, are relevant only for the first and last quartile regressions.

4.2 Inclusion probabilities

One of the significant advantages of BDQMA over traditional static model selection methods lies in its dynamic updating of posterior inclusion probabilities $\pi_{t|t-1}$. However, the superiority of BDQMA extends beyond dynamic model selection. Effective model selection procedures are also evaluated based on their ability to accurately describe the characteristics of the response variable while maintaining simplicity. When model selection is conducted for forecasting purposes, simplicity translates into avoiding overfitting and typically results in smaller forecasted confidence intervals. As suggested by Koop and Korobilis (2012) for dynamic model averaging (DMA), if the approach prioritizes models with fewer predictors, it

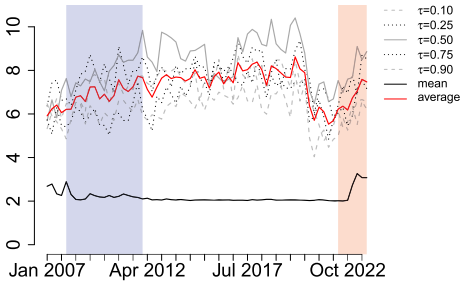


Fig. 5 US Inflation dataset. Expected number of predictors over time for each quantile confidence levels. The quantiles are denoted as follows: $\tau = 0.10$ *(gray dashed line)*, $\tau = 0.25$ *(black dotted line)*, $\tau = 0.50$ *(gray solid line)*, $\tau = 0.75$ *(black dotted line)*, $\tau = 0.90$ *(gray dashed line)*, and the mean *(black solid line)*. The mean (Gaussian regression) is depicted by the black solid line. Additionally, the red line indicates the average number of predictors across all quantiles. The shaded areas identify two significant periods: the US Great Financial Crisis from 2007-Q4 to 2011-Q4 *(blue)*, and the Russian-Ukraine crisis *(red)*

can preserve out-of-sample forecasting performance without compromising goodness-of-fit properties. This emphasizes the importance of parsimony in model selection for forecasting, where simpler models often yield more reliable predictions. The expected number of predictors included by the BDQMA model selection procedure can be analytically evaluated using the predicted inclusion probabilities $\pi^k_{t|t-1}$, for $t = 1, 2, \ldots, T$ and $k = 1, 2, \ldots, K$, such as:

$$\mathsf{E}\left(S^{\tau}_t\right) = \sum_{k=1}^{K} \pi^{(k)}_{t|t-1} S^{\tau,k}_t. \tag{29}$$

The expected number of regressors $\mathsf{E}(S_t)$, where S^k_t denotes the number of regressors included in model $k = 1, 2, \ldots, K$ at each point in time $t = 1, 2, \ldots, T$, provides insight into the average number of predictors included by BDQMA at time t. Figure 5 illustrates $\mathsf{E}(S_t)$ for all the quantile confidence levels $\tau = (0.10, 0.25, 0.50, 0.75, 0.90)$. It is important to note that the average number of predictors takes into account both significant and irrelevant regressors, while Figs. 3, 4 only consider regressors that should be included at least once in the dynamic regressor set. Hence, the expected value $\mathsf{E}(S_t)$ tends to be slightly larger on average than the number of relevant regressors included in Figs. 3, 4. Examining Fig. 5, we observe that the shrinkage increases as the quantile levels deviate further from the median. This observation aligns with intuition, indicating that the number of predictors relevant for explaining higher and lower inflation levels is, on average, lower than those for the median. Furthermore, we notice a noteworthy difference in the expected number of relevant predictors ($\mathsf{E}\left(S^{0.5}_t\right)$) between median and mean regression for the same dataset and period. Specifically, $\mathsf{E}\left(S^{0.5}_t\right)$ is considerably larger for the median regression. This disparity can be attributed to the robustness properties of the median dynamic regression, which underweights observations in the extreme tails, thereby reducing the variance that affects the Kalman filter predicting equations. Consequently, the proposed probability estimates are expected to exhibit greater efficiency compared to those derived from conditional mean regression. It is noteworthy that for all considered quantile levels, $\mathsf{E}\left(S^{0.25}_t\right)$ varies over time and eventually converges to a stable level after a few periods. This convergence underscores the dynamic nature of predictor relevance and highlights the importance of considering temporal dynamics in regression analysis. In conclusion, it is noteworthy that the average number of relevant predictors follows a distinct pattern. It increases, reaching its peak around mid-2019, then begins to decline until

Table 1 Backtesting results for the US inflation dataset. For each quantile level $\tau = (0.1, 0.25, 0.5, 0.75, 0.9)$, the table reports the p-values for three common back-testing methods: the conditional coverage test of Kupiec, (CC), the unconditional coverage test of (UC) and the CaViaR test of Engle and Manganelli (2004) (DQ)

τ	UC	CC	DQ
0.10	0.07	0.11	0.63
0.25	0.28	0.44	0.61
0.50	0.69	0.64	0.30
0.75	0.13	0.23	0.56
0.90	0.07	0.14	0.60

the end of 2022, and subsequently rises again during the Russian-Ukraine crisis for all quantile levels. Conversely, it remains relatively stable for the mean. This observation suggests that the BDQMA method effectively captures the intricate dynamics of the relationships between explanatory variables and the response variable over time, surpassing the capabilities of traditional mean-based approaches.

The flexibility of the BDQMA method results in good forecasting performances. Table 1 presents the outcomes of three common backtesting procedures for one-step ahead quantile forecasts of US inflation data. It is notable that both the conditional coverage and unconditional coverage, as well as the DQ test, fail to reject the null hypothesis of good quantile forecast ability at the 5% significance level.

5 Real estate forecasting with time-varying market conditions

We now turn to a second application where we seek to explain the monthly values of the REIT net-of-S&P500 return from January 1991 to September 2023, building upon the dataset utilized in Ling et al. (2000). The REIT net-of-S&P500 return represents the difference between the monthly NAREIT equity index return and the return to the S&P500 index for that month. A detailed dataset description is provided in Appendix C of the supplementary materials accompanying the paper.

5.1 A time-varying parameter prediction model

In this analysis, we employ the BDQMA model, e.g. a quantile-based dynamic regression approach which substantially differs in several aspects from the static mean-regression analysis with macro fundamental and financial variables utilized by Ling et al. (2000). This dynamic quantile regression method allows us to continuously update the sets of relevant regressors for predicting future REIT levels. While Ling et al. (2000) employed a combination of a rolling-window approach and a stochastic search variable selection method, our approach offers a more dynamic and flexible framework for analyzing the relationship between predictors and REIT returns.

Our analysis extends the work of Ling et al. (2000) in two key aspects. Firstly, we investigate the predictability of moderately large/low REIT returns. Secondly, we examine the median return as a function of the same set of explanatory variables as utilized by Ling et al. (2000). Intuitively, we anticipate that the determinants of high/low returns would exhibit broad similarities, yet substantial differences across various sub-periods due to evolving

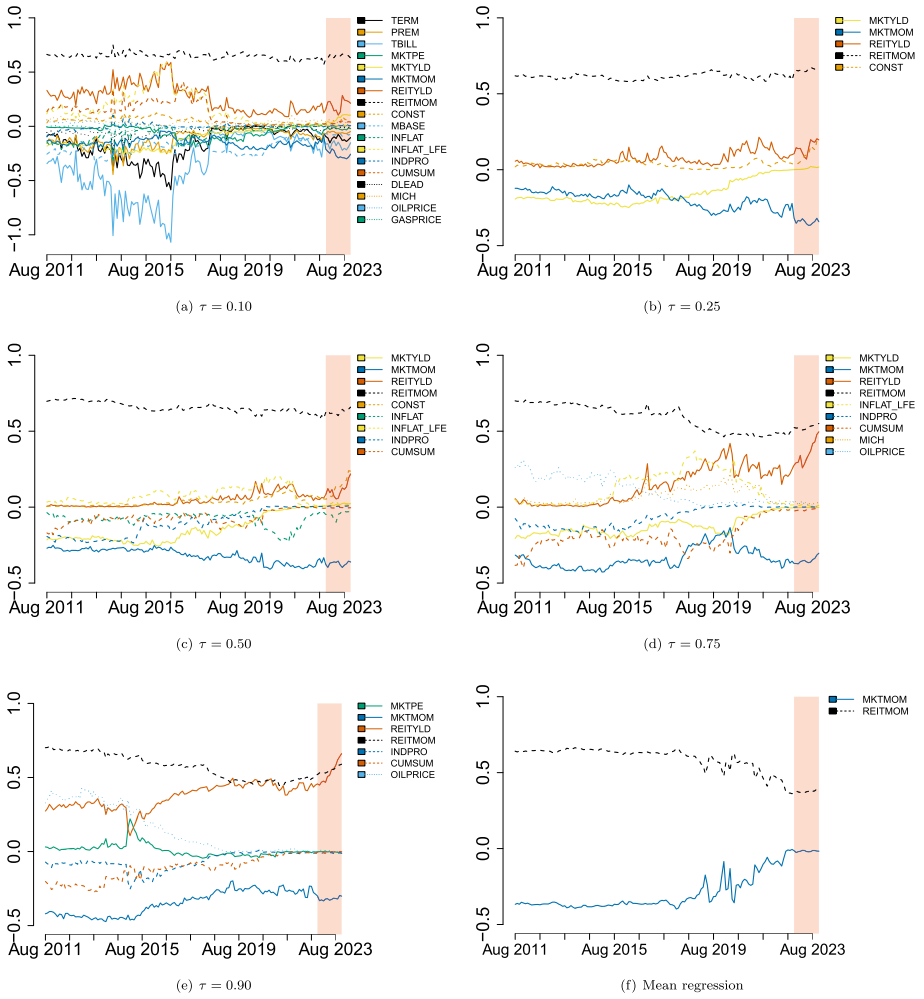
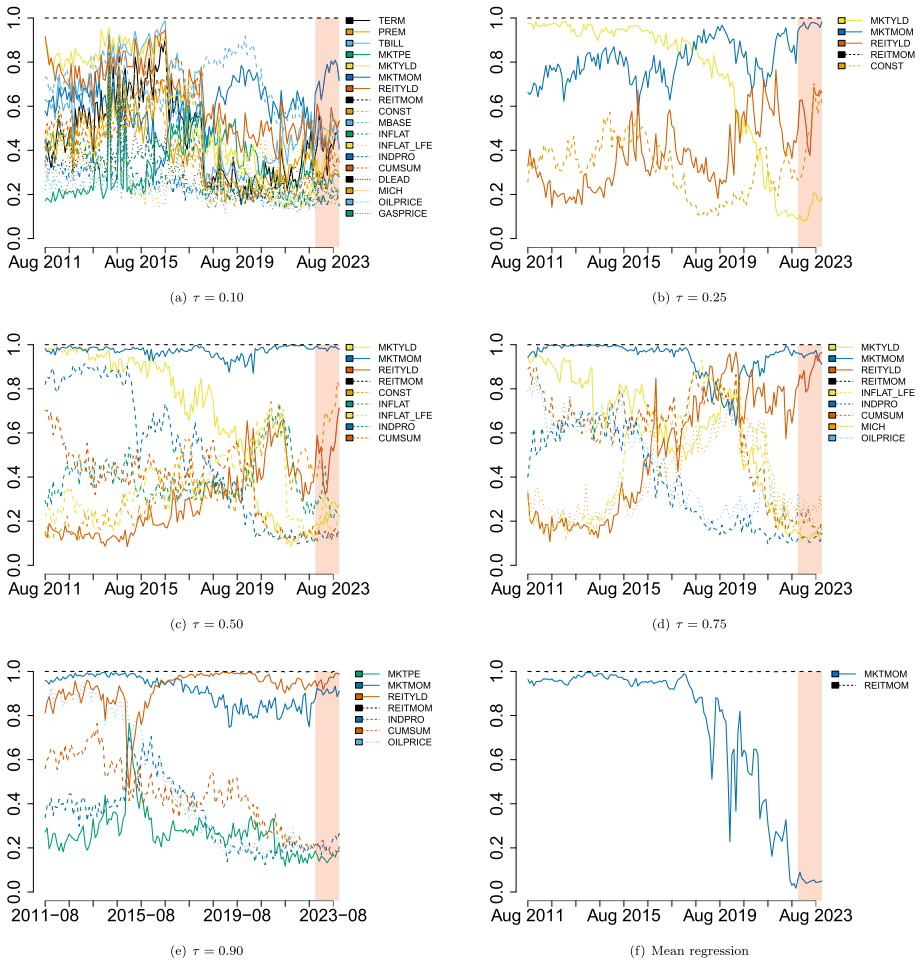


Fig. 6 US real estate data. Sequential estimates of the regression parameters by DMA (for the mean regression **f**) and BDQMA (for the quantile regression **a**–**e** for the REITMKT (monthly NAREIT equity return in excess of the monthly return on the S&P 500 stock index). For each quantile level τ the corresponding figure only reports those parameters having inclusion probability larger or equal to 0.7 for at least one quarter. See Table S.11 for a summary of the relevant covariate. The shaded area identifies the period from 2022-02 to 2023-10 of the Russian-Ukraine crisis

market conditions, underlying mechanisms, and economic forces. Indeed, our approach is inherently more robust and dynamic. In our BDQMA framework, the inclusion probabilities of relevant variables are continuously updated over time, and the parameters of the regressors follow a latent dynamic process. This dynamic updating of inclusion probabilities at different quantile levels offers insights into how evolving economic conditions influence housing bubble bursts, a phenomenon commonly observed in the market. The quantile regression model we employ is defined in Eq. (28), with lagged endogenous variables excluded from the set of covariates. This exclusion helps ensure the independence of current observations from past values, maintaining the integrity of the analysis.



(a) $\tau = 0.10$ (b) $\tau = 0.25$

(c) $\tau = 0.50$ (d) $\tau = 0.75$

(e) $\tau = 0.90$ (f) Mean regression

Fig. 7 US real estate data. Sequential update of the predicted inclusion probabilities $\pi_{t|t-1}$ by DMA (for the mean regression f12f) and BDQMA (for the quantile regression f12a–e) for the REITMKT (monthly NAREIT equity return in excess of the monthly return on the S&P 500 stock index). For each quantile level τ the corresponding figure only reports those parameters having inclusion probability larger or equal to 0.7 for at least one quarter. See Table S.11 for a summary of the relevant covariate. The shaded area identifies the period from 2022-02 to 2023-10 of the Russian-Ukraine crisis

The outcomes of the initial static quantile model applied across the complete dataset are showcased in Table S.13, Appendix E. Upon examination, it becomes apparent that although numerous variables are pertinent for elucidating high returns, as anticipated, only a select few delineate the conditional quantile above median confidence levels. Additionally, it is noteworthy that only three variables, namely MKTYLD, MKTMOM, and REITMIN, influence all quantiles. A visual inspection of Fig. 6 reveals compelling evidence of model variations over time across all confidence levels. For instance, TBILL, which gauges the opportunity cost of real estate investment, notably contributes to the prediction of lower real estate returns surpassed only during the period spanning from late 2012 to early 2018 only for the lowest quantile level ($\tau = 0.1$), after which its relevance diminishes. This suggests a

temporal aspect to its influence on real estate returns. At times, however, we observe that the inclusion probability of a specific regressor exhibits a dynamic pattern that remains largely consistent as we transition from lower to upper quantiles. This pattern is notably observed with the two most influential regressors (REITMOM, MKTMOM), which consistently appear across all quantile levels. These variables are also featured in the static quantile regression in Tab. S.13 across all quantile levels, indicating their importance in explaining real estate return dynamics.

From Fig. S.12, it becomes evident that certain variables, such as INFLAT (which measures the impact of inflation), are included as predictors solely for $\tau = 0.50$, indicating their relevance primarily in scenarios where real estate returns are moderate. Conversely, INFLAT_LFE (which measures inflation excluding food and energy prices) emerges as significant in predicting high real estate returns, even extending to the third quartile ($\tau = 0.75$). This observation suggests that accounting for price dynamics, particularly those related to essential commodities, may not be crucial for forecasting future house price levels. This finding prompts further inquiry into the underlying drivers of real estate returns and the factors that shape their dynamics. It raises questions about the relative importance of macroeconomic indicators versus specific inflation components in influencing housing market performance. Moreover, it underscores the nuanced nature of predictive modeling in real estate economics, where the inclusion of certain variables may vary across different quantiles, reflecting distinct market conditions and dynamics. Overall, these insights contribute to a deeper understanding of the complex relationship between economic factors and real estate returns, informing more nuanced and effective forecasting models. By discerning the varying impacts of different predictors across quantiles, researchers and policymakers can refine their strategies for mitigating risks and maximizing opportunities in the dynamic real estate market landscape.

5.2 Inclusion probabilities

Figure 7 illustrates the predicted inclusion probabilities $\pi_{t|t-1}$ of the regressors across various quantile levels. The dynamic progression of these probabilities elucidates the significance of variables in forecasting future quantile levels of real estate returns at each time point t. Figure 7 exclusively portrays those regressors deemed relevant at least once over time, indicated by a posterior inclusion probability of at least 0.7 for some date t. Unfortunately, Fig. 7 contains an excessive number of regressors and considerable noise, making it challenging to discern meaningful patterns. Therefore, we have chosen to present a smoothed version in Fig. S.12 in Appendix E of the supplementary materials accompanying the paper, where each inclusion probability has undergone local polynomial regression to mitigate noise and improve interpretability.

Interestingly, two significant variables in Table S.13, REITYLD and MKTYLD, demonstrate inclusion probability dynamics characterized by a sinusoidal shape with comparable amplitude and frequency across quantiles, transitioning to lower levels as the quantile increases. This sinusoidal behavior suggests a cyclical influence on real estate returns, warranting further investigation into its underlying drivers. Overall, the smoothed version of the inclusion probabilities provides a clearer understanding of the evolving dynamics of regressors in predicting real estate returns across different quantile levels, shedding light on both consistent influences and temporal variations in model predictors. Moreover, the percentage change in the monetary base (MBASE), which is also linked to inflation, exhibits low levels of inclusion probability across all quantile confidence levels. This suggests that the price dynamics are largely independent of the real estate market.

Another significant finding is that, BDQMA captured both gradual and abrupt changes in inclusion probabilities. For example, there is an abrupt change in the inclusion probability of the monetary base CUMSUM (which accounts for the growth rate of consumption expenditures for nondurable goods) from nearly 0.2 to 0.8 for $\tau = 0.75$. However, it is worth noting that there are also many instances where the inclusion probability of a specific predictor evolves smoothly over time. Our findings underscore the importance of adaptive policymaking strategies that are responsive to changing economic dynamics. Ultimately, the dynamic selection of relevant predictors enhances the accuracy of real estate return forecasts and provides invaluable insights for policymakers seeking to optimize their interventions in alignment with evolving economic conditions.

6 Conclusion

In this article, we propose a new Bayesian Dynamic Quantile Model Averaging (BDQMA) approach, which for the first time to our best knowledge, combines in a tractable way Bayesian time-varying quantile regressions and dynamic model averaging. The approach dynamically accounts for model risk and parameters uncertainty in the quantiles of the response variable in presence of time-varying features such as heteroskedasticity and volatility clustering, non-linearities, breaks and jumps, leptokurticity and unconditional non-normality. These input-data features represent a major challenge when predicting the variable of interest. The tractability of BDQMA relies on an original combination of the location-scale Gaussian mixture representation of the quantile regression error terms, a conditionally Normal state-space representation of the time-varying parameter model and a sequential Monte Carlo Markov chain sampler for posterior approximation. We demonstrate that our proposed combination of transition and jumping kernel, based on a series of Kalman filtering and smoothing steps, enables convergence to the posterior distribution. This framework is further enriched by incorporating an additional layer over the discrete space of regressors' combinations. While this resembles a conventional batch processing approach, it is typically impractical in its classical form, especially for small samples.

After rigorous testing with various simulation experiments, our approach has demonstrated effectiveness across different real-world data characteristics such as abrupt and smooth changes in the relationship between response and explanatory variables, in presence of heteroschedasticity in the obervation errors. We illustrate the potential of BDQMA through an application to predicting US inflation and real estate returns. In the two applications we found evidence of disparity between mean regression and its robust median version. Indeed, beyond the disparity in the coefficients of some regressors usually considered in the literature, the impact of energy prices became significant in the recent years in explaining low and high quantiles of inflation and real estates returns. Furthermore, comparing the results obtained from static quantiles with those from BDQMA, for most of the relevant variables, the significance, the magnitude and the sign of their coefficients fluctuate over time in response to prevailing economic conditions. When compared with the mean results, BDQMA also returns a larger number of significant covariates for a large part of the sample period. This disparity can be attributed to the robustness properties of BDQMA, which dynamically underweights observations in the extreme tails. The flexibility of the BDQMA approach results in good forecasting performances as confirmed by standard backtesting procedures.

Finally, we may think about several extensions of our study. Relating to the method, nontrivial interesting extensions include modelling and forecasting with multiple frequencies

(Eric Ghysels and Valkanov, 2007; Candila et al., 2023), horizons (Shackleton et al., 2010; Aastveit et al., 2014), and quantiles (Meng and Taylor, 2020; Taylor, 2022; Merlo et al., 2021; Zhang et al., 2023). Furthermiore, BDQMA can be combined with an ex-ante "haircut" correction (Boucher et al., 2014; Lazar and Zhang, 2019), in order to correct on an a priori basis the output estimates in a Bayesian setting. Finally, Machine Learning techniques can be used to embed textual media-news data (Huang et al., 2023) or to enhance the predictive power of the models (Medeiros et al., 2021; Barkan et al., 2023; Lenza et al., 2023a, b). Regarding the applications, other geographical areas, such as Europe, or markets, such as the bond and stock markets, can be considered. Tackling the problem of predicting key economic variables, namely for instance probability of financial distress and bankruptcies (Huang et al., 2023) might be also of interest.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10479-024-06378-7.

Acknowledgements We thank conference and seminar participants at the: SIRE Econometrics Workshop (Glasgow, 2014), the 8th CSDA International Conference on Computational and Financial Econometrics (Pisa, 2014), the 6th Italian Congress of Econometrics and Empirical Economics, and the FEM2022 Conference (Paris, 2022). An early-stage version of this article previously circulated under the title "Dynamic Model Averaging for Bayesian Quantile Regression". We also thank Jean-Luc Prigent for his positive comments and encouragement on the preliminary draft of this article. The usual disclaimer applies.

Funding This study was funded by the MUR— PRIN project "Discrete random structures for Bayesian learning and prediction" under g.a. n. 2022CLTYP4; the Next Generation EU— "GRINS—Growing Resilient, INclusive and Sustainable" project (PE0000018), National Recovery and Resilience Plan (NRRP)— PE9—Mission 4, C2, Intervention 1.3; the BERN BIRD2222 01—BIRD 2022 grant from the University of Padua, and by the European Union - Next Generation EU, Mission 4 Component 2—CUP C53D23002580006 via the MUR-PRIN grant 2022SMNNKY. The views and opinions expressed are only those of the authors and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them. This research used the HPC-VERA multiprocessor cluster system at Ca' Foscari University of Venice.

Declarations

Conflict of interest Authors declare that they have no Conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

Aastveit, K. A., Gerdrup, K. R., Jore, A. S., & Thorsrud, L. A. (2014). Nowcasting GDP in real-time: A density combination approach. *Journal of Business Economics & Statistics, 32*, 48–68.

Aastveit, K. A., Ravazzolo, F., & van Dijk, H. K. (2018). Combined density nowcasting in an uncertain economic environment. *Journal of Business Economics & Statistics, 36*(1), 131–145.

Alhamzawi, R., & Yu, K. (2012). Variable selection in quantile regression via Gibbs sampling. *Journal of Applied Statistics, 39*(4), 799–813.
Amédée-Manesme, C., & Barthélémy, F. (2018). Ex-ante real estate value at risk calculation method. *Annals of Operation Research, 262*, 257–285.
Anderson, B., & Moore, J. (1979). *Optimal filtering*. Prentice-Hall.
Barkan, Oren, Benchimol, Jonathan, Caspi, Itamar, Cohen, Eliya, Hammer, Allon, & Koenigstein, Noam. (2023). Forecasting CPI inflation components with Hierarchical recurrent neural networks. *International Journal of Forecasting, 39*(3), 1145–1162. https://doi.org/10.1016/j.ijforecast.2022.04.009
Bassetti, Federico, Casarin, Roberto, & Ravazzolo, Francesco. (2018). Bayesian nonparametric calibration and combination of predictive distributions. *Journal of the American Statistical Association, 113*(522), 675–685. https://doi.org/10.1080/01621459.2016.1273117
Behr, A. (2010). Quantile regression for robust bank efficiency score estimation. *European Journal of Operational Research, 200*(2), 568–581.
Bellini, F., Koch-Medina, P., Munari, C., & Svindland, G. (2021). Law-invariant functionals that collapse to the mean. *Insurance: Mathematics and Economics, 98*, 83–91.
Bellini, F., Rroji, E., & Sala, C. (2022). Implicit quantiles and expectiles. *Annals of Operation Research, 313*, 733–753.
Belmonte, M., & Koop, G. (2014). Model switching and model averaging in time-varying parameter regression models. *Advances in Econometrics, 34*, 45–69.
Ben Ameur, H., Ftiti, Z., Jawadi, F., & Louhichi, W. (2022). Measuring extreme risk dependence between the oil and gas markets. *Annals of Operation Research, 313*(2), 755–772.
Ben Bouheni, F., Obeid, H., & Margarint, E. (2022). Nonperforming loan of European Islamic banks over the economic cycle. *Annals of Operation Research, 313*(2), 773–808.
Benkraiem, R., & Zopounidis, C. (2021). Preface: Regression methods based on OR techniques and computational aspects in management, economics and finance. *Annals of Operations Research, 306*(1), 1–6.
Bernardi, M., Gayraud, G., & Petrella, L. (2015). Bayesian tail risk interdependence using quantile regression. *Bayesian Analysis, 10*(3), 553–603.
Billio, M., Casarin, R., Ravazzolo, F., & Van Dijk, H. (2013). Time-varying combinations of predictive densities using nonlinear filtering. *Journal of Econometrics, 177*, 213–232.
Billio, M., Casarin, R., Ravazzolo, F., & van Dijk, H. K. (2012). Combination schemes for turning point predictions. *The Quarterly Review of Economics and Finance, 52*(4), 402–412.
Bollerslev, T. (1986). Generalized autoregressive conditional heteroskesdaticity. *Journal of Econometrics, 31*, 307–327.
Bonaccolto, G., Caporin, M., & Maillet, B. (2022). Dynamic large financial networks via conditional expected shortfalls. *European Journal of Operational Research, 298*(1), 322–336.
Boucher, C. M., Daníelsson, J., Kouontchou, P. S., & Maillet, B. B. (2014). Risk models-at-risk. *Journal of Banking & Finance, 44*, 72–92.
Busetti, F. (2017). Quantile aggregation of density forecasts. *Oxford Bulletin of Economics and Statistics, 79*(4), 495–512.
Candila, V., Gallo, G. M., & Petrella, L. (2023). Mixed-frequency quantile regressions to forecast value-at-risk and expected shortfall. *Annals of Operations Research*, pp 1–34.
Casarin, R., Chang, C.-L., Jimenez-Martin, J.-A., McAleer, M., & Pérez-Amaral, T. (2013). Risk management of risk under the Basel accord: A Bayesian approach to forecasting Value-at-Risk of VIX futures. *Mathematics and Computers in Simulation, 94*, 183–204.
Casarin, R., Grassi, S., Ravazzolo, F., & van Dijk, H. K. (2023). A flexible predictive density combination for large financial data sets in regular and crisis periods. *Journal of Econometrics, 237*(2), 1–12.
Chalamandaris, G., & Vlachogiannakis, N. (2018). Are financial ratios relevant for trading credit risk? Evidence from the CDS market. *Annals of Operation Research, 266*, 395–440.
Chen, C. W., Dunson, D. B., Reed, C., & Yu, K. (2013). Bayesian variable selection in quantile regression. *Statistics and its Interface, 6*(2), 261–274.
Chun, S. Y., Shapiro, A., & Uryasev, S. (2012). Conditional value-at-risk and average value-at-risk: Estimation and asymptotics. *Operations Research, 60*(4), 739–756.
Durbin, J., & Koopman, S. (2012). *Time series analysis by state space methods*. Oxford University Press.
Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica, 50*(4), 987–1007.
Engle, R., & Manganelli, S. (2004). Caviar: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics, 22*, 367–381.
Eric Ghysels, A. S., & Valkanov, R. (2007). Midas regressions: Further results and new directions. *Econometric Reviews, 26*(1), 53–90.

Fawcett, N., Kapetanios, G., Mitchell, J., & Price, S. (2015). Generalised density forecast combinations. *Journal of Econometrics, 188*, 150–165.
George, E., & McCulloch, R. E. (1993). Variable selection via Gibbs sampling. *Journal of the American Statistical Association, 88*(423), 881–889.
Gneiting, T., & Ranjan, R. (2013). Combining predictive distributions. *Electronic Journal of Statistics, 7*, 1747–1782.
Guhaniyogi, R., Qamar, S., & Dunson, D. B. (2018). Bayesian conditional density filtering. *Journal of Computational and Graphical Statistics, 27*(3), 657–672.
Hall, S. G., & Mitchell, J. (2007). Combining density forecasts. *International Journal of Forecasting, 23*, 1–13.
Harvey, A. C. (1989). *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press.
Huang, B., Yao, X., Luo, Y., & Li, J. (2023). Improving financial distress prediction using textual sentiment of annual reports. *Annals of Operations Research, 330*(1), 457–484.
Huang, Q., Zhang, H., Chen, J., & He, M. (2017). Quantile regression models and their applications: A review. *Journal of Biometrics & Biostatistics, 8*(3), 1–6.
Ji, Y., Lin, N., & Zhang, B. (2012). Model selection in binary and tobit quantile regression using the Gibbs sampler. *Computational Statistics and Data Analysis, 56*(4), 827–839.
Jones, G. L., & Hobert, J. P. (2004). Sufficient burn-in for Gibbs samplers for a hierarchical random effects model. *Annals of Statistics, 32*(2), 784–817.
Kim, M. O. (2007). Quantile regression with varying coefficients. *Annals of Statistics, 35*, 92–108.
Koenker, R. (2005). *Quantile regression*. Cambridge University Press.
Koenker, R., & Basset, J. (1978). Regression quantiles. *Econometrica, 46*, 33–50.
Koop, G., & Korobilis, D. (2012). Forecasting inflation using dynamic model averaging. *International Economic Review, 53*(3), 867–886.
Koop, G., & Korobilis, D. (2013). Large time-varying parameter VARs. *Journal of Econometrics, 177*(2), 185–198.
Koop, G. & Onorante, L. (2012). Estimating Phillips curves in turbulent times using the ECB's survey of professional forecasters. Technical report, Working paper series, N. 1422, European Central Bank.
Koop, G., & Tole, L. (2013). Forecasting the European carbon market. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 176*(3), 723–741.
Kozumi, H., & Kobayashi, G. (2011). Gibbs sampling methods for Bayesian quantile regression. *Journal of Statistical Computation and Simulation, 81*, 1565–1578.
Kumar, S., & Wesselbaum, D. (2024). Contracts and firms' inflation expectations. *The Review of Economics and Statistics, 106*(1), 246–255.
Lazar, E., & Zhang, N. (2019). Model risk of expected shortfall. *Journal of Banking & Finance, 105*, 74–93.
Lee, E. R., Noh, H., & Park, B. U. (2014). Model selection via Bayesian information criterion for quantile regression models. *Journal of the American Statistical Association, 109*(505), 216–229.
Lenza, M., Moutachaker, I., & Paredes, J. (2023a). Density forecasts of inflation: A quantile regression forest approach. Working Paper Series 2830, European Central Bank.
Lenza, M., Moutachaker, I., & Paredes, J. (2023b). Forecasting Euro area inflation with machine-learning models. *Research Bulletin*, 112.
Lin, E., Sun, E., & Yu, M. (2018). Systemic risk, financial markets, and performance of financial institutions. *Annals of Operation Research, 262*, 579–603.
Ling, D. C., Naranjo, A., & Ryngaert, M. D. (2000). The predictability of equity REIT returns: Time variation and economic significance. *Journal of Real Estate Finance and Economics, 20*(2), 117–136.
Liu, J. (1994). The collapsed Gibbs sampler in Bayesian computations with applications to a gene regulation problem. *Journal of the American Statistical Association, 89*, 958–966.
Lum, K., & Gelfand, A. (2012). Spatial quantile multiple regression using the asymmetric Laplace process. *Bayesian Analysis, 7*, 1–24.
Mamatzakis, E. C., & Tsionas, M. G. (2021). Making inference of British household's happiness efficiency: A Bayesian latent model. *European Journal of Operational Research, 294*(1), 312–326.
Marin, J.-M., & Robert, C. (2007). *Bayesian core: A practical approach to computational statistics*. Springer.
McAlinn, K., & West, M. (2019). Dynamic Bayesian predictive synthesis in time series forecasting. *Journal of Econometrics, 210*, 155–169.
McCormick, T. H., Raftery, A. E., Madigan, D., & Burd, R. S. (2012). Dynamic logistic regression and dynamic model averaging for binary classification. *Biometrics, 68*(1), 23–30.
Medeiros, M. C., Vasconcelos, G. F., Veiga, Á., & Zilberman, E. (2021). Forecasting inflation in a data-rich environment: The benefits of machine learning methods. *Journal of Business & Economic Statistics, 39*(1), 98–119.

Meligkotsidou, L., Vrontos, I., & Vrontos, S. (2009). Quantile regression analysis of hedge fund strategies. *Journal of Empirical Finance, 16*(2), 264–279.
Meng, X., & Taylor, J. W. (2020). Estimating value-at-risk and expected shortfall using the intraday low and range data. *European Journal of Operational Research, 280*(1), 191–202.
Merlo, L., Petrella, L., & Raponi, V. (2021). Forecasting VaR and ES using a joint quantile regression and its implications in portfolio allocation. *Journal of Banking & Finance, 133*, 106248.
Meyn, S. P., & Tweedie, R. L. (1993). *Markov chains and stochastic stability*. Communications and Control Engineering Series. Springer-Verlag, London Ltd., London.
Noh, H., Chung, K., & Van Keilegom, I. (2012). Variable selection of varying coefficient models in quantile regression. *Electronic Journal of Statistics, 6*, 1220–1238.
Papaspiliopoulos, O., & Roberts, G. (2008). Stability of the Gibbs sampler for Bayesian hierarchical models. *Annals of Statistics, 36*(1), 95–117.
Park, T., & Van Dyk, D. (2009). Partially collapsed Gibbs samplers: Illustrations and applications. *Journal of Computational and Graphical Statistics, 18*, 528–550.
Phillips, A. W. (1958). The relation between unemployment and the rate of change of money wage rates in the United Kingdom, 1861–19571. *Economica, 25*(100), 283–299.
Pisarenko, V., & Lyubushin, A. (1997). Statistical estimation of maximum peak ground acceleration at a given point of a seismic region. *Journal of Seismology, 1*(4), 395–405.
Primiceri, G. (2005). Time varying structural vector auto regressions and monetary policy. *Review of Economic Studies, 72*, 821–852.
Raftery, A., Karny, M., & Ettler, P. (2010). Online prediction under model uncertainty via dynamic model averaging: Application to a cold rolling mill. *Technometrics, 52*, 52–66.
Risse, M., & Kern, M. (2016). Forecasting house-price growth in the Euro area with dynamic model averaging. *The North American Journal of Economics and Finance, 38*, 70–85.
Robert, C., & Casella, G. (2004). *Monte Carlo Statistical Methods*. Springer-Verlag.
Rockafellar, R. T. & Royset, J. O. (2013). Superquantiles and their applications to risk, random variables, and regression. In *Theory Driven by Influential Applications*, chapter Chapter 8, pages 151–167. INFORMS.
Román, J. C., & Hobert, J. P. (2012). Convergence analysis of the gibbs sampler for bayesian general linear mixed models with improper priors. *Annals of Statistics, 40*(6), 2823–2849.
Shackleton, M. B., Taylor, S. J., & Yu, P. (2010). A multi-horizon comparison of density forecasts for the S&P 500 using index returns and option prices. *Journal of Banking & Finance, 34*(11), 2678–2693.
Simon, D. (2006). *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. Wiley.
Steel, M. F. (2020). Model averaging and its use in economics. *Journal of Economic Literature, 58*(3), 644–719.
Stock, J., & Watson, M. (1999). Forecasting inflation. *Journal of Monetary Economics, 44*, 293–335.
Stock, J., & Watson, M. (2007). Why has U.S. inflation become harder to forecast? *Journal of Monetary Credit and Banking, 39*, 3–33.
Stone, M. (1961). The linear pool. *Annals of Mathematical Statistics, 2*, 1339–1342.
Taylor, J. W. (2020). Forecast combinations for value at risk and expected shortfall. *International Journal of Forecasting, 36*(2), 428–441.
Taylor, J. W. (2022). Forecasting value at risk and expected shortfall using a model with a dynamic omega ratio. *Journal of Banking & Finance, 140*, 106519.
Tian, H., Yim, A., & Newton, D. P. (2020). Tail-heaviness, asymmetry, and profitability forecasting by quantile regression. *Management Science, 67*(8), 1–30.
Tsionas, M. G. (2020). Quantile stochastic frontiers. *European Journal of Operational Research, 282*(3), 1177–1184.
Van Dyk, D., & Park, T. (2008). Partially collapsed Gibbs samplers: Theory and methods. *Journal of the American Statistical Association, 110*, 790–796.
Wang, L., & Wu, Y. (2012). Quantile regression for analyzing heterogeneity in ultra-high dimension. *Journal of the American Statistical Association, 107*, 214–222.
Wu, Y., & Liu, Y. (2009). Variable selection in quantile regression. *Statistica Sinica, 19*, 801–817.
Yang, Y. & Dunson, D. B. (2013). Sequential Markov chain Monte Carlo. arXiv preprint arXiv:1308.3861.
Yu, K., & Jones, M. C. (1998). Local linear quantile regression. *Journal of the American Statistical Association, 93*(441), 228–237.
Zhang, N., Su, X., & Qi, S. (2023). An empirical investigation of multiperiod tail risk forecasting models. *International Review of Financial Analysis, 86*, 102498.