Constructing Density Forecasts from Quantile Regressions: Multimodality in Macroeconomic Dynamics *

James Mitchell (Federal Reserve Bank of Cleveland)†
Aubrey Poon (University of Strathclyde)‡
Dan Zhu (Monash University)§

November 5, 2021

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 growth-at-risk and GDP density forecasts 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 and show how this affects the macroeconomic narrative about US GDP growth.

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, em-

*Thanks to Todd Clark, Gary Koop and Benjamin Wang 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.

†james.mitchell@clev.frb.org
‡aubrey.poon@strath.ac.uk
§Dan.Zhu@monash.edu

phasizing asymmetries in the distribution of (real) GDP growth.[1] A commonly adopted approach in this literature, following Adrian et al. (2019) [henceforth ABG], is to produce the density forecasts in two-steps. At 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, at a second step, the skewed-t density function of Azzalini & 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, like 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 re-considers the use of QRs when interests 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 (e.g., see Komunjer (2013)).

Our paper proposes and then contrasts with the aforementioned two-step ABG method, that has become so established, a simple nonparametric (strictly "semi-parametric") approach to the production of density forecasts from QRs. Unlike ABG, 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, only assuming 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 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

[1] On the use of QR methods to produce density nowcasts and forecasts, see e.g., Gaglianone & Lima (2012), Gaglianone & Lima (2014), Manzan & Zerom (2013), Manzan (2015), Korobilis (2017), Ferrara et al. (2021), Chen et al. (2021) and Mitchell et al. (2021). 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. (2020*b*), Carriero et al. (2020*a*), Reichlin et al. (2020), Brownlees & Souza (2021) and Figueres & Jarocinski (2020).

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 central macroeconomic narrative of the conditional predictive distribution of GDP growth.

The plan of the remainder of this paper is 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 & Machado (1999) established that likelihood-based inference using independently distributed asym-

[2]Recent research in macroeconomics has moved onto 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.

metric Laplace densities (ALD) is directly related to (2). Yu & Moyeed (2001) show how exact Bayesian inference using Markov Chain Monte Carlo (MCMC) 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 & 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):

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

where, for notational ease, we denote these quantile forecasts $\hat{Q}_{y_{t+h|x_t}}(\tau_j|x_t) = x_t'\hat{\beta}_\tau$; i.e., 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 it contradicting 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 & Capitanio (2003) to the quantile forecasts, (3). They minimize the distance (the ℓ_2 norm) between the (empirical) regression quantiles and the (theoretical) density 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 reduces to the Student t. When, in addition, $\upsilon=\infty$, (7) reduces to a Gaussian density, 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 density, below we also consider fitting the skewed-t density 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. (2021), 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), p53).

In practice, we follow Koenker & 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 & 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 density 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; i.e., the marginal benefits to the first-order approximation decline as k increases, an issue we explore below in the simulations and in the application. Unlike ABG, this approach does not superimpose a global (parametric, e.g., skewed-t) density 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 non-parametrically 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

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

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

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

   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 density (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 density 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 density:

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

   $$= \tau_j + f(y^*_{t+h,2}|x_t)(y_{t+h} - x_t\beta_{\tau_j}), \tag{19}$$

   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

[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.

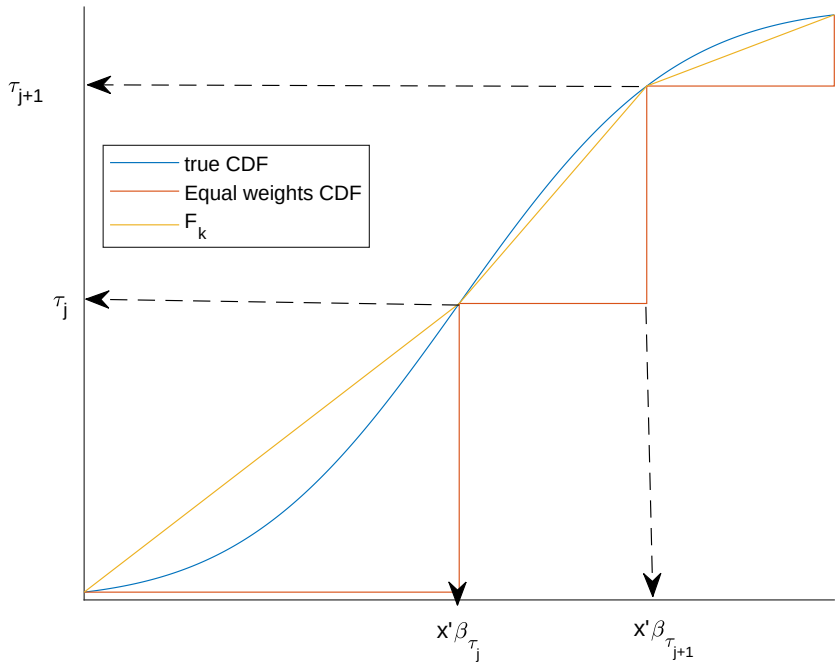


Figure 1: Illustrative comparison of the true CDF against Algorithm 1 (F_k) and the CDF assuming uniform (equal) weights between adjacent quantiles

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.

Algorithm 1, where the proposed density 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 nowcast $\widehat{Q}_{y_{T+h}}(\tau|x_t)^r$ across $\tau \in [0.05, 0.10, ..., 0.90, 0.95]$ and then constructing the full posterior density nowcast 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, like the conditional mean, can be obtained via numerical integration of:

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

where:

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

Algorithm 1, instead, relies on samples from the conditional distribution $\hat{f}(y_{t+h}|x_t)$, 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 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:

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:

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

- Obtain $\tilde{Q}_t$ by sorting Q_t based on the elements in the second column.
- 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 $\tilde{Q}_t$ 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: $\left[\boldsymbol{y}_{t+h,1}, \boldsymbol{y}_{t+h,2}, ..., \boldsymbol{y}_{t+h,k}, \boldsymbol{y}_{t+h,k+1}\right]$.

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).

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 $T=1000$ 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 $x_t=1$, i.e., 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]We note that in ABG's Matlab replication materials (available at `http://doi.org/10.3886/E113169V1`), they approximate integrals with discrete sums when matching the quantile forecasts to the skewed-t density. 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.

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

For all the Bayesian models, we estimate with 20,000 MCMC draws with a burn-in of 10,000 draws. With regards to the Bayesian QR and Algorithm 1, we save every 10th draw from the 10,000 draws. This yields 1000 draws (across k quantiles) which are then inputted by draw into Algorithm 1 where $N = 100$. This delivers a vector of 100,000 draws (1000*100) from each predictive forecast density.

Tables 1 and 2, for $T = 100$ and $T = 1000$, 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. Looking at the KL distance first, as a measure of overall density fit, we see 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. The equal-weighted Bayesian approach, EW(B), also performs very 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.

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

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

[7]To assess 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 $k = 4$ and $k = 99$; 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% tails as is the case when $k = 99$, was therefore found to deliver noisier estimates of the underlying conditional density. By contrast, due to its parametric assumption, increasing k did little to affect results for ABG.

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.

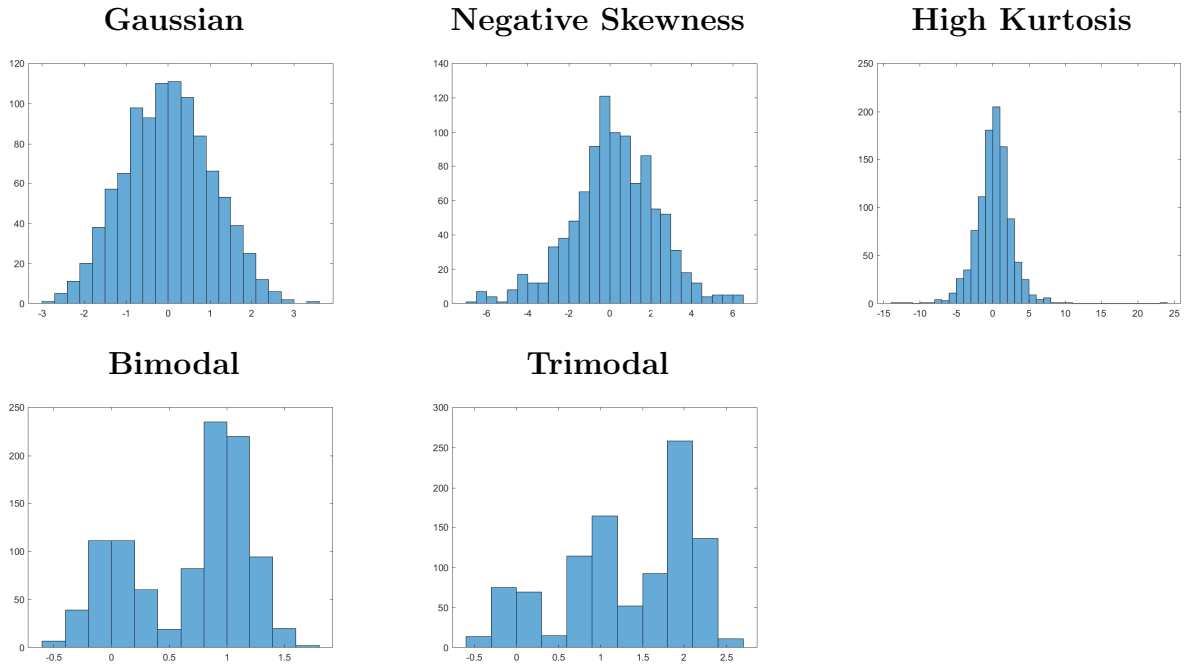


Figure 2: Simulated draws from the 5 DGP densities ($T = 1000$)

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 $T = 1000$

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 Net 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. (2020*a*) 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 but asymmetric unconditional forecast densities. These could be produced, for example, by Bayesian VAR models with stochastic volatility. Caldara et al. (2020) similarly suggest use of a parametric modeling framework that both rationalizes the empirical findings of ABG but maintains use of symmetric conditional densities. They capture nonlinear effects by 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, that 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 then lets us produce, via the aforementioned parametric and nonparametric

[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 and/or considered additional indicators; e.g., see Plagborg-Moller et al. (2020), Reichlin et al. (2020) and Brownlees & Souza (2021). 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 NFCI alone. We expect that adding in extra variables, and allowing for possible additional nonlinearities, will distinguish our approach further from ABG. Given their skewed-t assumption, ABG's densities cannot accommodate the

approaches, one-quarter-ahead and one-year-ahead forecast densities for GDP growth conditional firstly on both lagged GDP growth and NFCI, and secondly conditonal 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 consideration of 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 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. (2020*b*) 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. (2020*b*) 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 the 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 & Wallis (2011)), these PITS, at the minimum, should be uniformly distributed. As shown by Diebold et al. (1998), correctly calibrated forecast densities minimize specific loss functions. To supplement the PITS-based tests of calibration, and facilitate cross-model comparison, we also report logarithmic predictive scores and Cumulative Ranked Probability Scores (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 & Raftery (2007).

likely multimodalities associated with nonlinearity.

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 PITS 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% significance level, according to the PITS test of Rossi & 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 more 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 is explaining this result rather than use of Algorithm 1, we did produce 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 be perform similarly to those when $k = 4$. Thus we conclude that it is 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 when we estimate the QRs by Bayesian methods, whether with an uninformative or more informative prior. This tells us more about the relative merits of Bayesian versus frequentist QR. 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 worse than at $k = 19$ (as shown in Figures 3 and 4), with evidence from the cumulated PITS that the density forecast becomes too wide when $k = 99$.

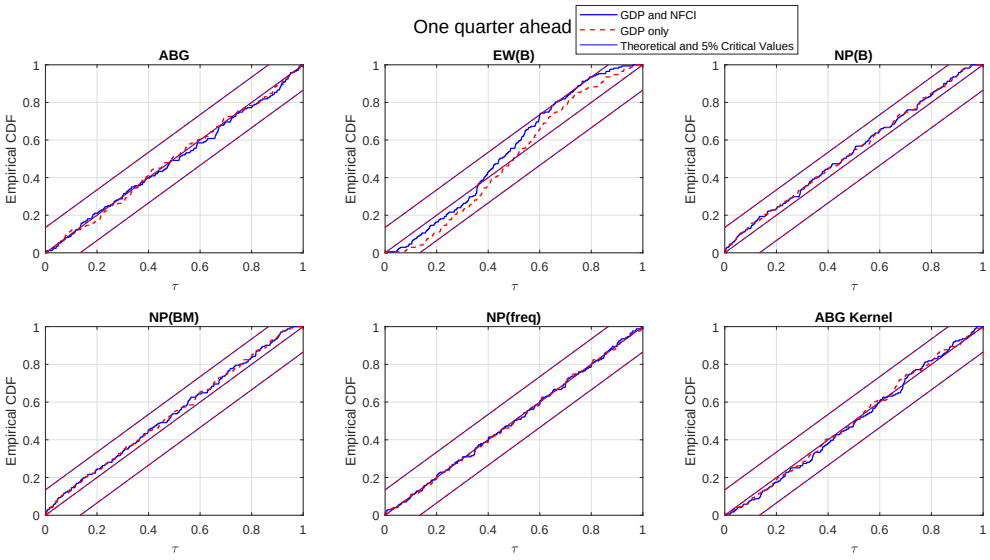


Figure 3: CDF of the in-sample PITS (one-quarter-ahead forecasts, 1973Q1-2015Q3) from the 6 density forecasts with and without NFCI. Note: the figures show the empirical CDF of the PITS (blue line) from the QR models with NFCI (and lagged GDP), the empirical CDF of the PITS (dashed red line) from the 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

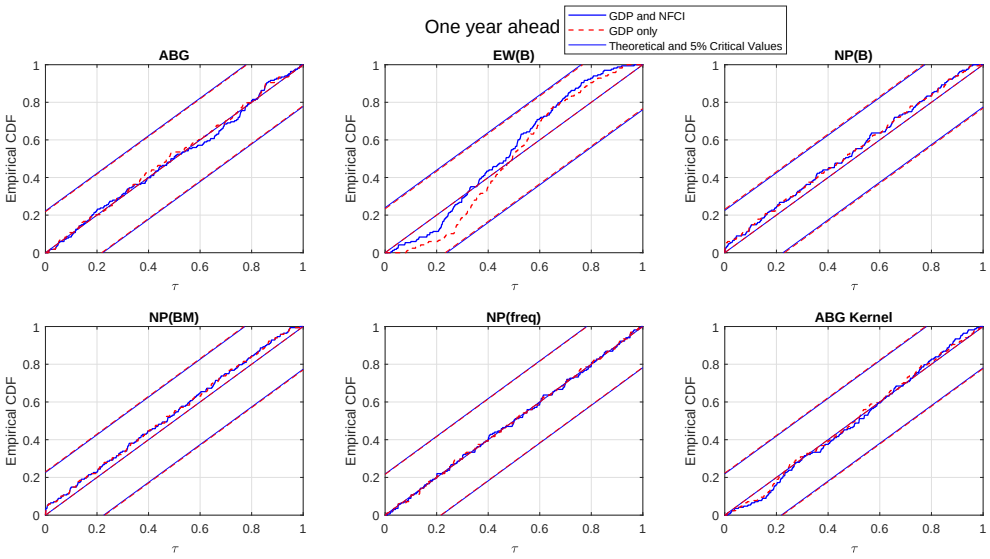


Figure 4: CDF of the in-sample PITS (one-year-ahead forecasts, 1974Q1-2015Q4) from the 6 density forecasts with and without NFCI. Note: the figures show the empirical CDF of the PITS (blue line) from the QR models with NFCI and lagged GDP, the empirical CDF of the PITS (dashed red line) from the QR models 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

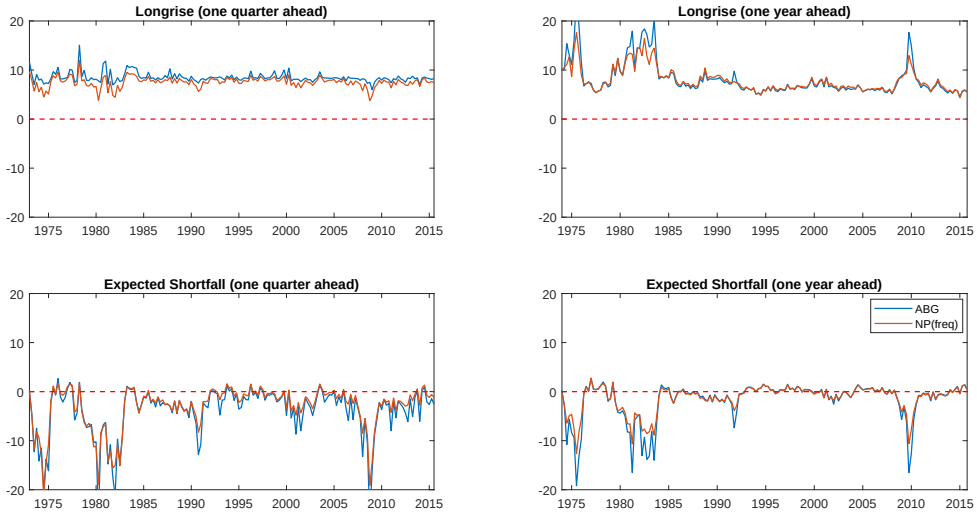


Figure 5: In-sample plots of the expected shortfall and expected longrise at $\tau = 0.05$ using ABG and NP(freq), from QRs with NFCI and lagged GDP

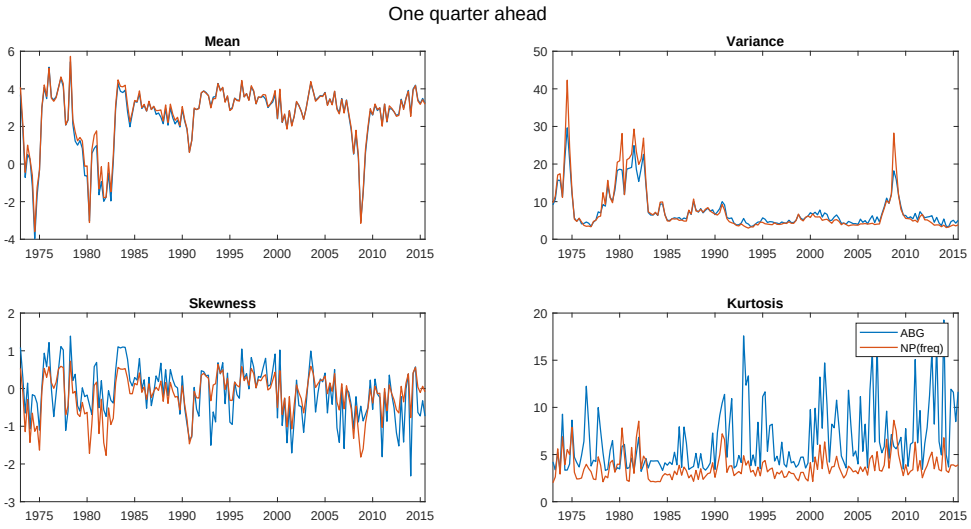


Figure 6: In-sample plots of the four moments of the ABG and NP(freq) forecast densities (one-quarter-ahead), from QRs with NFCI and lagged GDP

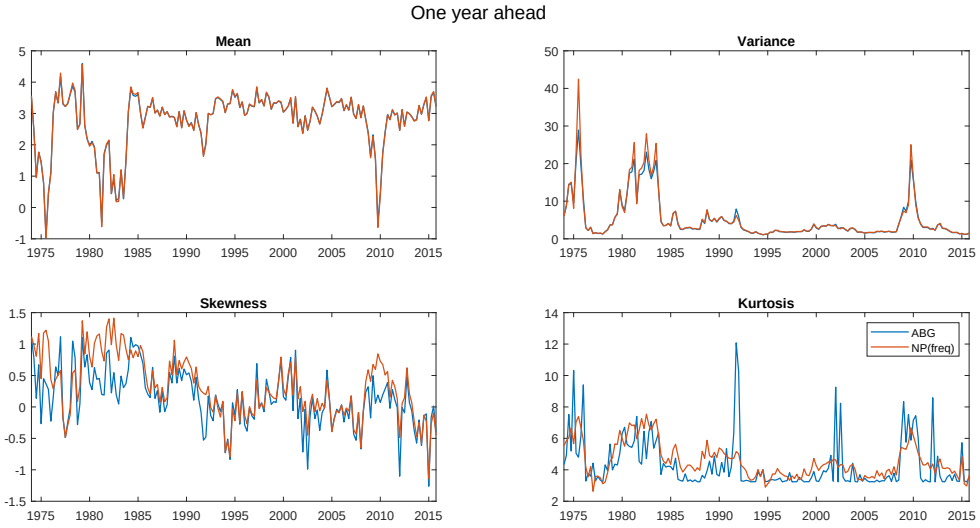


Figure 7: In-sample plots of the four moments of the ABG and NP(freq) forecast densities (one-year-ahead), from QRs with NFCI and lagged GDP

Figure 5 confirms that use of 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] 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:

$$SF_{t+h} = \frac{1}{\pi}\int_{0}^{\pi} \hat{F}_{y_{t+h}|x_t}(\tau|x_t)d\tau; \tag{28}$$

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

It is seen from Figure 5 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)

[12]This "stylized fact" has been confirmed using alternative modeling approaches to QR, such as the parametric time-varying skew-t model of Delle-Monache et al. (2021).

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. (2020*a*) 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 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. 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 Hartigan & Hartigan (1985) test is used, rejections of unimodality are far greater when we do condition on NFCI. These rejections are also especially prounounced during NBER recessionary periods, again confirming the finding of Adrian et al. (2021).

[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.

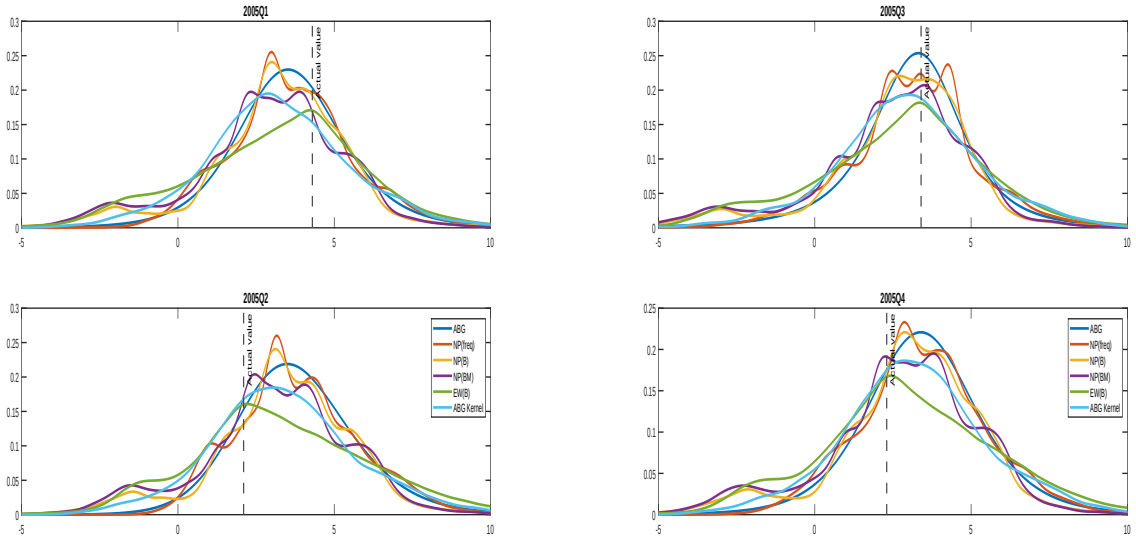


Figure 8: GDP growth density forecasts conditional on NFCI and lagged GDP for 2005 made one-quarter-ahead (in-sample)

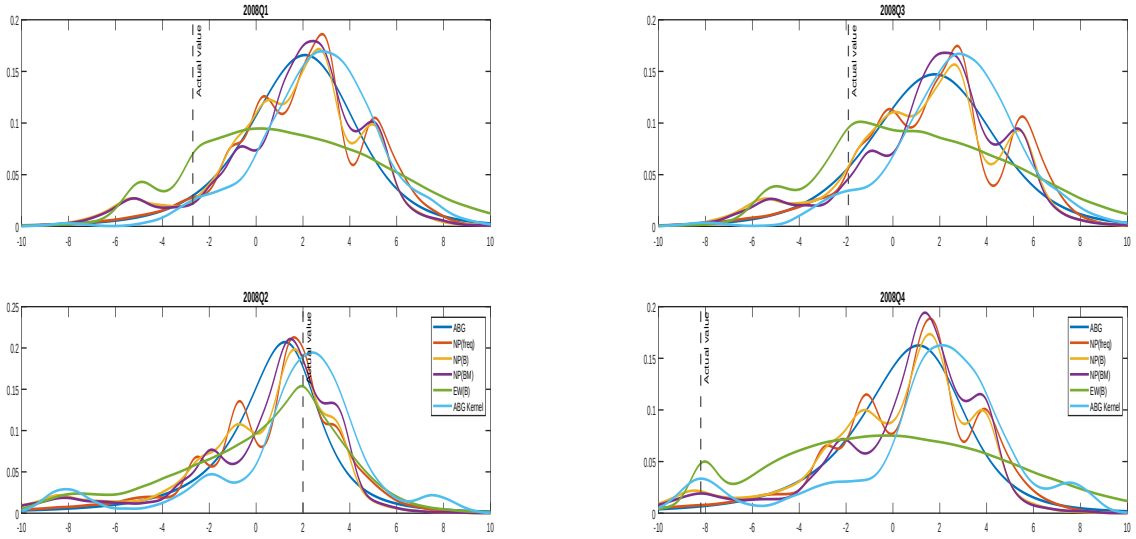
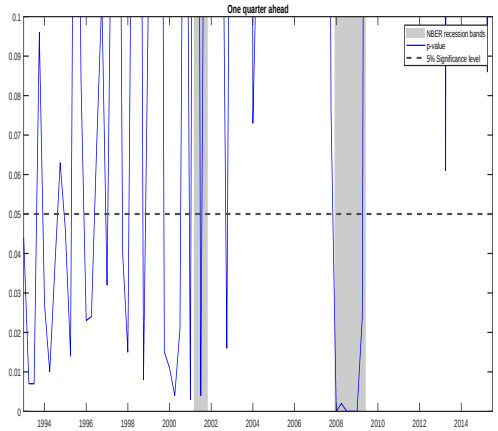
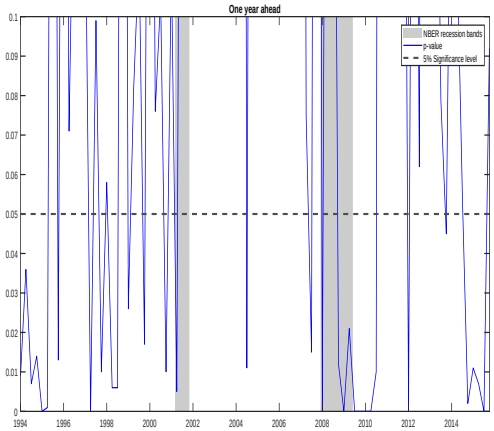


Figure 9: GDP growth density forecasts conditional on NFCI and lagged GDP for 2008 made one-quarter-ahead (in-sample)

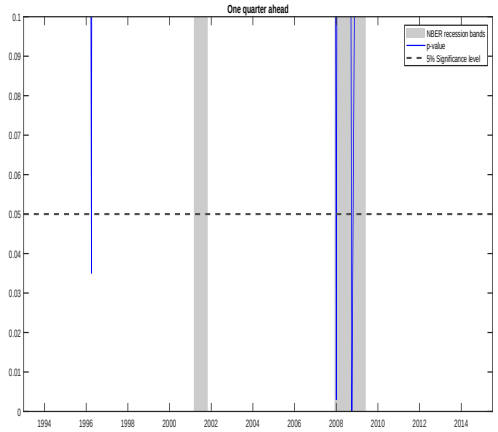
Panel A: NFCI and GDP



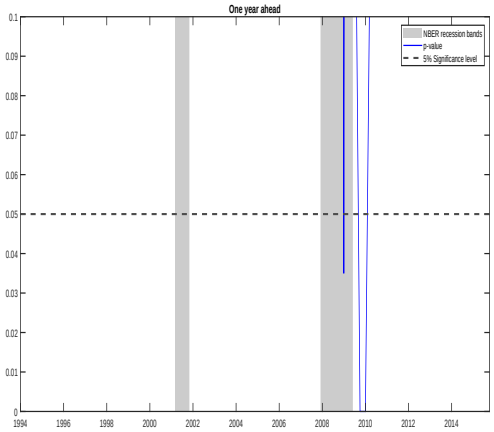
Panel B: NFCI and GDP



Panel C: GDP only



Panel D: GDP only



Notes: Panel A are 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 are 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 are 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 are 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 only lagged GDP.

Figure 10: P-values from the Hartigans' unimodality test over time, alongside NBER recessionary periods (shaded grey)

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. Figures 11 and 12 show that the accuracy of the forecast densities deteriorates out-of-sample. The null hypothesis of correct calibration is frequently rejected at a 95% significance level one-quarter-ahead, but not one-year-ahead. Comparison with the in-sample densities over this shorter evaluation period, from 1993/1994, indicates that this deterioration is in fact shared by the in-sample density forecasts over this shorter period.[16] Interestingly, the

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

PITS are closer to the 45 degree line when not conditioning on financial conditions. Table 3 shows that it is the Bayesian QR methods, with Algorithm 1, that out-of-sample tend to deliver the highest logarithmic predictive scores and the lowest CRPS. Importantly, in terms of this paper's focus on isolating the best means of constructing density forecasts from the same quantile forecasts, NP(freq) at least matches the accuracy of ABG, according to CRPS, at both forecast horizons. The average logarithmic score statistics are dominated by the forecasting failures at the time of the global financial crisis. So we prefer to emphasize the CRPS. Conditioning the GDP density forecasts on NFCI also now leads to improvements, especially one-quarter-ahead.

Despite the fact that the ABG densities are often beaten, this is not the takeaway. Instead, the bottom-line is that these alternative ways of constructing the predictive density from QRs match, and at times (albeit perhaps modestly) improve upon, the accuracy of the ABG densities. 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 feature of GDP density forecasts that should be emphasized when conditioning on financial conditions.

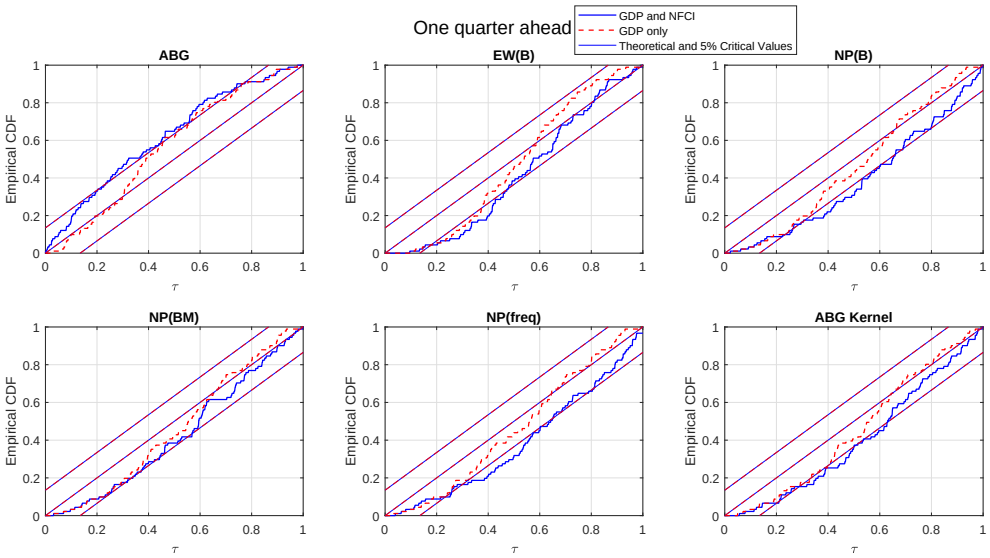


Figure 11: CDF of the out-of-sample PITS (one-quarter-ahead, 1993Q1-2015Q3) from the 6 density forecasts with NFCI and lagged GDP. Note: the figures show the empirical CDF of the PITS (red line), the CDF of the PITS under the null hypothesis of correct calibration (the 45 degree line) and the 5% critical value bands of the Rossi and Sekhposyan (2019) PITS test

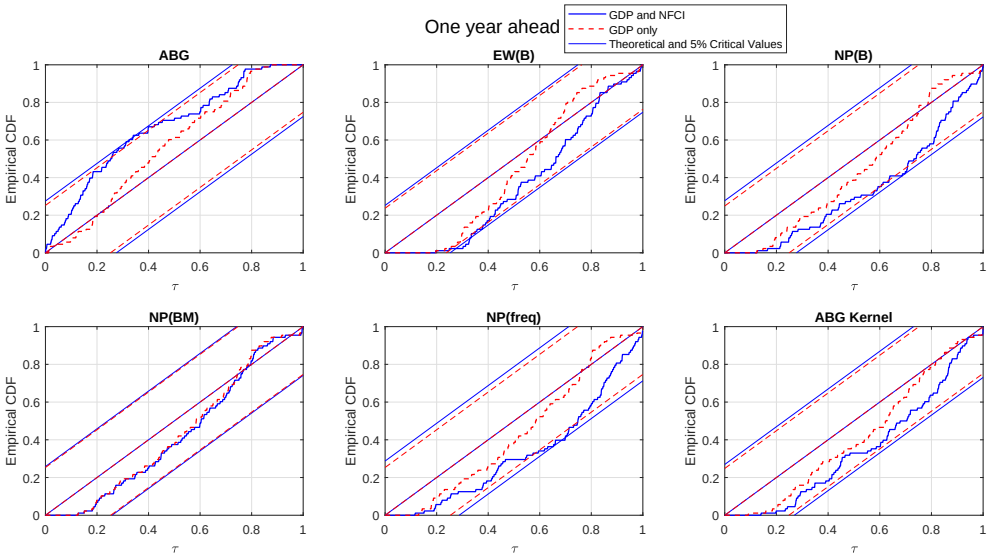


Figure 12: CDF of the out-of-sample PITS (one-year-ahead, 1994q4-2015Q4) from the 6 density forecasts with NFCI and lagged GDP. Note: the figures show the empirical CDF of the PITS (red line), the CDF of the PITS under the null hypothesis of correct calibration (the 45 degree line) and the 5% critical value bands of the Rossi and Sekhposyan (2019) PITS test

Table 3: Average Log Predictive Score (LPS) and Cumulative Ranked Probability Score (CRPS) for the one-quarter-ahead forecasts (out-of-sample, 1993Q1-2015Q3) and the one-year-ahead forecasts (out-of-sample, 1994q4-2015Q4)

	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, as one means of accommodating nonlinear relations, the specific issue of how to construct density forecasts from quantile regressions has received less attention. In the macroeconomic

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 construct predictive densities from quantile regressions, estimated either by frequentist or Bayesian methods, and compare these with the parametric approach. We suggest a simple simulation-based algorithm. Unlike the parametric approach of ABG, we find 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 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.

References

Adrian, T., Boyarchenko, N. & Giannone, D. (2019), 'Vulnerable growth', *American Economic Review* **109**(4), 1263–1289.

Adrian, T., Boyarchenko, N. & Giannone, D. (2021), 'Multimodality In Macrofinancial Dynamics', *International Economic Review* **62**(2), 861–886.

Azzalini, A. & Capitanio, A. (2003), 'Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution.', *Journal of the Royal Statistical Society: Series B* **65**, 367–389.

Brownlees, C. & Souza, A. B. (2021), 'Backtesting global growth-at-risk', *Journal of Monetary Economics* **118**, 312–330.

Caldara, D., Cascaldi-Garcia, D., Cuba-Borda, P. & Loria, F. (2020), Understanding Growth-at-Risk: A Markov Switching Approach. mimeo, Federal Reserve Board.

Carriero, A., Clark, T. E. & Marcellino, M. (2020*a*), Capturing Macroeconomic Tail Risks with Bayesian Vector Autoregressions, Working Papers 202002R, Federal Reserve Bank of Cleveland.

Carriero, A., Clark, T. E. & Marcellino, M. (2020*b*), Nowcasting Tail Risks to Economic Activity with Many Indicators, Working Papers 202013R3, Federal Reserve Bank of Cleveland.

Chen, L., Dolado, J. J. & Gonzalo, J. (2021), 'Quantile Factor Models', *Econometrica* **89**(2), 875–910.

Chernozhukov, V. (2005), 'Extremal quantile regression', *The Annals of Statistics* **33**(2), 806–839.

Chernozhukov, V., Fernández-Val, I. & Galichon, A. (2010), 'Quantile and probability curves without crossing', *Econometrica* **78**(3), 1093–1125.

Delle-Monache, D., De-Polis, A. & Petrella, I. (2021), Modelling and Forecasting Macroeconomic Downside Risk, EMF Research Papers 34, Economic Modelling and Forecasting Group, Warwick Business School.

Diebold, F. X., Gunther, A. & Tay, K. (1998), 'Evaluating density forecasts with application to financial risk management', *International Economic Review* **39**, 863–883.

Fernandes, M., Guerre, E. & Horta, E. (2021), 'Smoothing quantile regressions', *Journal of Business and Economic Statistics* **39**(1), 338–357.

Ferrara, L., Mogliani, M. & Sahuc, J.-G. (2021), 'High-frequency monitoring of growth at risk', *International Journal of Forecasting* . Forthcoming.

Figueres, J. M. & Jarocinski, M. (2020), 'Vulnerable growth in the euro area: Measuring the financial conditions', *Economics Letters* **191**.

Gaglianone, W. P. & Lima, L. R. (2012), 'Constructing Density Forecasts from Quantile Regressions', *Journal of Money, Credit and Banking* **44**(8), 1589–1607.

Gaglianone, W. P. & Lima, L. R. (2014), 'Constructing Optimal Density Forecasts From Point Forecast Combinations', *Journal of Applied Econometrics* **29**(5), 736–757.

Ghysels, E., Iania, L. & Striaukas, J. (2018), Quantile-based Inflation Risk Models, Working Paper Research 349, National Bank of Belgium.

Giglio, S., Kelly, B. & Pruitt, S. (2016), 'Systemic risk and the macroeconomy: An empirical evaluation', *Journal of Financial Economics* **119**(3), 457 – 471.

Gneiting, T. & Raftery, A. E. (2007), 'Strictly proper scoring rules, prediction, and estimation', *Journal of the American Statistical Association* **102**, 359–378.

Hartigan, J. A. & Hartigan, P. M. (1985), 'The Dip Test of Unimodality', *The Annals of Statistics* **13**(1), 70 – 84.

Koenker, R. (2005), *Quantile Regression*, Cambridge University Press.

Koenker, R. & Machado, J. A. (1999), 'GMM inference when the number of moment conditions is large', *Journal of Econometrics* **93**(2), 327–344.

Koenker, R. & Zhao, Q. (1996), 'Conditional quantile estimation and inference for ARCH models', *Econometric Theory* **12**(5), 793–813.

Komunjer, I. (2013), *Quantile Prediction*, Vol. 2 of *Handbook of Economic Forecasting*, Elsevier, pp. 961–994.

Korobilis, D. (2017), 'Quantile regression forecasts of inflation under model uncertainty', *International Journal of Forecasting* **33**, 11–20.

Korobilis, D., Landau, B., Musso, A. & Phella, A. (2021), The time-varying evolution of inflation risks, Working Paper Series 2600, European Central Bank.

Kozumi, H. & Kobayashi, G. (2011), 'Gibbs sampling methods for Bayesian quantile regression', *Journal of Statistical Computation and Simulation* **81**(11), 1565–1578.

Kruger, F., Lerch, S., Thorarinsdottir, T. & Gneiting, T. (2021), 'Predictive Inference Based on Markov Chain Monte Carlo Output', *International Statistical Review* **89**(2), 274–301.

Manzan, S. (2015), 'Forecasting the distribution of economic variables in a data-rich environment', *Journal of Business & Economic Statistics* **33**(1), 144–164.

Manzan, S. & Zerom, D. (2013), 'Are macroeconomic variables useful for forecasting the distribution of U.S. inflation?', *International Journal of Forecasting* **29**(3), 469–478.

Mitchell, J., Poon, A. & Mazzi, G.-L. (2021), 'Nowcasting Euro Area GDP growth using Bayesian quantile regression', *Advances in Econometrics: Essays in Honor of M. Hashem Pesaran* **Forthcoming**.

Mitchell, J. & Wallis, K. F. (2011), 'Evaluating density forecasts: Forecast combinations, model mixtures, calibration and sharpness', *Journal of Applied Econometrics* **26**(6), 1023–1040.

Parzen, E. (1979), 'Nonparametric statistical data modeling', *Journal of the American Statistical Association* **74**(365), 105–121.

Plagborg-Moller, M., Reichlin, L., Ricco, G. & Hasenzagl, T. (2020), 'When is growth at risk?', *Brookings Papers on Economic Activity* pp. 167–213.

Prasad, A., Elekdag, S., Jeasakul, P., Lafarguette, R., Alter, A., , Xiaochen Feng, A. & Wang, C. (2019), Growth at risk: Concept and application in IMF country surveillance, IMF Working Papers 2019/036, International Monetary Fund.

Reichlin, L., Ricco, G. & Hasenzagl, T. (2020), Financial variables as predictors of real growth vulnerability, Discussion Papers 05/2020, Deutsche Bundesbank.

Rossi, B. (2021), 'Forecasting in the presence of instabilities: How do we know whether economic models work and how to improve them', *Journal of Economic Literature* **Forthcoming**.

Rossi, B. & Sekhposyan, T. (2019), 'Alternative tests for correct specification of conditional predictive densities', *Journal of Econometrics* **208**(2), 638–657.

Yu, K. & Moyeed, R. A. (2001), 'Bayesian quantile regression', *Statistics & Probability Letters* **54**(4), 437 – 447.

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** $T = 1000$					
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	0.12	0.07	0.47	0.05
DGP2: Unimodal (Negative Skewness)	0.04	2.76	0.08	3.46	0.04
DGP3: Unimodal (High Kurtosis)	0.01	0.33	1.15	82.94	0.03
DGP4: Bimodal	0.00	0.00	0.01	0.02	0.07
DGP5: Trimodal	0.00	0.01	0.01	0.03	0.06
$k = 99$ **and** $T = 1000$					
DGP1: Unimodal (Gaussian)	0.00	0.11	0.02	0.34	0.02
DGP2: Unimodal (Negative Skewness)	0.00	3.02	0.01	1.70	0.02
DGP3: Unimodal (High Kurtosis)	0.00	0.37	0.86	85.31	0.00
DGP4: Bimodal	0.00	0.00	0.01	0.00	0.02
DGP5: Trimodal	0.00	0.01	0.00	0.00	0.04

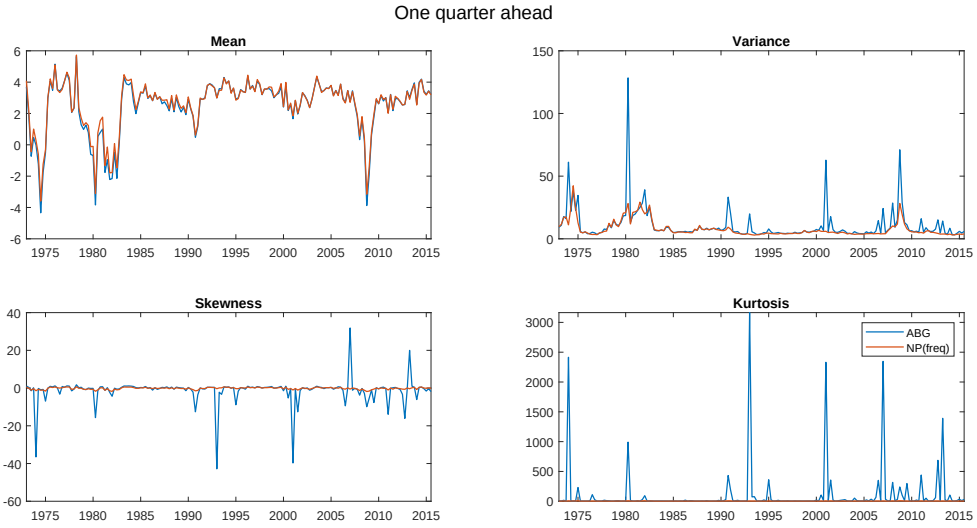


Figure 13: In-sample plots of the four moments of the ABG and NP(freq) forecast densities (one-quarter-ahead), when ABG's skewed-t density is simulated not truncated

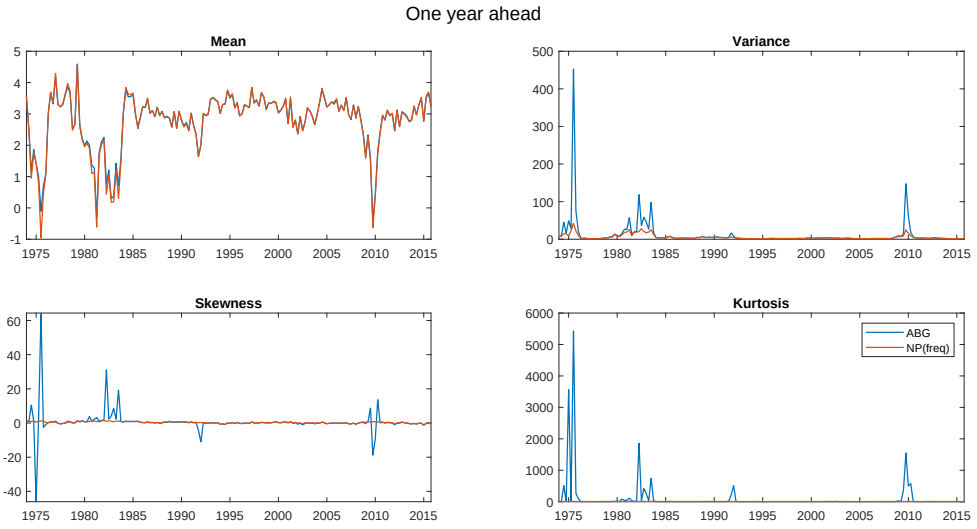


Figure 14: In-sample plots of the four moments of the ABG and NP(freq) forecast densities (one-year-ahead), when ABG's skewed-t density is simulated not truncated

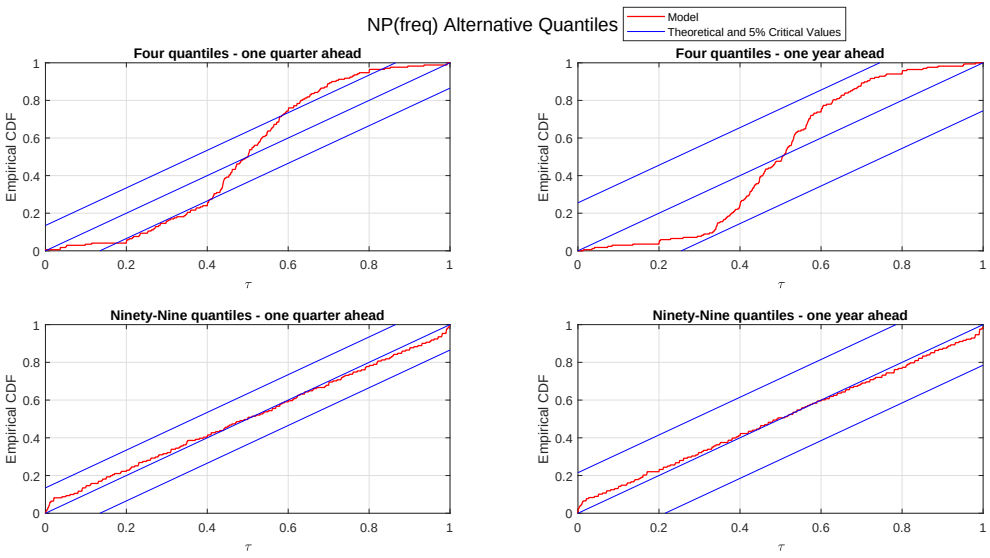


Figure 15: CDF of the in-sample PITS for NP(freq) when $k = 4$ and $k = 99$. Note: the figures show the empirical CDF of the PITS (red line), the CDF of the PITS under the null hypothesis of correct calibration (the 45 degree line) and the 5% critical value bands of the Rossi and Sekhposyan (2019) PITS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

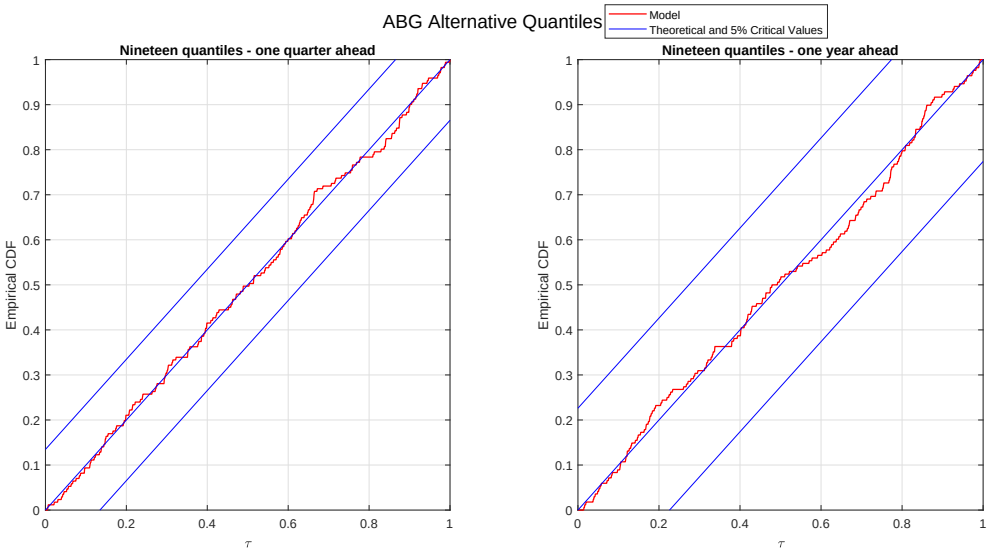


Figure 16: CDF of the in-sample PITS for ABG when $k = 19$. Note: the figures show the empirical CDF of the PITS (red line), the CDF of the PITS under the null hypothesis of correct calibration (the 45 degree line) and the 5% critical value bands of the Rossi and Sekhposyan (2019) PITS test

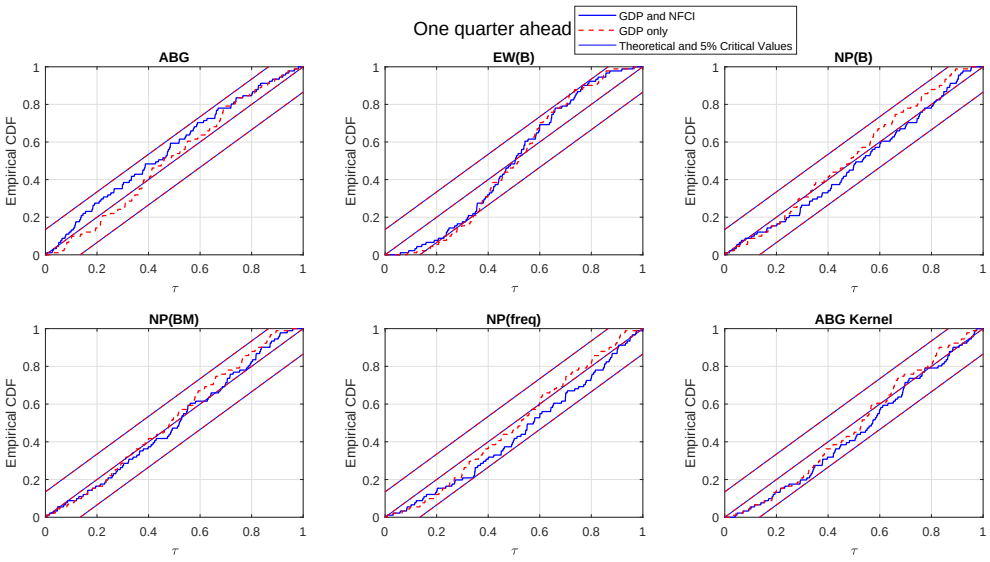


Figure 17: CDF of the in-sample PITS (one-quarter-ahead forecasts, 1993Q1-2015Q3) from the 6 density forecasts with and without NFCI. Note: the figures show the empirical CDF of the PITS (blue line) from the QR models with NFCI (and lagged GDP), the empirical CDF of the PITS (dashed red line) from the 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

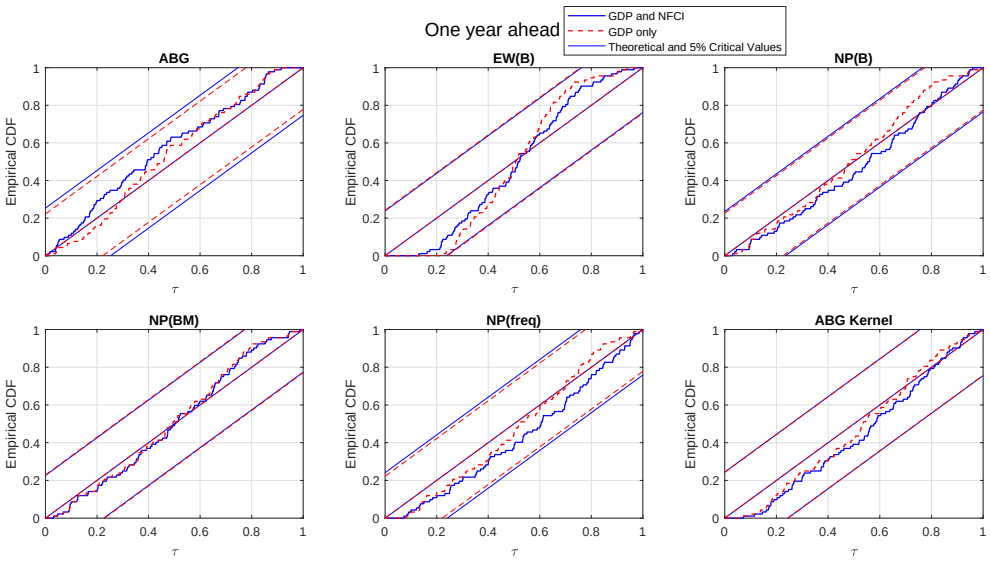


Figure 18: CDF of the in-sample PITS (one-year-ahead forecasts, 1994Q1-2015Q4) from the 6 density forecasts with and without NFCI. Note: the figures show the empirical CDF of the PITS (blue line) from the QR models with NFCI (and lagged GDP), the empirical CDF of the PITS (dashed red line) from the QR 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

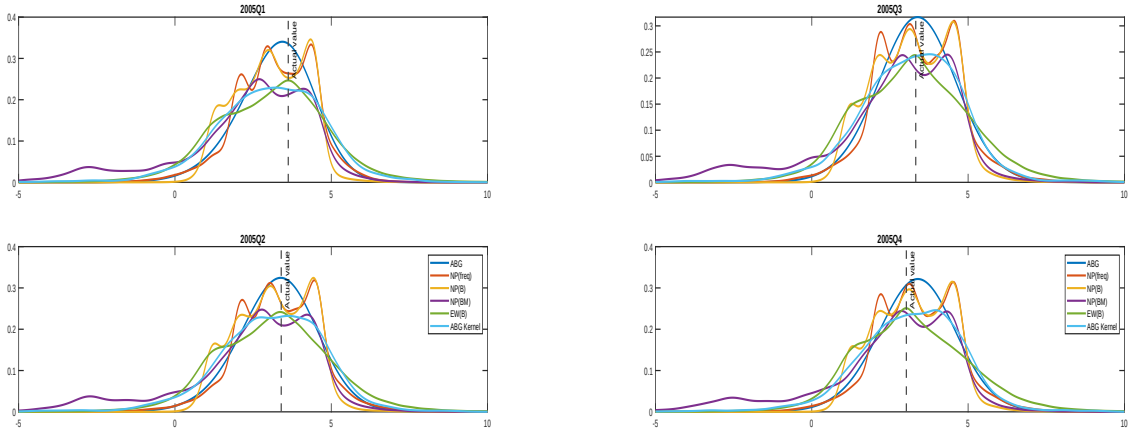


Figure 19: GDP growth density forecasts conditional on NFCI and lagged GDP for 2005 made one-year-ahead (in-sample)

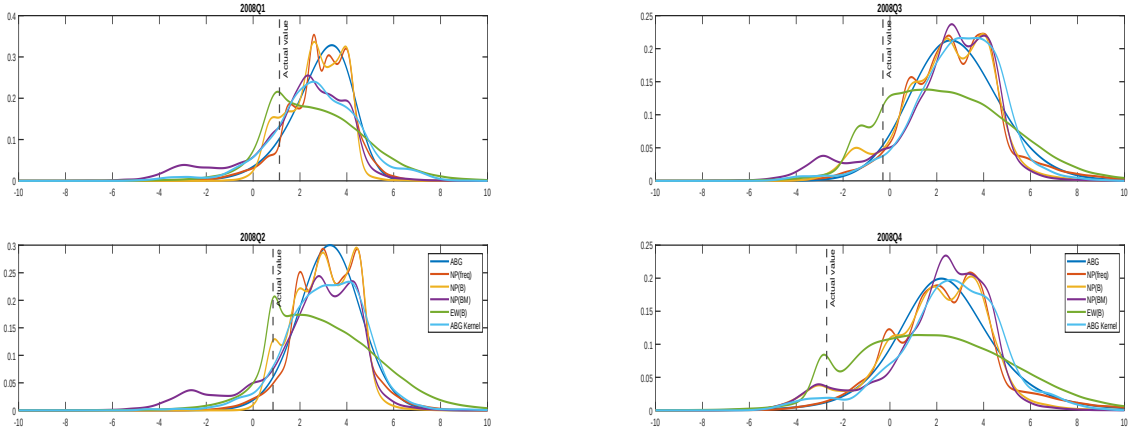


Figure 20: GDP growth density forecasts conditional on NFCI and lagged GDP for 2008 made one-year-ahead (in-sample)

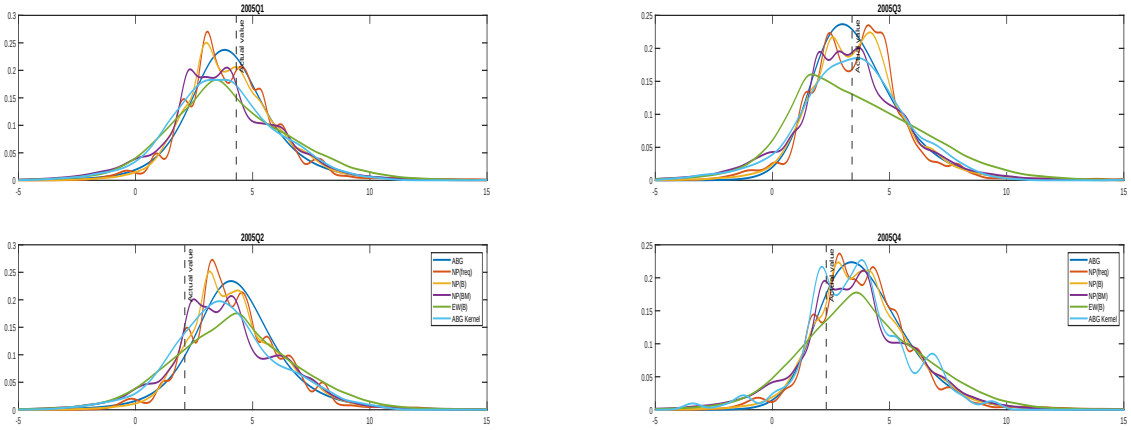


Figure 21: GDP growth density forecasts conditional on NFCI and lagged GDP for 2005 made one-quarter-ahead (out-of-sample)

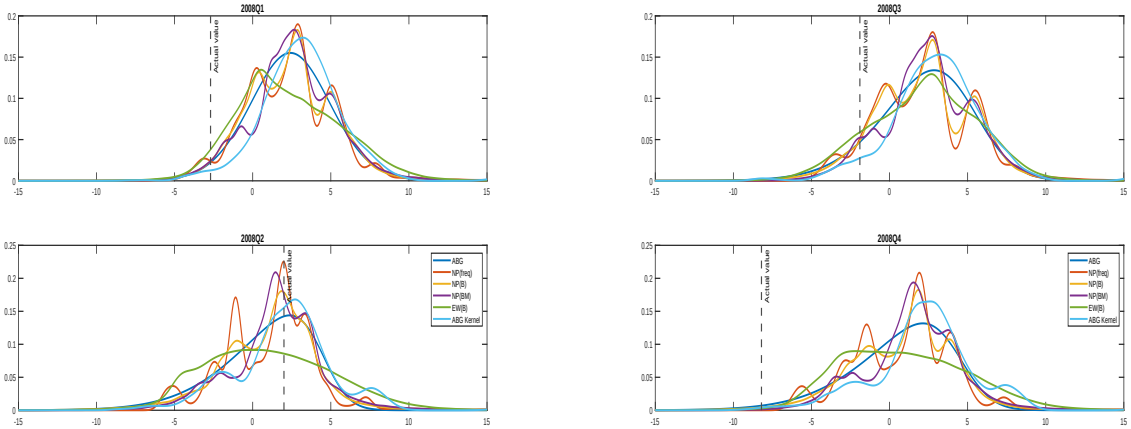


Figure 22: GDP growth density forecasts conditional on NFCI and lagged GDP for 2008 made one-quarter-ahead (out-of-sample)

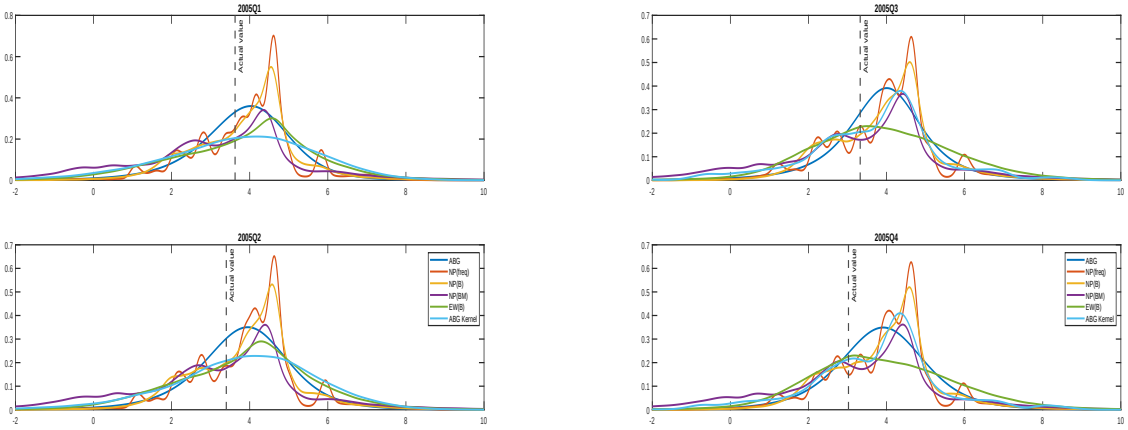


Figure 23: GDP growth density forecasts conditional on NFCI and lagged GDP for 2005 made one-year-ahead (out-of-sample)

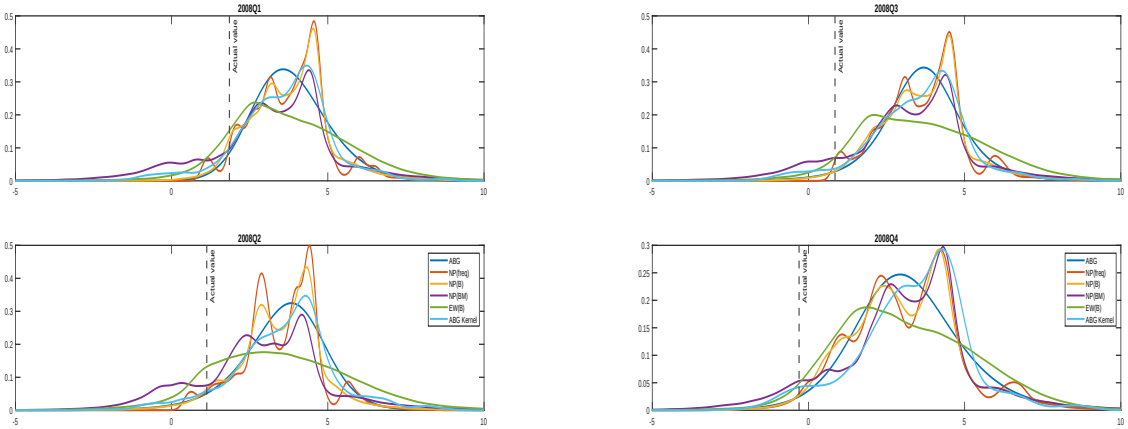


Figure 24: GDP growth density forecasts conditional on NFCI and lagged GDP for 2008 made one-year-ahead (out-of-sample)