Federal Reserve Bank of Cleveland Working Paper Series

Constructing Density Forecasts from Quantile Regressions: Multimodality in Macro-Financial Dynamics

James Mitchell, Aubrey Poon, and Dan Zhu

Working Paper No. 22-12

May 2022

Suggested citation: Mitchell, James, Aubrey Poon, and Dan Zhu. 2022. "Constructing Density Forecasts from Quantile Regressions: Multimodality in Macro-Financial Dynamics." Working Paper No. 22-12. Federal Reserve Bank of Cleveland. https://doi.org/10.26509/frbc-wp-202212.

Federal Reserve Bank of Cleveland Working Paper Series
ISSN: 2573-7953

Constructing Density Forecasts from Quantile Regressions: Multimodality in Macro-Financial Dynamics*

James Mitchell†, Aubrey Poon‡, and Dan Zhu§

April 25, 2022

Abstract

Quantile regression methods are increasingly used to forecast tail risks and uncertainties in macroeconomic outcomes. This paper reconsiders how to construct predictive densities from quantile regressions. We compare a popular two-step approach that fits a specific parametric density to the quantile forecasts with a nonparametric alternative that lets the "data speak." Simulation evidence and an application revisiting GDP growth uncertainties in the US demonstrate the flexibility of the nonparametric approach when constructing density forecasts from both frequentist and Bayesian quantile regressions. They identify its ability to unmask deviations from symmetrical and unimodal densities. The dominant macroeconomic narrative becomes one of the evolution, over the business cycle, of multimodalities rather than asymmetries in the predictive distribution of GDP growth when conditioned on financial conditions.

JEL Code: C53; E32; E37; E44

Keywords: Density Forecasts; Quantile Regressions; Financial Conditions

*Thanks to Aaron Amburgey, Todd Clark, Gergely Ganics, Domenico Giannone, Gary Koop, Ed Knotek, Mike McCracken, Tatevik Sekhposyan, Benjamin Wang, and participants at the ISF 2021, SNDE 2022, and an IIF MacroFor seminar for helpful comments. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Cleveland or the Federal Reserve System.

†Federal Reserve Bank of Cleveland; james.mitchell@clev.frb.org

‡Örebro University; aubrey.poon@oru.se

§Monash University; Dan.Zhu@monash.edu

1 Introduction

Recent research has used quantile regression (QR) methods both to produce density nowcasts and forecasts of macroeconomic and financial variables and to assess tail risks, emphasizing asymmetries in the distribution of (real) GDP growth when conditioned on financial conditions.[1] A commonly adopted approach in this literature, following Adrian et al. (2019) [henceforth ABG], is to produce the density forecasts in two steps. As a first step, the QRs are estimated. This means that the underlying conditional density is defined only at the chosen quantiles (typically four quantiles are chosen). As a result, as a second step, the skewed-t density function of Azzalini and Capitanio (2003) is fitted to these quantile forecasts by minimizing the distance (the ℓ_2 norm) between the (empirical) regression quantiles and the (theoretical) density-implied quantiles. This second step both smooths the estimated quantile functions and provides a complete density forecast, albeit one whose form is now controlled by the class of skewed-t density assumed. This second step, therefore, contrasts the nonparametric nature of the first-step quantile regressions. Policy institutions, such as the IMF, have also adopted this two-step approach to monitor international macroeconomic risks, such as growth-at-risk (GaR); see Prasad et al. (2019).

This paper reconsiders the use of QRs when interest rests with the production and subsequent evaluation of density forecasts, from which specific risk forecasts, such as GaR, can always be extracted. The attraction of producing density forecasts rather than specific point, quantile, or interval forecasts is that, given the forecast user's loss function, one can readily extract from the density forecast the features of specific interest to the user. Such a focus on the production of density forecasts is rare in the quantile regression literature (with the notable exceptions listed above). This is despite considerable attention having been paid to the production and evaluation of the quantile forecasts themselves (for example, see Komunjer (2013)).

Our paper proposes and then contrasts with the aforementioned two-step ABG method, which has become so established, a simple nonparametric (strictly "semi-parametric") approach to the production of density forecasts from QRs. Unlike ABG's, this approach does not superimpose a global density on specific quantile forecasts. Instead, the conditional quantile forecasts from the first-step QRs are mapped directly to a conditional density, assuming only local uniformity between the quantile forecasts. In an application to US GDP growth, we find that use of this nonparametric approach matches or slightly improves upon the accuracy

[1] On the use of QR methods to produce density nowcasts and forecasts, see e.g., Gaglianone and Lima (2012), Gaglianone and Lima (2014), Manzan and Zerom (2013), Manzan (2015), Korobilis (2017), Ferrara et al. (2022), Chen et al. (2021), and Mitchell et al. (2022). On the more specific but connected issue of the assessment of tail risks using QRs, see e.g., Giglio et al. (2016), Ghysels et al. (2018), Adrian et al. (2019), Carriero et al. (2020a), Carriero et al. (2020b), Reichlin et al. (2020), Brownlees and Souza (2021), and Figueres and Jarocinski (2020).

of the ABG densities. It also supports the much-cited finding of ABG that the left-tail of the conditional density of GDP growth moves with the tightness of financial conditions. But the nonparametric approach delivers conditional forecast densities with very different features than those when, following ABG, a skewed-t density is assumed globally. In particular, linking to Adrian et al. (2021), we find that the very same QRs used by ABG do, in fact, deliver multimodal GDP growth density forecasts. This is notably so at times of recession, when conditioning on a popular index of financial conditions. The evolution over the business cycle of multimodalities rather than asymmetries then becomes the dominant macroeconomic narrative of the conditional predictive distribution of GDP growth.

The focus in this paper is the construction of density forecasts from QRs, given their growing use in macroeconomics and finance since ABG. A large literature, of course, considers the production of density forecasts using other methods; see Aastveit et al. (2019) for a review. A literature has also grown up, in response to ABG, on the production of GaR and density forecasts using both parametric and nonparametric alternatives to QR; for example, see Carriero et al. (2020a), Caldara et al. (2021), Plagborg-Moller et al. (2020), De Polis et al. (2020), and Adrian et al. (2021). By contrast, we deliberately stick to the QR models of ABG. In so doing, we emphasize the empirical importance of moving beyond their skewed-t parametric assumption when fitting the density to these quantile forecasts.

The remainder of this paper is structured as follows. Section 2 considers the construction of density forecasts from quantile regressions, estimated via frequentist or Bayesian methods. It contrasts parametric and nonparametric methods for the production of the density forecast. Section 3 presents Monte Carlo evidence on the relative efficacy of the parametric and nonparametric approaches at fitting densities to distributions of various underlying shapes. Section 4 revisits the GaR application of ABG and contrasts empirical results using the parametric and nonparametric approaches. Section 5 concludes. An online appendix contains supplementary material.

2 Density forecasts from quantile regressions

Consider the QR relating the τ-th quantile of y_{t+h}, the variable of interest (GDP growth in our application), to x_t, a $d-$dimensional vector of conditioning variables including an intercept:

$$Q_{y_{t+h}|x_t}(\tau|x_t) = x_t'\beta_\tau, \tau \sim U(0,1), \tag{1}$$

with $t = 1, ..., T$ and where h is the forecast horizon and $U(.)$ is the uniform density. Note that, following ABG, we focus on QR models with time-invariant parameters.[2]

The QR slope, β_τ, is chosen to minimize the weighted absolute sum of errors:

$$\hat{\beta}_\tau = \arg\min_{\beta_\tau} \sum_{t=1}^{T} (y_{t+h} - x_t'\beta_\tau)(\tau - \mathbf{1}_{y_{t+h} \leq x_t'\beta_\tau}),\ \tau \in (0,1), \tag{2}$$

where $\mathbf{1}(.)$ denotes an indicator function. A perceived attraction of QR is that the informational importance of x_t for y_{t+h} can vary by quantile and thereby accommodate situations where conditioning variables have, for example, more or less informational content in the tails of the density.

The quantile forecasts from (2), conditional on x_t, are:

$$\hat{Q}_{y_{t+h|}x_t}(\tau|x_t) = x_t'\hat{\beta}_\tau. \tag{3}$$

Bayesian estimation of QRs has also gained attention recently. Koenker and Machado (1999) established that likelihood-based inference using independently distributed asymmetric Laplace densities (ALD) is directly related to (2). Yu and Moyeed (2001) show how exact Bayesian inference using Markov chain Monte Carlo (MCMC) methods can proceed by forming the likelihood function using the ALD; they emphasize the utility of the ALD, irrespective of the original distribution of the data. And Kozumi and Kobayashi (2011) propose a mixture representation of the ALD that renders the model conditionally Gaussian, facilitating estimation using more efficient MCMC methods. Unlike classical estimation methods, Bayesian methods naturally accommodate parameter uncertainty when forecasting.

Quantile forecasts can be constructed from the Bayesian QR, as per (3), from the posterior parameter distribution for β_τ. For the r-th MCMC draw, $\hat{\beta}_\tau^r$, these quantile forecasts are given as:

$$\hat{Q}_{y_{t+h|}x_t}(\tau|x_t)^r = x_t'\hat{\beta}_\tau^r. \tag{4}$$

In empirical applications, quantile regressions are estimated at a finite number of τ, i.e., $[\tau_1, ..., \tau_k]$, where $0 < \tau_1 < \tau_2 < .... < \tau_k < 1$. ABG, in fact, consider just $k = 4$. This means that the underlying conditional density is defined only at these k quantiles. To estimate the full conditional $h-$step predictive density, $\hat{f}(y_{t+h}|x_t)$, we therefore need to establish a mapping from the k quantile forecasts, as in (3) or (4):

[2]Recent research in macroeconomics has moved on to consider QR models with time-varying-parameters (e.g., see Korobilis et al. (2021)). The same issues, as discussed in this paper, arise when considering how to construct density forecasts from these QR models.

$$\left\{\hat{Q}_{y_{t+h}|x_t}(\tau_1|x_t), ..., \hat{Q}_{y_{t+h}|x_t}(\tau_k|x_t)\right\} \rightarrow \hat{f}(y_{t+h}|x_t), \forall [x_t', y_{t+h}]' \in \mathbb{R}^{dim(x)+1}, \tag{5}$$

where, for notational ease, we denote these quantile forecasts $\hat{Q}_{y_{t+h}|x_t}(\tau_j|x_t) = x_t'\hat{\beta}_\tau$; that is, we suppress dependence on the MCMC draw.

Below we set out two ways of establishing this mapping. We start with the parametric approach of ABG. As discussed in the introduction, this approach is used widely in macroeconomics, despite the contradiction with the nonparametric flavor of the first-step QRs.

2.1 ABG's parametric quantile-matching approach

To estimate the full continuous conditional density forecast of y_{t+h}, from the k quantile forecasts, ABG, in effect, combine them by fitting the skewed-t density function of Azzalini and Capitanio (2003) to the quantile forecasts, (3). They minimize the distance (the ℓ_2 norm) between the (empirical) regression quantiles and the (theoretical) distribution-implied quantiles:

$$\underset{\mu,\sigma,\alpha,\upsilon}{arg\ min} \sum_{\tau} \left(\hat{Q}_{y_{t+h}|x_t}(\tau|x_t) - \hat{F}^{-1}(\tau;\mu,\sigma,\alpha,\upsilon)\right)^2, \tag{6}$$

where F is the CDF of the skewed-t PDF, f, given as:

$$f(y;\mu,\sigma,\alpha,\upsilon) = \frac{2}{\sigma} t\left(\frac{y-\mu}{\sigma};\upsilon\right) T\left(\alpha\frac{y-\mu}{\sigma}\sqrt{\frac{\upsilon+1}{\upsilon+\left(\frac{y-\mu}{\sigma}\right)^2}};\upsilon+1\right), \tag{7}$$

where t and $T(.)$ respectively denote the PDF and CDF of the Student t-distribution, where μ is a location parameter, σ is the scale, υ is the fatness, and α is the shape. When $\alpha = 0$, the skewed-t reduces to the Student t. When, in addition, $\upsilon = \infty$, (7) reduces to a Gaussian distribution, with mean μ and standard deviation σ.

ABG focus on the exactly identified case of matching the 0.05, 0.25, 0.75, and 0.95 quantiles. But, in principle, as ABG discuss in a footnote but do not explore empirically, more quantiles could be used, allowing the four parameters of (7) to be over-identified. Since the choice of these $k = 4$ quantiles is somewhat arbitrary and may affect the shape of the fitted distribution, below we also consider fitting the skewed-t distribution to more quantiles.

While ABG used (6) on quantile forecasts, (3), produced from a frequentist QR, others have fitted the skewed-t-distribution to forecasts produced from a Bayesian QR. Ferrara et al. (2022), for example, use (6) on the mean (across $r = 1, ..., R$ MCMC draws) quantile forecasts, (4).

2.2 Constructing the density forecast nonparametrically

Rather than assume a parametric function for $\hat{f}(y_{t+h}|x_t)$, following Parzen (1979) and Koenker (2005), one can back out the conditional distribution directly from the conditional quantile function via the integral transforms:

$$\hat{F}(y_{t+h}|x_t) = \int_0^1 \mathbf{1}\{x_t'\hat{\beta}_\tau \leq y_{t+h}\}d\tau. \tag{8}$$

By considering all $\tau \in (0,1)$, one can approximate the true conditional quantile function arbitrarily well, when the true density is a smooth conditional density (Koenker (2005), p. 53).

In practice, we follow Koenker and Zhao (1996) and adopt a simple simulation-based approach, instead of relying on numerical integration. A random draw from the h-step-ahead forecast distribution is given by:

$$\hat{y}_{t+h} = \hat{Q}_{y_{t+h}|x_t}(U|x_t)^r, \tag{9}$$

where U is a uniformly distributed random variable on $[0,1]$ as in Koenker and Zhao (1996). Repeating across many random draws approximates $\hat{F}(y_{t+h}|x_t)$.

To operationalize, with a finite k, we smooth/interpolate across adjacent quantile forecasts by taking a first-order Taylor expansion of the CDF, (8), between the j-th and $j+1$-th quantiles:

$$\hat{F}_k(y_{t+h}|x_t) = \tau_j + \frac{\tau_{j+1}-\tau_j}{x_t'\hat{\beta}_{\tau_{j+1}} - x_t'\hat{\beta}_{\tau_j}}(y_{t+h} - x_t'\hat{\beta}_{\tau_j}) \tag{10}$$

$$= \tau_j + F'(y^*_{t+h,j}|x_t)(y_{t+h} - x_t'\hat{\beta}_{\tau_j}), \tag{11}$$

for $y^*_{t+h,j} \in \left(x_t'\hat{\beta}_{\tau_j}, y_{t+h}\right) \subset \left(x_t'\hat{\beta}_{\tau_j}, x_t'\hat{\beta}_{\tau_{j+1}}\right)$. Assuming that the interval between adjacent quantiles is relatively small, the implied distribution function is approximately linear within the interval.

Figure 1 provides an illustration, plotting the approximate CDF in yellow and the true CDF in blue. This illustration intuitively points to higher values of k delivering better approximations. That is, the marginal benefits of the first-order approximation decline as k increases, an issue we explore below both in the simulations and in the application. Unlike ABG's, this approach does not superimpose a global (parametric, such as a skewed-t) distribution on specific quantile forecasts. Instead, it assumes local uniformity between the k quantile

forecasts. Hence, it is best seen as a "semi-parametric" method, although for convenience we continue to refer to the method as nonparametric.

Algorithm 1 summarizes the mechanics of how the density forecast is formed nonparametrically from the QRs. Whether the QRs are estimated by frequentist or Bayesian methods, the empirical density forecast is constructed from the sample: $\left[\boldsymbol{y}_{t+h,1}, \boldsymbol{y}_{t+h,2}, ..., \boldsymbol{y}_{t+h,k}, \boldsymbol{y}_{t+h,k+1}\right]$. This vector can be used directly by the macroeconomist or a kernel could be fitted.[3]

We note four features of Algorithm 1:

1. Since:
$$Prob(F^{-1}(\tau_j|x_t) \leq y_{t+h} < F^{-1}(\tau_{j+1}|x_t)) = \tau_{j+1} - \tau_j, \tag{12}$$
to take a sample of length N from the conditional distribution $F(\cdot|X = x_t)$ requires $(\tau_{j+1} - \tau_j)N$ samples to be taken between:
$$\left(x_t'\hat{\beta}_{\tau_j}, x_t'\hat{\beta}_{\tau_{j+1}}\right). \tag{13}$$

2. The quantile forecasts are re-arranged as necessary (following Chernozhukov et al. (2010)) to avoid quantile crossing.

3. The "extreme" quantiles are approximated by a specified CDF, here assumed to be the Gaussian CDF, Φ, although any parametric CDF could be used.[4] This implies:
$$\Phi(x_t'\beta_{\tau_1}, \mu_1, \sigma_1) = \tau_1, \Phi(x_t'\beta_{\tau_2}, \mu_1, \sigma_1) = \tau_2 \tag{14}$$
$$\Phi(x_t'\beta_{\tau_{k-1}}, \mu_2, \sigma_2) = \tau_{k-1}, \Phi(x_t'\beta_{\tau_k}, \mu_2, \sigma_2) = \tau_k, \tag{15}$$
where we solve for $[\mu_1, \mu_2, \sigma_1, \sigma_2]$ to satisfy these 4 equations.

4. Algorithm 1 consistently estimates the true conditional distribution $F(y_{t+h}|x_t)$ as $T, k \to \infty$. This is understood by noting that there are two convergence aspects to consider in Algorithm 1: (a) statistical convergence, $T \to \infty$, and (b) convergence of the approximate distribution to the true distribution as the number of quantile levels, $k \to \infty$:

   (a) The consistency of the QR estimates $\hat{\beta_{\tau_j}}$ as $T \to \infty$ (see Chernozhukov et al. (2010) and Koenker (2005)), at the chosen quantile levels, j, implies that the

[3]See Krüger et al. (2021) for a discussion of the pros and cons of alternative methods for estimating the distribution from the underlying simulation output. Their analysis demonstrates that the empirical CDF-based approximation works well in many contexts.

[4]In our simulations and the application, we define "extreme" as those quantiles beyond 0.05 and 0.95 or 0.01 and 0.99. Following Chernozhukov (2005) extremal methods could be used instead.

approximate distribution $\hat{F}_k \to F_k$. That is, referring again to Figure 1, the approximate distribution converges to the piecewise-linear function (the yellow line) approximating the true CDF (the blue line) at the chosen quantile. For $\tau \in \{\tau_1, ..., \tau_k\}$:

$$F_k(x_t'\beta_\tau|x_t) = F(x_t'\beta_\tau|x_t), \tag{16}$$

i.e., the vertex of the function equals the true density at the finite sequence of quantile levels (and the blue and yellow lines equal each other).

(b) As $k \to \infty$, the piecewise-linear CDF (the yellow line in Figure 1) converges to the true distribution (the blue line in Figure 1) between these quantile levels. This is seen as follows. Given a smoothness assumption for the true density, by Taylor's theorem, rewrite the true distribution as:

$$F(y_{t+h}|x_t) = \tau_j + f(y^*_{t+h,1}|x_t)(y_{t+h} - x_t'\beta_{\tau_j}), \tag{17}$$

for any $y_{t+h} \in \left(x_t'\beta_{\tau_j}, x_t'\beta_{\tau_{j+1}}\right)$ and some $y^*_{t+h,1} \in \left(x_t'\beta_{\tau_j}, y_{t+h}\right)$. Then, by the mean value theorem, the approximate k quantile level distribution:

$$F_k(y_{t+h}|x_t) = \tau_j + \frac{\tau_{j+1} - \tau_j}{x_t'\beta_{\tau_{j+1}} - x_t'\beta_{\tau_j}}(y_{t+h} - x_t'\beta_{\tau_j}) \tag{18}$$

$$= \tau_j + f(y^*_{t+h,2}|x_t)(y_{t+h} - x_t'\beta_{\tau_j}), \tag{19}$$

for $y^*_{t+h,2} \in \left(x_t'\beta_{\tau_j}, x_t'\beta_{\tau_{j+1}}\right)$. Comparing (17) and (19), the only difference is between $y^*_{t+h,1}$ and $y^*_{t+h,2}$. Yet, note that:

$$x_t'\beta_{\tau_j} \leq y^*_{t+h,2} \leq x_t'\beta_{\tau_{j+1}} \tag{20}$$

$$x_t'\beta_{\tau_j} \leq y^*_{t+h,1} \leq y_{t+h} \leq x_t'\beta_{\tau_{j+1}}. \tag{21}$$

Further assume that the conditional quantiles are linear in the regressors, uniformly across all τ. Then, we can let $k \to \infty$. As $k \to \infty$, $\tau_{j+1} - \tau_j \to 0$, and the intervals in (20) and (21) converge by the sandwich theorem such that:

$$y^*_{t+h,1} = y^*_{t+h,2}.$$

Hence:

$$\lim_{k\to\infty} F_k(y_{t+h}|x_t) = F(y_{t+h}|x_t).$$

In the simulations and empirical application below, we consider how to choose k. We suggest, in effect, to choose k empirically to maximize forecasting performance.

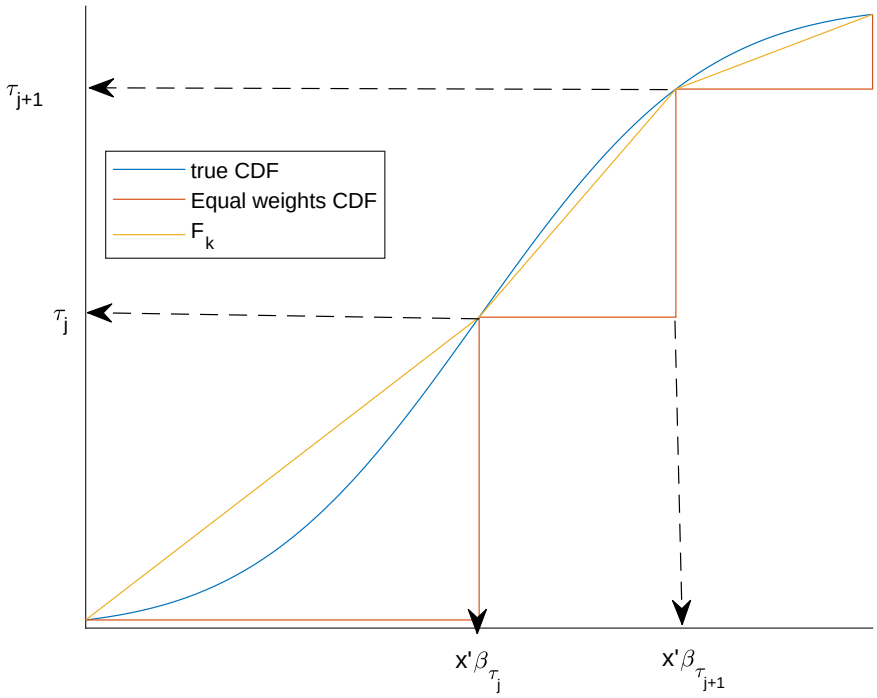


Figure 1: Illustrative comparison of the true CDF against Algorithm 1 (F_k) and the CDF assuming uniform (equal) weights between adjacent quantiles

In general, we find that intermediate values of k (such as $k = 19$) tend to work best. These balance the need for a value large enough to accurately trace out the shape of the underlying distribution, with the risk, especially in smaller samples, of introducing noise by estimating QRs in the tails of the distribution with too few observations.

Algorithm 1, where the proposed distribution is:

$$\hat{F}_k(y_{t+h}|x_t) = \tau_j + \frac{\tau_{j+1} - \tau_j}{x_t'\hat{\beta}_{\tau_{j+1}} - x_t'\hat{\beta}_{\tau_j}}(y_{t+h} - x_t'\hat{\beta}_{\tau_j}), \tag{22}$$

when $y_{t+h} \in \left(x_t'\hat{\beta}_{\tau_j}, x_t'\hat{\beta}_{\tau_{j+1}}\right)$, can be contrasted with an alternative of using equal weights between adjacent quantiles:

$$\hat{F}_{EW}(y_{t+h}|x_t) = \begin{cases} \tau_j & y_{t+h} \in (x_t'\hat{\beta}_{\tau_j}, x_t'\hat{\beta}_{\tau_{j+1}}) \\ 0 & y_{t+h} < x_t'\hat{\beta}_{\tau_1} \\ 1 & y_{t+h} \geq x_t'\hat{\beta}_{\tau_k} \end{cases}, \tag{23}$$

which amounts to a zero-order approximation of the CDF between quantiles j and *j+1*. We emphasize that this is, in effect, the approach used by Korobilis (2017) to produce density forecasts from Bayesian QRs. This approach involves collecting together the $r = 1, ..., R$ MCMC draws of the quantile forecast $\widehat{Q}_{y_{T+h}}(\tau|x_t)^r$ across $\tau \in [0.05, 0.10, ..., 0.90, 0.95]$ and then constructing the full posterior density forecast from this stacked vector - using a kernel to smooth.

Figure 1 also illustrates how equal weights differ from Algorithm 1. It shows how equal weights intuitively provide a worse approximation to the true CDF. Note that, given the estimated quantile levels, the straight lines that Algorithm 1 imposes between adjacent quantiles provide a piecewise-linear approximation to the CDF. Unlike the piecewise-constant function implied by equal weights, the piecewise-linear approximation benefits from smoothness in the estimated CDF. Statistics such as the conditional mean can be obtained via numerical integration of:

$$\int x_t \hat{f}_k(y_{t+h}|x_t)dx_t, \tag{24}$$

where:

$$\hat{f}_k(y_{t+h}|x_t) = \begin{cases} \phi(y_{t+h}|\hat{\mu_1}, \hat{\sigma}_1) & y_{t+h} \leq x_t'\hat{\beta}_{\tau_1} \\ \frac{\tau_{j+1}-\tau_j}{x_t'\hat{\beta}_{\tau_{j+1}} - x_t'\hat{\beta}_{\tau_j}} & x_t'\hat{\beta}_{\tau_j} < y_{t+h} \leq x_t'\hat{\beta}_{\tau_{j+1}} \\ \phi(y_{t+h}|\hat{\mu_2}, \hat{\sigma}_2) & y_{t+h} > x_t'\hat{\beta}_{\tau_k}. \end{cases} \tag{25}$$

Algorithm 1, instead, relies on samples from the conditional density $\hat{f}_k(y_{t+h}|x_t)$, which lets us readily construct the whole density.

3 Simulation results

To evaluate the performance of the nonparametric approach to construction of the predictive density from QRs, relative to extant alternatives including the approach of ABG, we conduct a set of Monte Carlo experiments. These experiments let us assess the ability of the different approaches to uncover a range of distributional forms. We consider five data-generating processes (DGPs) that yield densities for $\{y_t\}_{t=1}^T$ that are:

1. (DGP1) Gaussian: $N(0, 1)$.

2. (DGP2) Negatively skewed: $f(y; \mu = 1, \sigma = 2, \alpha = -0.5, \upsilon = 10)$, where $f(.)$ is as defined in (7).

Algorithm 1 A local-linear algorithm to construct density forecasts from quantile regressions

- Estimate the QR at τ_j $(j = 1, ..., k)$.
- Denote the QR estimates, $\hat{\beta}_{\tau_j}$, where for Bayesian estimation $\hat{\beta}_{\tau_j} = \{\hat{\beta^1_{\tau_j}}, ..., \hat{\beta^R_{\tau_j}}\}$ is a $d \times R$ dimensional matrix, where $r = 1, ..., R$, defined by stacking across the MCMC draws. In the frequentist case, $R = 1$. Define:

$$Q_t = \left[(x_t{}'\hat{\beta}_{\tau_1})', (x_t'\hat{\beta}_{\tau_2})', ..., (x_t'\hat{\beta}_{\tau_k})'\right] \in \mathbb{R}^{R \times k}.$$

- Rearrange the elements of the rth row of the matrix Q_t from smallest to largest in case they are not monotonic.
- for $j = 2 : k$
  - Obtain the sub-sample given random variables uniformly distributed on $[\tilde{Q}_{t,j-1}, \tilde{Q}_{t,j}]$:

$$\boldsymbol{y}_{t+h,j} = \tilde{Q}_{t,j-1}\mathbf{1}'_{(\tau_j - \tau_{j-1})N} + diag(\tilde{Q}_{t,j} - \tilde{Q}_{t,j-1})U_j$$

    where $\tilde{Q}_{t,j}$ denotes the jth column of Q_t and U_j is a matrix of dimension $R \times (\tau_j - \tau_{j-1})N$, with each element drawn from a standard uniform distribution similar to (9).
- end
- Fit a Gaussian (or some other) distribution via $\hat{\beta}_{\tau_1}$ and $\hat{\beta}_{\tau_2}$, and sample from the lower tail $F(y_{t+h}|x_t) < \tau_1$ to obtain $\boldsymbol{y}_{t+h,1}$
- Fit a Gaussian (or some other) distribution via $\beta_{\tau_{k-1}n}$ and $\beta_{\tau_k,n}$, and sample from the upper tail $F(y_{t+h}|x_t) > \tau_k$ to obtain $\boldsymbol{y}_{t+h,k+1}$

Finally, create the stacked vector of forecasts: $\left[\boldsymbol{y}_{t+h,1}, \boldsymbol{y}_{t+h,2}, ..., \boldsymbol{y}_{t+h,k}, \boldsymbol{y}_{t+h,k+1}\right]$.

3. (DGP3) High kurtosis: $f(y; \mu = 1, \sigma = 1, \alpha = 1, \upsilon = 5)$.

4. (DGP4) Bimodal (mixture of Gaussian) : $1/3N(0, .04) + 2/3N(1, .04)$.

5. (DGP5) Trimodal (mixture of Gaussian): $1/6N(0, 0.2) + 1/3N(1, 0.2) + 1/2N(2, 0.2)$.

For $\{y_t\}_{t=1}^T$ samples of size $T = 100$ and $T = 1,000$ drawn from each of these five DGPs (see Figure 2 for an illustrative visualization), we then estimate five alternative densities and compare their fit against the (true) DGP density. In all cases, when estimating the QR, we set $x_t = 1$, that is, we consider an intercept only.

The five densities we fit to the $\{y_t\}_{t=1}^T$ samples are:

1. NP(freq): estimate the QRs (where $k = 19$, such that $\tau \in [0.05, 0.10, ..., 0.90, 0.95]$) using frequentist methods, (2), and then construct the density nonparametrically via Algorithm 1. We also experiment, as summarized below, with $k = 4$ where $\tau \in [0.05, 0.25, 0.75, 0.95]$ (as in ABG) and $k = 99$ where $\tau \in [0.01, 0.02, ..., 0.99]$.

2. NP(B): estimate the QRs (where $k = 19$, such that $\tau \in [0.05, 0.10, ..., 0.90, 0.95]$) using Bayesian methods and then construct the density nonparametrically via Algorithm 1. At the first stage, the Bayesian QR is estimated using a standard normal uninformative prior for the $q-$vector of β_τ coefficients, centered on a zero mean:

$$\beta_\tau \sim N(0, \mathbf{V}_\beta), \tag{26}$$

   where $\mathbf{V}_\beta = 10\mathbf{I}_q$.

3. EW(B): estimate the QRs (where $k = 19$, such that $\tau \in [0.05, 0.10, ..., 0.90, 0.95]$) using Bayesian methods (as in NP(B)) but then construct the density using equal weights, (23).

4. ABG: follow ABG (using their replication material) and estimate the QRs (where $k = 4$, such that $\tau \in [0.05, 0.25, 0.75, 0.95]$) using frequentist methods and then construct the density parametrically via (7).[5]

5. ABG kernel: as a non-QR benchmark, follow ABG and nonparametrically estimate a kernel density.[6]

[5]We note that in ABG's Matlab replication materials (available at `http://doi.org/10.3886/E113169V1`), when matching the quantile forecasts to the skewed-t density they approximate integrals with discrete sums. Specifically, looking at ABG's `Step2match.m` file (line 100), we see that they evaluate the skewed-t density only over a grid from -15 to 10. Instead, we use an exact analytical solution. In the empirical section below we return to this issue, showing its empirical importance.

[6]See equation (8) of ABG for details of the specific kernel density estimator employed.

For all the Bayesian models, we estimate using 20,000 MCMC draws with a burn-in of 10,000 draws. With regard to the Bayesian QR and Algorithm 1, we save every 10th draw from the 10,000 draws. This yields 1,000 draws (across k quantiles) that are then inputted by draw into Algorithm 1 where $N = 100$. This delivers a vector of 100,000 draws (1,000*100) from each predictive forecast density.

Tables 1 and 2, for $T = 100$ and $T = 1,000$, respectively, report the mean squared error (across $R = 100$ parallelized chains) of the first four moments of the fitted densities relative to the true (DGP) density and the average Kullback-Leibler (KL) distance between the fitted and true densities. KL is constructed as the expected difference in their logarithmic scores. Looking at the KL distance first, as a measure of overall density fit, we see that the nonparametric (NP) estimators, whether NP(freq) or NP(B), consistently deliver the best-fitting densities irrespective of the shape of the true density.[7] As anticipated, ABG's parametric approach is competitive only when the true density is unimodal. Instabilities in estimation of the skewed-t density mean that ABG is not, however, always the best-fitting density even for DGP1 through DGP3, when the true density is unimodal, and we might expect the parametric nature of ABG to deliver gains. But for the multimodal densities (DGP4 and DGP5) use of Algorithm 1 is clearly preferable. Like ABG's, the equal-weighted Bayesian approach, EW(B), also performs poorly for multimodal densities and, in fact, under-performs relative to ABG for the three unimodal DGPs. The benchmark kernel density, like the NP estimators, can also accommodate multimodalities. However, the kernel density does not deliver as good-fitting densities as the NP approaches, in particular for the smaller sample size of $T = 100$.

Turning to the accuracy of the first four moments, as judged by the mean squared error (MSE) between the respective moment of the fitted and true densities, we observe a similar picture. The NP estimators dominate both ABG, EW(B), and kernel. We also note how explosive estimation, for some Monte Carlo replications, pushes up the MSE estimates in some instances, especially for EW(B) and ABG. When estimates of $\upsilon < 4$, not all of the first four moments of the skewed-t density are defined.

[7]To isolate the role of k in explaining this result, given $k = 4$ in ABG but $k = 19$ in NP(freq), we experimented with NP(freq) when $k = 4$ ($\tau \in [0.05, 0.25, 0.75, 0.95]$) and $k = 99$ ($\tau \in [0.01, 0.02, ..., 0.99]$); and we experimented with ABG when k was increased from its maintained value of 4. As Table 4 in the online appendix shows, decreasing k to $k = 4$ markedly lessens the accuracy of NP(freq) and increasing k to $k = 99$ also worsens accuracy. While we might expect increases in k to improve accuracy for NP(freq), as the local uniformity assumption becomes weaker, parameter estimation errors increase for more extreme quantiles. The objective function of the standard QR estimator is not smooth, and the QR estimates can experience jumps. Future work might consider the benefits of producing the density forecasts having first smoothed the objective function, e.g., as in Fernandes et al. (2021). Increasing k for NP(freq), well into the 5 percent tails as is the case when $k = 99$, was therefore found to deliver noisier estimates of the underlying conditional density, especially for the smaller $T = 100$. By contrast, due to its parametric assumption, increasing k did little to affect results for ABG.

In sum, the Monte Carlo evidence confirms that the choice of how to fit a density to quantile forecasts matters. While ABG's parametric assumption may work well, unsurprisingly it will only do so for true densities that are unimodal. Instead, it is relatively simple to let the "data speak," as they do when estimating the QRs in the first place, and use nonparametric approaches as detailed in Algorithm 1 to construct the forecast density from the quantile forecasts. While these simulations are, of course, just illustrative, they do indicate how the nonparametric approach of Algorithm 1 can flexibly accommodate a greater variety of distributional shapes than ABG, even for modest sample sizes. They also suggest that when using Algorithm 1 intermediate values of k (such as $k = 19$) best approximate the underlying density.

In principle, we anticipate a trade-off when selecting what k to use in Algorithm 1. Too small a value does not give NP sufficient flexibility to smoothly fit different distributional shapes. Too large a value for k, especially for smaller sample sizes, T, increasingly forces the QR into the tails of the density, where there are fewer observations. This may induce noise in the forecast density, and raises the risk of introducing erroneous spikes or modes (undersmoothing) in the forecast density when fitted using NP. To investigate this possible trade-off, in the empirical appendix we report supplementary simulation results (see Table 5). These involve, for DGP1 through DGP5, using the calibrated unimodality test of Hartigan and Hartigan (1985), as proposed by Cheng and Hall (1998), and reporting the proportion of rejections of unimodality. Table 5 confirms that while increasing k, when using NP(freq), does increase the chance of identifying false peaks in the unimodal densities of DGP1 through DGP3, this risk rapidly declines to zero for sample sizes of $T > 50$. This suggests that increasing k does not inject false peaks into the fitted densities, except for very small samples ($T = 25$). In turn, for the multimodal DGPs (DGP4 and DGP5), NP(freq) does a good job of rejecting unimodality, except for smaller values of k (specifically, $k = 4$ and $k = 9$). As long as k is at least 19, we see rejection rates in Table 5 of over 90 percent, even when $T = 25$. These rejection rates rise further as T increases. In short, these supplementary unimodality tests both support the use of intermediate values of k when using Algorithm 1 and provide confidence that Algorithm 1 does not identify false modes in the forecast density, unless T is especially small relative to k.

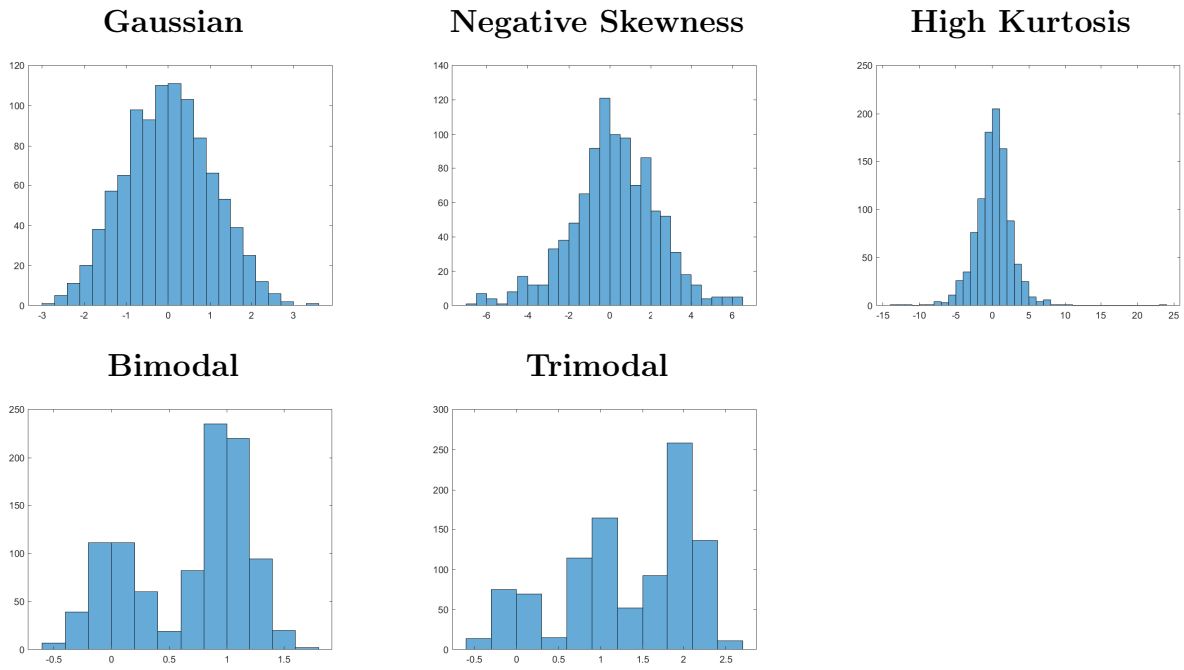


Figure 2: Simulated draws from the 5 DGP densities ($T = 1,000$)

Table 1: Average mean squared error and Kullback-Leibler (KL) distance for $T = 100$

Models	Mean	Variance	Skewness	Kurtosis	KL
DGP1: Unimodal (Gaussian)					
NP(freq)	0.01	0.03	0.12	0.40	0.02
NP(B)	0.01	0.03	0.12	0.60	0.06
EW(B)	0.01	0.48	0.01	0.05	0.06
ABG	0.01	0.05	Inf	Inf	0.02
ABG Kernel	0.01	0.07	0.04	0.10	0.02
DGP2: Unimodal (Negative Skewness)					
NP(freq)	0.05	0.73	0.14	1.51	0.02
NP(B)	0.05	0.70	0.18	1.02	0.05
EW(B)	0.04	7.32	0.03	0.61	0.06
ABG	0.05	Inf	Inf	Inf	0.03
ABG Kernel	0.05	1.32	0.10	0.82	0.04
DGP3: Unimodal (High Kurtosis)					
NP(freq)	0.01	0.12	1.11	80.41	0.02
NP(B)	0.01	0.08	0.46	49.13	0.06
EW(B)	0.01	0.34	0.82	72.22	0.07
ABG	0.01	Inf	Inf	Inf	0.03
ABG Kernel	0.01	0.30	0.66	59.62	0.12
DGP4: Bimodal					
NP(freq)	0.00	0.00	0.01	0.04	0.03
NP(B)	0.00	0.00	0.01	0.06	0.05
EW(B)	0.00	0.05	0.04	1.99	0.25
ABG	0.00	0.00	0.30	6.14	0.30
ABG Kernel	0.00	0.00	0.01	0.11	0.11
DGP5: Trimodal					
NP(freq)	0.00	0.00	0.01	0.04	0.05
NP(B)	0.00	0.00	0.01	0.07	0.06
EW(B)	0.00	0.23	0.07	1.23	0.32
ABG	0.00	0.01	0.31	5.20	0.26
ABG Kernel	0.00	0.01	0.02	0.07	0.21

Notes: Inf denotes infinity. NP(freq) uses $k = 4$. The 5 estimators and 5 DGPs are defined in Section 3.

Table 2: Average mean squared error and Kullback-Leibler (KL) distance for $T = 1,000$

Models	Mean	Variance	Skewness	Kurtosis	KL
DGP1: Unimodal (Gaussian)					
NP(freq)	0.00	0.01	0.06	0.15	0.00
NP(B)	0.00	0.00	0.01	0.03	0.01
EW(B)	0.00	0.52	0.00	0.04	0.06
ABG	0.00	0.00	0.02	0.21	0.00
ABG Kernel	0.00	0.01	0.01	0.02	0.01
DGP2: Unimodal (Negative Skewness)					
NP(freq)	0.01	0.31	0.06	1.04	0.00
NP(B)	0.01	0.09	0.02	0.56	0.01
EW(B)	0.00	7.37	0.02	0.63	0.06
ABG	0.00	0.15	0.04	Inf	0.00
ABG Kernel	0.00	0.18	0.02	0.30	0.01
DGP3: Unimodal (High Kurtosis)					
NP(freq)	0.00	0.10	0.99	82.64	0.00
NP(B)	0.00	0.03	0.28	60.95	0.02
EW(B)	0.00	0.28	0.79	72.63	0.07
ABG	0.00	0.03	Inf	Inf	0.00
ABG Kernel	0.00	0.03	0.60	142.25	0.05
DGP4: Bimodal					
NP(freq)	0.00	0.00	0.00	0.00	0.00
NP(B)	0.00	0.00	0.00	0.00	0.01
EW(B)	0.00	0.05	0.04	2.02	0.26
ABG	0.00	0.00	0.32	6.23	0.31
ABG Kernel	0.00	0.00	0.00	0.02	0.03
DGP5: Trimodal					
NP(freq)	0.00	0.00	0.00	0.01	0.03
NP(B)	0.00	0.00	0.00	0.01	0.02
EW(B)	0.00	0.27	0.07	1.22	0.33
ABG	0.00	0.01	0.30	5.09	0.25
ABG Kernel	0.00	0.00	0.00	0.01	0.09

4 Empirical results: Revisiting the growth-at-risk application of ABG

ABG established the empirical utility of quantile regressions for modeling the conditional density of US GDP growth. They found that deteriorating financial conditions, as captured by the Chicago Fed's National Financial Conditions Index (NFCI), have an asymmetric effect on GDP growth.[8] In particular, they link GDP growth tail risks to poor financial conditions. Recessions are associated with left-skewed conditional forecast densities. Carriero et al. (2020a) challenge this view, noting that ABG's empirical finding that downside risk varies more than upside risk could equally well be explained by symmetric conditional forecast densities as by asymmetric unconditional forecast densities. These could be produced, for example, by Bayesian VAR models with stochastic volatility. Caldara et al. (2021) similarly suggest use of a parametric modeling framework that rationalizes the empirical findings of ABG but maintains use of symmetric conditional densities. They capture nonlinear effects with a Markov-switching model, in which the transition probabilities depend, *inter alia*, on financial conditions. Adrian et al. (2021) also jettison use of QR and instead use kernel-based estimators to support their finding that the forecast density of GDP growth is approximately Gaussian and unimodal during normal periods, but becomes multimodal during periods of tight financial conditions. They also make the theoretical case for multimodality, explaining how it arises in macrofinancial intermediary models with occasionally binding financial constraints.

Given the degree to which ABG's empirical findings, based on their parametric quantile-matching approach, have influenced the subsequent literature, as we have just selectively reviewed, we emphasize the importance of letting the "data speak" about the nature of the conditional density forecast for GDP growth when mapping the quantile forecasts to the density forecasts. Accordingly, we revisit ABG's application. But we compare their skewed-t conditional density forecasts, which assume unimodality but allow for asymmetry, with those conditional density forecasts formed when we make no such assumption and, via Algorithm 1, better let the data inform this mapping.

Specifically, to facilitate comparison with ABG's parametric approach to constructing forecast densities from QRs, we use their data, sample periods, and preferred models. Specifically, we estimate QR models relating GDP growth to both lagged GDP growth and NFCI.[9] This

[8] The NFCI aggregates a large set of variables capturing credit quality, risk, and leverage.

[9] A subsequent literature has also used QRs to model GaR and construct GDP growth density forecasts. But it has examined the benefits of disaggregating the Chicago Fed's NFCI, using real-time NFCI vintages, and/or considered additional indicators; e.g., see Plagborg-Moller et al. (2020), Reichlin et al. (2020), Brownlees and Souza (2021), Kohns and Szendrei (2021), and Amburgey and McCracken (2022). Given the importance of the original modeling strategy in shaping the ongoing research agenda, as summarized in our introduction, we return to ABG's model space and consider (latest-vintage) NFCI alone. We expect that adding in extra

then lets us produce, via the aforementioned parametric and nonparametric approaches, one-quarter-ahead and one-year-ahead forecast densities for GDP growth conditional first on both lagged GDP growth and NFCI and second on just lagged GDP growth. Thereby, we isolate the role that NFCI plays in driving results. We re-assess ABG's claim that financial conditions are critical when density forecasting GDP growth in the US. Our focus, in common with much of the literature, is assessing the in-sample fit of the conditional densities. Thus we provide guidance on the importance of considering how to fit a density to the quantile forecasts. But we do provide some out-of-sample evaluation evidence too, although the latter arguably tells us more about the instabilities faced out-of-sample (see Rossi (2021)) than about the relative merits of different ways of constructing predictive densities from QRs. Nevertheless, in anticipation of the known benefits of shrinkage when forecasting out-of-sample, we do consider a variant of NP(B) that imposes a more informative prior. That is, we estimate Bayesian QRs with Minnesota priors. We follow Carriero et al. (2020b) and set V_i, the i-diagonal elements of $\mathbf{V}_\beta$, as follows:

$$V_i = \begin{cases} \lambda_1 \lambda_2 \frac{\sigma_{GDP}}{\sigma_j} & \text{for the coefficients other than the lag } l \text{ of GDP,} \\ \frac{\lambda_1}{l^{\lambda_3}} & \text{for the coefficients on the lag } l \text{ of GDP,} \\ 1000\sigma_{GDP} & \text{for the intercept,} \end{cases} \tag{27}$$

where σ_{GDP} and σ_j are the standard deviations from an AR(4) model for GDP growth and the j-th regressor (other than GDP growth), estimated with data available at the forecast origin. We follow Carriero et al. (2020b) and set $\lambda_1 = \lambda_2 = 0.2$, and $\lambda_3 = 1$. In terms of the in-sample fit, the prior variance on the coefficient on the lag of GDP is 0.2 for both the one-quarter- and one-year-ahead forecasts. On the other hand, the prior variance on the coefficient for NFCI differs. One-quarter-ahead, its prior variance is 0.25, while one-year-ahead it is 0.08. Let NP(BM) denote the forecast densities produced using this Minnesota prior and Algorithm 1.

Given this paper's emphasis on construction of the entire predictive density rather than just estimating GaR, we focus on assessing the overall fit of the competing forecast densities using probability integral transforms (PITs), i.e., the CDF of the forecast evaluated at the subsequent realization of GDP growth. For correctly calibrated forecast densities (see Diebold et al. (1998) and Mitchell and Wallis (2011)), these PITs, at the minimum, should be uniformly distributed. As shown by Diebold et al. (1998), correctly calibrated forecast densities will be preferred by all users, irrespective of their loss function. Nevertheless, to supplement

variables and allowing for possible additional nonlinearities will distinguish our approach further from ABG's. Given their skewed-t assumption, ABG's densities cannot accommodate the likely multimodalities associated with nonlinearity.

PITs-based tests of calibration and to facilitate cross-model comparison, we also report two commonly used scoring rules for density evaluation: the average logarithmic predictive score and the average continuous ranked probability score (CRPS). The CRPS is a popular density forecast-based scoring rule that offers greater robustness to outliers than the logarithmic score used by ABG; see Gneiting and Raftery (2007).

Figures 3 and 4 plot the cumulated PITs, respectively, for the one-quarter-ahead and one-year-ahead forecast densities produced using the 5 models of Section 3 plus NP(BM). These models consider both NFCI and lagged GDP growth as conditioning information, as favored by ABG. We also plot the PITs dropping NFCI from the QR, to isolate the importance of conditioning on financial conditions when density forecasting GDP growth.[10] Looking at these cumulated PIT plots across these 2 figures, it is apparent that NP(freq) appears to deliver the best-calibrated forecast densities. Its cumulated PITs are closest to the 45-degree line. Interestingly, the densities are well-calibrated at a 95 percent significance level, according to the PITs test of Rossi and Sekhposyan (2019), irrespective of whether NFCI is included in the QR.[11] The ABG densities perform second best, a very close second to NP(freq), but with a few extra little deviations from the 45 degree line. While based on the same frequentist QR as ABG, this indicates that fitting the skewed-t density to these same quantile forecasts is not as beneficial as using Algorithm 1. To investigate whether it is the higher value of $k = 19$ in NP(freq), relative to ABG (where $k = 4$), that explains this result rather than the use of Algorithm 1, we produced predictive densities from ABG assuming $k = 19$ (see Figure 16 in the online appendix). As in the Monte Carlo experiments, these alternative ABG densities are found to perform similarly to those when $k = 4$. Thus, we conclude that it is the use of Algorithm 1, rather than a different sized k, that yields the forecasting gains.

Algorithm 1 does not work quite as well (in-sample) when we estimate the QRs by Bayesian methods, whether with an uninformative or more informative prior. Accommodating parameter estimation uncertainties and shrinking the QR coefficients to zero (as with NP(BM)) does not help, at least in-sample. EW(B), in contrast, understates forecast uncertainty, as evidenced by S-shaped cumulated PITs.

[10] We emphasize how when constructing the ABG densities we use ABG's replication code. Therefore, as discussed in Section 3, we approximate integrals with discrete sums. We return later to the empirical applications of this.

[11] Figure 15 in the online appendix again shows how the choice of k in NP(freq) matters. From the S-shaped nature of the cumulated PITs, we can infer that the density forecast is too narrow at $k = 4$. Calibration is better at $k = 99$, but not obviously better than when $k = 19$ (as shown in Figures 3 and 4). This is consistent with the Monte Carlo evidence in Section 3 that a "medium-sized" value for k appears sufficient. The critical value bands of Rossi and Sekhposyan (2019) should be taken as "general guidance," to quote ABG, since they are derived assuming a rolling window of estimation, while, like ABG, we use an expanding window.

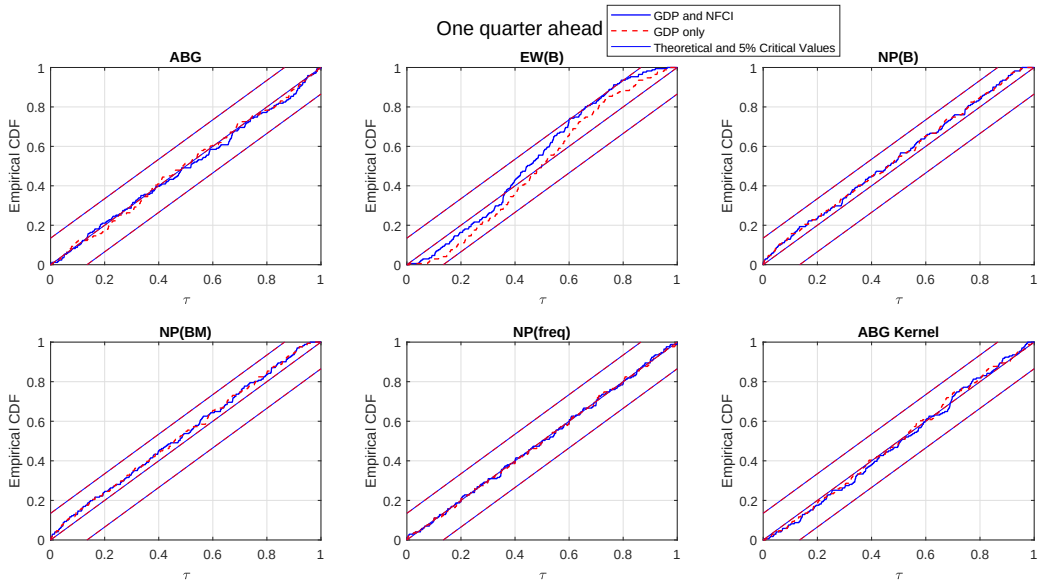


Figure 3: CDF of the in-sample PITs (one-quarter-ahead forecasts, 1973Q1-2015Q4) from the 6 density forecasts with and without NFCI. Note: the figures show the empirical CDF of the PITs (blue line) from the QR models with NFCI (and lagged GDP), the empirical CDF of the PITs (dashed red line) from the QR models without NFCI, the CDF of the PITs under the null hypothesis of correct calibration (the 45-degree line), and the 5% critical value bands of the Rossi and Sekhposyan (2019) PITs test.

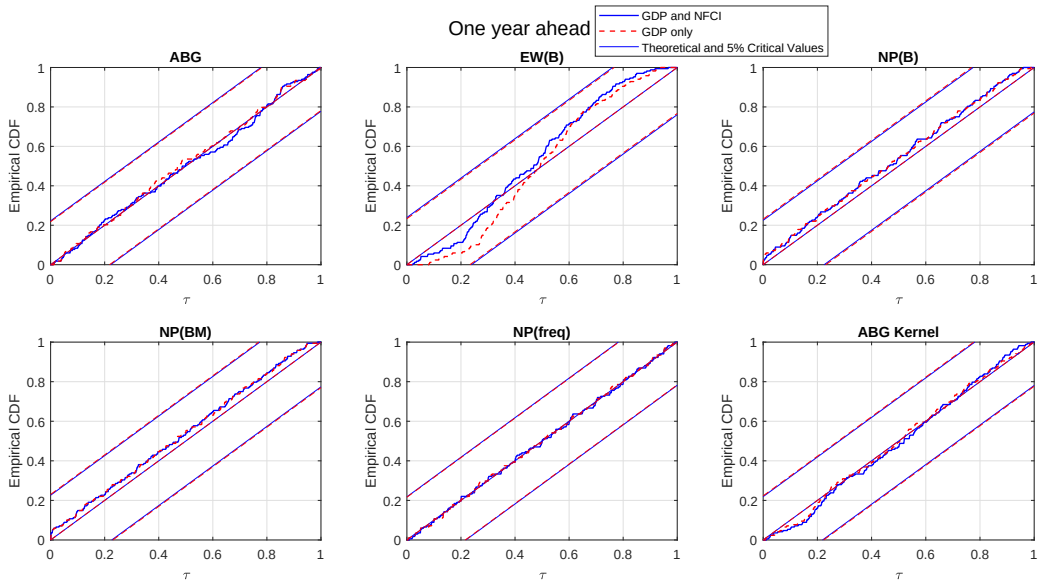


Figure 4: CDF of the in-sample PITs (one-year-ahead forecasts, 1973Q4-2015Q4) from the 6 density forecasts with and without NFCI. Note: the figures show the empirical CDF of the PITs (blue line) from the QR models with NFCI and lagged GDP, the empirical CDF of the PITs (dashed red line) from the QR models without NFCI, the CDF of the PITs under the null hypothesis of correct calibration (the 45-degree line), and the 5% critical value bands of the Rossi and Sekhposyan (2019) PITs test.

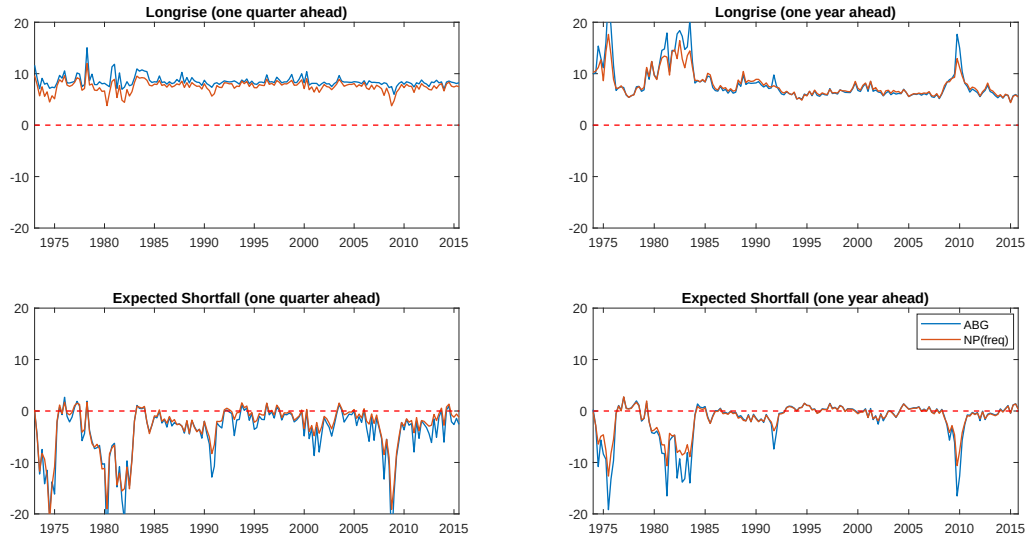


Figure 5: In-sample plots of the expected shortfall and expected longrise at $\tau = 0.05$ using ABG and NP(freq), from QRs with NFCI and lagged GDP

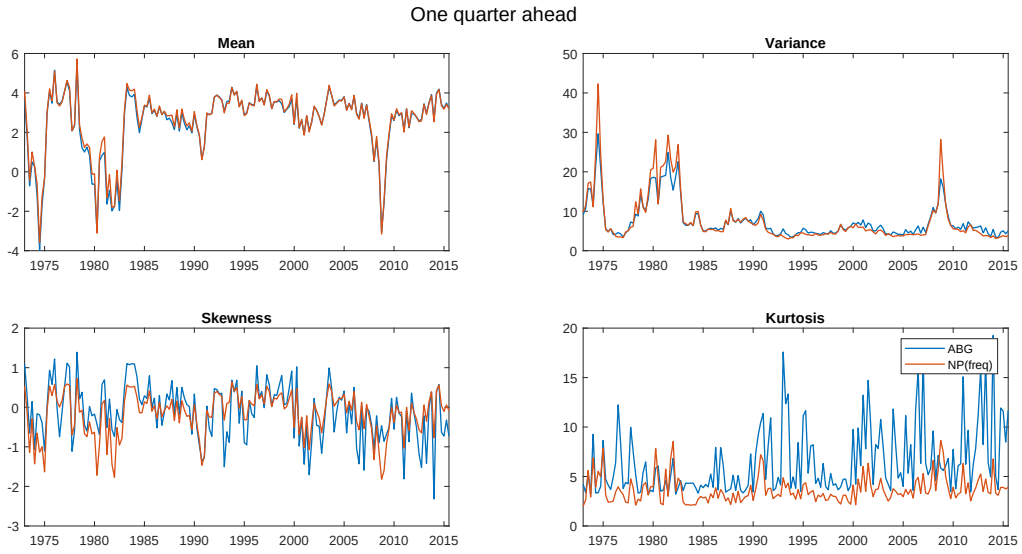


Figure 6: In-sample plots of the four moments of the ABG and NP(freq) forecast densities (one-quarter-ahead), from QRs with NFCI and lagged GDP

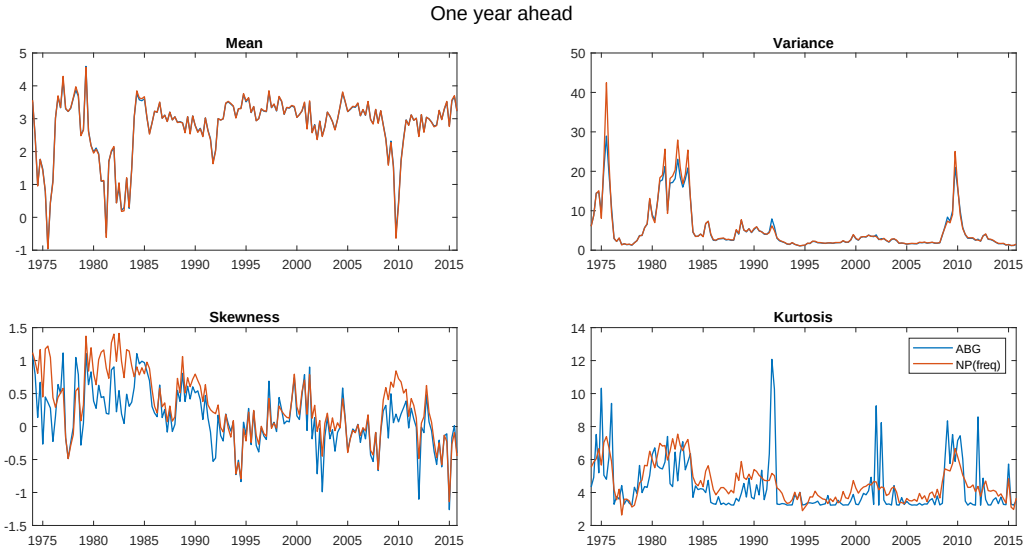


Figure 7: In-sample plots of the four moments of the ABG and NP(freq) forecast densities (one-year-ahead), from QRs with NFCI and lagged GDP

Figure 5 confirms that using our preferred density, NP(freq), when conditioned on both NFCI and lagged GDP growth, does not change the central narrative of ABG: the left tail of the conditional density of GDP growth moves with the tightness of financial conditions.[12]

[12]This "stylized fact" has been confirmed using alternative modeling approaches to QR, such as the

And the right tail is relatively invariant. Figure 5 evidences this by plotting, over time, the expected shortfall and longrise estimates from both ABG and NP(freq). Expected shortfall (SF_{t+h}) and longrise (LR_{t+h}) summarize downside and upside risk, respectively. They measure the total probability mass that the conditional distribution assigns to the left and right tails of the distribution:

$$SF_{t+h} = \frac{1}{\pi} \int_0^{\pi} \hat{F}_{y_{t+h}|x_t}(\tau|x_t) d\tau; \tag{28}$$

$$LR_{t+h} = \frac{1}{\pi} \int_{1-\pi}^{1} \hat{F}_{y_{t+h}|x_t}(\tau|x_t) d\tau. \tag{29}$$

Figure 5 shows that the expected shortfall and longrise estimates from ABG and NP(freq) track each other very closely. Expected shortfall is far more volatile than expected longrise, as the narrative of ABG emphasizes.

However, despite this similarity, when we look more deeply at the densities underlying these estimates we start to appreciate that the choice of how to construct the density from the quantile forecasts does still matter. It can reveal further features of economic interest. Figures 6 and 7 show this by plotting, over time, for the one-quarter-ahead and one-year-ahead forecasts, respectively, the first 4 moments of the ABG and NP(freq) densities. While the first two moments from ABG and NP(freq) are similar, the third and especially fourth moments differ, albeit they share some commonalities. In particular, we note how the evidence for or against skewness in GDP growth varies over time. This is consistent with Carriero et al. (2020a), who find, using alternative tests, weak evidence for skewness. Figure 6, in particular, shows that NP(freq) points to less negative skewness during the period of the global financial crisis.[13] This disagreement between ABG and NP(freq) is also consistent with the finding in Plagborg-Moller et al. (2020) that only the lower moments of the GDP growth conditional density are well-estimated.[14]

Next we provide some illustrative in-sample plots of our predictive densities. In Figure 8 we zoom in on a relatively stable period: 2005. Then, in Figure 9, we look at 2008, during the global financial crisis, a period also emphasized in ABG and Adrian et al. (2021). We focus on the one-quarter-ahead in-sample densities, with the analogous one-year-ahead and out-of-sample plots in the online appendix.[15] Confirming the findings of Adrian et al. (2021), who

parametric time-varying skewed-t model of De Polis et al. (2020).

[13]This is consistent with modest falls in the degree of asymmetry when NP(freq) rather than ABG is used in Figure 5. That is, while following the same general patterns, expected shortfall and longrise are more volatile, over time, when ABG rather than NP(freq) is consulted.

[14]Figures 13 and 14 in the online appendix indicate how ABG's coding choice to assess the skewed-t density over a finite grid is important. If, instead, we assess the skewed-t density analytically, instead of relying on ABG's approximation, we observe far more extreme estimates for the higher moments.

[15]Figures 19 through 24 in the online appendix qualitatively confirm the impression from Figures 8 and 9.

use kernel methods, clear evidence of multimodality emerges at the time of the global financial crisis when we use Algorithm 1 to construct the density forecast from the QR.[16] If, as in ABG, we assume a skewed-t density we obscure this important macroeconomic feature. Instead, we would simply infer more evidence for a skewed density. The evidence of multimodality during the global financial crisis, gleaned from NP(freq), is somewhat more muted when we look at the out-of-sample density forecasts as plotted in the online appendix. But, as shown by Figure 10, when the calibrated unimodality test of Hartigan and Hartigan (1985) as proposed by Cheng and Hall (1998) is used, rejections of unimodality are far greater when we do condition on NFCI. These rejections are especially pronounced during NBER recessionary periods, again confirming the finding of Adrian et al. (2021).

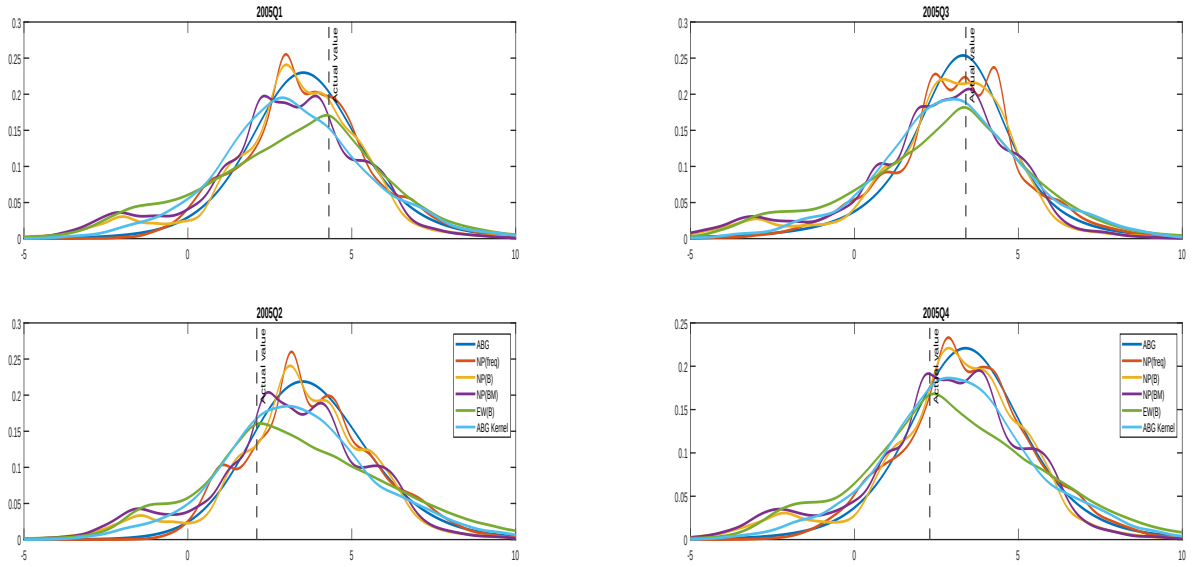


Figure 8: GDP growth density forecasts conditional on NFCI and lagged GDP for 2005 made one-quarter-ahead (in-sample)

[16]There is also recent evidence that professional forecasters' density forecasts for GDP growth are best acknowledged, at certain points in time, as multimodal. Ganics et al. (2020), who study the Survey of Professional Forecasters (SPF) in the US, find that multimodalities in their combined GDP growth densities emerge around business cycle turning points, such as the Great Recession.

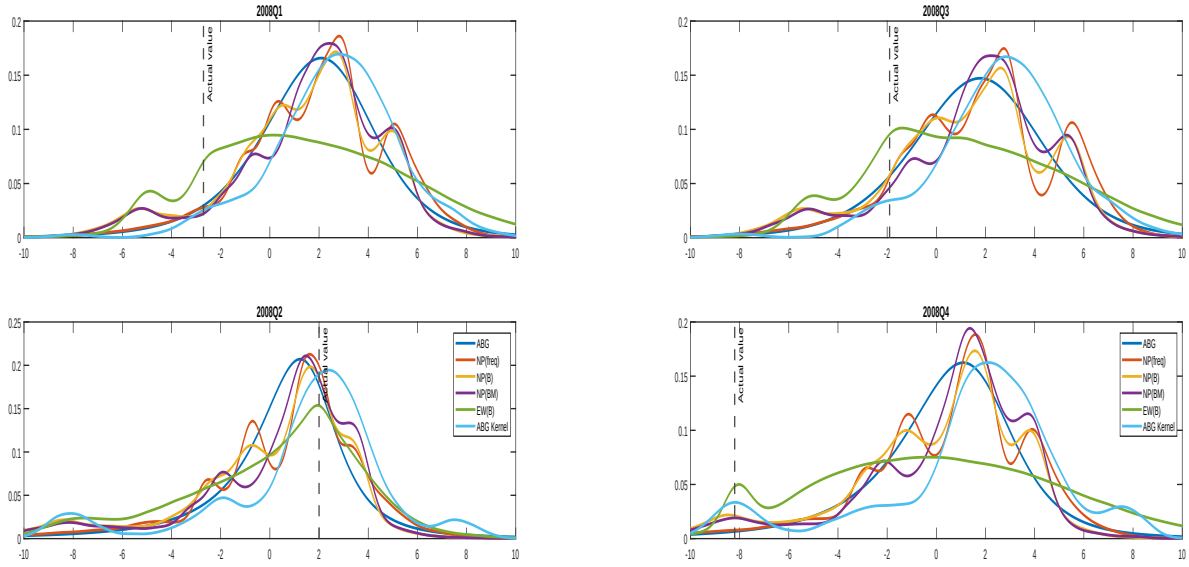
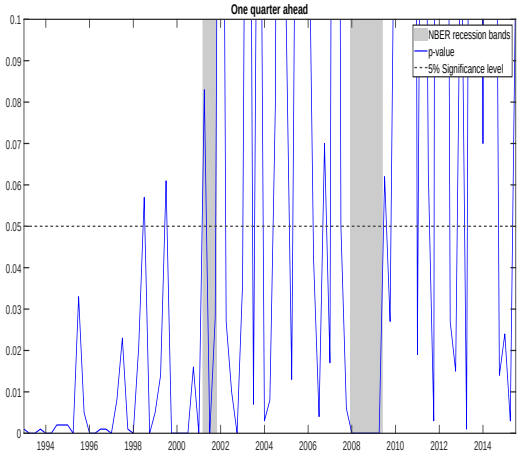
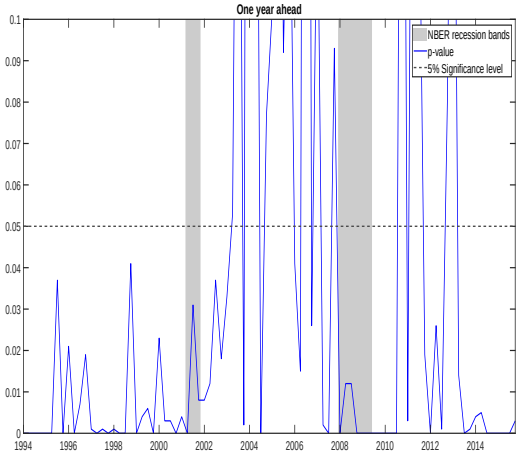


Figure 9: GDP growth density forecasts conditional on NFCI and lagged GDP for 2008 made one-quarter-ahead (in-sample)

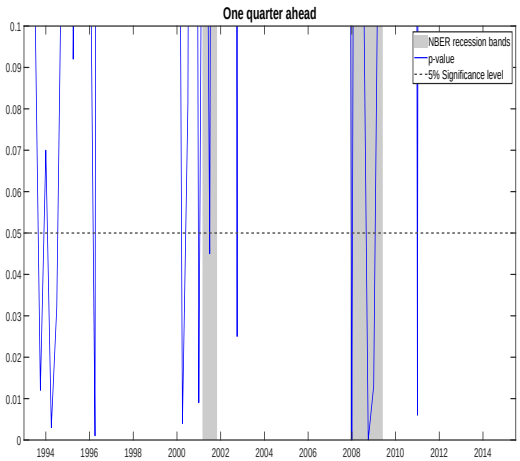
Panel A: NFCI and GDP



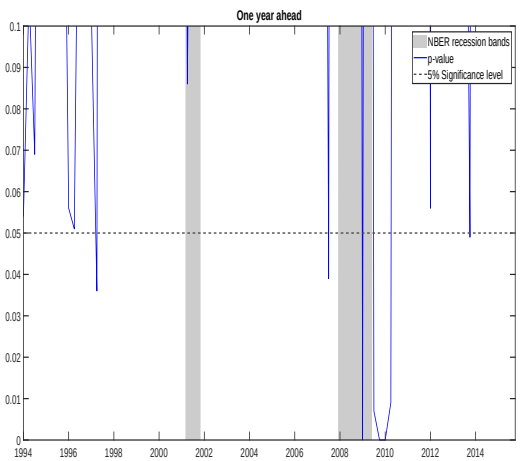
Panel B: NFCI and GDP



Panel C: GDP only



Panel D: GDP only



Notes: Panel A reports the p-values from the Hartigans' unimodality test (one-quarter-ahead) for the NP(freq) in-sample GDP growth density forecasts conditional on NFCI and lagged GDP. Panel B reports the p-values from the Hartigans' unimodality test over time (one-year-ahead) for the NP(freq) in-sample GDP growth density forecasts conditional on NFCI and lagged GDP. Panel C reports the p-values from the Hartigans' unimodality test over time (one-quarter-ahead) for the NP(freq) in-sample GDP growth density forecasts conditional on only lagged GDP. Panel D reports the p-values from the Hartigans' unimodality test over time (one-year-ahead) for the NP(freq) in-sample GDP growth density forecasts conditional only on lagged GDP. NBER recessionary periods are shaded gray.

Figure 10: P-values over time from the calibrated Hartigans' unimodality test

Finally, we turn to out-of-sample evaluation of the forecast densities over the sample period 1993Q1-2015Q4. Again this is the same evaluation period as ABG, and we follow ABG in recursively producing the predictive densities from QRs estimated on expanding windows of data dating back to 1973Q1. Figures 11 and 12 show that the accuracy of the forecast densities is worse out-of-sample. The null hypothesis of correct calibration is frequently rejected at a

95 percent significance level one-quarter-ahead, but not one-year-ahead. Comparison with the in-sample densities indicates that they too deteriorate in accuracy when evaluated on the sub-sample from 1993.[17] Interestingly, the PITs are closer to the 45-degree line when not conditioning on financial conditions, reminding us that autoregressive models can be hard to beat when forecasting out-of-sample.

Table 3 shows that out-of-sample the Bayesian QR methods, with Algorithm 1, tend to deliver the highest average logarithmic predictive scores and the lowest CRPSs. But the average logarithmic score statistics, in particular, are dominated by the forecasting failures at the time of the global financial crisis. So we prefer to emphasize the CRPS, given it is more robust to large but rare forecasting errors. Conditioning the GDP density forecasts on NFCI also tends to lead to improvements in the CRPS, especially one-quarter-ahead. Importantly, in terms of this paper's focus on isolating the best means of constructing density forecasts from the same quantile forecasts, Table 3 shows that NP(freq) at least matches the accuracy of ABG, according to CRPS, at both forecast horizons.

Despite the fact that the accuracy of the ABG densities is often improved upon, both in-sample and out-of-sample, this is not the key takeaway we wish to emphasize. Instead, the bottom line is that these alternative, nonparametric ways of constructing the predictive density from QRs *on average* match, and at times (albeit perhaps modestly) improve upon, the statistical accuracy of the ABG densities.[18] But in so doing they unmask deviations from unimodality lost by ABG. In turn, they suggest that multimodalities, rather than deviations from symmetry, are the primary *economic* feature of GDP density forecasts that should be emphasized, particularly when conditioning on financial conditions. But, as also emphasized by Ganics et al. (2020) in their analysis of the density forecasts from the SPF, periods when multimodalities emerge tend to be rare and short-lived. This means that accommodating them does not make a big difference when evaluating *the average* statistical performance of the models. But it affects the economic narrative.

[17] See Figures 17 and 18 in the online appendix.

[18] Giacomini and White (2006) tests confirm that the differences between the average scores seen in Table 3 are not statistically significant at traditional significance levels.

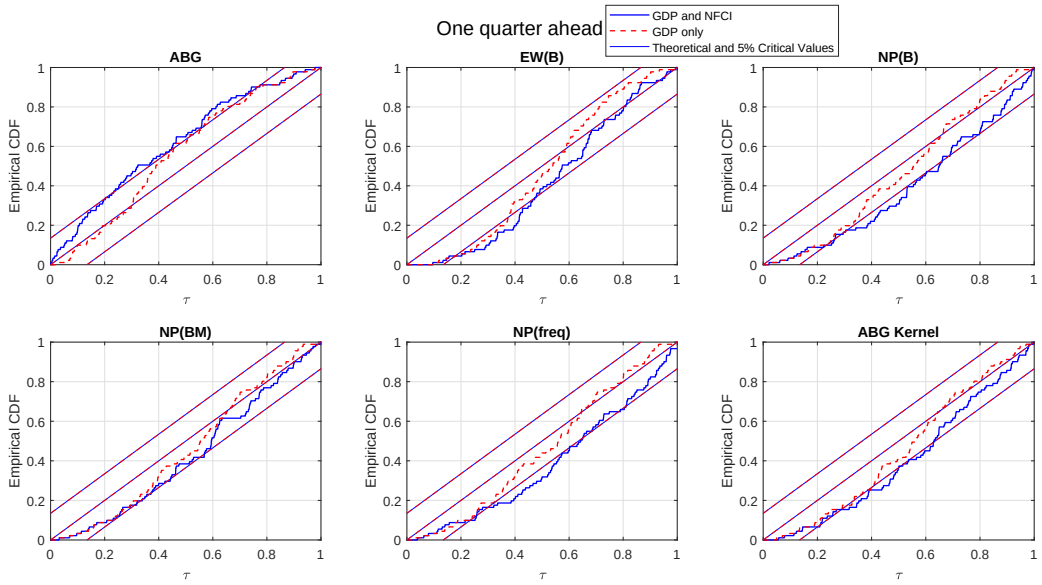


Figure 11: CDF of the out-of-sample PITs (one-quarter-ahead, 1993Q1-2015Q4) from the 6 density forecasts with NFCI and lagged GDP. Note: the figures show the empirical CDF of the PITs (red line), the CDF of the PITs under the null hypothesis of correct calibration (the 45-degree line), and the 5% critical value bands of the Rossi and Sekhposyan (2019) PITs test.

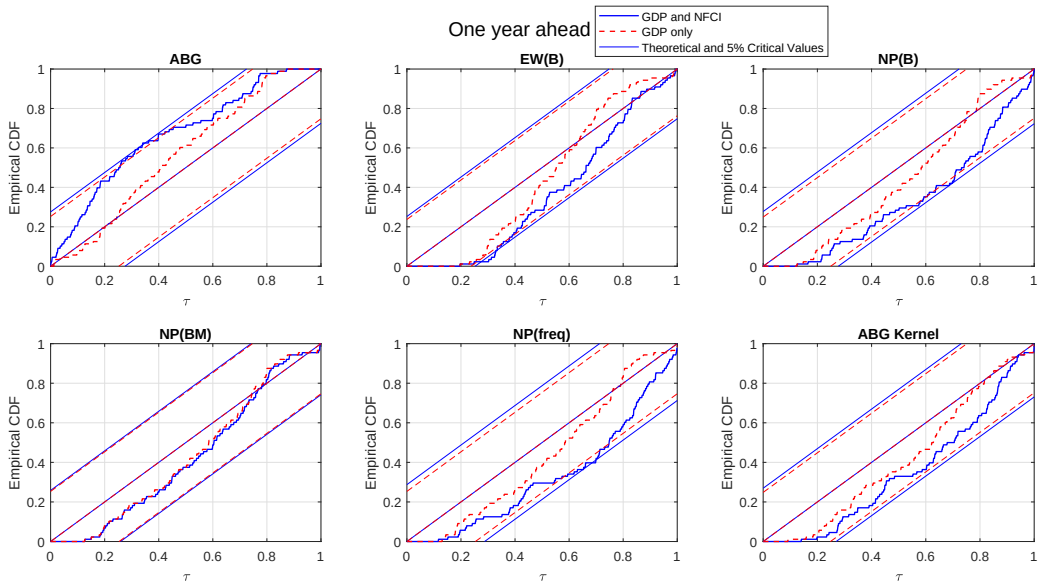


Figure 12: CDF of the out-of-sample PITs (one-year-ahead, 1993q4-2015Q4) from the 6 density forecasts with NFCI and lagged GDP. Note: the figures show the empirical CDF of the PITs (red line), the CDF of the PITs under the null hypothesis of correct calibration (the 45-degree line), and the 5% critical value bands of the Rossi and Sekhposyan (2019) PITs test.

Table 3: Average log predictive score (LPS) and continuous ranked probability score (CRPS) for the one-quarter-ahead forecasts (out-of-sample, 1993Q1-2015Q4) and the one-year-ahead forecasts (out-of-sample, 1993q4-2015Q4)

	With NFCI & GDP				With lagged GDP only			
	One-quarter-ahead		One-year-ahead		One-quarter-ahead		One-year-ahead	
	LPS	CRPS	LPS	CRPS	LPS	CRPS	LPS	CRPS
ABG	-2.24	1.27	-2.02	0.98	-2.31	1.32	-1.99	0.96
EW(B)	-2.32	1.30	-2.00	0.97	-2.45	1.38	-2.14	1.01
NP(B)	-2.25	1.25	-1.99	0.96	-2.31	1.29	-2.02	0.96
NP(BM)	-2.23	1.24	-2.01	0.95	-2.31	1.29	-2.02	0.96
NP(freq)	-2.47	1.26	-2.05	0.97	-2.33	1.29	-2.08	0.96
ABG Kernel	-2.27	1.31	-2.12	1.01	-2.34	1.32	-2.10	0.99

5 Conclusion

This paper reconsiders how to construct density forecasts from quantile regressions. While quantile regression methods are finding increasing application in macroeconomics and finance, as one means of accommodating nonlinear relationships, the specific issue of how to construct density forecasts from quantile regressions has received less attention. In the macroeconomic and finance literature, following ABG, it has become popular to assume a specific parametric form when matching the quantile forecasts to a density forecast. We reconsider nonparametric approaches to constructing predictive densities from quantile regressions, estimated either by frequentist or by Bayesian methods, and compare these with the parametric approach. We suggest a simple simulation-based algorithm. Unlike the parametric approach of ABG, we find that it can flexibly accommodate various distributional shapes.

In an application revisiting ABG, our proposed nonparametric approach corroborates the finding of Adrian et al. (2021) that the conditional density of GDP growth in the US can exhibit multimodality, especially during recessionary periods. These multimodalities in GDP growth are found to be increasingly prominent when the density forecasts, as suggested by ABG, are conditioned on financial conditions. But while Adrian et al. (2021) are forced to move away from the QR framework of ABG to document this novel empirical fact, we show that this finding is indeed shared by QR-based density forecasts - as long as we let the "data speak." However, we need to let the "data speak" not just when we model GDP growth with respect to financial conditions, via the first-step quantile regressions, but also when we

subsequently construct the forecast densities from the quantile forecasts.

Accordingly, this paper supports the addition of QR methods to the toolkit of the macro modeler. But it suggests that, when constructing density forecasts from quantile forecasts, it is better to respect the nonparametric flavor of QR by also using non- (or semi-) parametric methods to construct the density. Importantly, these methods provide similarly accurate, even slightly improved, density forecasts for US GDP growth. The methods are also operational irrespective of whether the first-step QRs are estimated via frequentist or Bayesian methods. Relative to ABG and their assumption that the forecast density is skewed-t, our nonparametric approach unmasks deviations from unimodality in GDP growth forecast densities when conditioned on financial conditions. The evolution of multimodalities, rather than asymmetries, then becomes the central macroeconomic narrative for the conditional predictive distribution of GDP growth. Following Adrian et al. (2021), this calls for structural macroeconomic models able to replicate these new empirical features, such as the nonlinear dynamic stochastic general equilibrium model of Rottner (2021) that allows for excessive leverage accumulation and endogenous financial crises.

References

Aastveit, Knut Are, James Mitchell, Francesco Ravazzolo, and Herman K. van Dijk (2019). "The Evolution of Forecast Density Combinations in Economics." In *Oxford Research Encyclopedia of Economics and Finance*. Oxford University Press. doi:10.1093/acrefore/9780190625979.013.381.

Adrian, Tobias, Nina Boyarchenko, and Domenico Giannone (2019). "Vulnerable Growth." *American Economic Review*, 109(4), pp. 1263–1289. doi:10.1257/aer.20161923.

Adrian, Tobias, Nina Boyarchenko, and Domenico Giannone (2021). "Multimodality in Macrofinancial Dynamics." *International Economic Review*, 62(2), pp. 861–886. doi:10.1111/iere.12501.

Amburgey, Aaron and Michael W. McCracken (2022). "On the Real-Time Predictive Content of Financial Conditions Indices for Growth." Working Paper 2022-003, Federal Reserve Bank of St. Louis. doi:10.20955/wp.2022.003.

Azzalini, Adelchi and Antonella Capitanio (2003). "Distributions Generated by Perturbation of Symmetry with Emphasis on a Multivariate Skew t-Distribution." *Journal of the Royal Statistical Society: Series B*, 65, pp. 367–389. doi:10.1111/1467-9868.00391.

Brownlees, Christian and André B.M. Souza (2021). "Backtesting Global Growth-at-Risk." *Journal of Monetary Economics*, 118, pp. 312–330. doi:10.1016/j.jmoneco.2020.11.003.

Caldara, Dario, Danilo Cascaldi-Garcia, Pablo Cuba-Borda, and Francesca Loria (2021). "Understanding Growth-at-Risk: A Markov Switching Approach." URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3992793, mimeo, Federal Reserve Board.

Carriero, Andrea, Todd E. Clark, and Massimiliano Marcellino (2020a). "Capturing Macroeconomic Tail Risks with Bayesian Vector Autoregressions." Working Paper 20-02, Federal Reserve Bank of Cleveland. doi:10.26509/frbc-wp-202002r.

Carriero, Andrea, Todd E. Clark, and Massimiliano Marcellino (2020b). "Nowcasting Tail Risks to Economic Activity with Many Indicators." Working Paper 20-13, Federal Reserve Bank of Cleveland. doi:10.26509/frbc-wp-202013.

Chen, Liang, Juan J. Dolado, and Jesús Gonzalo (2021). "Quantile Factor Models." *Econometrica*, 89(2), pp. 875–910. doi:10.3982/ECTA15746.

Cheng, Ming-Yen and Peter Hall (1998). "Calibrating the Excess Mass and Dip Tests of Modality." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 60(3), pp. 579–589. doi:10.1111/1467-9868.00141.

Chernozhukov, Victor (2005). "Extremal Quantile Regression." *The Annals of Statistics*, 33(2), pp. 806–839. doi:10.1214/009053604000001165.

Chernozhukov, Victor, Iván Fernández-Val, and Alfred Galichon (2010). "Quantile and Probability Curves Without Crossing." *Econometrica*, 78(3), pp. 1093–1125. doi:10.3982/ECTA7880.

De Polis, Andrea, Davide Delle Monache, and Ivan Petrella (2020). "Modeling and Forecasting Macroeconomic Downside Risk." CEPR Discussion Paper 15109. URL https://ideas.repec.org/p/cpr/ceprdp/15109.html.

Diebold, Francis X., Todd A. Gunther, and Anthony S. Tay (1998). "Evaluating Density Forecasts with Application to Financial Risk Management." *International Economic Review*, 39, pp. 863–883. doi:10.2307/2527342.

Fernandes, Marcelo, Emmanuel Guerre, and Eduardo Horta (2021). "Smoothing Quantile Regressions." *Journal of Business and Economic Statistics*, 39(1), pp. 338–357. doi:10.1080/07350015.2019.1660177.

Ferrara, Laurent, Matteo Mogliani, and Jean-Guillaume Sahuc (2022). "High-Frequency Monitoring of Growth at Risk." *International Journal of Forecasting*, 38(2), pp. 582–595. doi:10.1016/j.ijforecast.2021.06.010.

Figueres, Juan Manuel and Marek Jarocinski (2020). "Vulnerable Growth in the Euro Area: Measuring the Financial Conditions." *Economics Letters*, 191. doi:10.1016/j.econlet.2020.109126.

Gaglianone, Wagner Piazza and Luiz Renato Lima (2012). "Constructing Density Forecasts from Quantile Regressions." *Journal of Money, Credit and Banking*, 44(8), pp. 1589–1607. doi:10.1111/j.1538-4616.2012.00545.x.

Gaglianone, Wagner Piazza and Luiz Renato Lima (2014). "Constructing Optimal Density Forecasts From Point Forecast Combinations." *Journal of Applied Econometrics*, 29(5), pp. 736–757. doi:10.1002/jae.2352.

Ganics, Gergely, Barbara Rossi, and Tatevik Sekhposyan (2020). "From Fixed-event to Fixed-horizon Density Forecasts: Obtaining Measures of Multi-horizon Uncertainty from

Survey Density Forecasts." CEPR Discussion Papers 14267. URL https://cepr.org/active/publications/discussion_papers/dp.php?dpno=14267.

Ghysels, Eric, Leonardo Iania, and Jonas Striaukas (2018). "Quantile-based Inflation Risk Models." Working Paper Research 349, National Bank of Belgium. URL https://ideas.repec.org/p/nbb/reswpp/201810-349.html.

Giacomini, Raffaella and Halbert White (2006). "Tests of Conditional Predictive Ability." *Econometrica*, 74, pp. 1545–1578. doi:10.1111/j.1468-0262.2006.00718.x.

Giglio, Stefano, Bryan Kelly, and Seth Pruitt (2016). "Systemic Risk and the Macroeconomy: An Empirical Evaluation." *Journal of Financial Economics*, 119(3), pp. 457–471. doi:10.1016/j.jfineco.2016.01.010.

Gneiting, Tilmann and Adrian E Raftery (2007). "Strictly Proper Scoring Rules, Prediction, and Estimation." *Journal of the American Statistical Association*, 102, pp. 359–378. doi:10.1198/016214506000001437.

Hartigan, J. A. and P. M. Hartigan (1985). "The Dip Test of Unimodality." *The Annals of Statistics*, 13(1), pp. 70 – 84. doi:10.1214/aos/1176346577.

Koenker, Roger (2005). *Quantile Regression*. Cambridge University Press. doi:10.1017/CBO9780511754098.

Koenker, Roger and José António Machado (1999). "GMM Inference when the Number of Moment Conditions is Large." *Journal of Econometrics*, 93(2), pp. 327–344. doi:10.1016/S0304-4076(99)00014-7.

Koenker, Roger and Quanshui Zhao (1996). "Conditional Quantile Estimation and Inference for ARCH Models." *Econometric Theory*, 12(5), pp. 793–813. doi:10.1017/S0266466600007167.

Kohns, David and Tibor Szendrei (2021). "Horseshoe Prior Bayesian Quantile Regression." URL https://ideas.repec.org/p/arx/papers/2006.07655.html, arXiv.org: 2006.07655.

Komunjer, Ivana (2013). *Quantile Prediction*, volume 2 of *Handbook of Economic Forecasting*, pp. 961–994. Elsevier. doi:10.1016/B978-0-444-62731-5.00017-8.

Korobilis, Dimitris (2017). "Quantile Regression Forecasts of Inflation Under Model Uncertainty." *International Journal of Forecasting*, 33, pp. 11–20. doi:10.1016/j.ijforecast.2016.07.005.

Korobilis, Dimitris, Bettina Landau, Alberto Musso, and Anthoulla Phella (2021). "The Time-varying Evolution of Inflation Risks." Working Paper Series 2600, European Central Bank. URL https://www.ecb.europa.eu//pub/pdf/scpwps/ecb.wp2600~8dae8e832f.en.pdf.

Kozumi, Hideo and Genya Kobayashi (2011). "Gibbs Sampling Methods for Bayesian Quantile Regression." *Journal of Statistical Computation and Simulation*, 81(11), pp. 1565–1578. doi:10.1080/00949655.2010.496117.

Krüger, Fabian, Sebastian Lerch, Thordis Thorarinsdottir, and Tilmann Gneiting (2021). "Predictive Inference Based on Markov Chain Monte Carlo Output." *International Statistical Review*, 89(2), pp. 274–301. doi:10.1111/insr.12405.

Manzan, Sebastiano (2015). "Forecasting the Distribution of Economic Variables in a Data-Rich Environment." *Journal of Business and Economic Statistics*, 33(1), pp. 144–164. doi:10.1080/07350015.2014.937436.

Manzan, Sebastiano and Dawit Zerom (2013). "Are Macroeconomic Variables Useful for Forecasting the Distribution of U.S. Inflation?" *International Journal of Forecasting*, 29(3), pp. 469–478. doi:10.1016/j.ijforecast.2013.01.005.

Mitchell, James, Aubrey Poon, and Gian-Luigi Mazzi (2022). "Nowcasting Euro Area GDP Growth using Bayesian Quantile Regression." *Advances in Econometrics: Essays in Honor of M Hashem Pesaran*, 43A, pp. 51–72. doi:10.1108/S0731-90532021000043A004.

Mitchell, James and Kenneth F. Wallis (2011). "Evaluating Density Forecasts: Forecast Combinations, Model Mixtures, Calibration and Sharpness." *Journal of Applied Econometrics*, 26(6), pp. 1023–1040. doi:10.1002/jae.1192.

Parzen, Emanuel (1979). "Nonparametric Statistical Data Modeling." *Journal of the American Statistical Association*, 74(365), pp. 105–121. doi:10.2307/2286734.

Plagborg-Moller, Mikkel, Lucrezia Reichlin, Giovanni Ricco, and Thomas Hasenzagl (2020). "When is Growth at Risk?" *Brookings Papers on Economic Activity*, pp. 167–213. URL https://www.brookings.edu/bpea-articles/when-is-growth-at-risk/.

Prasad, Ananthakrishnan, Selim Elekdag, Phakawa Jeasakul, Romain Lafarguette, Adrian Alter, Alan X. Feng, and Changchun Wang (2019). "Growth at Risk: Concept and Application in IMF Country Surveillance." IMF Working Paper 19/36, International Monetary Fund. URL https://www.imf.org/en/Publications/WP/Issues/2019/02/21/Growth-at-Risk-Concept-and-Application-in-IMF-Country-Surveillance-46567.

Reichlin, Lucrezia, Giovanni Ricco, and Thomas Hasenzagl (2020). "Financial Variables as Predictors of Real Growth Vulnerability." Discussion Papers 05/2020, Deutsche Bundesbank. URL https://www.bundesbank.de/resource/blob/827682/6bd9b43c07b6cf1cc998b9bf12b4d7c4/mL/2020-03-05-dkp-05-data.pdf.

Rossi, Barbara (2021). "Forecasting in the Presence of Instabilities: How We Know Whether Models Predict Well and How to Improve Them." *Journal of Economic Literature*, 59(4), pp. 1135–90. doi:10.1257/jel.20201479.

Rossi, Barbara and Tatevik Sekhposyan (2019). "Alternative Tests for Correct Specification of Conditional Predictive Densities." *Journal of Econometrics*, 208(2), pp. 638–657. doi:10.1016/j.jeconom.2018.07.008.

Rottner, Matthias (2021). "Financial Crises and Shadow Banks: A Quantitative Analysis." Economics Working Papers EUI ECO 2021/02, European University Institute. URL https://ideas.repec.org/p/eui/euiwps/eco2021-02.html.

Yu, Keming and Rana A. Moyeed (2001). "Bayesian Quantile Regression." *Statistics and Probability Letters*, 54(4), pp. 437 – 447. doi:10.1016/S0167-7152(01)00124-9.

6 Online Appendix

This appendix contains supplementary tables and figures referred to in the main paper.

Table 4: Average mean squared error and Kullback-Leibler (KL) distance for NP(freq) using $k = 4$ and $k = 99$

Models	**Mean**	**Variance**	**Skewness**	**Kurtosis**	**KL**
$k = 4$ **and** $T = 100$					
DGP1: Unimodal (Gaussian)	0.02	1.19	0.05	1.84	0.31
DGP2: Unimodal (Negative Skewness)	0.17	20.87	0.02	5.42	0.32
DGP3: Unimodal (High Kurtosis)	0.03	1.21	1.26	105.97	0.35
DGP4: Bimodal	0.02	0.03	0.16	0.28	0.32
DGP5: Trimodal	0.04	0.36	0.00	0.06	0.36
$k = 4$ **and** $T = 1,000$					
DGP1: Unimodal (Gaussian)	0.01	1.02	0.04	1.85	0.30
DGP2: Unimodal (Negative Skewness)	0.06	18.61	0.01	5.42	0.32
DGP3: Unimodal (High Kurtosis)	0.00	0.76	1.23	106.43	0.35
DGP4: Bimodal	0.02	0.03	0.15	0.27	0.32
DGP5: Trimodal	0.04	0.36	0.00	0.06	0.36
$k = 99$ **and** $T = 100$					
DGP1: Unimodal (Gaussian)	0.01	0.02	0.07	0.35	0.03
DGP2: Unimodal (Negative Skewness)	0.05	0.59	0.39	18.53	0.04
DGP3: Unimodal (High Kurtosis)	0.01	0.16	1.16	61.56	0.04
DGP4: Bimodal	0.00	0.00	0.01	0.02	0.03
DGP5: Trimodal	0.00	0.00	0.01	0.02	0.05
$k = 99$ **and** $T = 1,000$					
DGP1: Unimodal (Gaussian)	0.00	0.00	0.01	0.06	0.00
DGP2: Unimodal (Negative Skewness)	0.00	0.07	0.02	0.45	0.00
DGP3: Unimodal (High Kurtosis)	0.00	0.02	0.37	59.98	0.00
DGP4: Bimodal	0.00	0.00	0.00	0.00	0.01
DGP5: Trimodal	0.00	0.00	0.00	0.00	0.03

DGP	$T = 25$	$T = 50$	$T = 250$	$T = 1,000$	$T = 10,000$
			$k = 4$		
Gaussian	28.70%	30.00%	23.30%	15.20%	2.00%
Negative Skewness	18.00%	16.60%	5.30%	1.00%	0.00%
High Kurtosis	14.00%	9.70%	0.20%	0.00%	0.00%
Bimodal	94.50%	98.70%	100.00%	100.00%	100.00%
Trimodal	16.80%	8.50%	0.20%	0.00%	0.00%
			$k = 9$		
Gaussian	27.80%	13.00%	0.00%	0.00%	0.00%
Negative Skewness	21.30%	8.30%	0.00%	0.00%	0.00%
High Kurtosis	20.50%	8.00%	0.10%	0.00%	0.00%
Bimodal	93.30%	96.60%	99.90%	100.00%	100.00%
Trimodal	71.50%	18.80%	0.90%	0.00%	0.00%
			$k = 19$		
Gaussian	34.20%	11.70%	0.00%	0.00%	0.00%
Negative Skewness	31.80%	9.50%	0.00%	0.00%	0.00%
High Kurtosis	30.10%	7.90%	0.00%	0.00%	0.00%
Bimodal	91.30%	86.30%	94.20%	99.50%	100.00%
Trimodal	95.60%	94.90%	97.40%	99.90%	100.00%
			$k = 99$		
Gaussian	65.10%	31.70%	0.00%	0.00%	0.00%
Negative Skewness	62.60%	28.50%	0.00%	0.00%	0.00%
High Kurtosis	59.20%	20.30%	0.00%	0.00%	0.00%
Bimodal	96.00%	96.40%	98.70%	99.90%	100.00%
Trimodal	99.10%	99.30%	99.90%	100.00%	100.00%

Table 5: Rejection rates (across 1,000 replications) of unimodality at 95% using the calibrated Hartigan test

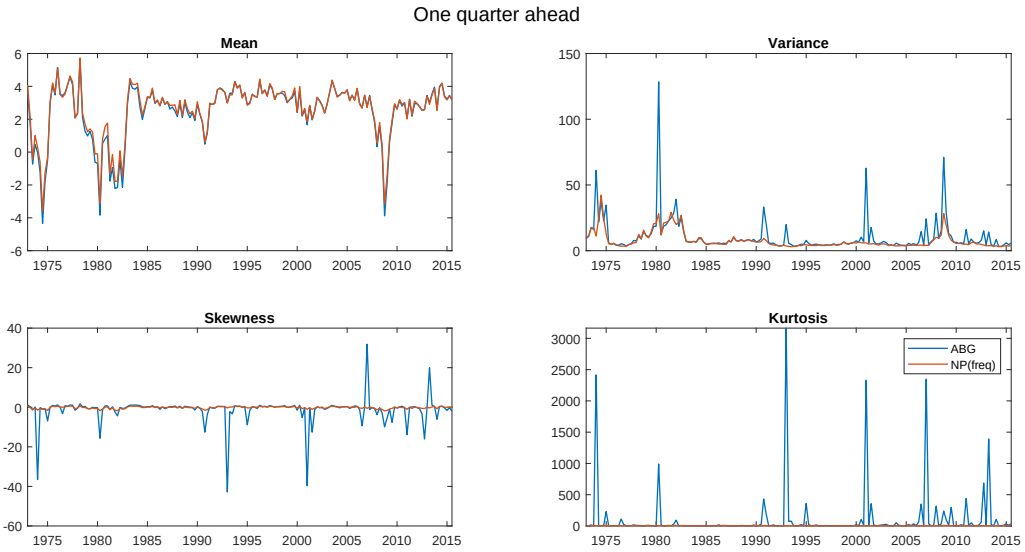


Figure 13: In-sample plots of the four moments of the ABG and NP(freq) forecast densities (one-quarter-ahead), when ABG's skewed-t density is simulated not truncated

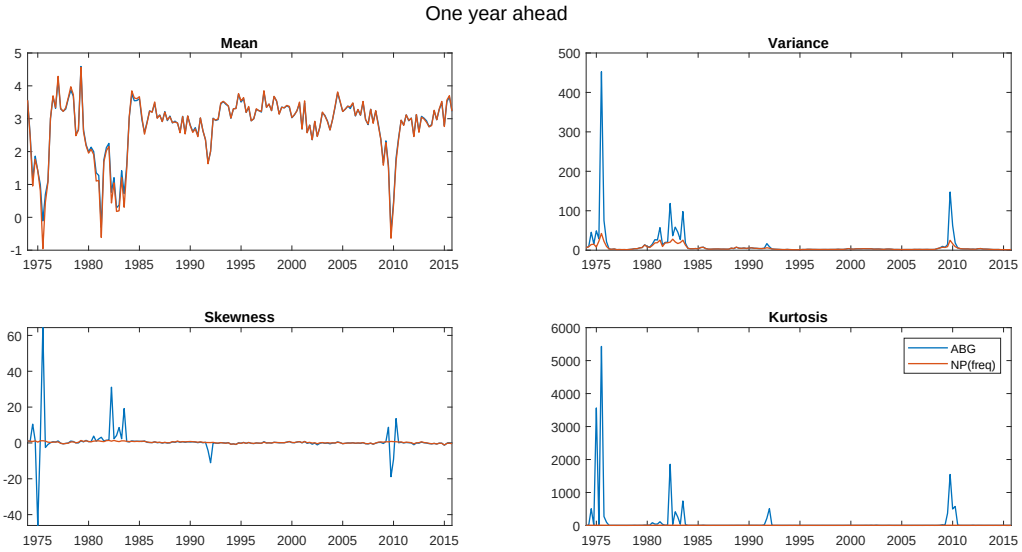


Figure 14: In-sample plots of the four moments of the ABG and NP(freq) forecast densities (one-year-ahead), when ABG's skewed-t density is simulated not truncated

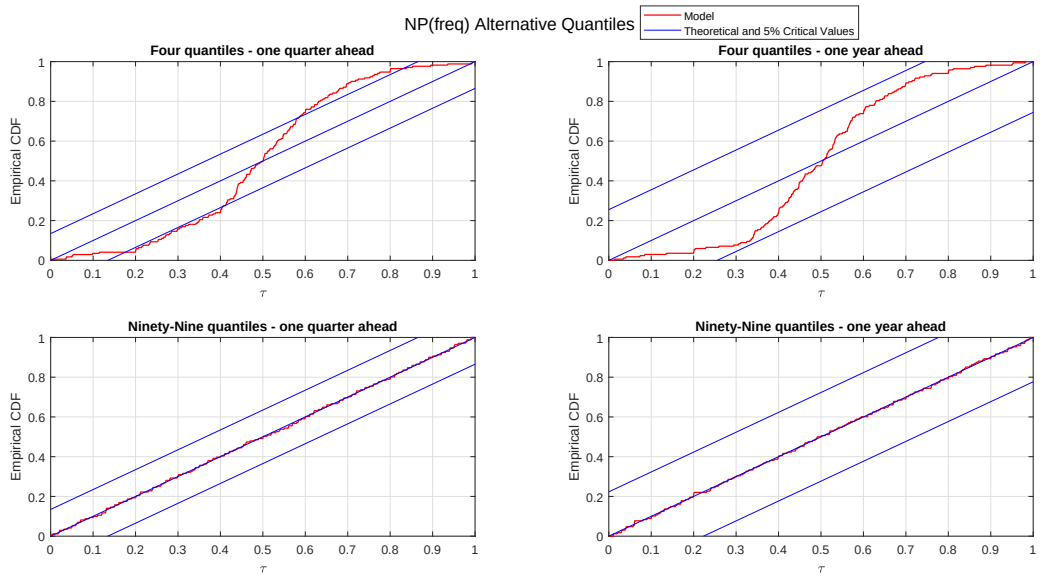


Figure 15: CDF of the in-sample PITs for NP(freq) when $k = 4$ and $k = 99$. Note: the figures show the empirical CDF of the PITs (red line), the CDF of the PITs under the null hypothesis of correct calibration (the 45-degree line), and the 5% critical value bands of the Rossi and Sekhposyan (2019) PITs test.

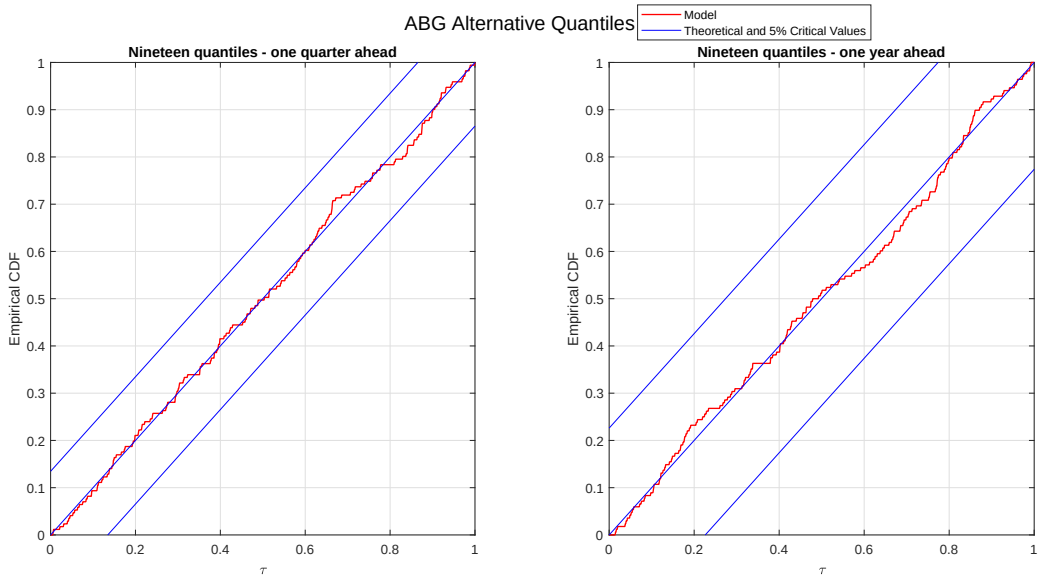


Figure 16: CDF of the in-sample PITs for ABG when $k = 19$. Note: the figures show the empirical CDF of the PITs (red line), the CDF of the PITs under the null hypothesis of correct calibration (the 45-degree line), and the 5% critical value bands of the Rossi and Sekhposyan (2019) PITs test.

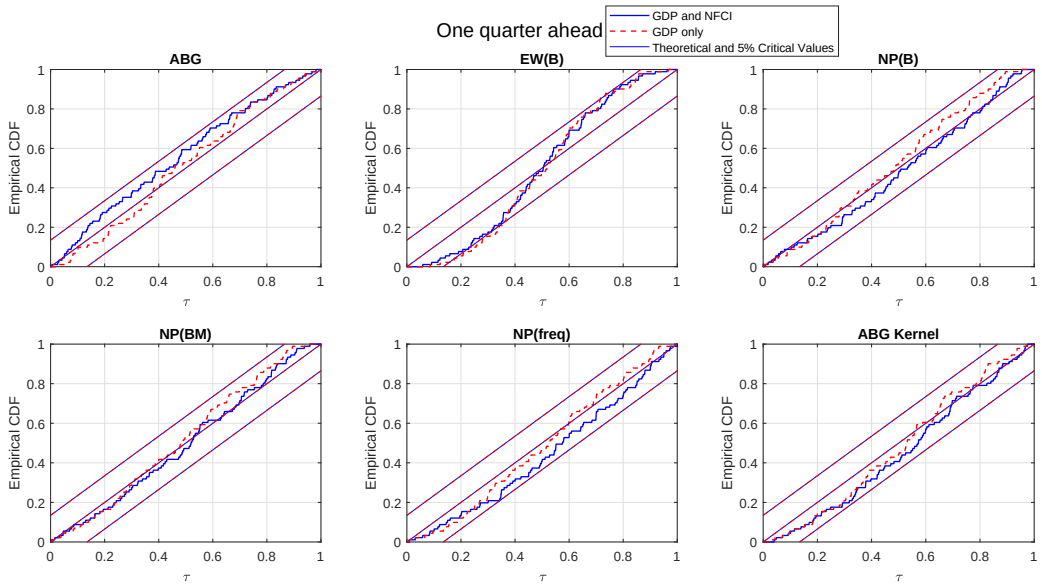


Figure 17: CDF of the in-sample PITs (one-quarter-ahead forecasts, 1993Q1-2015Q4) from the 6 density forecasts with and without NFCI. Note: the figures show the empirical CDF of the PITs (blue line) from the QR models with NFCI (and lagged GDP), the empirical CDF of the PITs (dashed red line) from the QR models without NFCI, plus the CDF of the PITs under the null hypothesis of correct calibration (the 45-degree line), and the 5% critical value bands of the Rossi and Sekhposyan (2019) PITs test.

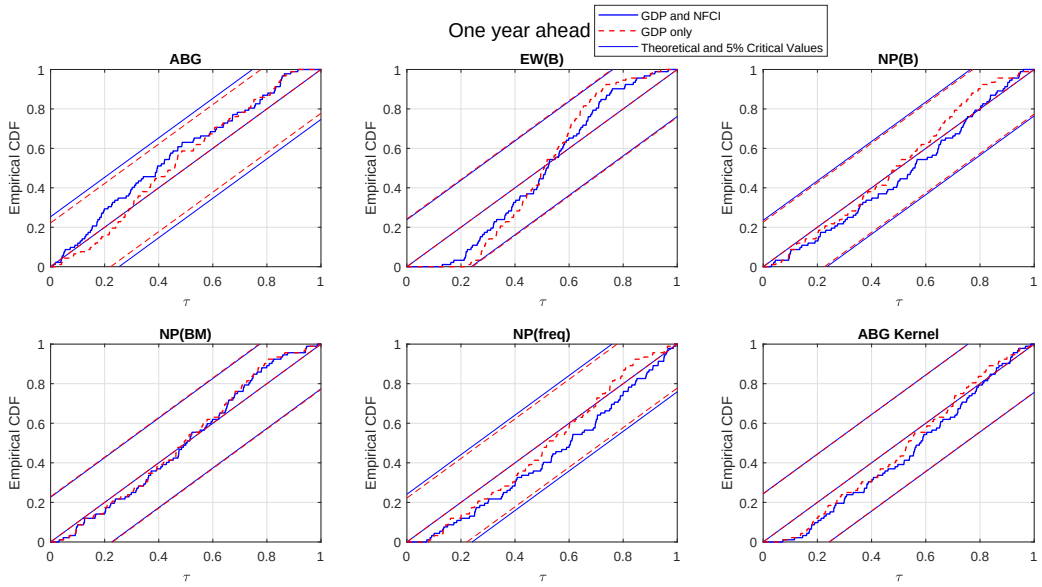


Figure 18: CDF of the in-sample PITs (one-year-ahead forecasts, 1993Q4-2015Q4) from the 6 density forecasts with and without NFCI. Note: the figures show the empirical CDF of the PITs (blue line) from the QR models with NFCI (and lagged GDP), the empirical CDF of the PITs (dashed red line) from the QR models without NFCI, plus the CDF of the PITs under the null hypothesis of correct calibration (the 45-degree line), and the 5% critical value bands of the Rossi and Sekhposyan (2019) PITs test.

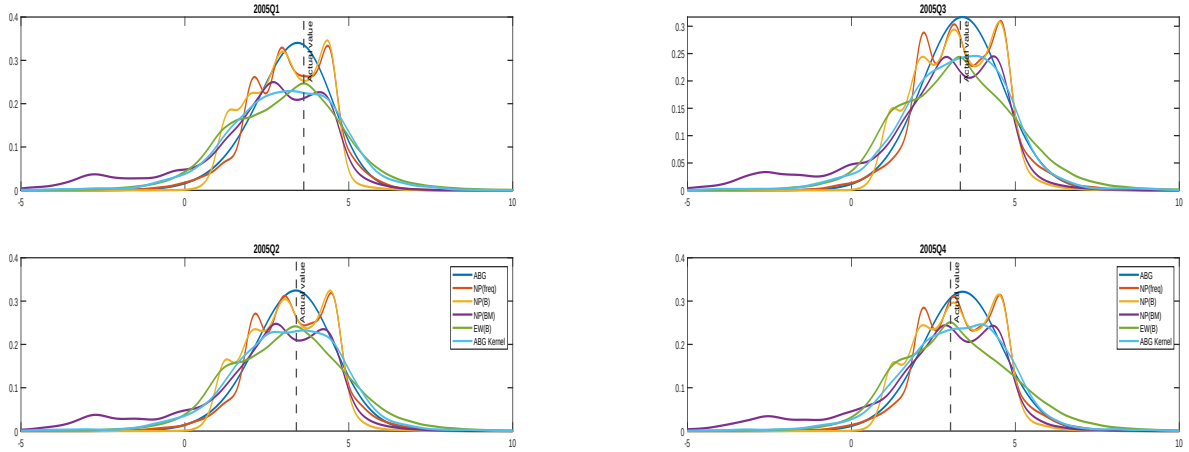


Figure 19: GDP growth density forecasts conditional on NFCI and lagged GDP for 2005 made one-year-ahead (in-sample)

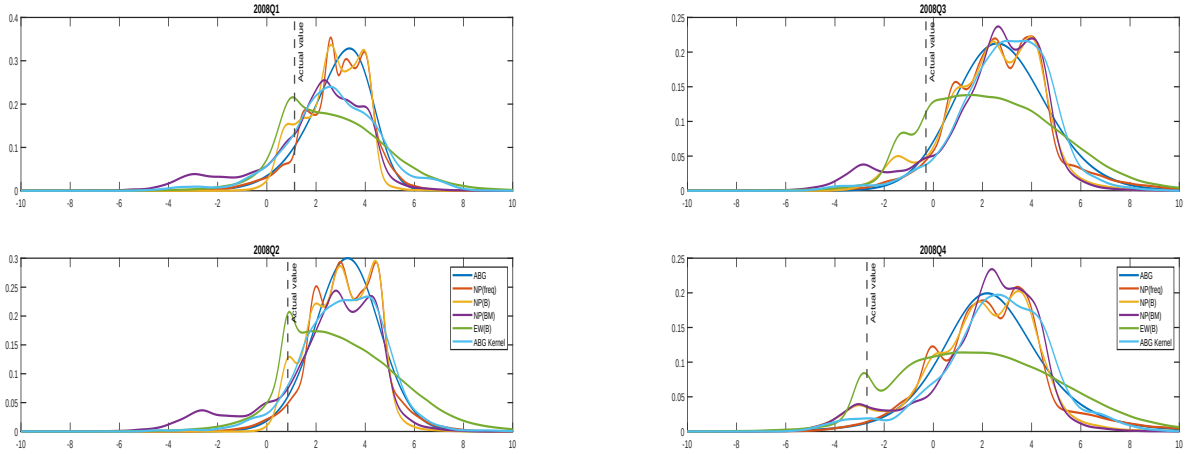


Figure 20: GDP growth density forecasts conditional on NFCI and lagged GDP for 2008 made one-year-ahead (in-sample)

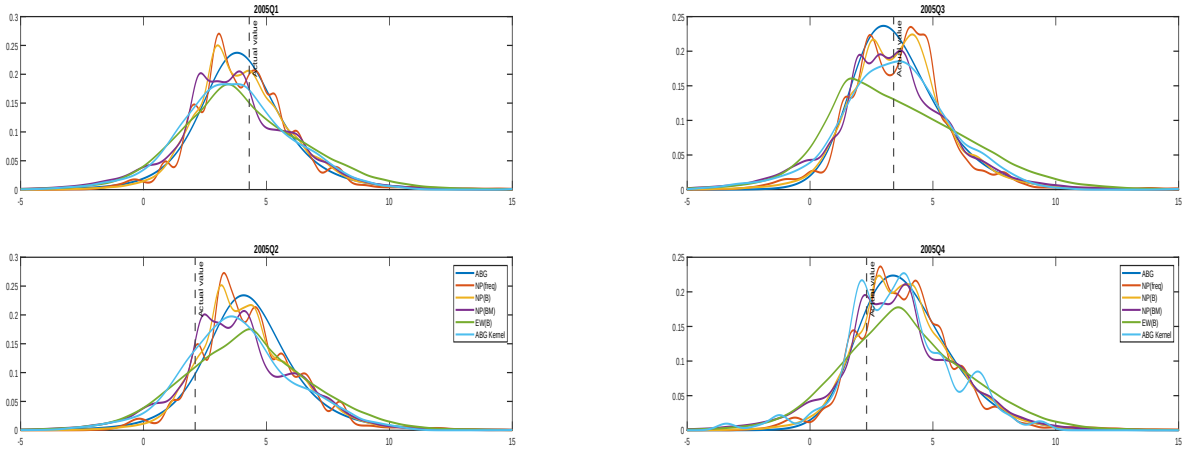


Figure 21: GDP growth density forecasts conditional on NFCI and lagged GDP for 2005 made one-quarter-ahead (out-of-sample)

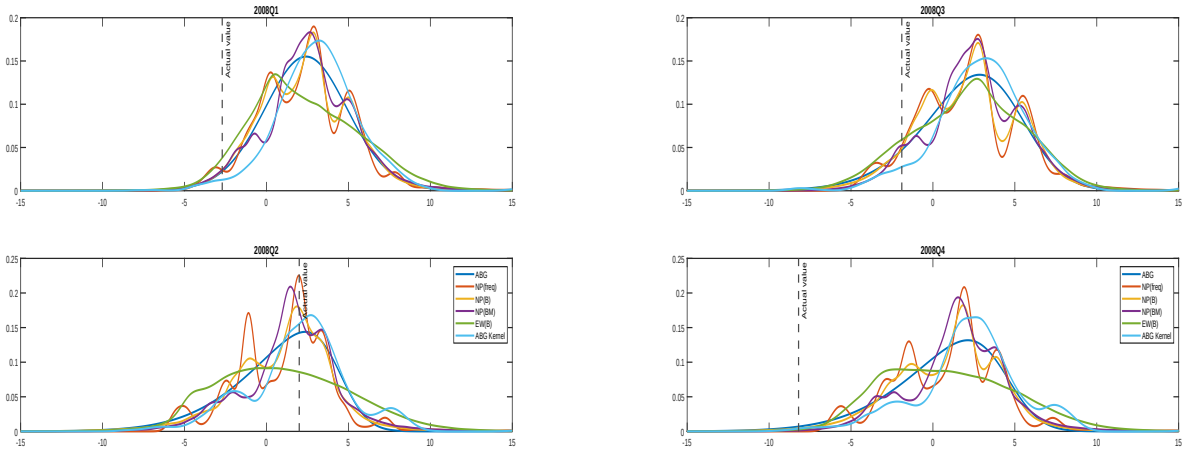


Figure 22: GDP growth density forecasts conditional on NFCI and lagged GDP for 2008 made one-quarter-ahead (out-of-sample)

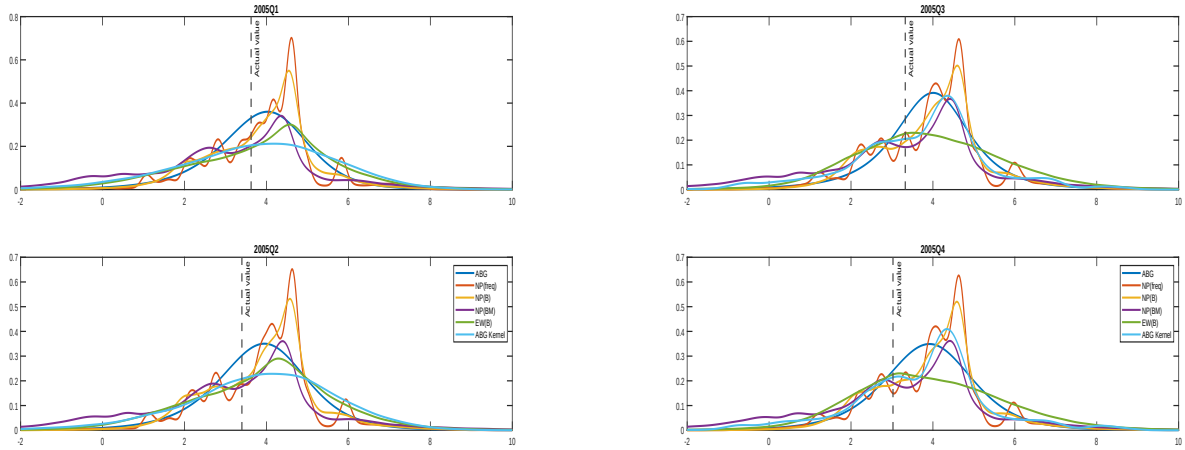


Figure 23: GDP growth density forecasts conditional on NFCI and lagged GDP for 2005 made one-year-ahead (out-of-sample)

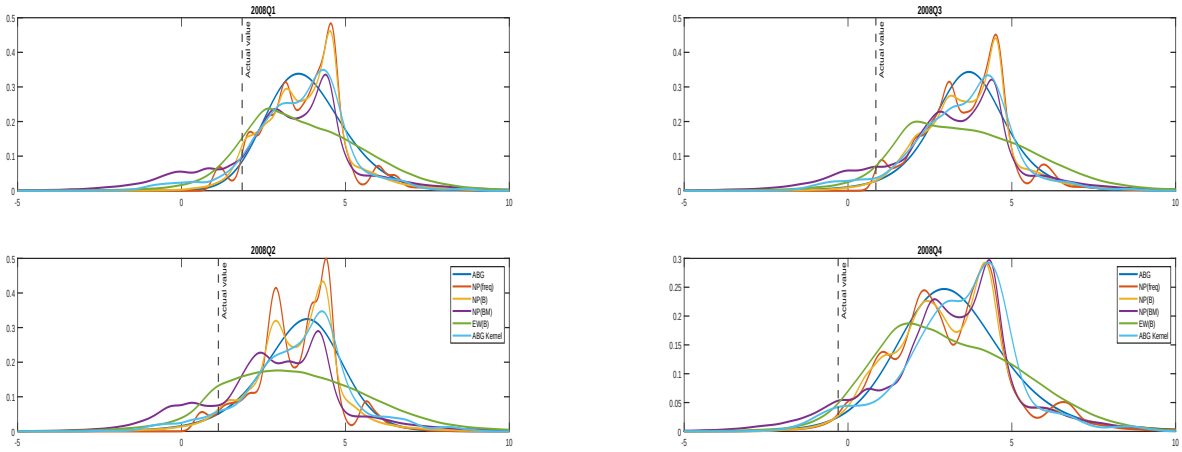


Figure 24: GDP growth density forecasts conditional on NFCI and lagged GDP for 2008 made one-year-ahead (out-of-sample)