Staff Working Paper/Document de travail du personnel—2025-4

Last updated: January 17, 2025

Quantile VARs and Macroeconomic Risk Forecasting

by Stéphane Surprenant

Canadian Economic Analysis Department
Bank of Canada
SSurprenant@bankofcanada.ca

Bank of Canada staff working papers provide a forum for staff to publish work-in-progress research independently from the Bank's Governing Council. This research may support or challenge prevailing policy orthodoxy. Therefore, the views expressed in this paper are solely those of the authors and may differ from official Bank of Canada views. No responsibility for them should be attributed to the Bank.

DOI: https://doi.org/10.34989/swp-2025-4 | **ISSN** 1701-9397 ©2025 Bank of Canada

Acknowledgements

The views expressed in this paper are solely those of the author and may differ from official Bank of Canada views. No responsibility for them should be attributed to the Bank.

Abstract

Recent rises in macroeconomic volatility have prompted the introduction of quantile vector autoregression (QVAR) models to forecast macroeconomic risk. This paper provides an extensive evaluation of the predictive performance of QVAR models in a pseudo-out-of-sample experiment spanning 112 monthly US variables over 40 years, with horizons of 1 to 12 months. We compare QVAR with three parametric benchmarks: a Gaussian VAR, a generalized autoregressive conditional heteroskedasticity VAR and a VAR with stochastic volatility. QVAR frequently, significantly and quantitatively improves upon the benchmarks and almost never performs significantly worse. Forecasting improvements are concentrated in the labour market and interest and exchange rates. Augmenting the QVAR model with factors estimated by principal components or quantile factors significantly enhances macroeconomic risk forecasting in some cases, mostly in the labour market. Generally, QVAR and the augmented models perform equally well. We conclude that both are adequate tools for modeling macroeconomic risks.

Topics: Econometrics and statistical methods; Business fluctuations and cycles
JEL codes: C53, E37, C55

Résumé

Les récentes hausses de la volatilité macroéconomique ont suscité l'adoption de modèles vectoriels autorégressifs quantiles (QVAR) pour prévoir les risques macroéconomiques. Cette étude présente une évaluation approfondie de la qualité des prévisions découlant d'un modèle QVAR dans une expérience en pseudo hors échantillon couvrant 112 variables mensuelles américaines recueillies sur 40 ans, avec des horizons de 1 à 12 mois. Nous comparons ce modèle avec trois modèles paramétriques de référence : un modèle vectoriel autorégressif (VAR) gaussien, un modèle VAR à hétéroscédasticité conditionnelle autorégressive généralisée et un modèle VAR à volatilité stochastique. Nous constatons que les résultats du modèle QVAR sont fréquemment, significativement et quantitativement supérieurs à ceux des modèles de référence et ne leur sont presque jamais significativement inférieurs. L'amélioration se constate particulièrement dans les prévisions relatives au marché du travail, aux taux d'intérêt et aux taux de change. L'ajout de facteurs estimés par des composantes principales ou des facteurs quantiles au modèle QVAR améliore considérablement la prévision des risques macroéconomiques dans certains cas, principalement en ce qui concerne le marché du travail. En général, le modèle QVAR et ses versions enrichies fonctionnent tout aussi bien, et nous concluons qu'ils constituent tous des outils adéquats pour modéliser les risques macroéconomiques.

Sujets : Méthodes économétriques et statistiques; Cycles et fluctuations économiques
Codes JEL : C53, E37, C55

1. Introduction

The rise in macroeconomic volatility experienced during the 2007 financial crisis and the COVID-19 pandemic ended the Great Moderation and increased interest in modeling macroeconomic risk. Work by Giglio, Kelly, and Pruitt (2016) and Adrian, Boyarchenko, and Giannone (2019) popularized the use of quantile regressions in this context, finding evidence that financial stress leads to asymmetry in output growth. Many studies applied those methods in a single equation framework, focusing on the predictive power of financial indicators for risk to output growth (e.g., Figueres and Jarociński (2020), Adams et al. (2021) and Iseringhausen (2024)) and inflation (e.g., Manzan and Zerom (2013), Manzan (2015) and López-Salido and Loria (2020)). Others have proposed using quantile regressions as part of a structural analysis studying the effects of shocks on the conditional distribution of output growth (Loria, Matthes, and Zhang (2024)) or to distinguish between shocks to upside, downside and total uncertainty (Forni, Gambetti, and Sala (2024)). Against this background, several researchers (White, Kim, and Manganelli (2015), Chavleishvili and Manganelli (2021), Chavleishvili et al. (2021) and Ruzicka (2021)) have recently proposed a quantile VAR (QVAR) model for forecasting, scenario analysis, macroprudential risk management and quantile impulse responses. However, the forecasting performance of the QVAR model has yet to be assessed.

The use of linear quantile regression models is primarily motivated by their robustness as approximations to conditional quantiles and distributions. Economic theory can justify a wide variety of VAR processes for modeling conditional distributions,[1] but all of them require committing to a particular functional form. Linear quantile regressions provide a weighted least square optimal linear approximation to the true conditional quantiles (Angrist, Chernozhukov, and Fernández-Val (2006)). As a result they have been employed to produce forecasts or insights regarding macroeconomic risks in ways that are hopefully robust to the unknown form of the underlying data-generating process.

The first contribution of this paper is to provide an extensive evaluation of the predictive performance of the QVAR model. Other papers have explored a similar comparison in a single equation setting between quantile regression models and AR-

[1] Occasionally binding collateral constraints (Aiyagari and Gertler (1999)) or a kinked Phillips curve (Benigno and Eggertsson (2023)) suggests using a threshold VAR. The model in Acemoglu and Scott (1997) implies a smooth transition process for output where the transition function emerges from firm heterogeneity as only some firms opt to invest at a given point in time. Real options arguments (Bernanke (1983) and McDonald and Siegel (1986)) and frictions to the supply of credit (e.g., Adrian and Boyarchenko (2012) and Brunnermeier and Sannikov (2014)) can motivate the use of volatility-in-means effects (e.g., Elder and Serletis (2010)).

GARCH models (e.g., Brownlees and Souza (2021), Iseringhausen (2024) and Kipriyanov (2022)). Still others have compared quantile regression models with more sophisticated parametric VAR alternatives (e.g., Carriero, Clark, and Marcellino (2021) and Caldara et al. (2021)), but the QVAR model has yet to be compared to parametric alternatives. Throughout this paper, we target conditional densities with a focus on both tails of conditional distributions. The comparison features 112 US monthly macroeconomic variables and an out-of-sample period of over 40 years with forecasting horizons of between a month and a year. This contrasts with the typical forecasting evaluation in this literature that focuses on a just few targets. We supplement this comparative analysis with some specification tests used in the financial literature to evaluate value-at-risk models. This allows us to evaluate the 5th and 95th quantile forecasts produced by the QVAR model independently of the choice of benchmark models and to inspect the contexts in which we can find evidence of misspecification.

The forecasting experiment is built around bivariate VAR models that pair the target variable with the National Financial Conditions Index (NFCI). This is perhaps the most interesting comparison as it is the most commonly used predictor in the growth-at-risk literature following Adrian, Boyarchenko, and Giannone (2019). There is also some evidence that credit shocks are important drivers of macroeconomic fluctuations for a large number of variables (Boivin, Giannoni, and Stevanović (2020)). Financial stress is therefore relevant to many of our target variables insofar as it captures this type of shock. On this basis, we compare the QVAR model with three parametric alternatives. The first alternative is a Gaussian VAR (VAR-N), which allows us to evaluate when and how much gain there is to moving beyond iid disturbances. We also include a VAR-GARCH model as in Normandin and Phaneuf (2004), Bouakez and Normandin (2010) or Bouakez, Chihi, and Normandin (2014), and a VAR-SV similar to those used by Cogley and Sargent (2005), Primiceri (2005) and Chan and Eisenstat (2018). This offers two common and relatively simple ways to introduce parametric changes in volatility. However, unlike these authors, we do not pursue time-varying parameters in an effort to limit our deviation from the iid setting to changes in volatility. Moreover, as we explain in Section 2, all four models (QVAR, VAR-N, VAR-GARCH and VAR-SV) impose a linear functional form on conditional expectations at all future horizons.

We find that the QVAR model provides statistically significant improvements in tail-density forecasting accuracy over the VAR-N model in close to half of all variables considered. Those improvements are frequently quantitatively important, with reductions in density scores on the order of 10% to 30% in many cases. These are particularly important for labor market variables across all horizons considered and for interest and exchange rates at shorter horizons. The QVAR model also offers improvements over VAR-GARCH and VAR-SV models, albeit in fewer cases that are concentrated in those same groups of variables. More importantly, the QVAR model almost never does statistically significantly and substantially quantitatively worse than any of the parametric alternatives: it is therefore a robust way to model macroeconomic risk. Those results surprisingly turn out to not be driven by the QVAR model doing exceptionally better than the parametric alternatives during NBER recessions. Finally, specification tests reveal evidence of misspecification. In particular, realized values that fall below or above the 5th and 95th quantile forecasts, respectively, tend to be serially correlated, whereas such "tail event" should be unpredictable under a correctly specified model.

The second contribution of this paper is to extend the analysis to a data-rich environment by augmenting QVAR models with latent factors estimated from our set of 112 target variables. Applications featuring principal component estimates (PCA) (e.g., Manzan (2015) and Goulet Coulombe et al. (2022)) and the recently introduced iterative quantile regression (IQR) estimates of quantile factors (Chen, Dolado, and Gonzalo (2021)) have been considered in the past, but all of them involve direct forecasting models in a univariate setting. In contrast, factor augmented QVAR models jointly model the dynamic between observed variables and latent factor estimates.

We find that QFAVAR and QVAR models tend to perform equally well at forecasting macroeconomic risks across all variable categories. PCA and IQR factors may carry information that significantly overlaps with the NFCI. However, QFAVAR models provide statistically significant improvements in about 13% of cases, most of them in the labor market across all horizons. Specification tests reveal that introducing IQR factors into the set of variables available to QVAR models reduces the frequency of misspecification and the incidence of serially correlated "tail events" This suggests the specifications issues reported in both cases may be due to the small set of variables we considered. We conclude that QVAR and QFAVAR models are appropriate tools for modeling macroeconomic risk.

The paper is organized as follows. Section 2 introduces the QVAR model, details some of its properties and explains how to use it for forecasting. Section 3 details the forecasting experiment, the parametric alternatives and the tests used for evaluating QVAR and QFAVAR models. Section 4 presents and discusses the results. Section 5 concludes.

2. Quantile VAR Models

The QVAR model considered in this paper has been studied for scenario analysis and structural analysis by Chavleishvili and Manganelli (2021), Montes-Rojas (2021) and Ruzicka (2021). For a $K \times 1$ vector $\boldsymbol{y_t}$ of time series, the conditional quantile $\tau_k \in [0, 1]$ of the k-th variable takes the form

$$\mathbb{Q}_{y_{k,t}}\left(\tau_k | \tilde{\boldsymbol{x}}_{\boldsymbol{t}}^{(\boldsymbol{k})}\right) = \sum_{i \leq k} a_{0,k,i}(\tau_k)\, y_{i,t} + \sum_{i=1}^{K} \sum_{j=1}^{p} a_{j,k,i}(\tau_k)\, y_{i,t-j} + \epsilon_k(\tau_k) \tag{1}$$

where $\tilde{\boldsymbol{x}}_{\boldsymbol{t}}^{(\boldsymbol{k})}$ contains the regressors for this equation. It is well known in this literature that quantile regressions admit a (restricted) random coefficient representation. In this way, data can be simulated by uniformly sampling parameters over a grid of quantiles one equation at a time, one period at a time. This leads to

$$\begin{aligned} y_{k,t} &= \sum_{i \leq k} a_{0,k,i}(u_{k,t})\, y_{i,t} + \sum_{i=1}^{K} \sum_{j=1}^{p} a_{j,k,i}(u_{k,t})\, y_{i,t-j} + \epsilon_k(u_{k,t}) \\ \Leftrightarrow \boldsymbol{y_t} &= \boldsymbol{A_0}(\boldsymbol{u_t})\, \boldsymbol{y_t} + \sum_{j=1}^{p} \boldsymbol{A_j}(\boldsymbol{u_t})\, \boldsymbol{y_{t-j}} + \epsilon(\boldsymbol{u_t}) \end{aligned} \tag{2}$$

where $\boldsymbol{u_t} \sim U[0, 1]^K$ and $\boldsymbol{A_0(u_t)}$ is a lower triangular matrix with a null diagonal. Contemporary terms are included to ensure coefficients across equations do not depend on multiple $u_{k,t}$'s, but are instead independent to eliminate the need for a notion of multivariate quantiles.[2] The triangular structure simplifies estimation and is applied to all models compared in the forecasting experiment.

Before turning to estimation and forecasting, we consider a few properties of QVAR processes. The QVAR model in (2) admits the following SVAR representation

$$\boldsymbol{y_t} = \bar{\boldsymbol{A}}_{\boldsymbol{0}}\, \boldsymbol{y_t} + \sum_{j=1}^{p} \bar{\boldsymbol{A}}_{\boldsymbol{j}}\, \boldsymbol{y_{t-j}} + \bar{\boldsymbol{\epsilon}}_{\boldsymbol{t}} \tag{3}$$

where $\bar{\boldsymbol{\epsilon}}_{\boldsymbol{t}} := \left(\boldsymbol{A_0}(\boldsymbol{u_t}) - \bar{\boldsymbol{A}}_{\boldsymbol{0}}(\boldsymbol{u_t})\right) y_t + \sum_{j=1}^{p} \left(\boldsymbol{A_j}(\boldsymbol{u_t}) - \bar{\boldsymbol{A}}_{\boldsymbol{j}}\right) \boldsymbol{y_{t-j}} + \epsilon(\boldsymbol{u_t})$ and $\bar{\boldsymbol{A}}_{\boldsymbol{j}} := \mathbb{E}\left(\boldsymbol{A_j}(\boldsymbol{u_t})\right)$ under technical conditions spelled out in Proposition 1.5 of Ruzicka (2021). This es-

[2] The interested reader can also find a technical explanation in the Theorem 1 of Chavleishvili and Manganelli (2021).

tablishes that VAR and QVAR processes impose the same linear functional form for conditional expectations. Moreover, when the QVAR processes admit a VAR representation, results in Lütkepohl (2005) concerning linear transformations of the form $\boldsymbol{F}\boldsymbol{y_t}$ apply. In particular, if a large set of variables follow a QVAR(p) process (and thus a VAR(p) process), then a subset of it will generally follow a VARMA($\tilde{p}$,$\tilde{q}$) process (with possibly some heteroskedasticity or other higher order dependence). We should therefore expect that QVAR and VAR models offer similar mean forecasts under fairly general conditions.

Equation (3) also shows that QVAR processes capture such things as changes in conditional heteroskedasticity by allowing slope parameters to vary across quantiles. If the coefficient matrices were constant across quantiles (i.e., $\boldsymbol{A_j}(\boldsymbol{u_t}) = \bar{\boldsymbol{A_j}}$ for $j = 0, \ldots, p$), then we would have a linear model with iid shocks. This observation is also the reason why model (2) implies that the support of $\boldsymbol{y_t}$ must generally be bounded. Otherwise, quantile crossing would occur even in large samples. It is best seen in the simpler QAR(1) case (i.e., $y_t = a_1(u_t)\, y_{t-1} + \epsilon(u_t)$) where variation in the slope parameters means the conditional quantiles of y_t must cross somewhere along the real line and bounding the process makes visiting this region a zero probability event.[3] When this condition is violated, the approximation the QVAR model provides to the conditional distribution of $\boldsymbol{y_t}$ in part of its domain may be poor.[4] However, as we explain below, the quantile regression estimator we use enjoys an optimality property that should limit this process to a small region. How much each of these points matter is an empirical question. Finally, the same univariate QAR(1) process is useful to intuitively understand the technical condition under which a QVAR process is ergodic, as well as both weakly and strongly stationary. In this simple case, the condition is $\mathbb{E}\ \left(a_1(u_t)^2\right) < 1$, which allows for unit and explosive roots for some subsets of conditional quantiles.

2.1. Estimation and Forecasting

The parameters of the QVAR process (1) can be estimated by linear quantile regression (Koenker and Bassett (1978)) one equation at a time for a grid of quantiles. Let $\beta^{(\boldsymbol{k})}(\tau_k)$ be all parameters for regression k, including the intercept $\epsilon_k(\tau_k)$, and $\boldsymbol{x_t^{(k)}} = \left(1, \tilde{\boldsymbol{x}}_{\boldsymbol{t}}^{(\boldsymbol{k})\prime}\right)^{\prime}$

[3] See discussions in Koenker and Xiao (2006) and Hallin and Werker (2006) or Ruzicka (2021) for the multivariate case.

[4] Ruzicka (2021) mentions that one could mitigate this problem by using nonlinear transformations of regressions in quantile local projection setting, but this lies beyond the scope of the present paper.

be the corresponding vector of regressors. Then the estimator is given by

$$\hat{\beta}^{(k)}(\tau_k) := \underset{\beta \in \mathbb{R}^{(k+Kp)}}{\operatorname{argmin}} \sum_{t=p+1}^{T} \rho_{\tau_k}\left(y_{k,t} - \beta' \boldsymbol{x}_t^{(k)}\right) \tag{4}$$

where $\rho_{\tau_k}(u) := u(\tau_k - \mathbb{I}\{u < \tau_k\})$ is the quantile loss function. Under some technical conditions that guarantee, among other things, that the process is strongly stationary and ergodic, Ruzicka (2021) has established the asymptotic normality of this estimator.[5] This estimator further enjoys a similar property to ordinary least squares under misspecification as it offers the optimal linear approximation to conditional quantiles in a weighted least square sense (Angrist, Chernozhukov, and Fernández-Val (2006)). This "robustness" property is one of the primary motivations behind its use for macroeconomic risk modeling.

In this paper, we produce all forecasts for QVAR models by simulating future sample paths from iteratively applying the random coefficient representation (2) using estimates obtained from (4). Specifically, at each point in time the parameters are selected by choosing the point on the quantile grid that falls closest to a uniform random draw $u_{k,t}$ for each equation k. Iterating this forward allows us to draw a sample path for $\boldsymbol{y}_{t+1}, \ldots, \boldsymbol{y}_{t+12}$, and repeating this a large number of times allows us to compute a variety of statistics at each point in time (quantile forecasts, mean forecasts, etc.).

This algorithm contrasts with the approach introduced by Adrian, Boyarchenko, and Giannone (2019) in a univariate context whereby the skewed t distribution of Azzalini and Capitanio (2003) is fitted to closely match a handful of conditional quantile forecasts produced using quantile regression estimates. On the other hand, it is closer in spirit to the method used by Chavleishvili and Manganelli (2021) for stress testing and Chavleishvili et al. (2021) for risk management in a macroprudential context as we can condition forecasts on scenarios by simply imposing predetermined sequences of quantiles. It also mirrors Ruzicka (2021)'s approach for obtaining quantile impulse responses. Considering this is how the QVAR model was introduced, we limit our attention to this approach.

An important detail concerns the choice of a grid of quantiles. We opted to use a relatively fine grid of 100 equally spaced quantiles, but note that some of those quantiles

[5] Using weights based on its asymptotic covariance matrix, $\hat{\beta}^{(k)}(\tau_k)$ viewed as a process over $\tau_k \in [0, 1]$ converges to a $Kp + k$-dimensional standard Brownian Bridge. The interested reader can also find some results for the quantile regression estimator under unit roots and cointegration in Koenker (2004), Xiao (2009) or Cho, Kim, and Shin (2015).

may not be well estimated. Chernozhukov, Fernández-Val, and Kaji (2017) suggested using extreme value methods for quantiles beyond $\tau T/(Kp+K) \leq 15$ where $Kp+K$ is the number of parameters in the last equation. For example, a bivariate QVAR model with a single lag estimated on 400 observations gives us the interval [0.15, 0.85], whereas adding a second lag reduces it to [0.225, 0.775]. Parsimony may thus be even more important when dealing with quantiles in the tails. For this reason, we follow Chavleishvili and Manganelli (2021) and Chavleishvili et al. (2021) and use a QVAR model with one lag throughout the paper. This also obviates the need to implement necessarily different lag selection procedures across models. Moreover, while information criteria to choose the number of lags in each equation separately have been proposed in the literature (e.g., Koenker and Machado (1999)), there currently is no counterpart for the entire QVAR process and this question is thus left open to future research.

3. Forecasting Experiment

In this section, we conduct an out-of-sample forecasting experiment in which we target many monthly US variables obtained from the FRED-MD data set (McCracken and Ng (2016)) spanning the period between January 1959 to June 2022. Since all our models will also feature the National Financial Conditions Index (NFCI), which is observed from January 1971 to June 2022, we select all target variables from FRED-MD that started at least as early as the NFCI and did not feature any missing values in the July 2022 version of the data. This leaves us with a subset of 112 target variables. To obtain many cycles of recessions and expansions, we set the start of the out-of-sample period to January of 1980, giving us six NBER recessions and a total of 510 periods for model comparison.

All target variables are transformed to induce stationarity.[6] We target the resulting values in $h = 1, \ldots, 12$ months rather than h period averages as forecasts are produced iteratively through simulations for all models.[7] Finally, given our focus on forecasting tails, a difficult balance must be struck between allowing a large sample size for estimation and allowing the model to adapt to structural changes. We opt for a rolling window of 400 observations, allowing the window to initially expand to this size to include the two recessions from the early 1980s in the analysis.

[6]We follow McCracken and Ng (2016), except that we do not take second differences on interest rates, unemployment rates, monetary aggregates and prices as in Bernanke, Boivin, and Eliasz (2005). All transformations are given in Table A1 Appendix.

[7]Results in Goulet Coulombe et al. (2021) suggests averaging single period forecasts *ex post* is generally preferable to directly targeting averages when point forecasts are of primary interests, but this question lies beyond the scope of this paper.

3.1. Models

The forecasting experiment includes four bivariate models with the targeted variable ordered first, followed by the NFCI. These models are the QVAR, as well as three parametric alternatives: a VAR-N, a VAR-GARCH and a VAR-SV. The VAR-N is a useful benchmark insofar as it is not obvious that modeling moments beyond the mean is meaningful for macroeconomic data (Plagborg-Moller et al. (2020)). The VAR-GARCH and VAR-SV models are interesting as common tools in the structural VAR literature, which relaxes the iid assumption of the VAR-N by allowing conditional volatility to change over time. We further consider two additional variations on the baseline QVAR model by introducing latent factor and latent quantile factor estimates as regressors, a set of hitherto unexplored extensions we call a factor augmented QVAR or QFAVAR model.

VAR-N. This model takes the form

$$\boldsymbol{y}_{t+1} = \boldsymbol{\nu} + \boldsymbol{A}_1 \boldsymbol{y}_t + \boldsymbol{u}_{t+1}, \quad \boldsymbol{u}_{t+1} \sim N(\boldsymbol{0}, \boldsymbol{\Sigma}). \tag{5}$$

We estimate mean parameters $\boldsymbol{\nu}$ and $\boldsymbol{A}_1$ by ordinary least squares. The covariance matrix of innovations is estimated as $\hat{\boldsymbol{\Sigma}} = \sum_{t=2}^{T} \hat{\boldsymbol{u}}_t \hat{\boldsymbol{u}}_t' / (T - Kp - 2)$ where $K = 2$, $p = 1$ and $\hat{\boldsymbol{u}}_t$ are residuals.

VAR-GARCH. We follow the structural VAR literature (e.g., Normandin and Phaneuf (2004); Bouakez and Normandin (2010); Bouakez, Chihi, and Normandin (2014)) and create a multivariate GARCH process by imposing that each "structural" shock follows its own GARCH(1,1) process. Hence, we replace the normal for the vector of innovations with

$$\begin{aligned} \boldsymbol{u}_{t+1} &= \boldsymbol{D}\boldsymbol{\epsilon}_{t+1} \\ \epsilon_{k,t+1} &= \sqrt{h_{k,t+1}} z_{k,t+1}, \quad z_{k,t+1} \sim N(0,1) \\ h_{k,t+1} &= (1 - \alpha_k - \beta_k) + \alpha_k \epsilon_{k,t}^2 + \beta_k h_{k,t}. \end{aligned} \tag{6}$$

where $\boldsymbol{D}$ is lower triangular to use the same restriction as in the QVAR model. We use the same parameter estimates for $\boldsymbol{\nu}$, $\boldsymbol{A}_1$ and $\boldsymbol{\Sigma}$ as we do for the Gaussian VAR case. $\hat{\boldsymbol{D}}$ is obtained from a Cholesky factorization of $\hat{\boldsymbol{\Sigma}}$. Series of "structural residuals" $\hat{\epsilon}_{\boldsymbol{k,t}}$ are then obtained on which individual GARCH(1,1) processes are fitted by maximum likelihood.

(Bayesian) VAR-SV. We use one of the restricted models featured in Chan and Eisenstat (2018), which essentially replaces individual GARCH processes featured in the VAR-GARCH model shown above by (random walk) stochastic volatility processes.

$$
\begin{aligned}
\boldsymbol{B_0}\, \boldsymbol{y}_{t+1} &= \mu + \boldsymbol{B_1}\, \boldsymbol{y}_t + \epsilon_{t+1} \\
\epsilon_{k,t+1} &= \exp\left(h_{k,t+1}/2\right) z_{k,t+1},\ \ z_{k,t+1} \sim N(0,1) \\
h_{k,t+1} &= h_{k,t} + \sigma_k \zeta_{k,t+1},\ \ \zeta_{k,t+1} \sim N(0,1)
\end{aligned}
\tag{7}
$$

We impose recursive short-run restrictions as with the QVAR and VAR-GARCH models such that $\boldsymbol{B}_0$ is set to a lower triangular matrix with a unit diagonal. It is a common choice (e.g., Cogley and Sargent (2005) and Primiceri (2005)). The model is estimated using Bayesian methods with the following priors:

$$
\theta := \left(\boldsymbol{vec}\left((\mu, \boldsymbol{B_1})'\right)', b_{0,2,1}\right)' \sim N(\boldsymbol{b}_\theta, \boldsymbol{V}_\theta), \quad \boldsymbol{h_0} \sim N(\boldsymbol{b_h}, \boldsymbol{V_h}) \ \text{ and } \ \sigma_k \sim IG(\nu_k, S_k).
$$

We set $\boldsymbol{b}_\theta$ and $\boldsymbol{V}_\theta$ as a Minnesota-type prior with common hyperparameter values centered on a random walk, except for the growth rates of consumption, exchange rates and stock market indexes, which we center on white noise. We center the value for $b_{0,2,1}$ at 0 with a relatively large variance of 10 and likewise for the initial log variance ($\boldsymbol{b_h} = \boldsymbol{0}$ and $\boldsymbol{V}_h = 10$) following Chan and Eisenstat (2018). We use the shape $\nu_k = 5$ and scale $S_k = 0.1(\nu_k - 1)$ as in Chan and Eisenstat (2018), reflecting a relatively diffuse prior centered on a small value (here, 0.1).

Their Gibb Sampling algorithm has two particular features. First, it jointly samples mean parameters θ for each equation whereas other algorithms would sample free elements in $\boldsymbol{B_0}$ separately. Second, while it applies the common auxiliary mixture sampler proposed by Kim, Shephard, and Chib (1998), which allows using methods for linear Gaussian state-space models, it also samples the sequence of log variances $(\boldsymbol{h_t})_{t=1}^{T}$ in a single step for each equation using the precision sampler of Chan and Hsiao (2014). These features make the algorithm fairly efficient.

QFAVAR. As a means of exploring the value of a data-rich environment for macroeconomic forecasting, we introduce latent factor estimates as part of the vector of variables $\boldsymbol{y_t}$ in (2). This is similar in spirit to the FAVAR model of Boivin and Ng (2005), although we do not impose restrictions that would strictly justify treating the target variable and NFCI as "observed" factors. In all cases, latent factors are recursively estimated using the in-sample data window. We collect our 112 variables into a matrix X and let variable i obey

$$x_{i,t} = \boldsymbol{\lambda}_{\boldsymbol{i}}' \boldsymbol{f}_{\boldsymbol{t}} + \nu_{i,t} \tag{8}$$

where $\boldsymbol{f}_{\boldsymbol{t}}$ is a $r \times 1$ vector of factors and $\boldsymbol{\lambda}_{\boldsymbol{i}}$ is the corresponding vector of loadings. Following common practice since Stock and Watson (2002a,b), we obtain factor estimates $\hat{\boldsymbol{f}}_{\boldsymbol{t}}$ by principal component. We set $r = 1$ factor out of concern for parsimony so our vector of time series becomes $\boldsymbol{y}_{\boldsymbol{t}} = (y_{1,t}, \hat{f}_t, NFCI_t)'$. A natural alternative would be to consider doing the same thing using the quantile latent factors recently introduced by Chen, Dolado, and Gonzalo (2021). In this case, we have

$$\mathbb{Q}_{x_{i,t}}\left(\tau | \boldsymbol{f}_{\boldsymbol{t}}(\tau)\right) = \boldsymbol{\lambda}_{\boldsymbol{i}}(\tau)' \boldsymbol{f}_{\boldsymbol{t}}(\tau) + \nu_{i,t}(\tau) \tag{9}$$

with $f_t(\tau)$ being $r(\tau) \times 1$. We obtain estimates $\tilde{\boldsymbol{f}}_{\boldsymbol{t}}(\tau)$ for the 5th and 95th quantiles using the IQR algorithm (Chen, Dolado, and Gonzalo (2021)). Again, we set $r(\tau) = 1$ for parsimony and use $\boldsymbol{y}_{\boldsymbol{t}} = (y_{1,t}, \tilde{f}_t(0.05), \tilde{f}_t(0.95), NFCI_t)'$.

3.2. Relative Forecasting Evaluation

To perform the model comparison, we follow Carriero, Clark, and Marcellino (2024) and Carriero, Clark, and Marcellino (2022) in our evaluation of density forecasts and adopt the quantile weighted continuous ranked probability score (CRPS) introduced by Gneiting and Ranjan (2011). For model m and variable v, we define the h period ahead quantile forecasts as

$$\hat{q}_{t+h,v,m}(\tau) := \hat{\mathbb{Q}}^{(m)}_{y_{t+h,v}}(\tau | \mathcal{F}_t) .$$

and quantile scores as

$$QS_\tau\left(\hat{q}_{t+h,v,m}(\tau), y_{t+h,v}\right) := \rho_\tau\left(\hat{q}_{t+h,v,m}(\tau) - y_{t+h,v}\right) \tag{10}$$

where we recall that $\rho_\tau(u) := u(\tau - \mathbb{I}\{u < \tau\})$. For a grid of N quantiles, the quantile weighted CRPS is defined as

$$qwCRPS\left(\hat{\boldsymbol{q}}_{\boldsymbol{t+h,v,m}}, \boldsymbol{\nu}, y_{t+h,v}\right) = \frac{2}{N-1} \sum_{j=1}^{N} \nu(\tau_j) QS_{\tau_j}\left(\hat{q}_{t+h,v,m}(\tau_j), y_{t+h,v}\right) \tag{11}$$

where $\boldsymbol{\nu} := \left(\nu(\tau_j)\right)_{j=1}^{N}$ is a vector of weights and $\boldsymbol{\hat{q}_{t+h,v,m}} := \left(\hat{q}_{t+h,v,m}(\tau_j)\right)_{j=1}^{N}$ stacks quantile forecasts into a vector. Gneiting and Ranjan (2011) proposed using the functions $\nu(\tau_j) = \tau_j^2$, $\nu(\tau_j) = (1-\tau_j)^2$ and $\nu(\tau_j) = (2\tau_j - 1)^2$ to put more weight on the right tail, left tail or both tails jointly. The use of this scoring rule is motivated by the fact that it is minimized in expectation by the true conditional density (Gneiting and Raftery (2007)).

Diebold and Mariano (1995) tests allow us to evaluate the null hypothesis of equal forecasting performance between models m_1 and m_2 using the following regression

$$qwCRPS\left(\boldsymbol{\hat{q}_{t+h,v,m_1}}, \boldsymbol{\nu}, y_{t+h,v}\right) - qwCRPS\left(\boldsymbol{\hat{q}_{t+h,v,m_2}}, \boldsymbol{\nu}, y_{t+h,v}\right) = \alpha_{v,h,m_1,m_2} + v_{t+h,v,m_1,m_2}$$

for each forecasting horizon h and variable v where $\alpha_{v,h,m_1,m_2} = 0$ under the null.[8] In this context, note that $\alpha_{v,h,m_1,m_2} < 0$ means that model m_1 is performing better than model m_2 (i.e., its average score is lower).

3.3. Absolute Forecasting Evaluation

In an effort to mitigate concerns with the choice of benchmark models, we supplement the model comparison with specification tests used in finance for evaluating value-at-risk models.

Quantile Mincer-Zarnowitz Tests. Gaglianone et al. (2011) proposed a test of quantile forecast optimality in the spirit of Mincer and Zarnowitz (1969) based on a quantile regression by adapting an idea from Christoffersen, Hahn, and Inoue (2001). Let $\mathbb{Q}_{y_{i,t+h}}(\tau|\mathcal{F}_t)$ and $\hat{\mathbb{Q}}_{y_{i,t+h}}(\tau|\mathcal{F}_t)$ be the h-period ahead τ-th quantile of variable i conditional on information $\mathcal{F}_t$ and its forecast by some model, respectively. If the model is correctly specified, we should have

$$\mathbb{Q}_{y_{i,t+h}}(\tau|\mathcal{F}_t) = \alpha_0(\tau) + \alpha_1(\tau)\hat{\mathbb{Q}}_{y_{i,t+h}}(\tau|\mathcal{F}_t) := \boldsymbol{x_t'}\boldsymbol{\alpha}(\tau)$$

with $\boldsymbol{\alpha}(\tau) := (\alpha_0(\tau), \alpha_1(\tau))' = (0,1)'$ and $\boldsymbol{x_t} := \left(1, \hat{\mathbb{Q}}_{y_{i,t+h}}(\tau|\mathcal{F}_t)\right)'$. These parameters can be estimated by a quantile regression of realized values $y_{i,t+h}$ on the quantile forecasts $\hat{\mathbb{Q}}_{y_{i,t+h}}(\tau|\mathcal{F}_t)$ at the corresponding quantile τ. Under mild regularity conditions, the Wald statistic testing the null of correct specification has a χ_2^2 asymptotic distribution.[9]

[8] The constant is estimated by OLS and HAC standard errors are used in all cases.

[9] For the implementation, we follow the authors' suggestion and use Koenker and Machado (1999)'s estimator for the covariance matrix.

Note that simulation evidence in Gaglianone et al. (2011) suggests this test suffers from size distortion in small sample (the true size tends to be larger than the nominal size), but it tends to enjoy as much or more power than the more common alternatives tests based on dummy variables such as Kupiec (1995), Christoffersen (1998) or Engle and Manganelli (2004).

Coverage Tests. We begin by defining $I_t := \mathbb{1}\left\{y_{i,t} \in \left[\hat{\mathbb{Q}}_{y_{i,t+1}}(0.05|\mathcal{F}_t), \hat{\mathbb{Q}}_{y_{i,t+1}}(0.95|\mathcal{F}_t)\right]\right\}$ as a dummy variable indicating when observations fall inside this symmetric 90% interval.[10] Following Kupiec (1995), Christoffersen (1998) leverages the idea that a correctly specified model implies I_t should be an iid Bernoulli variable with a $\rho = 0.9$ success rate. The likelihood is thus given by

$$\mathcal{L}(\rho) = \Pi_{t=1}^{T}(1-\rho)^{1-I_t}\tau^{I_t} = (1-\rho)^{T_0}\rho^{T_1}$$

where $T_0 = \sum_{t=1}^{T}(1-I_t)$ and $T_1 = \sum_{t=1}^{T} I_t$ where T is the size of the pseudo-out-of-sample period. The unconditional coverage test is based on a likelihood ratio statistic that compares this likelihood evaluated at the nominal coverage rate $\rho = 0.9$ and its sample counterpart, $\hat{\rho} = T^{-1}\sum_{t=1}^{T} I_t$, which is the maximum likelihood estimator. Under the null, Christoffersen (1998) shows the likelihood ratio statistic satisfies $LR_{uc} = -2\log\left(\mathcal{L}(\rho)/\mathcal{L}(\hat{\rho})\right) \xrightarrow{d} \chi_1^2$. The conditional coverage test changes the alternative hypothesis by modelling possible serial dependence in I_t as a first order Markov Chain with transition matrix

$$\Pi_1 := \begin{bmatrix} 1-\pi_{01} & \pi_{01} \\ 1-\pi_{10} & \pi_{11} \end{bmatrix}$$

where $\pi_{ij} = \mathbb{P}(I_{t+1} = j | I_t = i)$. The likelihood function is then given by

$$\mathcal{L}(\Pi_1) = (1-\pi_{01})^{T_{00}}\pi_{01}^{T_{01}}(1-\pi_{11})^{T_{10}}\pi_{11}^{T_{11}}$$

where $T_{ij} := \sum_{t=2}^{T}\mathbb{1}\{I_{t+1} = j\}\mathbb{1}\{I_t = i\}$ counts transition cases with the maximum likelihood estimator being again the sample shares of the relevant events, that is $\hat{\pi}_{i1} = T_{i1}/(T_{i0} + T_{i1})$. Under the null, Christoffersen (1998) shows the likelihood ratio statistic satisfies $LR_{cc} = -2\log\left(\mathcal{L}(\rho)/\mathcal{L}(\hat{\Pi}_1)\right) \xrightarrow{d} \chi_2^2$. Note that for both coverage tests, we follow

[10] Since these are binary events, this is equivalent to jointly testing coverage in the 5% tail on each side of the distribution.

Christoffersen (2004) and adopt the Monte Carlo testing approach of Dufour (2006) and obtain exact finite sample p-values instead of relying on asymptotic approximations.[11] The ability to control the size of these tests exactly in small sample is an advantage they possess over the previous specification tests. However, the power of those tests varies with sample size[12] and they require non overlapping forecasts, so we only perform these test for $h = 1$. Nevertheless, as QVAR (and QFAVAR) models produce forecasts iteratively, misspecification at $h = 1$ would naturally propagate forward and may pose problems with quantile forecasting accuracy at longer horizons.

4. Discussion

As explained in the previous sections, we rely on Diebold and Mariano (1995) tests to evaluate the forecasting accuracy of the QVAR model relative to the three parametric alternatives, VAR-N, VAR-GARCH and VAR-SV. This raises the problem of concisely reporting a very large number of results. We proceed in a manner similar to Stock and Watson (1998) who report test rejection counts. We use Diebold and Mariano (1995) tests as a means of categorizing variables. Specifically, given that we seek to minimize the tail weighted quantile CRPS, we consider that the QVAR model "wins" against a given benchmark, at a given horizon and for a given variable when it has a lower average score and the null of equal forecasting performance is rejected at 5%. The QVAR model "loses" if it has a higher average score and the null of equal forecasting performance is rejected at 5%. In all other cases, we consider that the models have equal forecasting performance. A similar idea is applied to build figures for the Gaglianone et al. (2011) and Christoffersen (1998) tests and the same figures are presented for QFAVAR models.

4.1. QVAR Results

Figure 1 features two panels that each display the number of variables in each of the eight groups featured in FRED-MD for which the QVAR model wins and loses. Group results are stacked so that the total number of wins and losses correspond to the top of area for group one. The figure shows results for each decade of the out-of-sample period, and the whole out-of-sample period and for each of the parametric benchmark. As an example of how to read the figure, consider the area plot displayed in the first row and first column of panel A. The counts refer to the number of variables for which the

[11] The procedure is detailed in Section A of the Appendix.
[12] See, for example, simulation evidence in Gaglianone et al. (2011).

QVAR model statistically significantly outperformed the VAR-N models in the 1980s. At horizon 3, there are about 40 variables out of 112 and about 10 of those variables are in the labor market (group 2) category. The same entry in panel B shows the VAR-N models statistically significantly outperform the QVAR models for fewer than 10 variables out of 112 at all horizons.

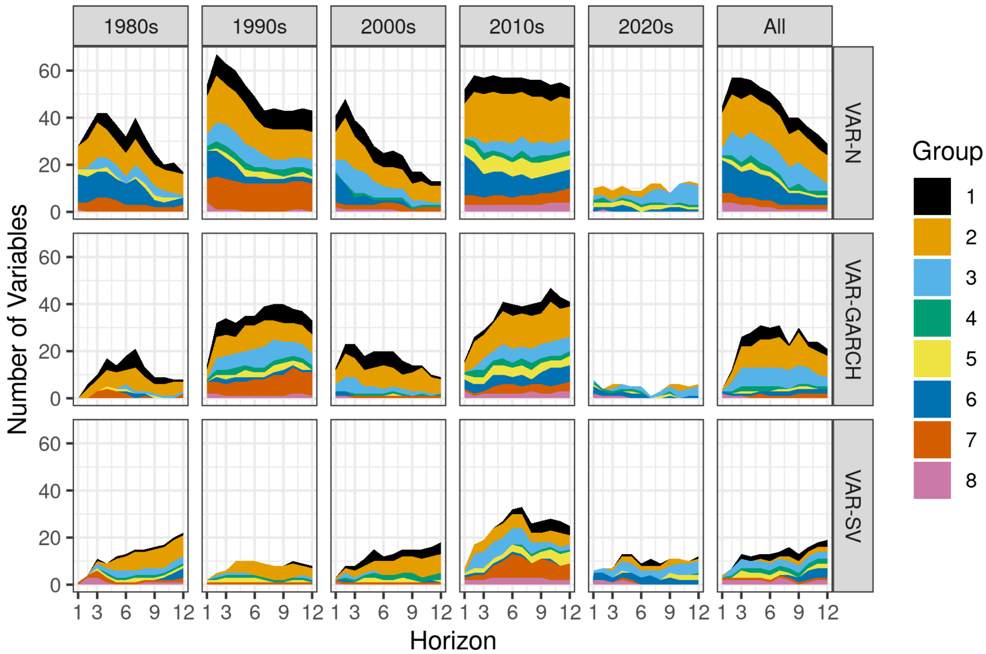
We begin by focusing on the last column of each panel for the average relative performance across the whole out-of-sample period. For this period, panel A shows that the QVAR models significantly outperform the VAR-N models in 25% to 50% of cases depending on the horizon. It also significantly improves upon the VAR-GARCH and VAR-SV models in about 25% and 10% of cases, respectively. Importantly, panel B reveals that the QVAR model rarely does significantly worse than any of the three benchmark models considered, losing in around 18% of cases at a horizon of one month against the VAR-GARCH and VAR-SV models and in only 5% to 10% of cases at all other horizons. Breaking down results across different categories of variables, the QVAR model appears to perform best relative to benchmark models when applied to the labor market (group 2) at all horizons and to interest and exchange rates (group 6), especially at shorter horizons. We note in the few cases where the QVAR model is outperformed across all benchmark models, most are prices (group 6). The bulk of issues of relative performances identified in the short 2020s sub-sample included in the forecasting experiment are also related to prices.

Shifting our attention across the first five columns of each panel, we can get a sense the stability of those results. The patterns of relative performance appear to vary slightly over time, but the broad qualitative message remains the same. Across each decade, the QVAR model rarely does worse than benchmark models, more frequently improves upon them, and both of these observations are concentrated in the same categories of variables.

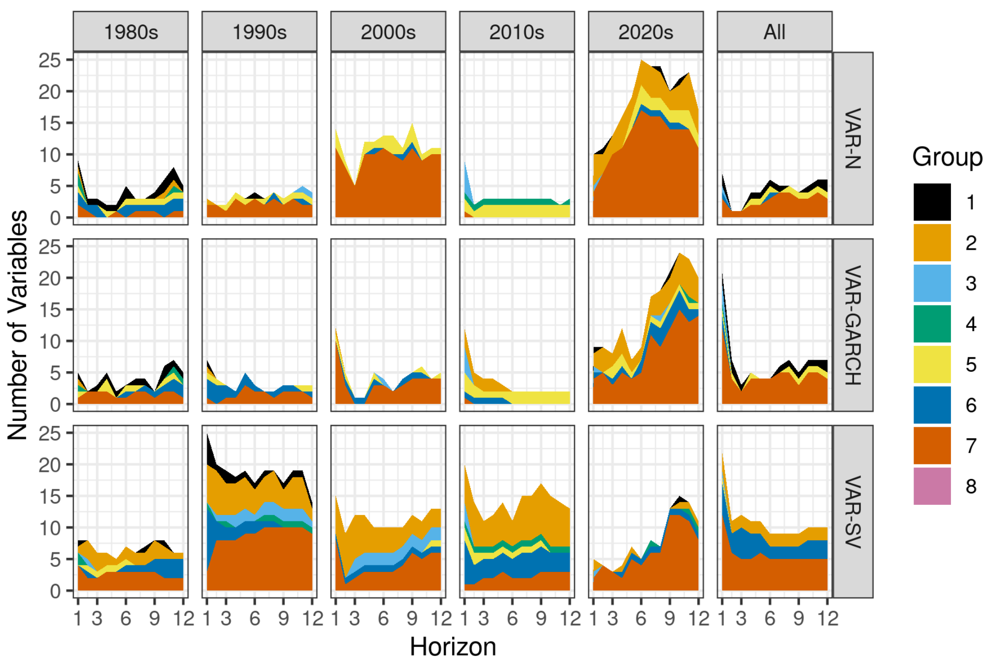
Of course, Figure 1 does not tell us whether the statistically significant differences in performance between models are meaningfully large. To this end, Figure 2 displays the average log differences in scores between the QVAR and VAR-GARCH and VAR-SV [13] models over the whole out-of-sample period for each variable in each of the eight groups of variables in FRED-MD. Values below zero indicate that the QVAR models have a smaller average score than the benchmark and a rejection of the null in the corresponding Diebold and Mariano (1995) test in either direction is indicated by the

[13]Panel B omits the results for the oil price variable in group 7 because the VAR-SV model performs too poorly and it hindered visualizing the rest of the results. The random walk process for stochastic volatilities seems to be the culprit.

color yellow. For example, take the plot in the first row and first column of panel A. It shows that there is a variable in the output and income (group 1) for which the QVAR model is about 15% more accurate at forecasting the tails than the VAR-GARCH model at all horizons, and this difference is statistically significant at 5%.



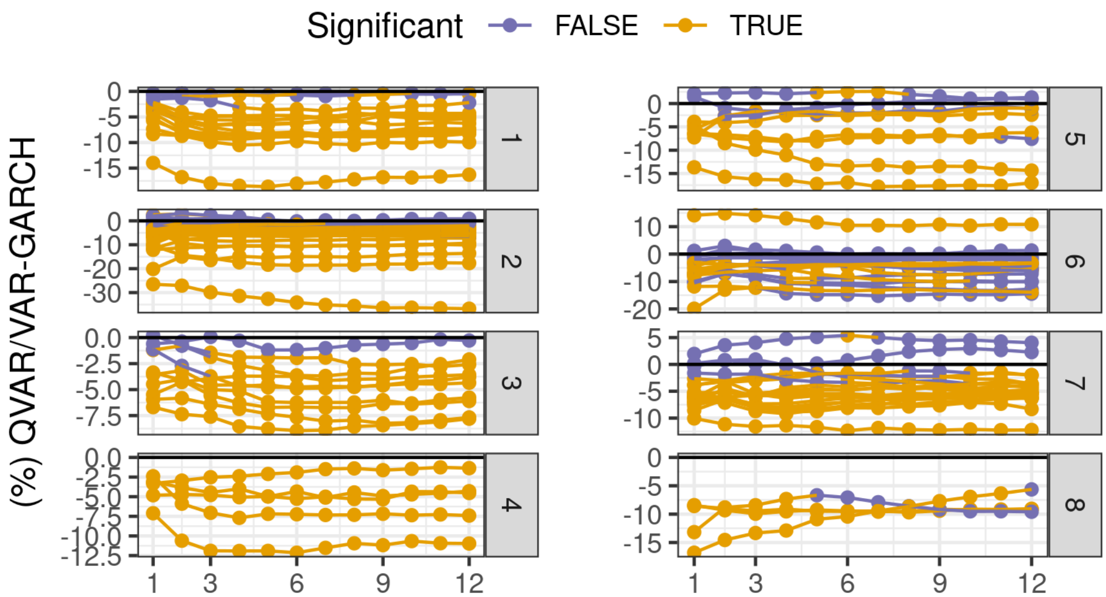
A. Number of Cases where QVAR Beats the Benchmark



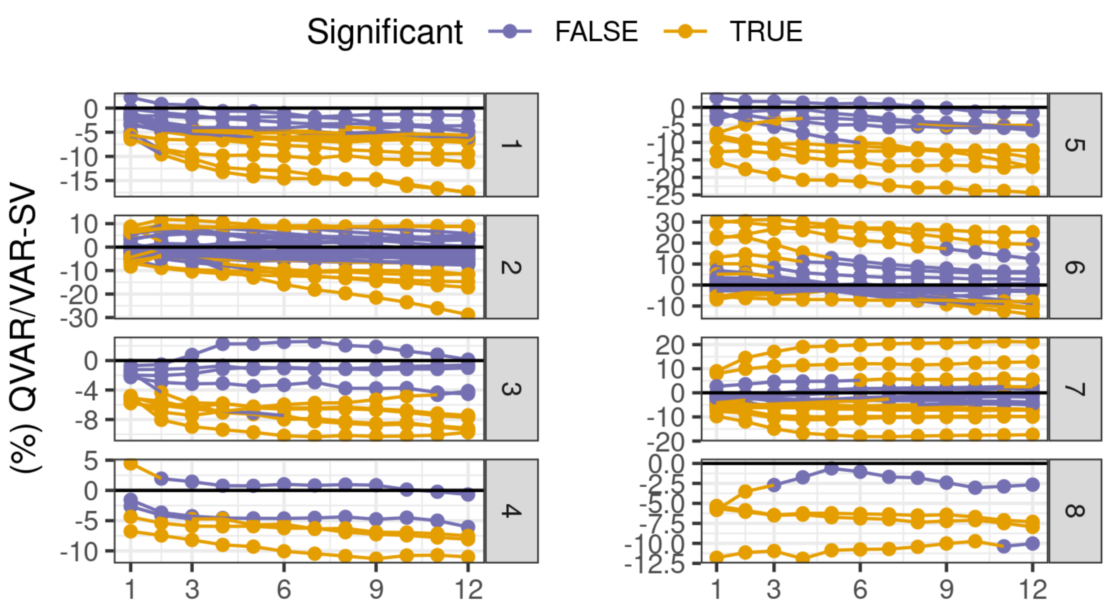
B. Number of Cases where QVAR Loses to the Benchmark

FIGURE 1. QVAR Diebold-Mariano Tests (tail-weighted CRPS)

Note: The QVAR model wins (loses) when it has a lower (higher) average score and the Diebold-Mariano statistic is significant at the 5% level. Columns are periods and rows are different benchmark models. Colors indicate FRED groups: **(1)** Output and income, **(2)** Labor market, **(3)** Housing, **(4)** Consumption, orders and inventories, **(5)** Money and credit, **(6)** Interest and exchange rate, **(7)** Prices and **(8)** Stock market.



A. QVAR and VAR-GARCH



B. QVAR and VAR-SV

FIGURE 2. QVAR Relative Scores (tail-weighted CRPS)

Note: Negative values are improvements. Yellow corresponds to rejecting the null of equal scores at 5%. Yellow corresponds to not rejecting the null of equal scores at 5%. FRED groups are: **(1)** Output and income, **(2)** Labor market, **(3)** Housing, **(4)** Consumption, orders and inventories, **(5)** Money and credit, **(6)** Interest and exchange rates, **(7)** Prices and **(8)** Stock market.

We find that the statistically significant improvements provided by QVAR over the VAR-GARCH and VAR-SV benchmarks routinely exceed 10%, even sometimes 20% or more, across horizons. These gains tend to concentrate in the categories of output and income (group 1), labor market (group 2) and money and credit (group 5). Importantly, the QVAR model almost never does much worse than the VAR-GARCH. However, it does underperform the VAR-SV model by statistically significant margins in excess of 10%. This can be seen in interest and exchange rates (group 6), as well as prices (group 7), even as it happens infrequently. Figure A1 in the appendix displays the corresponding comparison between the QVAR and VAR-N models, where the QVAR model is considerably more favorable: it virtually never does worse, it almost always does better, the improvements can be large or even very large and it's frequently statistically significant.

Figures 1, 2 and A1 all point to the QVAR model performing relatively better against the VAR-N model than the VAR-GARCH and VAR-SV models. Moreover, this pattern seems to hold over time. Since all four models imply the same linear form for the conditional expectation of the target variable at all horizons, these results suggest that there is enough information in macroeconomic data to meaningfully capture variation in conditional volatility. Overall, these figures also offers evidence in favor of the purported robustness of the QVAR model in the sense that it tends to perform as well as or better than, but almost never much worse than, competing parametric alternatives.

Perhaps surprisingly, Figure 3 reveals that this good relative performance of the QVAR model is generally not driven by recessions. When a forecasted value is realized in what the NBER later determines to be a recession month, the QVAR model performance is statistically indistinguishable from that of benchmark models in 75% to 85% of cases depending on the model and horizon. There is a slightly greater advantage during recessions against the VAR-SV than the VAR-N and VAR-GARCH models at longer horizons. Given that this pattern does not hold against the VAR-GARCH models, this may be due to the fact that the random walk in stochastic volatility may overstate the persistence of uncertainty in those circumstances. This finding points to the presence of important variations in macroeconomic risks during periods of economic expansion that aren't as well captured by the parametric alternatives we considered. One possible explanation is that the binary discrete approximation to what is an otherwise continuous state variable we call "the business cycle" is neglecting meaningful information and more than two states should be considered. Alternatively, we can also note that the concept of a recession is fuzzy and the NBER recession dates are up for debate.

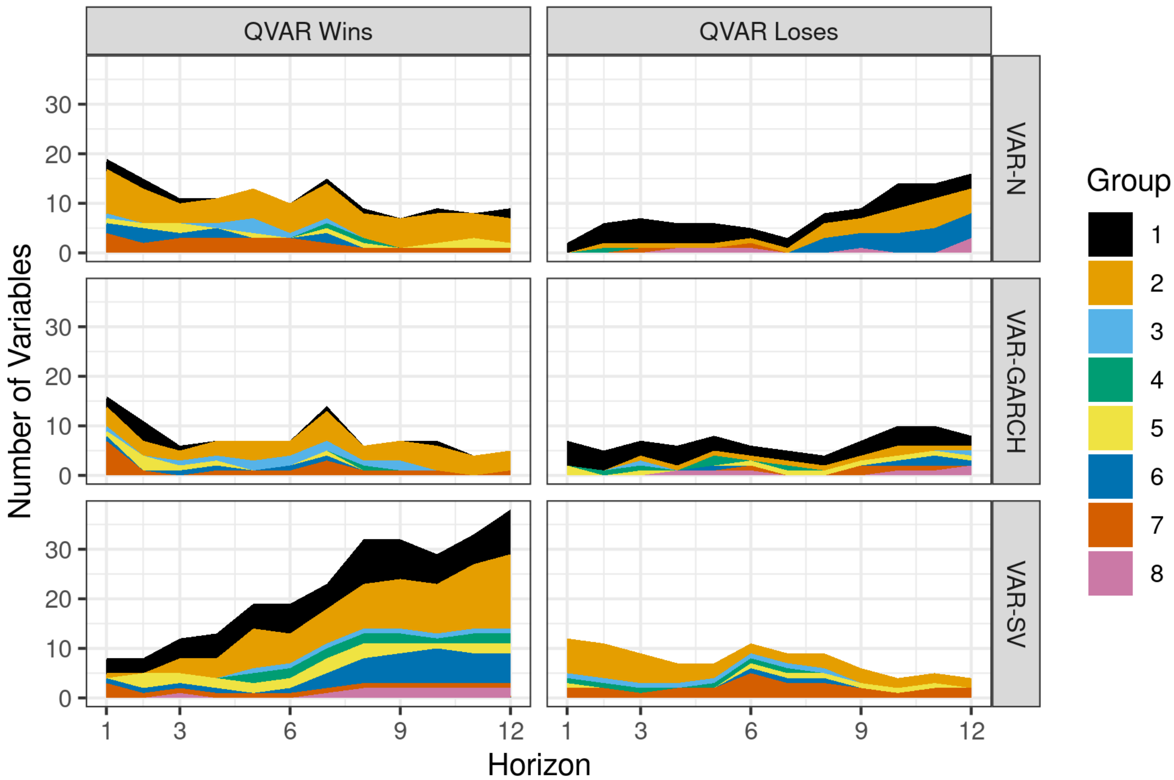


FIGURE 3. QVAR Recesssion Comparison (tail-weighted CRPS)

Note: The QVAR model wins (loses) when it has a lower (higher) average score and the Diebold-Mariano statistic is significant at the 5% level. Rows are different benchmark models. Colors indicate FRED groups: **(1)** Output and income, **(2)** Labor market, **(3)** Housing, **(4)** Consumption, orders and inventories, **(5)** Money and credit, **(6)** Interest and exchange rates, **(7)** Prices and **(8)** Stock market.

We now turn to specification tests. Figure 4 displays results for the quantile Mincer-Zarnowitz tests of Gaglianone et al. (2011) for forecasts at the 5th and 95th quantiles seperately. The figure counts the number of variables for which the null hypothesis of a correctly specified quantile forecast cannot be rejected at a level of 5% and breaks these results down for each decade in the out-of-sample period, as well as the whole period, and each group of variables in FRED-MD. As an example of how to read the figure, the area plot in the column of the first row shows that at a horizon of 1 and 2 months, we cannot reject the hypothesis that the QVAR forecast is well specified at 5% during the 1980s for about 75 out of 112 variables. This is also true for over 20 labor market variables (group 2).

Looking across all columns and rows, we see that the null hypothesis of optimal quantile forecasts cannot be rejected for 25% to 50% of cases, depending on the horizon and period covered. However, we consistently find greater evidence of misspecification at longer than at shorter horizons. Of all variable types, the test again singles out labor market (group 2) and interest and exchange rates (group 6) variables as cases where the model performs particularly well. This picture is also relatively stable over time.

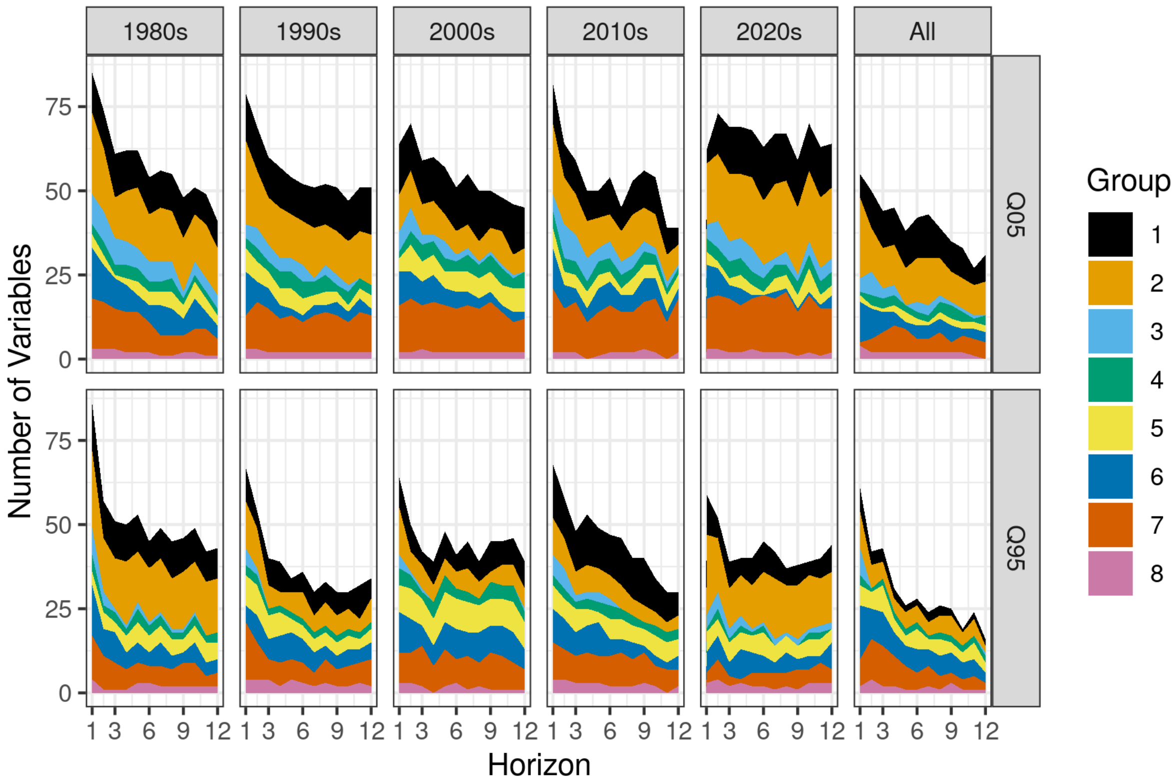


FIGURE 4. Number of Optimal QVAR Forecasts

Note: Number of cases where we obtain a non-rejection of the null of optimal forecast at 5% for the Gaglianone et al. (2011) test. Columns are periods and rows are the 5th (Q05) and 95th (Q95) quantile forecasts, respectively. Colors indicate FRED groups: **(1)** Output and income, **(2)** Labor market, **(3)** Housing, **(4)** Consumption, orders and inventories, **(5)** Money and credit, **(6)** Interest and exchange rates, **(7)** Prices and **(8)** Stock market.

Results for coverage tests are presented in Figure 5. Here we focus on one month ahead 5th and 95th quantile forecasts,[14] so an observation that falls outside of the 90% interval between them will be our notion of a "tail event." If the QVAR model is correctly specified, the unconditional probability of a tail event would be 10%. That's the null of the unconditional coverage tests. Moreover, "tail events" should be "unpredictable" and, in particular, they shouldn't be serially correlated. The null of correct conditional

[14] Recall that the tests are carried out only for horizon $h = 1$ for reasons discussed in Section 3.

coverage jointly tests both of these restrictions. The figure shows the shares of non-rejection of the null for each of those tests for each group of variables and each decade of the out-of-sample period.

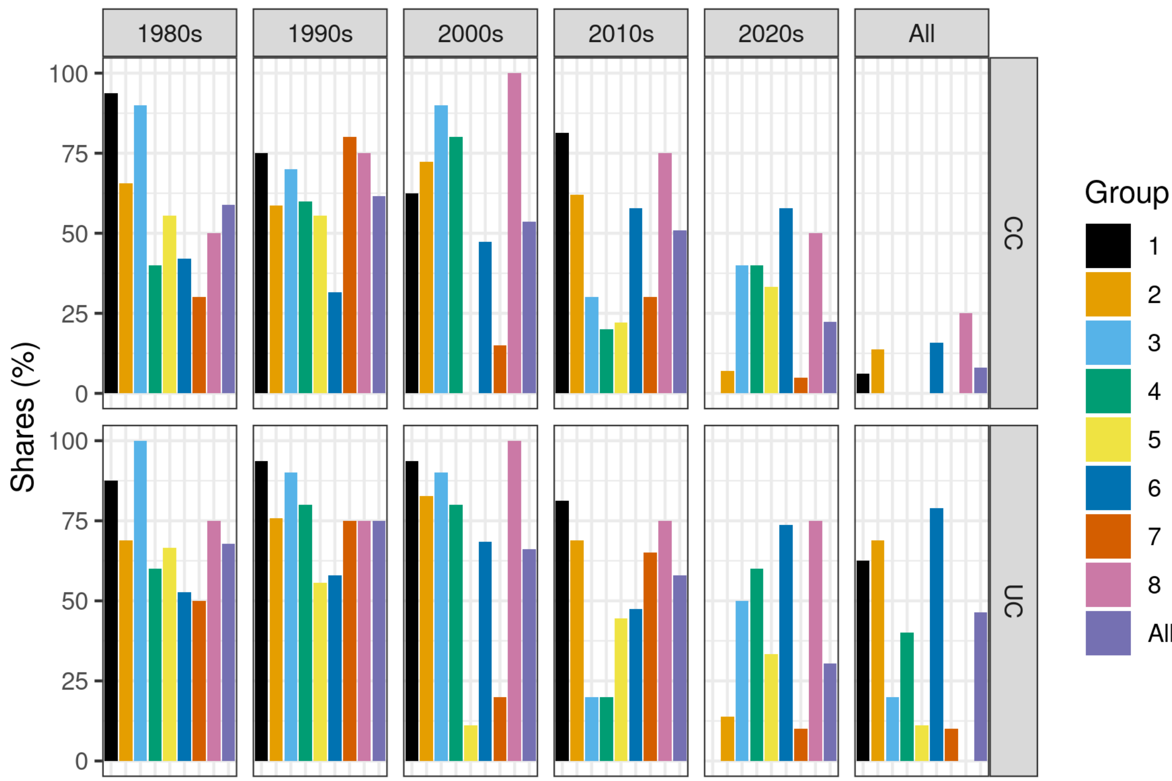


FIGURE 5. Unconditional and Conditional Coverage Tests on QVAR (90% Interval)

Note: Shares of non-rejection of the null at 5% using Monte Carlo p-values (Dufour (2006)). UC is the unconditional coverage test and CC is the conditional covargae test. Columns are periods. Colors indicate FRED groups: **(1)** Output and income, **(2)** Labor market, **(3)** Housing, **(4)** Consumption, orders and inventories, **(5)** Money and credit, **(6)** Interest and exchange rates, **(7)** Prices and **(8)** Stock market and **(All)** 112 variables.

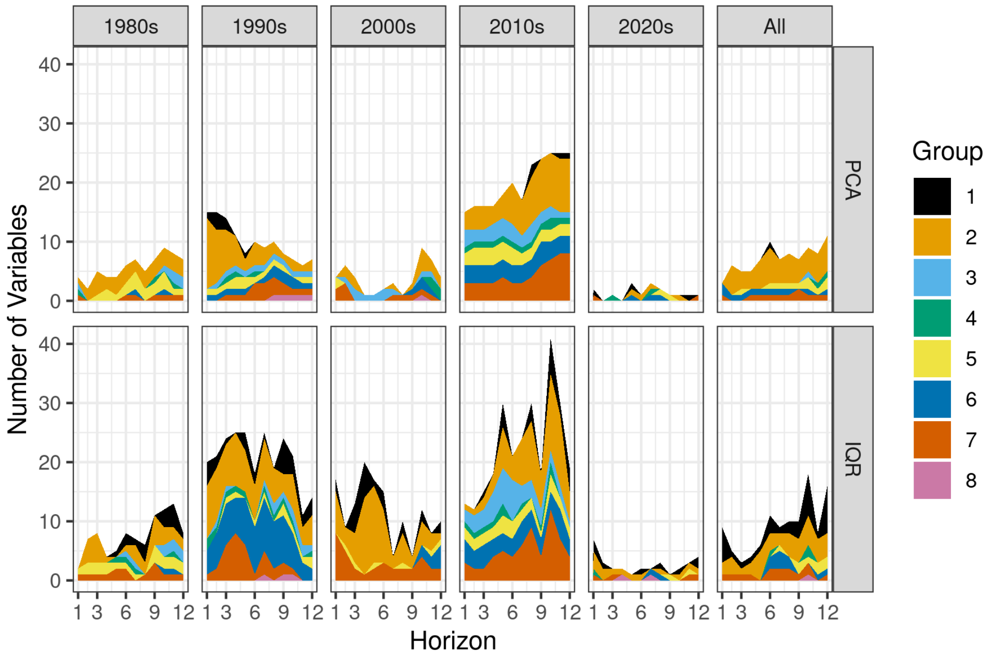
The last column shows results for the whole pseudo-out-of-sample period. We can see in the bottom row that the test fails to reject the null of correct unconditional coverage in almost 50% of cases and, likewise, for more than half of output and income (group 1), labor market (group 2) and interest and exchange rates (group 6). However, we can see that the null of correct conditional coverage is rejected over 90% of times across all variable groups. This means that, on average, QVAR quantile forecasts lead to a correct probability of 10% for observing tail events, but it leads to tail events that are serially correlated. Since forecasts at longer horizons are produced iteratively, this may explain the consistent pattern in Figure 4 where the null of optimal quantile forecasts can always be rejected more frequently at longer than at shorter horizons.

4.2. QFAVAR Results

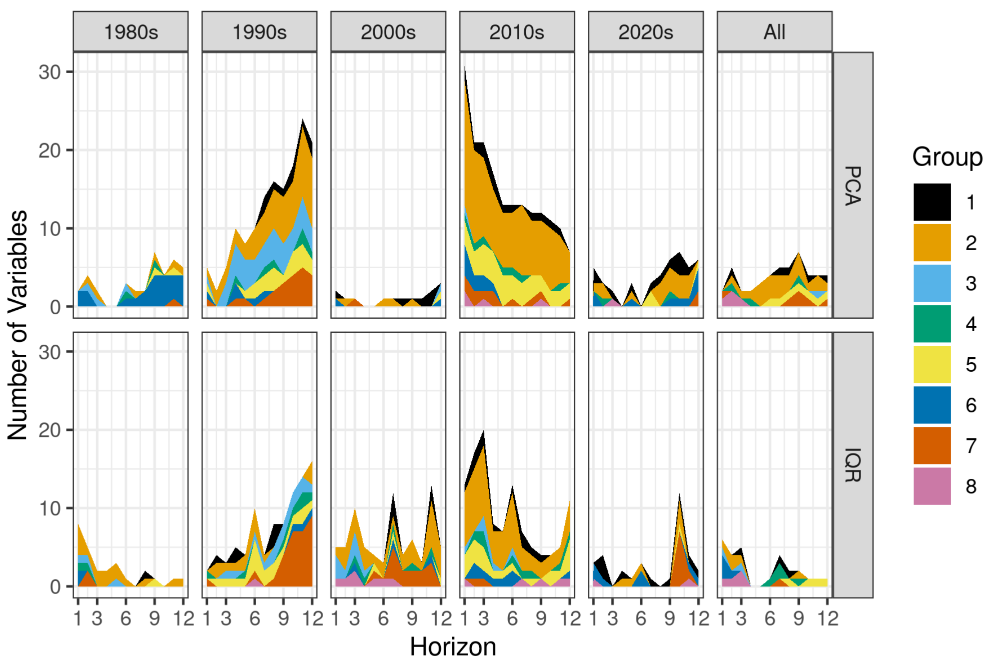
Figure 6 presents the results for the Diebold and Mariano (1995) tests comparing both QFAVAR models with the QVAR model. The last column of both panels shows that on average over the whole pseudo-out-of-sample period, the QVAR model and both QFAVAR models have statistically indistinguishable tail-density forecasting performances in 85% to 90% of cases depending on the horizon and the type of factor included. More differences can be noticed during the 1990s and 2010s where the addition of either PCA or IQR factors improves performance for prices (group 7) in particular. Recall that the few cases in which the VAR-SV model significantly outperforms the QVAR model with relatively large magnitudes are concentrated in this category.

Figure 7 displays the average log differences in scores, as well as whether these differences are statistically significant. As the QFAVAR models serve as a benchmark in these comparisons, adding factors is found to be helpful when the values displayed are positive. We can see that the changes in accuracy resulting from the addition of either type of factors are relatively small, with the vast majority below 5% in either direction. This corroborates the main finding from the previous figure and suggests that adding factors usually doesn't substantially affect tail-density forecasting accuracy. That being said, panel A does show a few moderate improvements obtained from the addition of a PCA factor in interest and exchange rates (group 6) and prices (group 7) variables. At the same time, introducing a PCA factor can be costly, as we can see in panel A, with some moderately negative values in money and credit (group 5), as well as in some of the interest and exchange rate (group 6) variables.

Perhaps where differences are most striking is during NBER recessions as can be seen in Figure 8. QFAVAR models outperform the QVAR model in 12% to 18% of cases depending on the horizon and type of factors considered. It also appears to be rarely costly to add factors when the realized value turns out to fall during a recession. It is especially visible with the labor market (group 2) where adding a PCA factor helps at all horizons, while the IQR factors seem to be most helpful at shorter horizons.



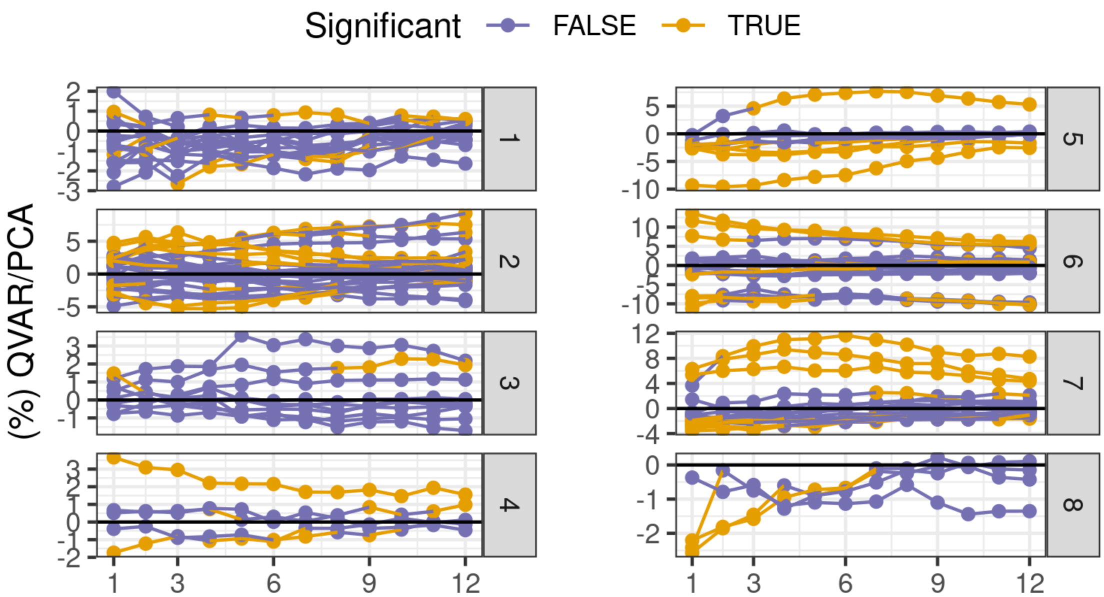
A. Number of Cases where QVAR Loses to QFAVAR



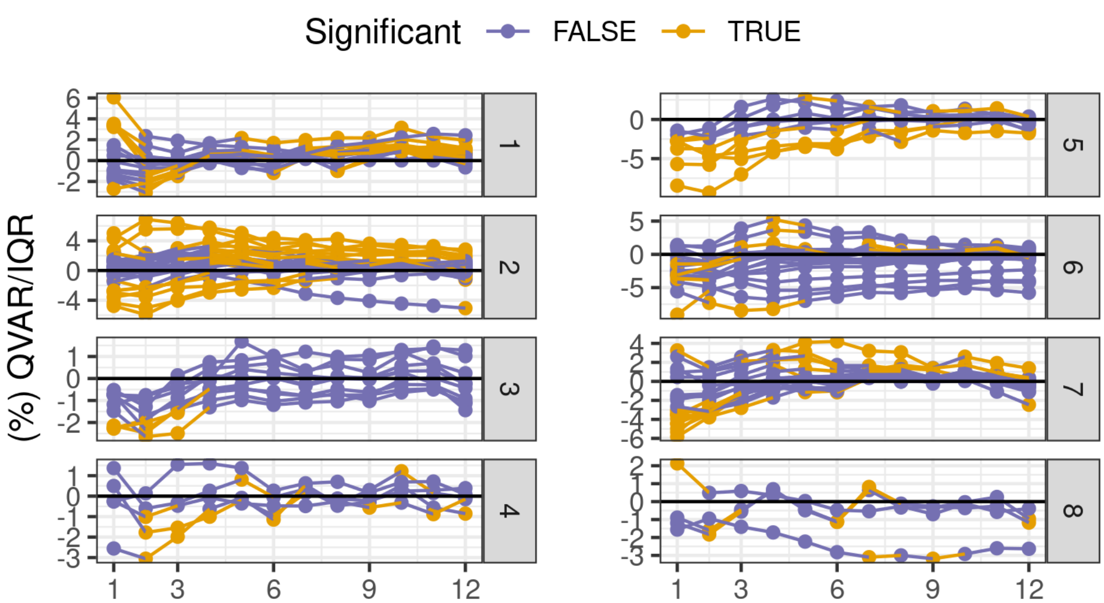
B. Number of Cases where QVAR Beats QFAVAR

FIGURE 6. QFAVAR Diebold-Mariano Tests (tail-weighted CRPS)

Note: The QVAR model wins (loses) when it has a lower (higher) average score and the Diebold-Mariano statistic is significant at the 5% level. Columns are periods and rows are different benchmark models. Colors indicate FRED groups: **(1)** Output and income, **(2)** Labor market, **(3)** Housing, **(4)** Consumption, orders and inventories, **(5)** Money and credit, **(6)** Interest and exchange rate, **(7)** Prices and **(8)** Stock market.



A. QVAR and QFAVAR (PCA)



B. QVAR and QFAVAR (IQR)

FIGURE 7. QFAVAR Relative Scores (tail-weighted CRPS)

Note: Positive values are improvements over the QVAR model. Yellow corresponds to rejecting the null of equal scores at 5%. Yellow corresponds to not rejecting the null of equal scores at 5%. FRED groups are: **(1)** Output and income, **(2)** Labor market, **(3)** Housing, **(4)** Consumption, orders and inventories, **(5)** Money and credit, **(6)** Interest and exchange rate, **(7)** Prices and **(8)** Stock market.

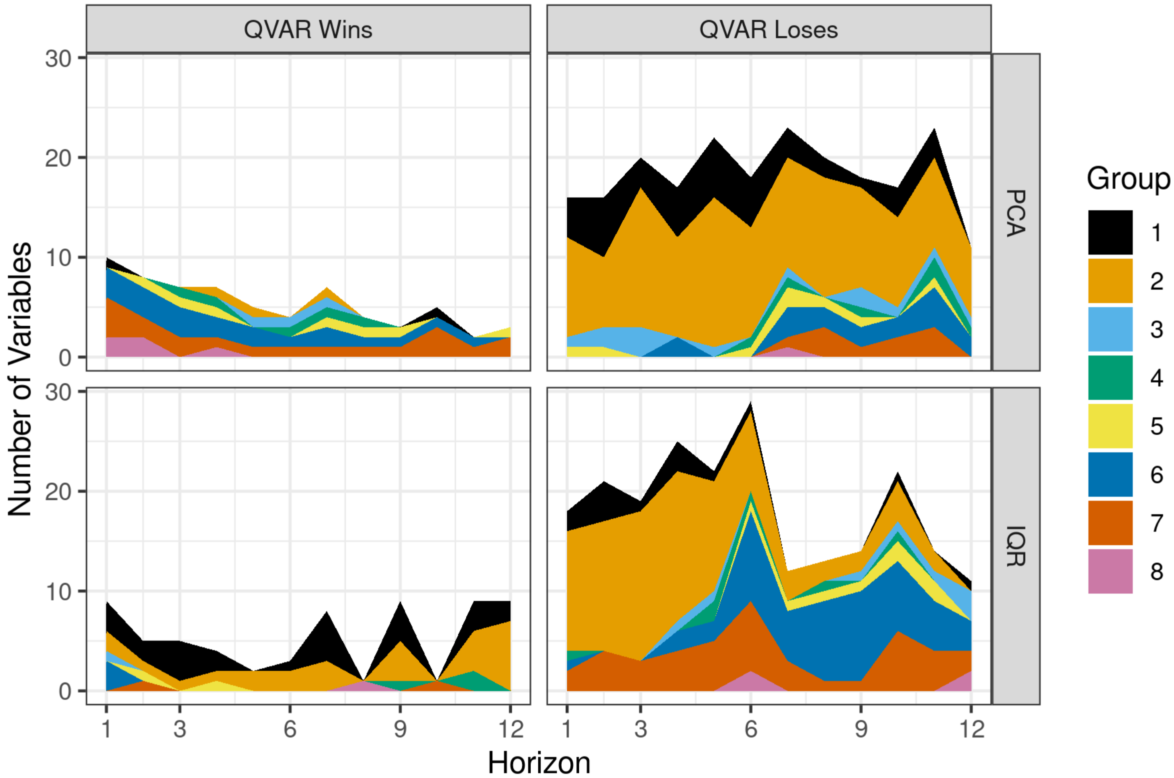
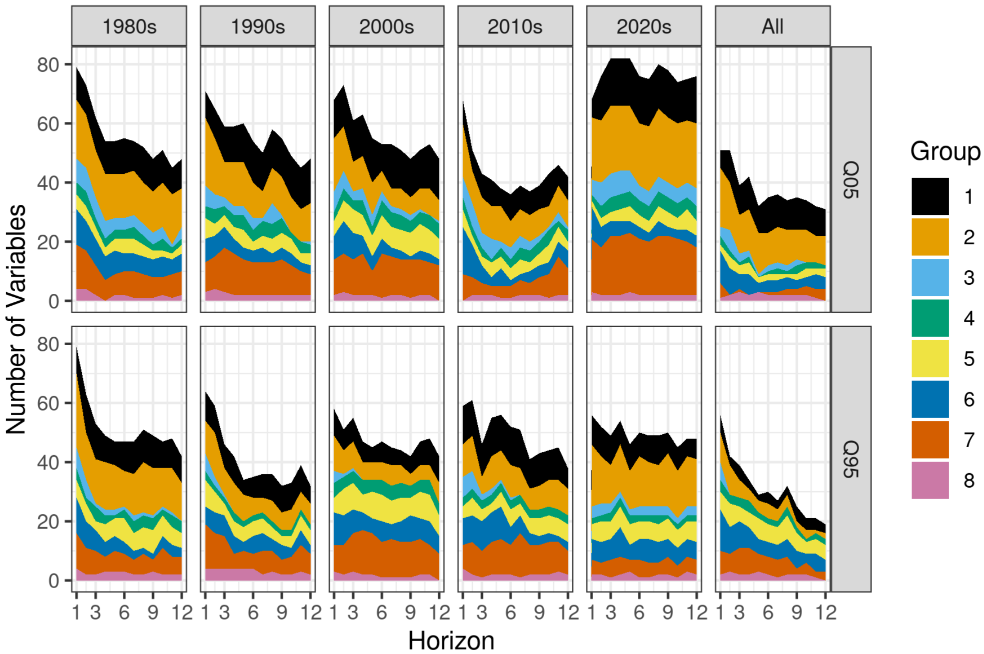


FIGURE 8. QVAR and QFAVAR Recesssion Comparison (tail-weighted CRPS)

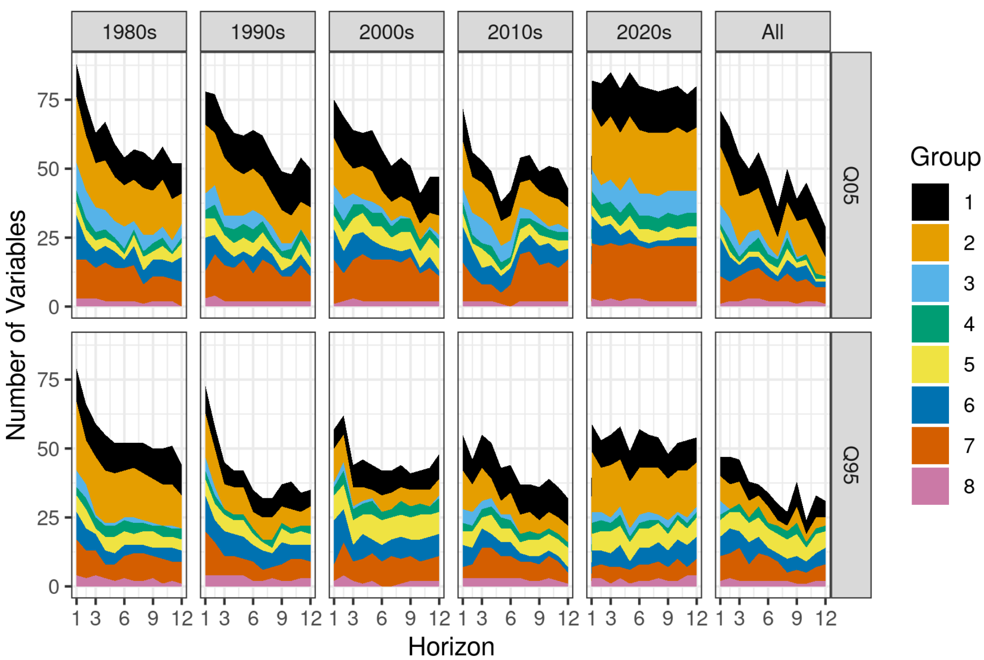
Note: The QVAR model wins (loses) when it has a lower (higher) average score and the Diebold-Mariano statistic is significant at the 5% level. Rows are different benchmark models. Colors indicate FRED groups: **(1)** Output and income, **(2)** Labor market, **(3)** Housing, **(4)** Consumption, orders and inventories, **(5)** Money and credit, **(6)** Interest and exchange rate, **(7)** Prices and **(8)** Stock market.

Introducing latent factor estimated by either PCA or IQR in the set of variables used by the QVAR model to produce forecasts has little effect on its tail forecasting accuracy. However, there remains the question of whether we can find evidence that QFAVAR models tend to be less frequently misspecified than the QVAR model. Figure 9 displays the number of cases where the null of a correctly specified quantile forecast at the 5th and 95th quantiles could not be rejected at the 5% level. We can see in the last column that using IQR factors rather than a PCA factor results in slightly fewer rejections at both quantiles over all horizons across the whole pseudo-out-of-sample period. Adding factors does not meaningfully impact conclusions we previously reached for the QVAR model, nor their stability over time, except that including either PCA and IQR factors slightly reduces the number of rejections.

Finally, Figure 5 presents the shares of non-rejection of the null hypothesis of correct unconditional and conditional coverage, respectively, for both QFAVAR models. On average across variable categories and time, introducing a PCA factor slightly increases the cases in which the model is found to have incorrect coverage. Moreover, it does not address the issue of excessive clustering of violations of the 5th and 95th conditional quantile bounds. Two notable exceptions are with interest and exchange rates (group 6) and the stock market (group 8) where both the issues with coverage and clustering are improved. The story is quite different when we introduce IQR factors. This QFAVAR model has more frequently correct unconditional and conditional coverage across time and variable categories. This provides some suggestive evidence that factors specifically targeting tail behavior in large data sets carry useful information that allows for improving the timing of changes in risk such that the model less frequently leads to serially correlated tail events (i.e., observations that lie in the tails of its forecasts). This would be worth exploring in future research.



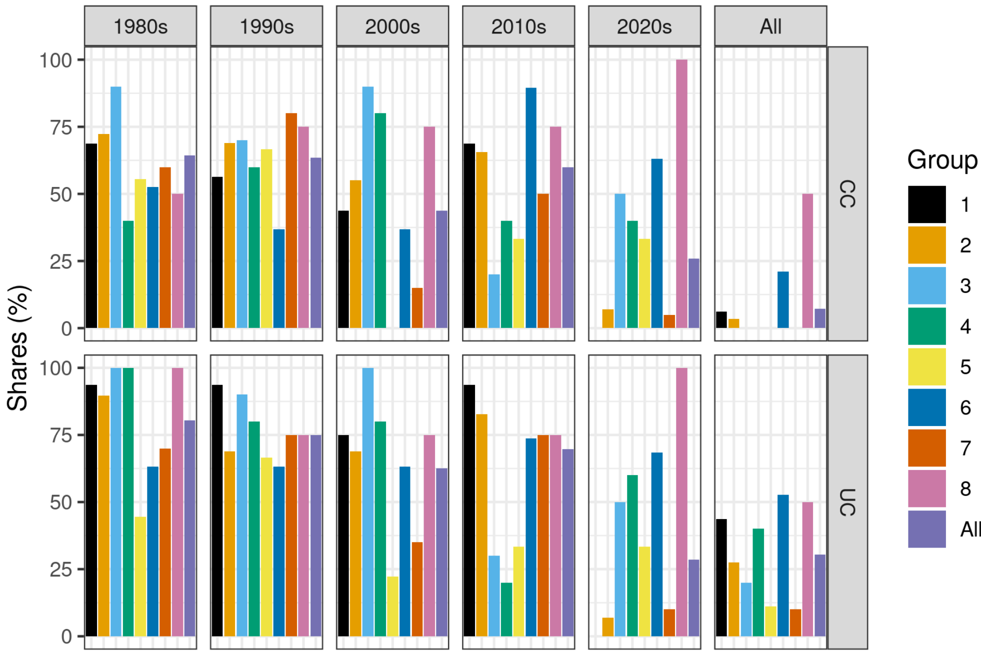
A. Number of Optimal QFAVAR (PCA) Forecasts



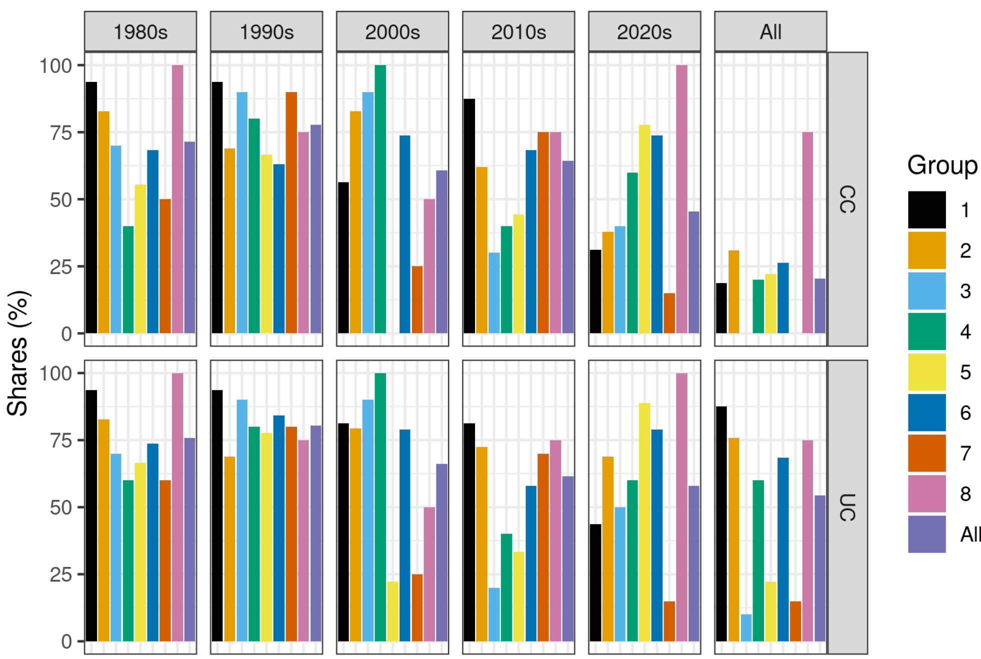
B. Number of Optimal QFAVAR (IQR) Forecasts

FIGURE 9. QFAVAR Relative Scores (tail-weighted CRPS)

Note: Number of cases where we obtain a non-rejection of the null of optimal forecast at 5% for the Gaglianone et al. (2011) test. Columns are periods and rows are the 5th (Q05) and 95th (Q95) quantile forecasts, respectively. Colors indicate FRED groups: **(1)** Output and income, **(2)** Labor market, **(3)** Housing, **(4)** Consumption, orders and inventories, **(5)** Money and credit, **(6)** Interest and exchange rate, **(7)** Prices and **(8)** Stock market.



A. PCA Factor



B. IQR Factors

FIGURE 10. Unconditional and Conditional Coverage Tests on QFAVARs (90% Interval)

Note: Shares of non-rejection of the null at 5% using Monte Carlo p-values (Dufour 2006). UC is the unconditional coverage test and CC is the conditional covargae test. Columns are periods. Colors indicate FRED groups: **(1)** Output and income, **(2)** Labor market, **(3)** Housing, **(4)** Consumption, orders and inventories, **(5)** Money and credit, **(6)** Interest and exchange rate, **(7)** Prices and **(8)** Stock market and **(All)** 112 variables.

5. Conclusion

In this paper, we evaluated the performance of the QVAR model for forecasting macroeconomic risk. To this end, we conducted a large out-of-sample forecasting experiment on US monthly variables using VAR-N, VAR-GARCH and VAR-SV models as parametric benchmarks. All models are specified as bivariate models featuring the target variables and the NFCI.

We find that the QVAR model provides significant improvements in tail-density forecasting accuracy over the VAR-N model in close to half of all variables considered. Those improvements are frequently quantitatively important, reaching 10% to 30% in many cases. The QVAR model also offers improvements over VAR-GARCH and VAR-SV models, albeit in fewer cases. For all three benchmark models, improvements are concentrated in the labor market as well as interest and exchange rate variables. However, we also find evidence that observations falling in the tails of QVAR forecasts tend to be serially correlated, which points to misspecification.

We then extend the QVAR model to a data-rich environment by introducing PCA and IQR factors as additional predictors. The resulting QFAVAR model significantly improves upon the QVAR model for forecasting macroeconomic risks in 13% of our target variables. Most of the improvements are tied to labor market variables. Interestingly, adding IQR factors also reduces the incidence of serial correlations with observations that fall in the tails of density forecasts. This suggests the specification issue with the QVAR model may be alleviated by adding information in the model and that IQR factors, in particular, carry information that helps improve the timing of predicted changes in macroeconomic risks.

In summary, we find that QVAR and QFAVAR models are adequate tools for modeling macroeconomic risk. This is relevant from a macroprudential risk management perspective, as in Chavleishvili et al. (2021), since the relevance of conclusions drawn from such studies requires reliable and accurate models of risk.

References

Acemoglu, Daron, and Andrew Scott. 1997. "Asymmetric business cycles: Theory and time-series evidence I." *Journal of Monetary Economics* 40: 501–533.

Adams, Patrick A., Tobias Adrian, Nina Boyarchenko, and Domenico Giannone. 2021. "Forecasting macroeconomic risks." *International Journal of Forecasting* 37: 1173–1191.

Adrian, Tobias, and Nina Boyarchenko. 2012. "Intermediary leverage cycles and financial stability." Working Paper 2012-010, Becker Friedman Institute for Research in Economics.

Adrian, Tobias, Nina Boyarchenko, and Domenico Giannone. 2019. "Vulnerable growth." *The American Economic Review* 109: 1263–1289.

Aiyagari, S. Rao, and Mark Gertler. 1999. ""Overreaction" of asset prices in general equilibrium." *Review of Economic Dynamics* 2: 3–35.

Angrist, Joshua, Victor Chernozhukov, and Iván Fernández-Val. 2006. "Quantile regression under misspecification, with an application to the U.S. wage structure." *Econometrica* 74: 539–563.

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 B* 65: 367–389.

Benigno, Pierpaolo, and Gauti B Eggertsson. 2023. "It's Baaack: The Surge in Inflation in the 2020s and the Return of the Non-Linear Phillips Curve." Working Paper 31197, National Bureau of Economic Research.

Bernanke, Ben S. 1983. "Irreversibility, uncertainty, and cyclical investment." *Source: The Quarterly Journal of Economics* 98: 85–106.

Bernanke, Ben S., Jean Boivin, and Piotr Eliasz. 2005. "Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach." *Quarterly Journal of Economics* 120: 387–422.

Boivin, Jean, Marc P. Giannoni, and Dalibor Stevanović. 2020. "Dynamic effects of credit shocks in a data-rich environment." *Journal of Business and Economic Statistics* 38: 272–284.

Boivin, Jean, and Serena Ng. 2005. "Understanding and Comparing Factor-Based Forecasts." Working Paper 11285, National Bureau of Economic Research.

Bouakez, Hafedh, Foued Chihi, and Michel Normandin. 2014. "Measuring the effects of fiscal policy." *Journal of Economic Dynamics and Control* 47: 123–151.

Bouakez, Hafedh, and Michel Normandin. 2010. "Fluctuations in the foreign exchange market: How important are monetary policy shocks?" *Journal of International Economics* 81: 139–153.

Brownlees, Christian, and André B.M. Souza. 2021. "Backtesting global growth-at-risk." *Journal of Monetary Economics* 118: 312–330.

Brunnermeier, Markus K., and Yuliy Sannikov. 2014. "A macroeconomic model with a financial sector." *American Economic Review* 104: 379–421.

Caldara, Dario, Danilo Cascaldi-Garcia, Pablo Cuba-Borda, and Francesca Loria. 2021. "Understanding growth-at-risk: A markov-switching approach."

Carriero, Andrea, Todd E. Clark, and Massimiliano Marcellino. 2021. "Using time-varying volatility for identification in vector autoregressions: An application to endogenous uncertainty." *Journal of Econometrics* 225: 47–73.

Carriero, Andrea, Todd E. Clark, and Massimiliano Marcellino. 2022. "Specification choices in quantile regression for empirical macroeconomics." Working Paper WP 22-25, Federal Reserve Bank of Cleveland.

Carriero, Andrea, Todd E Clark, and Massimiliano Marcellino. 2024. "Capturing macroeconomic tail risks with Bayesian vector autoregressions." *Journal of Money, Credit and Banking* 56 (5): 1099–1127.

Chan, Joshua C. C., and Cody Y. L. Hsiao. 2014. *Estimation of stochastic volatility models with heavy tails and serial dependence*. 155–176: John Wiley and Sons Inc..

Chan, Joshua C.C., and Eric Eisenstat. 2018. "Bayesian model comparison for time-varying parameter VARs with stochastic volatility." *Journal of Applied Econometrics* 33: 509–532.

Chavleishvili, Sulkhan., Robert F. Engle, Stephan. Fahr, Manfred. Kremer, Simone. Manganelli, and Bernd. Schwaab. 2021. "The risk management approach to macro-prudential policy." Working Paper 2565, European Central Bank.

Chavleishvili, Sulkhan, and Simone Manganelli. 2021. "Forecasting and stress testing with quantile vector autoregression." Working Paper 2330, European Central Bank.

Chen, Liang, Juan J. Dolado, and Jesús Gonzalo. 2021. "Quantile factor models." *Econometrica* 89: 875–910.

Chernozhukov, Victor, Iván Iván Fernández-Val, and Tetsuya Kaji. 2017. *Extremal quantile regression*. 1st ed. 333–362: CRC Press.

Cho, Jin Seo, Tae Hwan Kim, and Yongcheol Shin. 2015. "Quantile cointegration in the autoregressive distributed-lag modeling framework." *Journal of Econometrics* 188: 281–300.

Christoffersen, Peter F. 1998. "Evaluating interval forecasts." *International Economic Review* 39: 841–862.

Christoffersen, Peter F. 2004. "Backtesting value-at-risk: A duration-based approach." *Journal of Financial Econometrics* 2: 84–108.

Christoffersen, Peter, Jinyong Hahn, and Atsushi Inoue. 2001. "Testing and comparing value-at-risk measures." *Journal of Empirical Finance* 8: 325–342.

Cogley, Timothy, and Thomas J. Sargent. 2005. "Drifts and volatilities: Monetary policies and outcomes in the post WWII US." *Review of Economic Dynamics* 8: 262–302.

Diebold, Francis X., and Roberto S. Mariano. 1995. "Comparing predictive accuracy." *Journal of Business and Economic Statistics* 13: 253–263.

Dufour, Jean Marie. 2006. "Monte Carlo tests with nuisance parameters: A general approach to finite-sample inference and nonstandard asymptotics." *Journal of Econometrics* 133: 443–477.

Elder, John Robert, and Apostolos Serletis. 2010. "Oil price uncertainty." *Journal of Money, Credit and Banking* 42: 1137–1159.

Engle, Robert F., and Simone Manganelli. 2004. "CAViaR: Conditional autoregressive value at risk by regression quantiles." *Journal of Business and Economic Statistics* 22: 367–381.

Figueres, Juan Manuel, and Marek Jarociński. 2020. "Vulnerable growth in the euro area: Measuring the financial conditions." *Economics Letters* 191.

Forni, Mario, Luca Gambetti, and Luca Sala. 2024. "Downside and upside uncertainty shocks." *Journal of the European Economic Association*: jvae024.

Gaglianone, Wagner Piazza, Luiz Renato Lima, Oliver Linton, and Daniel R. Smith. 2011. "Evalu-

ating value-at-risk models via quantile regression." *Journal of Business and Economic Statistics* 29: 150–160.
Giglio, Stefano, Bryan Kelly, and Seth Pruitt. 2016. "Systemic risk and the macroeconomy: An empirical evaluation." *Journal of Financial Economics* 119: 457–471.
Gneiting, Tilmann, and Adrian E. Raftery. 2007. "Strictly proper scoring rules, prediction, and estimation." *Journal of the American Statistical Association* 102: 359–378.
Gneiting, Tilmann, and Roopesh Ranjan. 2011. "Comparing density forecasts using thresholdand quantile-weighted scoring rules." *Journal of Business and Economic Statistics* 29: 411–422.
Goulet Coulombe, Philippe, Maxime Leroux, Dalibor Stevanovic, and Stéphane Surprenant. 2021. "Macroeconomic data transformations matter." *International Journal of Forecasting* 37: 1338–1354.
Goulet Coulombe, Philippe, Maxime Leroux, Dalibor Stevanovic, and Stéphane Surprenant. 2022. "How is machine learning useful for macroeconomic forecasting?" *Journal of Applied Econometrics* 37: 920–964.
Hallin, Marc, and Bas J.M. Werker. 2006. "Comment." *Journal of the American Statistical Association* 101: 996–998.
Iseringhausen, Martin. 2024. "A time-varying skewness model for growth-at-risk." *International Journal of Forecasting* 40 (1): 229–246.
Kim, Sangjoon, Neil Shephard, and Siddhartha Chib. 1998. "Stochastic volatility : Likelihood inference and comparison with ARCH Models." *Review of Economic Studies* 65: 361–393.
Kipriyanov, Aleksei. 2022. "Comparison of models for growth-at-risk forecasting." *Russian Journal of Money and Finance* 81: 23–45.
Koenker, Roger. 2004. "Quantile regression for longitudinal data." *Journal of Multivariate Analysis* 91: 74–89.
Koenker, Roger, and Gilbert Bassett. 1978. "Regression quantiles." *Econometrica* 46: 33.
Koenker, Roger, and José A.F. Machado. 1999. "Goodness of Fit and Related Inference Processes for Quantile Regression." *Journal of the American Statistical Association* 94: 1296–1310.
Koenker, Roger, and Zhijie Xiao. 2006. "Rejoinder." *Journal of the American Statistical Association* 101: 1002–1006.
Kupiec, Paul H. 1995. "Techniques for verifying the accuracy of risk measurement models." *Journal of Derivatives* 3: 73–84.
Loria, Francesca, Christian Matthes, and Donghai Zhang. 2024. "Assessing Macroeconomic Tail Risk." *The Economic Journal*: 1–21.
López-Salido, David, and Francesca Loria. 2020. "Inflation at Risk." Working Paper 2020-013, Board of Governors of the Federal Reserve System.
Lütkepohl, Helmut. 2005. *A new introduction to multiple time series.*: 764: Springer-Verlag.
Manzan, Sebastiano. 2015. "Forecasting the Distribution of Economic Variables in a Data-Rich Environment." *Journal of Business and Economic Statistics* 33: 144–164.
Manzan, Sebastiano, and Dawit Zerom. 2013. "Are macroeconomic variables useful for forecasting the distribution of U.S. inflation?" *International Journal of Forecasting* 29: 469–478.
McCracken, Michael W., and Serena Ng. 2016. "FRED-MD: A Monthly Database for Macroeconomic Research." *Journal of Business and Economic Statistics* 34: 574–589.

McDonald, Robert, and Daniel Siegel. 1986. "The value of waiting to invest." *The Quarterly Journal of Economics* 101: 707–728.
Mincer, Jacob A., and Victor Zarnowitz. 1969. *The evaluation of economic forecasts and expectations.* 3–46: NBER.
Montes-Rojas, Gabriel. 2021. "Estimating impulse-response functions for macroeconomic models using directional quantiles." *Journal of Time Series Econometrics*: 1–27.
Normandin, Michel, and Louis Phaneuf. 2004. "Monetary policy shocks: Testing identification conditions under time-varying conditional volatility." *Journal of Monetary Economics* 51: 1217–1243.
Plagborg-Moller, Mikkel, Lucrezia Reichlin, Giovanni Ricco, and Thomas Hasenzagl. 2020. "When is growth at risk?" *Brookings Papers on Economic Activity*: 167–213.
Primiceri, Giorgio E. 2005. "Time varying structural vector autoregressions and monetary policy." *Review of Economic Studies* 72: 821–852.
Ruzicka, Josef. 2021. "Dynamic quantile causal inference and forecasting." PhD dissertation, Universidad Carlos III de Madrid.
Stock, James H., and Mark W. Watson. 2002a. "Forecasting using principal components from a large number of predictors." *Journal of the American Statistical Association* 97: 1167–1179.
Stock, James H., and Mark W. Watson. 2002b. "Macroeconomic forecasting using diffusion indexes." *Journal of Business and Economic Statistics* 20: 147–162.
Stock, James, and Mark Watson. 1998. "A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series." Working Paper 6607, National Bureau of Economic Research.
White, Halbert, Tae-Hwan Kim, and Simone Manganelli. 2015. "VAR for VaR: Measuring tail dependence using multivariate regression quantiles." *Journal of Econometrics* 187: 169–188.
Xiao, Zhijie. 2009. "Quantile cointegrating regression." *Journal of Econometrics* 150: 248–260.

Appendix A. Monte Carlo Tests

The idea behind the Dufour (2006) Monte Carlo testing framework is that exact finite sample p-values can be obtained, even when an analytic formula for the distribution of a test statistic under the null is unavailable, provided we can simulate it. In our case, we draw N samples of T observations of $I_t \sim iid\ B(0.9)$ and compute likelihood ratio statistics LR_i for $i = 1, \ldots, N$. If LR_0 is the corresponding statistic we computed on the actual data, then the p-value is given by

$$\hat{p}_n(LR_0) = \frac{N\hat{G}_N(LR_0) + 1}{N + 1}$$

where $\hat{G}_N(LR_0) = \sum_{i=1}^{N} \mathbb{I}\left\{LR_i > LR_0\right\}/N$. As noted by Christoffersen (2004), the distribution is discrete such that ties can happen and need to be handled. They propose breaking ties by drawing $N+1$ uniform random variables $U_i \sim U[0,1]$ where $i = 0, \ldots, N$ and using

$$\hat{G}_N(LR_0) = 1 - \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}\left\{LR_i \leq LR_0\right\} + \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}\left\{LR_i = LR_0\right\} \mathbb{1}\left\{U_i \geq U_0\right\}$$

in the above formula.

Appendix B. Additional Results

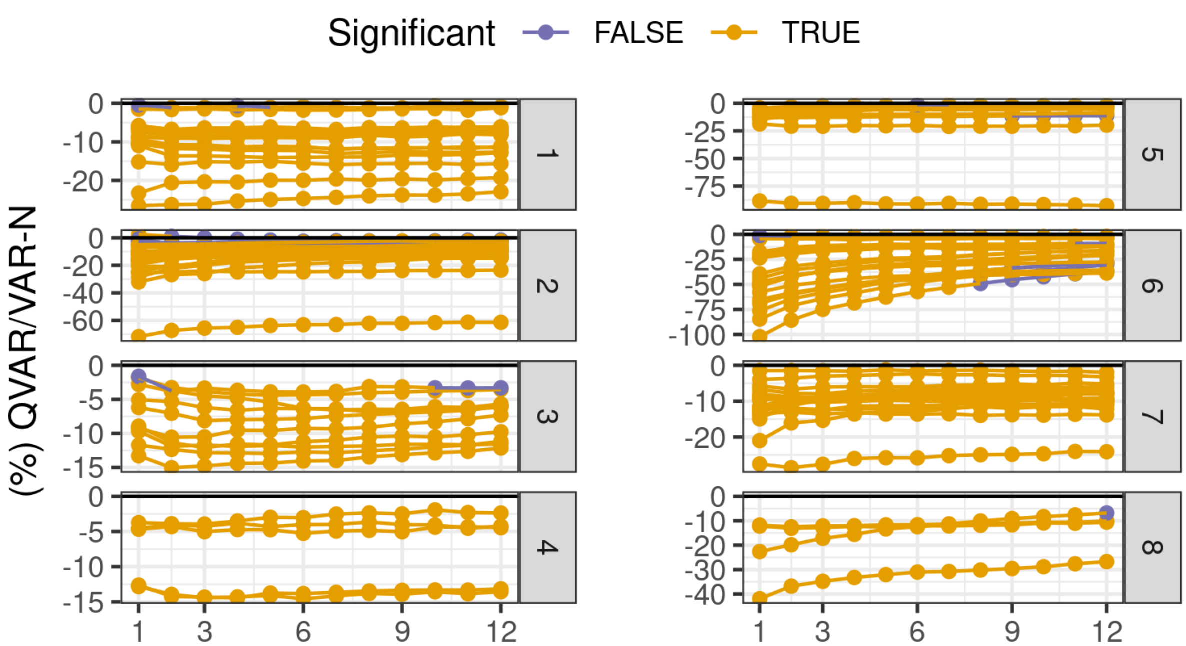
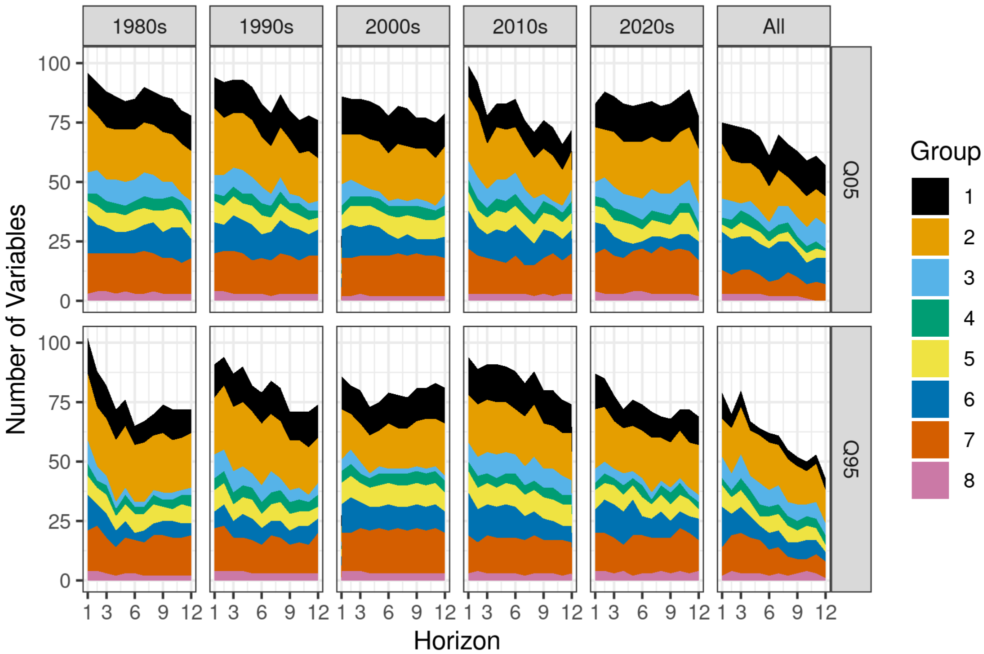
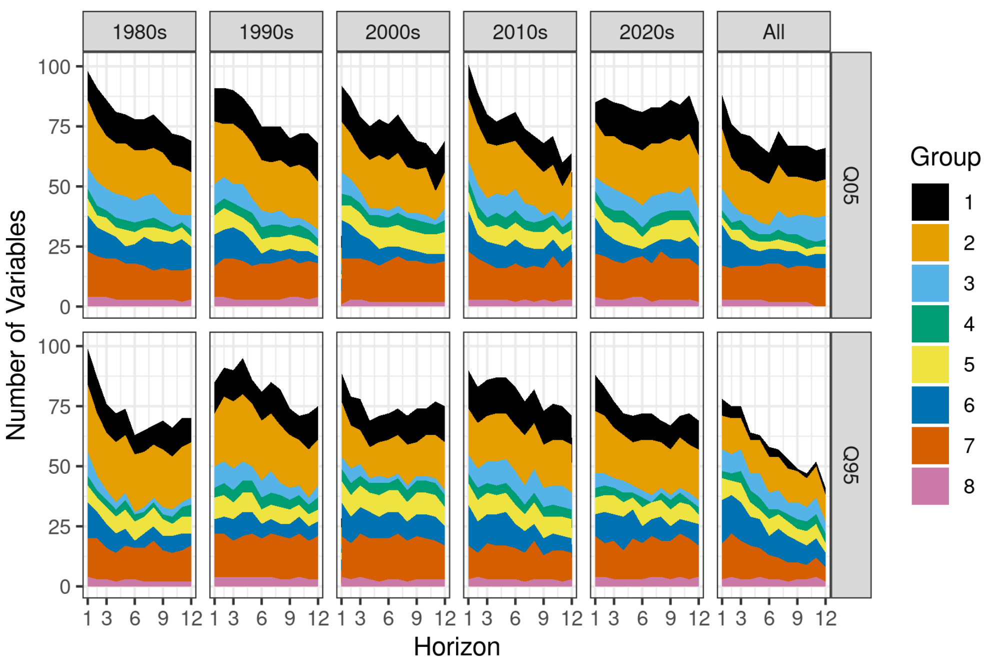


FIGURE A1. QVAR and VAR-N Relative Scores (tail-weighted CRPS)

Note: Negative values are improvements. Yellow corresponds to rejecting the null of equal scores at 5%. Yellow corresponds to not rejecting the null of equal scores at 5%. FRED groups are: **(1)** Output and income, **(2)** Labor market, **(3)** Housing, **(4)** Consumption, orders and inventories, **(5)** Money and credit, **(6)** Interest and exchange rate, **(7)** Prices and **(8)** Stock market.



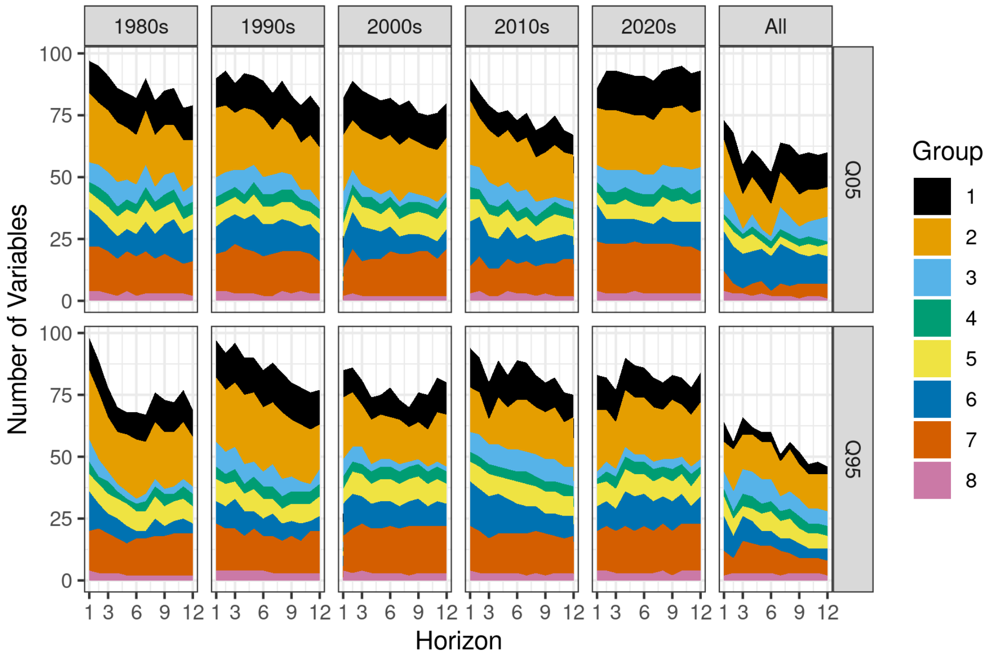
A. Null Constant



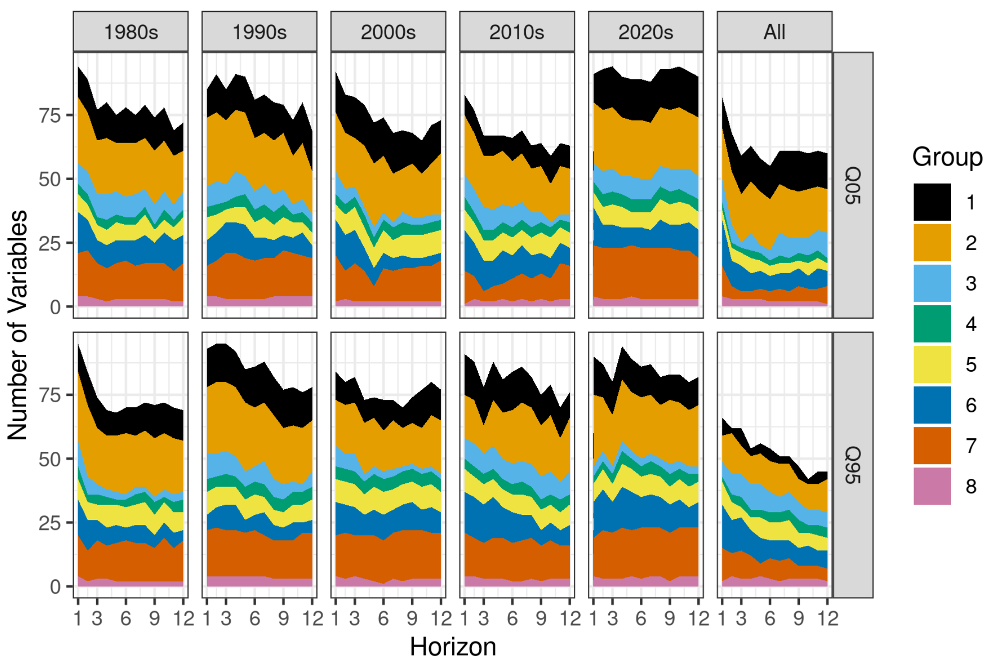
B. Unit Slope

FIGURE A2. Number of Optimal QVAR Forecasts

Note: Number of cases where we obtain a non-rejection of the null at 5% for the Gaglianone et al. (2011) test. Columns are periods and rows are the 5th (Q05) and 95th (Q95) quantile forecasts, respectively. Colors indicate FRED groups: **(1)** Output and income, **(2)** Labor market, **(3)** Housing, **(4)** Consumption, orders and inventories, **(5)** Money and credit, **(6)** Interest and exchange rate, **(7)** Prices and **(8)** Stock market.



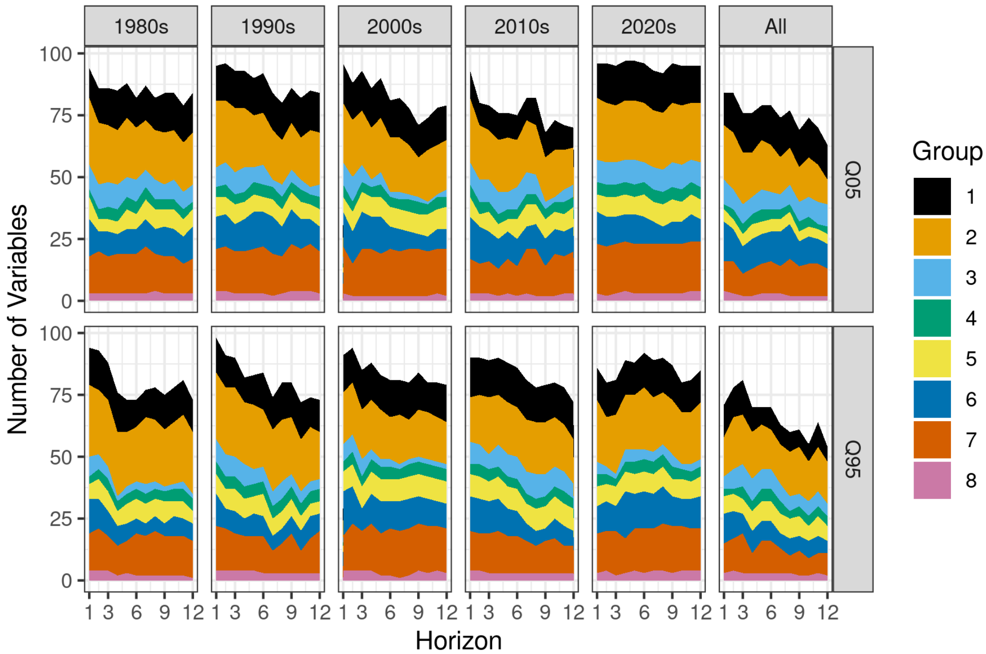
A. Null Constant



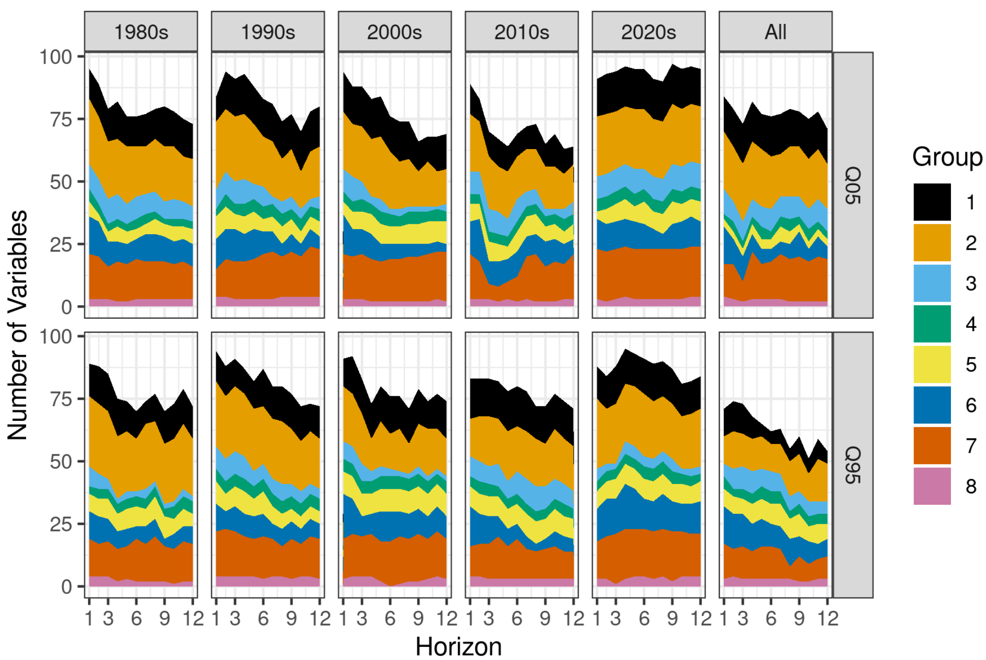
B. Unit Slope

FIGURE A3. Number of Optimal QFAVAR (PCA) Forecasts

Note: Number of cases where we obtain a non-rejection of the null at 5% for the Gaglianone et al. (2011) test. Columns are periods and rows are the 5th (Q05) and 95th (Q95) quantile forecasts, respectively. Colors indicate FRED groups: **(1)** Output and income, **(2)** Labor market, **(3)** Housing, **(4)** Consumption, orders and inventories, **(5)** Money and credit, **(6)** Interest and exchange rate, **(7)** Prices and **(8)** Stock market.



A. Null Constant



B. Unit Slope

FIGURE A4. Number of Optimal QFAVAR (IQR) Forecasts

Note: Number of cases where we obtain a non-rejection of the null at 5% for the Gaglianone et al. (2011) test. Columns are periods and rows are the 5th (Q05) and 95th (Q95) quantile forecasts, respectively. Colors indicate FRED groups: **(1)** Output and income, **(2)** Labor market, **(3)** Housing, **(4)** Consumption, orders and inventories, **(5)** Money and credit, **(6)** Interest and exchange rate, **(7)** Prices and **(8)** Stock market.

Appendix C. Data Transformation

As in the reference documentation of FRED-MD, the transformation codes are (1) y_t, (2) Δy_t, (3) $\Delta^2 y_t$, (4) $ln y_t$, (5) $\Delta ln y_t$, (6) $\Delta^2 ln y_t$ and (7) $y_t / y_{t-1} - 1$.

TABLE A1. Data Transformation

ID	Description	Used	FRED
RPI	Real Personal Income	5	5
W875RX1	Real personal income ex transfer receipts	5	5
INDPRO	IP Index	5	5
IPFPNSS	IP: Final Products and Nonindustrial Supplies	5	5
IPFINAL	IP: Final Products (Market Group)	5	5
IPCONGD	IP: Consumer Goods	5	5
IPDCONGD	IP: Durable Consumer Goods	5	5
IPNCONGD	IP: Nondurable Consumer Goods	5	5
IPBUSEQ	IP: Business Equipment	5	5
IPMAT	IP: Materials	5	5
IPDMAT	IP: Durable Materials	5	5
IPNMAT	IP: Nondurable Materials	5	5
IPMANSICS	IP: Manufacturing (SIC)	5	5
IPB51222s	IP: Residential Utilities	5	5
IPFUELS	IP: Fuels	5	5
CUMFNS	Capacity Utilization: Manufacturing	1	2
HWI	Help-Wanted Index for United States	5	2
HWIURATIO	Ratio of Help Wanted/No. Unemployed	4	2
CLF16OV	Civilian Labor Force	5	5
CE16OV	Civilian Employment	5	5
UNRATE	Civilian Unemployment Rate	1	2
UEMPMEAN	Average Duration of Unemployment (Weeks)	1	2
UEMPLT5	Civilians Unemployed - Less Than 5 Weeks	1	5
UEMP5TO14	Civilians Unemployed for 5-14 Weeks	1	5
UEMP15OV	Civilians Unemployed - 15 Weeks & Over	1	5
UEMP15T26	Civilians Unemployed for 15-26 Weeks	1	5
UEMP27OV	Civilians Unemployed for 27 Weeks and Over	1	5
CLAIMSx	Initial Claims	5	5
PAYEMS	All Employees: Total nonfarm	5	5
USGOOD	All Employees: Goods-Producing Industries	5	5

Table A1. Data Transformation (Continued)

ID	Description	Used	FRED
CES1021000001	All Employees: Mining and Logging: Mining	5	5
USCONS	All Employees: Construction	5	5
MANEMP	All Employees: Manufacturing	5	5
DMANEMP	All Employees: Durable goods	5	5
NDMANEMP	All Employees: Nondurable goods	5	5
SRVPRD	All Employees: Service-Providing Industries	5	5
USTPU	All Employees: Trade, Transportation & Utilities	5	5
USWTRADE	All Employees: Wholesale Trade	5	5
USTRADE	All Employees: Retail Trade	5	5
USFIRE	All Employees: Financial Activities	5	5
USGOVT	All Employees: Government	5	5
CES0600000007	Avg Weekly Hours : Goods-Producing	1	1
AWOTMAN	Avg Weekly Overtime Hours : Manufacturing	1	2
AWHMAN	Avg Weekly Hours : Manufacturing	1	1
CES0600000008	Avg Hourly Earnings : Goods-Producing	5	6
CES2000000008	Avg Hourly Earnings : Construction	5	6
CES3000000008	Avg Hourly Earnings : Manufacturing	5	6
HOUST	Housing Starts: Total New Privately Owned	4	4
HOUSTNE	Housing Starts, Northeast	4	4
HOUSTMW	Housing Starts, Midwest	4	4
HOUSTS	Housing Starts, South	4	4
HOUSTW	Housing Starts, West	4	4
PERMIT	New Private Housing Permits (SAAR)	4	4
PERMITNE	New Private Housing Permits, Northeast (SAAR)	4	4
PERMITMW	New Private Housing Permits, Midwest (SAAR)	4	4
PERMITS	New Private Housing Permits, South (SAAR)	4	4
PERMITW	New Private Housing Permits, West (SAAR)	4	4
DPCERA3M086SBEA	Real personal consumption expenditures	5	5
CMRMTSPLx	Real Manu. and Trade Industries Sales	5	5
RETAILx	Retail and Food Services Sales	5	5
ACOGNO	New Orders for Consumer Goods	5	5
AMDMNOx	New Orders for Durable Goods	5	5
ANDENOx	New Orders for Nondefense Capital Goods	5	5
AMDMUOx	Unfilled Orders for Durable Goods	5	5
BUSINVx	Total Business Inventories	5	5

TABLE A1. Data Transformation (Continued)

ID	Description	Used	FRED
ISRATIOx	Total Business: Inventories to Sales Ratio	2	2
UMCSENTx	Consumer Sentiment Index	2	2
M1SL	M1 Money Stock	5	6
M2SL	M2 Money Stock	5	6
M2REAL	Real M2 Money Stock	5	5
BOGMBASE	Monetary Base	5	6
TOTRESNS	Total Reserves of Depository Institutions	5	6
NONBORRES	Reserves Of Depository Institutions	7	7
BUSLOANS	Commercial and Industrial Loans	5	6
REALLN	Real Estate Loans at All Commercial Banks	5	6
NONREVSL	Total Nonrevolving Credit	5	6
CONSPI	Nonrevolving consumer credit to Personal Income	5	2
DTCOLNVHFNM	Consumer Motor Vehicle Loans Outstanding	5	6
DTCTHFNM	Total Consumer Loans and Leases Outstanding	5	6
INVEST	Securities in Bank Credit at All Commercial Banks	5	6
FEDFUNDS	Effective Federal Funds Rate	1	2
CP3Mx	3-Month AA Financial Commercial Paper Rate	1	2
TB3MS	3-Month Treasury Bill:	1	2
TB6MS	6-Month Treasury Bill:	1	2
GS1	1-Year Treasury Rate	1	2
GS5	5-Year Treasury Rate	1	2
GS10	10-Year Treasury Rate	1	2
AAA	Moody's Seasoned Aaa Corporate Bond Yield	1	2
BAA	Moody's Seasoned Baa Corporate Bond Yield	1	2
COMPAPFFx	3-Month Commercial Paper Minus FEDFUNDS	1	1
TB3SMFFM	3-Month Treasury C Minus FEDFUNDS	1	1
TB6SMFFM	6-Month Treasury C Minus FEDFUNDS	1	1
T1YFFM	1-Year Treasury C Minus FEDFUNDS	1	1
T5YFFM	5-Year Treasury C Minus FEDFUNDS	1	1
T10YFFM	10-Year Treasury C Minus FEDFUNDS	1	1
AAAFFM	Moody's Aaa Corporate Bond Minus FEDFUNDS	1	1
BAAFFM	Moody's Baa Corporate Bond Minus FEDFUNDS	1	1
TWEXAFEGSMTHx	Trade Weighted U.S. Dollar Index	5	5
EXSZUSx	Switzerland / U.S. Foreign Exchange Rate	5	5
EXJPUSx	Japan / U.S. Foreign Exchange Rate	5	5

Table A1. Data Transformation (Continued)

ID	Description	Used	FRED
EXUSUKx	U.S. / U.K. Foreign Exchange Rate	5	5
EXCAUSx	Canada / U.S. Foreign Exchange Rate	5	5
WPSFD49207	PPI: Finished Goods	5	6
WPSFD49502	PPI: Finished Consumer Goods	5	6
WPSID61	PPI: Intermediate Materials	5	6
WPSID62	PPI: Crude Materials	5	6
OILPRICEx	Crude Oil, spliced WTI and Cushing	5	6
PPICMM	PPI: Metals and metal products:	5	6
CPIAUCSL	CPI : All Items	5	6
CPIAPPSL	CPI : Apparel	5	6
CPITRNSL	CPI : Transportation	5	6
CPIMEDSL	CPI : Medical Care	5	6
CUSR0000SAC	CPI : Commodities	5	6
CUSR0000SAD	CPI : Durables	5	6
CUSR0000SAS	CPI : Services	5	6
CPIULFSL	CPI : All Items Less Food	5	6
CUSR0000SA0L2	CPI : All items less shelter	5	6
CUSR0000SA0L5	CPI : All items less medical care	5	6
PCEPI	Personal Cons. Expend.: Chain Index	5	6
DDURRG3M086SBEA	Personal Cons. Exp: Durable goods	5	6
DNDGRG3M086SBEA	Personal Cons. Exp: Nondurable goods	5	6
DSERRG3M086SBEA	Personal Cons. Exp: Services	5	6
S&P 500	S&P's Common Stock Price Index: Composite	5	5
S&P: indust	S&P's Common Stock Price Index: Industrials	5	5
S&P div yield	S&P's Composite Common Stock: Dividend Yield	1	2
S&P PE ratio	S&P's Composite Common Stock: Price-Earnings Ratio	1	5
VIXCLSx	VIX	1	1