Quantile-Covariance Three-Pass Regression Filter

Pedro Isaac Chavez-Lopez* Tae-Hwy Lee†

January 12, 2025

Abstract

We propose a factor model for quantile regression using *quantile-covariance*(*qcov*), which will be called the Quantile-Covariance Three-Pass Regression Filter (Qcov3PRF). Inspired by the Three-Pass Regression Filter (3PRF, Kelly and Pruitt, 2015), our method selects the relevant factors from a large set of predictors to forecast the conditional quantile of a target variable. The measure *qcov* is implied by the first order condition from a univariate linear quantile regression. Our approach differs from the Partial Quantile Regression (PQR, Giglio et al., 2016) as Qcov3PRF successfully allows the estimation of more than one relevant factor using *qcov*. In particular, *qcov* permits us to run time series least squares regressions of each regressor on a set of transformations of the variables, indexed for a specific quantile of the forecast target, known as proxies that only depend on the relevant factors. This is not possible to be executed using quantile regressions as regressing each predictor on the proxies refers to the conditional quantile of the predictor and not the quantile corresponding to the target. As a consequence of running a quantile regression of the target or proxy on each predictor, only one factor is recovered with PQR. By capturing the correct number of the relevant factors, the Qcov3PRF forecasts are consistent and asymptotically normal when both time and cross sectional dimensions become large. Our simulations show that Qcov3PRF exhibits good finite sample properties compared to alternative methods. Finally, three applications are presented: forecasting the Industrial Production Growth, forecasting the Real GDP growth, and forecasting the global temperature change index.

Keywords: *Quantile regression, factor models, quantile-covariance.*

*Economics Research Division, Bank of Mexico, pedro.chavez@banxico.org.mx
†Department of Economics, University of California-Riverside, taelee@ucr.edu

1 Introduction

In economics and finance, forecasting a time series variable through factor models has been widely used as a method of dimension reduction under the presence of a very large set of predictors, which may be highly correlated. Some relevant works include Stock and Watson, 2002a, Stock and Watson, 2002b, Bai and Ng, 2002, Bai, 2003, and Bai and Ng, 2006, among many others. However, this has been usually applied in a setup where in the first step one decomposes a set of N covariates into K orthogonal factors using Principal Components Analysis (PCA); and then, the researcher runs a linear regression of the target variable on the set of the K factor estimates. This two steps conform the method Principal Components Regression (PCR, Stock and Watson, 2002a), and although the forecasts resulting from using PCR are shown to be consistent, for finite samples the estimates can bring poor and/or inefficient results as the target variable may depend only on a subset of factors out of those factors selected in the first step. For this reason, it is important to consider a method where we can determine only the relevant factors before the forecasting step. This can be done by considering the forecast target in steps of computing the factors, for example, the target-PCA of Bai and Ng, 2008, the Three Pass Regression Filter (3PRF) of Kelly and Pruitt, 2015, and the scaled-PCA of Huang et al., 2022, among others.

Specifically, in 3PRF one can obtain the relevant factors to forecast the target variable for its conditional mean. Interestingly, Kelly and Pruitt, 2015 showed that Partial Least Squares (PLS) (Wold, 1966) is a particular case of 3PRF. Applications found in Kelly and Pruitt, 2013, Huang et al., 2015, Lyle and Wang, 2015, Light et al., 2017, Gu et al., 2020, and Huang et al., 2021, among others, showed that 3PRF (or PLS) is successful at extracting factors for predicting stock returns and economic activities in time series and cross-sectional data. On the other hand, methods involving factor models regarding the prediction of other functionals are much more scarce. In particular, for the conditional quantile prediction of the target we have the Principal Components Quantile Regression (PCQR) and Partial Quantile Regression (PQR), both studied in Giglio et al., 2016.

The method 3PRF assumes the researcher has K_f proxy variables that contain information of the relevant factors such that, through simple cross sectional and time series linear regressions, one can obtain the K_f relevant factors. PQR shares some similarities with 3PRF when a proxy is the target itself, however, it only obtains one relevant factor for each quantile by construction. The reason is that extending 3PRF to the PQR approach would rely on running time series quantile regressions of each predictor on the proxies, but this refers to the conditional quantile of the predictor and not the one corresponding to the target/proxies. Then, as a consequence of running a quantile regression of the target or proxy on each predictor, only one factor is recovered with PQR.

Motivated by this limitation, we propose a new supervised method to forecast the conditional quantile of a target variable that depends on the latent factors obtained from a high dimension set of covariates. We call the new method the Quantile Three Pass Regression Filter (Qcov3PRF). This method is similar in nature to PLS and 3PRF as it exploits the covariance between the target variable and the predictors, but we instead incorporate the *quantile-covariance* (*qcov*) definition of Li et al., 2015. The measure *qcov* is implied by the first order condition from a linear univariate quantile regression. In contrast to PQR, our method can successfully estimate more than one relevant factor and it allows the inclusion of proxies which, by assumption, only depend on the factors that forecast the conditional quantile of the target variable. Our algorithm is easy to implement and only requires least square regressions to extract the relevant factors. Illustrating that our approach can be extended in various ways, similar to 3PRF, we also propose a simple modification of Qcov3PRF with mixed frequency data, low frequency in the proxies and the target, and high frequency in the predictors. This extension of Qcov3PRF follows the approach of Hepenstrick and Marcellino, 2019 and is an example of other potential extensions of Qcov3PRF similar to existent extensions of 3PRF (for example, the Markov-Switching 3PRF proposed by Guérin et al., 2020).

We are not the first to incorporate a *quantile-covariance* concept in a conditional quantile forecast factor model focusing on the estimation of the factors which are relevant for the target. Dodge and Whittaker, 2009 and Méndez-Civieta et al., 2022 proposed PLS modifications for quantile regression based on a different definition of quantile-covariance and *qcov*, respectively. However, there are some shortcomings compared to Qcov3PRF. The method of Dodge and Whittaker, 2009 provides no background on what is the optimization problem that their PQR algorithm is solving, the quantile-covariance they use does not have properties that theoretically justifiy the estimation of more than one factor, and its implementation can be computationally expensive. Regarding the method of Méndez-Civieta et al., 2022, the authors do not provide asymptotic results, their approach does not allow the inclusion of proxies, and their empirical applications and simulations focus on the median.

In this paper, asymptotic properties for the estimated factors and quantile forecasts obtained through Qcov3PRF are studied. We show the consistency of the estimated forecasts towards the infeasible best forecasts, i.e., the population conditional quantile forecasts. We also show the asymptotic normality of the forecasts. In particular, the consistency and asymptotic normality hold under the use of *automatic proxies*, which are proxies that can always be generated with the target and the set of predictors only. In addition, the finite sample performance of Qcov3PRF through simulations is verified. We show how poorly PCQR performs under the presence of irrelevant factors and a high degree of cross sectional and serial correlation in the idiosyncratic component.

Finally, we provide three empirical applications. The first one involves the real growth vulnerability from a set of financial variables based on Adrian et al., 2019). The second one considers an economic activity index as a proxy variable to determine the quantile interactions between financial risk and Real GDP growth. The third application is related to climate change and is focused on how well the distribution of the global change in temperatures can be predicted with the carbon dioxide (CO_2)

emissions from a large set of countries, we call this relationship as Climate at Risk (CaR). Our results suggest that our method surpasses PCQR, PQR and other similar alternatives in forecasting accuracy.

The rest of the paper is organized as follows. Section 2 presents the related literature by introducing the 3PRF and some other related methods focused on quantile prediction. In Section 3 we present the methodology for quantile-covariance and Qcov3PRF. Section 4 establishes the consistency and asymptotic normality of the estimated infeasible conditional quantile forecasts. Section 5 presents a study of the finite sample performance of our method through simulations. Section 6 provides three empirical applications, and Section 7 concludes the paper. All mathematical proofs are included in the Appendix.

2 Related literature

2.1 The Three Pass Regression Filter

Proposed by Kelly and Pruitt, 2015, the 3PRF is a forecasting method that focuses on predicting a target variable *in conditional mean* by extracting the relevant factors from a set of predictors. This is a supervised method in the sense that the method only determines the relevant factors for the target and not *all* the factors driven by the set of predictors. Specifically, 3PRF considers the following data generating process (DGP):

$$\boldsymbol{x}_t = \boldsymbol{\phi}_0 + \boldsymbol{\Phi}\boldsymbol{F}_t + \boldsymbol{\varepsilon}_t, \tag{1}$$

$$y_{t+1} = \beta_0 + \boldsymbol{\beta}'\boldsymbol{F}_t + u_{t+1}, \tag{2}$$

$$\boldsymbol{z}_t = \boldsymbol{\lambda}_0 + \boldsymbol{\Lambda}\boldsymbol{F}_t + \boldsymbol{\omega}_t, \tag{3}$$

where $\boldsymbol{x}_t$ is a vector of predictors with a large dimension N. These predictors are approximated through a factor model such that we can obtain K latent unobservable factors $\boldsymbol{F}_t$ that predict a target variable y_{t+1}. Without Eq.(3), forecasts for y_{t+1} can be obtained through PCR applying PCA in the set of predictors (based on Eq.(1)) to obtain estimates for the latent factors $\hat{\boldsymbol{F}}_t^{PCA}$, and the prediction stage (based on Eq.(2)) involves running a least squares regression of y_{t+1} on $\hat{\boldsymbol{F}}_t^{PCA}$ and a constant.

Suppose that y_{t+1} is only affected by $K_f \leq K$ factors, denoted by $\boldsymbol{f}_t$ such that $\boldsymbol{F}_t = (\boldsymbol{f}_t', \boldsymbol{g}_t')'$. The factors $\boldsymbol{f}_t$ ($\boldsymbol{g}_t$) are *relevant* (*irrelevant*) for the target in the sense that $\boldsymbol{\beta} = (\boldsymbol{\beta}_f', \boldsymbol{0}')'$ with $\boldsymbol{\beta}_f \neq \boldsymbol{0}$. These factors are also the only relevant factors to proxies $\boldsymbol{z}_t$ such that $\boldsymbol{\Lambda} = (\boldsymbol{\Lambda}_f, \boldsymbol{0})$. Then, Eq.(3) is what allows the recovery of the relevant factors by using $\boldsymbol{z}_t$.

A relevant factor of y_{t+1} can be potentially omitted from the factors estimated via PCA since the relevant factors do not necessarily have to come from the eigenvectors associated to the largest eigenvalues. In contrast, 3PRF has an advantage of only estimating the factors that affect the target. This nice feature comes from the fact that the factor loadings in Eq.(1) depend on the target (or the proxies) such that the DGP becomes $\boldsymbol{x}_t = \boldsymbol{\phi}_0 + \boldsymbol{\Phi}(y_{t+1})\boldsymbol{F}_t + \tilde{\boldsymbol{\varepsilon}}_t$ as the factor loadings $\boldsymbol{\Phi}(y_{t+1}) = [\boldsymbol{\Phi}_f \quad \boldsymbol{0}]$, where the matrix $\boldsymbol{\Phi}_f$ corresponds to the loadings of the factors $\boldsymbol{f}_t$, and are identified by Eq.(3) and $\tilde{\boldsymbol{\varepsilon}}_t$ contains the remaining (irrelevant) factors, given by $\boldsymbol{g}_t$, and the idiosyncratic component $\boldsymbol{\varepsilon}_t$. The method 3PRF is presented in Algorithm 1.

Algorithm 1 Three-Pass Regression Filter (3PRF)

Pass 1: Run a time series least squares regression of $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{iT})$ on $\boldsymbol{z} = (\boldsymbol{z}_1', \ldots, \boldsymbol{z}_T')'$ ($\boldsymbol{z}_t = (z_{1t}, \ldots, z_{K_f t})$, $t = 1, \ldots, T$) for $i = 1, \ldots, N$,

$$x_{it} = \phi_{0,i} + \boldsymbol{\phi}_i' \boldsymbol{z}_t + \epsilon_{it},$$

and retain the estimate $\hat{\boldsymbol{\phi}}_{\boldsymbol{i}}$.

Pass 2: Run a cross section least squares regression of $\boldsymbol{x}_t = (x_{1t}, \ldots, x_{Nt})$ on $\hat{\boldsymbol{\phi}} = (\hat{\boldsymbol{\phi}}_1', \ldots, \hat{\boldsymbol{\phi}}_N')'$ ($\hat{\boldsymbol{\phi}}_i = (\hat{\phi}_{1i}, \ldots, \hat{\phi}_{K_f i})$, $i = 1, \ldots, N$) for $t = 1, \ldots, T$,

$$x_{it} = \phi_{0,t} + \boldsymbol{f}_t' \hat{\boldsymbol{\phi}}_i + \varepsilon_{it},$$

and retain the estimate $\hat{\boldsymbol{f}}_t$.

Pass 3: Run a time series least squares regression of $\boldsymbol{y} = (y_2, \ldots, y_{T+1})$ on the estimated factors $\hat{\boldsymbol{f}} = (\hat{\boldsymbol{f}}_1', \ldots, \hat{\boldsymbol{f}}_T')'$ ($\hat{\boldsymbol{f}}_t = (\hat{f}_{1t}, \ldots, \hat{f}_{K_f t})$ for $t = 1, \ldots, T$),

$$y_{t+1} = \beta_0 + \boldsymbol{\beta}_f' \hat{\boldsymbol{f}}_t + u_{t+1}.$$

This gives the forecast for the conditional mean of y_{t+1}.

Pass 1 and Pass 2 determine the relevant factors through $\boldsymbol{z}_t$[1], while Pass 3 is the same prediction stage compared to PCR. In general, the proxies can be difficult to get since they require a justified economic/financial theory and/or additional data. To remediate this, Kelly and Pruitt, 2015 proposed an algorithm to determine automatic-proxies by only using the data for $\boldsymbol{x}_t$ and y_{t+1}.[2] We present Algorithm 2, the procedure to determine the automatic proxies for 3PRF.

Algorithm 2 Automatic proxy selection for 3PRF

Initialize $\boldsymbol{r}_0 = \boldsymbol{y}$.

for $\ell = 1, \ldots, K_f$ **do**

Step 0: Let the ℓth automatic proxy be $\boldsymbol{r}_{\ell-1}$. Stop if $\ell = K_f$; otherwise proceed.

Step 1: Apply 3PRF using the $T \times N$ matrix of predictors, $\boldsymbol{X}$, and the automatic proxies $\boldsymbol{r}_0, \ldots, \boldsymbol{r}_{\ell-1}$. Denote the forecast obtained as $\hat{\boldsymbol{y}}_{\ell-1}$.

Step 2: Let the residual $\boldsymbol{r}_{\ell,\tau} = \boldsymbol{y} - \hat{\boldsymbol{y}}_{\ell-1}$ be the $(\ell+1)$th automatic proxy.

Kelly and Pruitt, 2015 showed that these proxies are linearly independent and uncorrelated with the

[1] The idiosyncratic component $\boldsymbol{\omega}_t$ in Eq.(3) is crucial for 3PRF in the sense that, without it, the proxies are perfect for the unobservable factors, i.e., $\boldsymbol{z}_t = \boldsymbol{\lambda}_0 + \boldsymbol{\Lambda}\boldsymbol{F}_t$, then Eq.(1) becomes irrelevant. Hence, the forecasts for y_{t+1} are given by simply running a linear regression of y_{t+1} on $\boldsymbol{z}_t$. Otherwise, when $\boldsymbol{\omega}_t$ is present, running a linear regression of y_{t+1} on $\boldsymbol{z}_t$ would result in biased coefficients.

[2] Note that PLS is a particular case of 3PRF when the proxies are automatic and the set of predictors $\boldsymbol{x}_t$ are standardized over time.

irrelevant factors. In particular:

$$
\begin{aligned}
\hat{\text{cov}}(\boldsymbol{y}, \boldsymbol{g}_j) &= \hat{\text{cov}}(\boldsymbol{r}_0, \boldsymbol{g}_j) = 0, \\
\hat{\text{pcov}}(\boldsymbol{y}, \boldsymbol{g}_j | \hat{\boldsymbol{f}}_1^{(1)}) &= \hat{\text{cov}}(\boldsymbol{y} - \hat{\beta}_{1f}^{(1)} \hat{\boldsymbol{f}}_1^{(1)}, \boldsymbol{g}_j) = \hat{\text{cov}}(\boldsymbol{r}_1, \boldsymbol{g}_j) = 0, \\
&\vdots \\
\hat{\text{pcov}}(\boldsymbol{y}, \boldsymbol{g}_j | \hat{\boldsymbol{f}}_1^{(K_f-1)}, \ldots, \hat{\boldsymbol{f}}_{K_f-1}^{(K_f-1)}) &= \hat{\text{cov}}(\boldsymbol{y} - \hat{\beta}_{1f}^{(K_f-1)} \hat{\boldsymbol{f}}_1^{(K_f-1)} - \cdots - \hat{\beta}_{(K_f-1)f}^{(K_f-1)} \hat{\boldsymbol{f}}_{K_f-1}^{(K_f-1)}, \boldsymbol{g}_j) \\
&= \hat{\text{cov}}(\boldsymbol{r}_{K_f-1}, \boldsymbol{g}_j) = 0,
\end{aligned} \tag{4}
$$

for $j = K_f + 1, \ldots, K$, where $\hat{\text{cov}}(\boldsymbol{y}, \boldsymbol{g}_j)$ and $\hat{\text{pcov}}(\boldsymbol{y}, \boldsymbol{g}_j | \hat{\boldsymbol{f}}_1^{(1)}, \ldots, \hat{\boldsymbol{f}}_\ell^{(\ell)})$ denote the sample covariance between $\boldsymbol{y}$ and $\boldsymbol{g}_j$, and the sample partial covariance between $\boldsymbol{y}$ and $\boldsymbol{g}_j$ conditioning on the variables $\hat{\boldsymbol{f}}_1^{(1)}, \ldots, \hat{\boldsymbol{f}}_\ell^{(\ell)}$, respectively, where (ℓ) denotes the estimates obtained in the ℓth iteration, and $\hat{\boldsymbol{f}}_\ell$ is uncorrelated with $\boldsymbol{g}_j$ for all $\ell = 1, \ldots, K_f$, $j = K_f + 1, \ldots, K$. Then, it is natural to think of an extension of 3PRF for the conditional quantile prediction by using quantile-covariance (*qcov*) instead of the standard covariance (*cov*).

2.2 Related methods that forecast the conditional quantile of the target

Giglio et al., 2016 extended 3PRF to predict conditional quantiles of the target variable. However, they were only successful at extracting one relevant factor. The method developed in that article is known as Partial Quantile Regression (PQR), whose algorithm, compared to 3PRF, replaces the prediction stage (Pass 3) with τ-quantile regression. More importantly, PQR replaces Pass 1 by running a time series τ-quantile regression of y_{t+1} on x_{it} and a constant, and keep the estimates, say, $\hat{\phi}_{i\tau}$, for $i = 1, \ldots, N$.[3] Pass 1 in PQR is associated to a quantile-covariance concept considered by Dodge and Whittaker, 2009 for a PLS extension with conditional quantile prediction (also called PQR). The reason why this method only captures one relevant factor is because a quantile regression can not be switched as it can be for the least squares case. This is, ommiting a constant term, a linear coefficient equal to zero by running a quantile regression of $\boldsymbol{y}$ on $\boldsymbol{x}$ is not necessarily zero if we run a quantile regression of $\boldsymbol{x}$ on $\boldsymbol{y}$, whereas for the conditional mean case running a least squares regression $\boldsymbol{y}$ on $\boldsymbol{x}$ leads to a coefficient equal to zero if and only if the coefficient is zero when running a least squares regression of $\boldsymbol{x}$ on $\boldsymbol{y}$. Comparing Pass 1 in Algorithm 1 for the conditional mean with $z_t = y_{t+1}$ and $K_f = 1$, we should run a quantile regression of $\boldsymbol{y}$ on $\boldsymbol{x}_i$ for the conditional quantile forecasts of $\boldsymbol{y}$ but we cannot run a quantile regression of $\boldsymbol{x}_i$ on $\boldsymbol{y}$. Extending Pass 1 for the quantile forecast with $K_f > 1$ is not even possible as it involves simultaneous linear relationships between $\boldsymbol{x}_i$ and the τ-quantile of $\boldsymbol{z}$.

Giglio et al., 2016 also presented the Principal Components Quantile Regression (PCQR). This

[3] In addition, they omitted the constant in Pass 2. Adding a constant is an important forecast improvement in 3PRF (with automatic proxies) compared to PLS when forecasting the conditional mean. See Kelly and Pruitt, 2015 for details.

method simply substitutes the prediction stage of PCR with quantile regression. The authors showed that both PQR and PCQR give consistent forecasts. As in the case of PCR when forecasting the conditional mean, PCQR is consistent for any number of relevant factors when forecasting conditional quantiles, as the estimated loadings for the irrelevant factors will converge to zero. However, in finite samples, the resulting forecasts may be inefficient when including several irrelevant factors as is the case when predicting the conditional mean.

Another supervised method for quantile prediction involving factor models is the fast PQR (fPQR) developed by Méndez-Civieta et al., 2022 where they followed a similar approach PQR but instead of the quantile covariance of Dodge and Whittaker, 2009, they maximized the *qcov* (Li et al., 2015) between the target and the predictors in a PLS framework.

3 Quantile-covariance Three Pass Regression Filter

3.1 Quantile-covariance (*qcov*) and Quantile partial-covariance (*qpcov*)

As we have noticed on the methods described above, a key point to successfully extend 3PRF to conditional quantile prediction relies on exploiting a quantile-covariance relationship between the target and the predictors such that the relevant factors for a conditional τ-quantile can be recovered. This quantile covariance needs to be used to allow a kind of "inverse" form for the quantile regression, as it is required in Pass 1 of 3PRF.

The quantile-covariance used in our forecasting method is the *qcov* presented in Li et al., 2015. Specifically, for two random variables f and y, let Q_y^τ be the τth unconditional quantile of y and $Q_{y|f}^\tau$ be the τth conditional quantile of y given f. The first order condition of quantile regression implies that $Q_{y|f}^\tau$ is independent of f with probability 1 if and only if the random variables $I(y - Q_y^\tau > 0)$ and f are independent, where $I(\cdot)$ is the indicator function. From this result, the *qcov* for $0 < \tau < 1$ is defined as follows:

$$\begin{aligned} \text{qcov}_\tau(y, f) &= \text{cov}(I(y - Q_y^\tau > 0), f) \\ &= \mathbb{E}[\psi_\tau(y - Q_y^\tau)(f - \mathbb{E}(f))], \end{aligned} \tag{5}$$

where $\psi_\tau(v) = \tau - I(v < 0)$. It is important to mention that *qcov* does not possess a symmetry property, i.e., $\text{qcov}_\tau(y, f) \neq \text{qcov}_\tau(f, y)$. However, $\text{cov}(I(y - Q_y^\tau > 0), f) = \text{cov}(f, I(y - Q_y^\tau > 0))$. This is the critical property for us to construct Pass 1 of the Qcov3PRF algorithm presented in Section 3.2. We note that this is different from the quantile-covariance considered implicitly in PQR of Giglio et al., 2016.[4]

Now, let us refer to the quantile regression $y = b_{0\tau} + b_{j\tau} f_j + u^\tau$, where $j = 1, \ldots, K_f$ and look for

[4]Specifically, the quantile-covariance considered in PQR is equal to the linear coefficient obtained after running a quantile regression of y on f and a constant, where f has mean 0 and variance 1.

the following minimizers:

$$(b_{0\tau}^*, b_{j\tau}^*) = \underset{b_{0\tau}, b_{j\tau}}{\arg\min}\, \mathbb{E}\left[\rho_\tau\left(y - b_{0\tau} - b_{j\tau} f_j\right)\right], \tag{6}$$

where $\rho_\tau(v) = v\left[\tau - I(v < 0)\right]$ is the check loss function. Li et al., 2015 found a nice result such that $\text{qcov}_\tau(y, f_j) = \nu(b_{j\tau}^*)$ where $\nu(\cdot)$ is a continuous and increasing function, and, more importantly, $\nu(b_{j\tau}^*) = 0$ if and only if $b_{j\tau}^* = 0$.

The previous result only holds for the case of one regressor at a time. However, as we will see in the next subsection, Qcov3PRF involves a quantile regression with multiple predictors (multiple relevant factors) in the prediction stage. So we need further tools that extend the result presented above for the case of multiple predictors. To see how *qcov* and the multivariate quantile regression are related, let us consider a linear quantile model for the target variable y and the set of predictors $\boldsymbol{f} = (f_1, \ldots, f_{K_f})$:

$$y = \beta_{0\tau} + \beta_{1\tau} f_1 + \cdots + \beta_{K_f\tau} f_{K_f} + u^\tau, \tag{7}$$

where the error term satisfies $P(u^\tau < 0|\boldsymbol{f}) = \tau$. This implies that the conditional quantile of y given $\boldsymbol{f}$ is $Q_{y|\boldsymbol{f}}^\tau = \beta_{0\tau} + \beta_{1\tau} f_1 + \cdots + \beta_{K_f\tau} f_{K_f}$. Without loss of generality, let us assume that $\mathbb{E}(f_j) = 0$ and $\text{Var}(f_j) = 1$ for all $j = 1, \ldots, K_f$, and denote $h_u(u|\boldsymbol{f})$ and $h_y(y|\boldsymbol{f})$ as the conditional density of u and y given $\boldsymbol{f}$, respectively.

Angrist et al., 2006 described a procedure to obtain the coefficient $b_{j\tau}^*$ alternatively as the minimizer of the following weighted least squares problem:

$$b_{j\tau}^* = \underset{b_{j\tau} \in \mathbb{R}}{\arg\min}\, \mathbb{E}\left[\tilde{w}_\tau(\boldsymbol{f}) \cdot \left(Q_{y|\boldsymbol{f}}^\tau - b_{j\tau} f_j\right)^2\right], \tag{8}$$

where $\tilde{w}_\tau(\boldsymbol{f}) = \frac{1}{2}\int_0^1 h_u\left(u \cdot \boldsymbol{\Delta}_\tau(f_j, b_{j\tau}^*)|\boldsymbol{f}\right) du$, and $\boldsymbol{\Delta}_\tau(f_j, b_{j\tau}^*) = b_{j\tau}^* + b_{j\tau}^* f_j - Q_{y|\boldsymbol{f}}^\tau$ for $j = 1, \ldots, K_f$. Then,

$$b_{j\tau}^* = \mathbb{E}(\tilde{w}_\tau(\boldsymbol{f}) f_j^2)^{-1} \mathbb{E}(\tilde{w}_\tau(\boldsymbol{f}) f_j Q_{y|\boldsymbol{f}}^\tau) = \beta_{j\tau}^* + d_{j\tau},$$

where $\beta_{j\tau}^*$ is the coefficient estimate of f_j obtained from running quantile regression of y on $\boldsymbol{f}$, and $d_{j\tau} = \sum_{k \neq j} \beta_{k\tau}^* \mathbb{E}(\tilde{w}_\tau(\boldsymbol{f}) f_j^2)^{-1} \mathbb{E}(\tilde{w}_\tau(\boldsymbol{f}) f_j f_k)$. The term $d_{j\tau}$ can be interpreted as the *bias* of the quantile estimator. It is equal to zero when $\mathbb{E}(\tilde{w}_\tau(\boldsymbol{F}) f_j f_k) = 0$, however, in contrast to the conditional mean case, orthogonality among the predictors is not enough to ensure $d_{j\tau} = 0$ due to the presence of the term $\tilde{w}_\tau(\boldsymbol{f})$. Therefore, the relationship between $\text{qcov}_\tau(y, f_j) = \nu(b_{j\tau}^*)$ and $\beta_{j\tau}^*$ is affected by the term $d_{j\tau}$. Li et al., 2015 and Ma et al., 2017 approached this problem by using the quantile partial covariance (*qpcov*). Similar to the partial covariance for the conditional mean, *qpcov* is defined by partialling out the confunding effects of the predictors $\boldsymbol{f}_{-j} = (1, \{f_k, k \neq j\})$ on f_j and on y. Specifically,

$$\text{qpcov}_\tau(y, f_j | \boldsymbol{f}_{-j}) = \text{cov}(I(y - \boldsymbol{\gamma}'_{0j}\boldsymbol{f}_{-j} > 0), f_j - \boldsymbol{\theta}'_{0j}\boldsymbol{f}_{-j}), \tag{9}$$

where $\boldsymbol{\theta}_{0j} = \arg\min_{\boldsymbol{\theta}_j} \mathbb{E}[(f_j - \boldsymbol{\theta}'_j \boldsymbol{f}_{-j})^2]$ and $\boldsymbol{\gamma}_{0j} = \arg\min_{\boldsymbol{\gamma}_j} \mathbb{E}[\rho_\tau(y - \boldsymbol{\gamma}'_j \boldsymbol{f}_{-j})]$. Based on the results in Ma et al., 2017 with the estimates obtained from solving:

$$(\delta^*_{0\tau}, \delta^*_{j\tau}) = \underset{\delta_{0\tau}, \delta_{j\tau}}{\arg\min} \, \mathbb{E}[\rho_\tau(y - \boldsymbol{\gamma}'_{0j}\boldsymbol{f}_{-j} - \delta_\tau - \delta_{j\tau} f_j)], \tag{10}$$

we have that $\text{qpcov}_\tau(y, f_j | \boldsymbol{f}_{-j}) = \varrho(\delta^*_{j\tau})$, where $\varrho(\cdot)$ is a continuous and increasing function and $\varrho(\delta^*_{j\tau}) = 0$ if and only if $\delta^*_{j\tau} = 0$. It can be shown that $\beta^*_{j\tau} = 0$ if and only if $\delta^*_{j\tau} = 0$. Hence, $\beta^*_{j\tau} = 0$ if and only if $\varrho(\delta^*_{j\tau}) = 0$.

3.2 Quantile-covariance Three Pass Regression Filter

Based on the work of Kelly and Pruitt, 2015 and Giglio et al., 2016 we develop a new method of Qcov3PRF. This forecasting method relies on the quantile-covariance *qcov* defined in the previous subsection.

We first establish the environment considered for this method. There is a target variable that we wish to forecast its conditional τ-quantile. There are many highly correlated predictors that can contain useful information to predict the conditional quantile. The number of predictors N can be large and its magnitude can be greater than or equal to the time series observations T, this is $N \geq T$. This complicates the estimation using quantile regression (Koenker and Bassett Jr, 1978). Then, we look to reduce the dimension of the covariates assuming that the covariates can be approximated using a (linear) factor model. The proxies that we consider in our method have a similar interpretation than for 3PRF. These are variables driven by the relevant factors that affect the target variable in the τth quantile. The complication now, in contrast to 3PRF, is that the proxies considered are not the variables themselves but transformations motivated by *qcov* (see Eq.(13), which will be justified later on in this subsection). Hence, the effects on the target come from the transformed proxies and *not* from the original variables.

Much more useful in practice is that these proxies can be obtained in an automatic manner, as we will describe when we present Algorithm 4. The target variable depends on those factors, but given that these factors are unobserved, the predictions obtained from the true factors are known as the *infeasible best forecasts*. For the DGP we first present the following models presented in Assumption 1.

Assumption 1. *The data for a fixed level $\tau \in (0,1)$ is generated as follows:*

$$\begin{aligned} \boldsymbol{x}_t &= \boldsymbol{\phi}_0 + \boldsymbol{\Phi}\boldsymbol{F}_t + \boldsymbol{\varepsilon}_t, & (11) \\ y_{t+1} &= \beta_{0,\tau} + \boldsymbol{\beta}'_\tau \boldsymbol{F}_t + u^\tau_{t+1}, & (12) \end{aligned}$$

where y_{t+1} denotes the target time series variable, $\boldsymbol{x}_t$ is an $N \times 1$ vector of predictors which are standardized to have unit time series variance, $\boldsymbol{F}_t$ is a $K \times 1$ vector of latent factors, and $\boldsymbol{\Phi}$ is an $N \times K$ matrix with $K < \min(N,T)$. In addition, $\boldsymbol{\beta}_\tau$ is a $K \times 1$ vector, whose dependence on τ is because the conditional quantile of y_{t+1} is $Q^\tau_{y_{t+1}|\boldsymbol{F}_t} = \beta_{\tau,0} + \boldsymbol{\beta}'_\tau \boldsymbol{F}_t$ where $\mathbb{E}_t[\tau - I(u^\tau_{t+1} < 0)] = \mathbb{E}[\tau - I(u^\tau_{t+1} < 0)|y_t, \boldsymbol{F}_t, y_{t-1}, \boldsymbol{F}_{t-1}, \dots] = 0$. Also, $\boldsymbol{\beta}_\tau = (\boldsymbol{\beta}'_{f,\tau}, \boldsymbol{0}')'$ with $\boldsymbol{\beta}_{f,\tau} \neq \boldsymbol{0}$ vector of dimension $K_f \times 1$.

Eq.(11) assumes a standard factor structure for a large set of predictors (Bai and Ng, 2002), Eq.(12) is the prediction stage which focuses on estimating the conditional quantile of the target. By only considering these models, there is no need to make a distinction between irrelevant or relevant factors. We know by Ando and Tsay, 2011 and Giglio et al., 2016 that PCQR provides consistent conditional quantile forecasts, however, in finite sample cases, under the presence of high cross section and/or serial correlation in ε_t and the existence of several irrelevant factors with high variance, the estimation can result in forecasts with poor performance.

On the other hand, referring to 3PRF with automatic-proxies (see Eqs.(1)-(3) and Algorithm 2), there are K_f proxies linearly independent and uncorrelated with the irrelevant factors $\boldsymbol{g}$ (see Eq.(4)). If now we instead incorporate *qcov* and *qpcov* in the automatic proxies framework we can formulate K_f proxies determined sequentially given by $I(\boldsymbol{y} - \hat{Q}^\tau_{\boldsymbol{y}} > 0)$, $I(\boldsymbol{y} - \hat{Q}^\tau_{\boldsymbol{y}|\hat{\boldsymbol{f}}^{(1)}_1} > 0) = I(\boldsymbol{y} - \hat{\beta}^{(1)}_{1,\tau}\hat{\boldsymbol{f}}^{(1)}_1 > 0)$, $\dots$, $I(\boldsymbol{y} - \hat{Q}^\tau_{\boldsymbol{y}|\hat{\boldsymbol{f}}^{(K_f-1)}_1,\dots,\hat{\boldsymbol{f}}^{(K_f-1)}_{K_f-1}} > 0) = I(\boldsymbol{y} - \hat{\beta}^{(K_f-1)}_{1f,\tau}\hat{\boldsymbol{f}}^{(K_f-1)}_1 - \dots - \hat{\beta}^{(K_f-1)}_{(K_f-1)f,\tau}\hat{\boldsymbol{f}}^{(K_f-1)}_{K_f-1} > 0)$, where (ℓ) denotes the relevant factors obtained in the ℓth iteration and $\hat{\boldsymbol{\beta}}^{(\ell)}_\tau$ are obtained by running quantile regression of $\boldsymbol{y}$ on $\hat{\boldsymbol{f}}^{(\ell)}$.

If we further generalize to any linearly independent group of K_f proxies by letting $\boldsymbol{z}^*_{j\tau} = I(\boldsymbol{z}_{j\tau} - Q^\tau_{z_j} > 0) = I(\boldsymbol{y} - \hat{Q}^\tau_{\boldsymbol{y}|\hat{\boldsymbol{f}}_1,\dots,\hat{\boldsymbol{f}}_j} > 0)$, and stacking over $j = 1, \dots, K_f$ we have:

$$\boldsymbol{z}^*_\tau = \boldsymbol{\lambda}_{0,\tau}\boldsymbol{\iota}' + \boldsymbol{\Lambda}_{f,\tau}\boldsymbol{f} + \boldsymbol{\omega}, \qquad (13)$$

where $\boldsymbol{\iota}$ is a column vector of ones and $\boldsymbol{\Lambda}_{f,\tau}$ is a $K_f \times K_f$ matrix. Eq.(13) is the proxy equation in our setting, which extends Eq.(3) to the quantile case. Eq.(13) is presented as Assumption 2 and it is associated to τ-quantile through $Q^\tau_{z_j}$, which makes it possible to extend Pass 1 in 3PRF to the conditional quantile case via least squares time series regressions. Note that when $z_t = y_{t+1}$ we go back to the estimation of only one relevant factor, which is covered in PQR.

Assumption 2. *The data for a fixed level $\tau \in (0,1)$ is generated as follows:*

$$\boldsymbol{z}^*_{t,\tau} \quad = \quad \boldsymbol{\lambda}_{\tau,0} + \boldsymbol{\Lambda}_\tau \boldsymbol{F}_t + \boldsymbol{\omega}_t, \tag{14}$$

where $\boldsymbol{z}^*_{t,\tau} = (z^*_{1t,\tau}, \ldots, z^*_{K_f t,\tau})$ with $z^*_{\ell t,\tau} = I(z_{\ell t} - Q^\tau_{z_{\ell t}} > 0)$, $\boldsymbol{z}_t$ is a $K_f \times 1$ vector of proxy data, $\boldsymbol{F}_t = (\boldsymbol{f}'_t, \boldsymbol{g}'_t)'$, $\boldsymbol{\Lambda}_\tau = (\boldsymbol{\Lambda}_{f,\tau}, \boldsymbol{\Lambda}_{g,\tau})$ is a $K_f \times K$ matrix. Let $K_f > 0$ and $K_g \geq 0$ the dimension of $\boldsymbol{f}_t$ and $\boldsymbol{g}_t$, respectively, such that $K_f + K_g = K$ where $K_f \ll min(N,T)$. $Q^\tau_{z_{\ell t}}$ denotes the unconditional τ quantile of $z_{\ell t}$. In addition, $\boldsymbol{\Lambda}_\tau = [\boldsymbol{\Lambda}_{f,\tau} \quad \boldsymbol{0}]$, where $\boldsymbol{\Lambda}_{f,\tau}$ is nonsingular.

Based on Assumptions 1 and 2 , the variable y_{t+1} depends only on a subset of factors $\boldsymbol{f}_t$, called the *relevant factors*. In the 3PRF for the conditional mean, $\boldsymbol{z}^*_{t,\tau} = \boldsymbol{z}_t$, and Eq.(14) gives a link to extract the relevant factors. However, for the conditional quantile case the implementation is not direct when there are more than one relevant factor. The reason is because, while in Pass 1 of 3PRF we run time series least squares regressions of x_{it} on $\boldsymbol{z}_t$, for the quantile case the proxies correspond to the conditional quantile and would require running time series quantile regressions of x_{it} on $\boldsymbol{z}_t$ focused on the τth quantile of $\boldsymbol{z}_t$ and not on the quantile of x_{it}. This is not possible to be implemented directly through quantile regressions. Hence, we make use of the fact that $\text{qcov}_\tau(z_{\ell t}, x_{it}) = \text{cov}(I(z_{\ell t} - Q^\tau_{z_{\ell t}} > 0), x_{it}) = \text{cov}(z^*_{\ell t,\tau}, x_{it}) = \text{cov}(x_{it}, z^*_{\ell t,\tau})$ along with the discussion that resulted in Eq.(13) (Assumption 2).

Qcov3PRF is implemented through Algorithm 3. As an initial step, the unconditional quantile for each proxy has to be estimated.[5] Pass 1 and Pass 2 estimate the K_f relevant factors $\hat{\boldsymbol{f}}_t$. In Pass 3 we obtain the estimated parameters $\hat{\beta}_{0f,\tau}$ and $\hat{\boldsymbol{\beta}}_{f,\tau}$ from the underlying predictors $\boldsymbol{x}_t$ for $t = 1, \ldots, T$. Then, we construct the forecasts one period ahead[6] for the observation y_{T+2} with the updated set of predictors $\boldsymbol{x}_t$, $t = 1, \ldots, T+1$ to obtain $\hat{\boldsymbol{f}}_{T+1}$ and compute $\hat{y}_{T+2} = \hat{\beta}_{0,\tau} + \hat{\boldsymbol{\beta}}'_{f,\tau} \hat{\boldsymbol{f}}_{T+1}$.

As we mentioned above, quantile regression cannot be implemented directly in Pass 1 under the presence of more than 1 relevant factor. This is because we are focusing on predicting the conditional quantile of the target variable and not on the conditional quantile of the predictors. Instead, we run least squares regressions given Eq.(14).

In contrast to 3PRF, our Qcov3PRF method does not provide a complete closed form estimator for the linear coefficients $\boldsymbol{\beta}_\tau$ because the linear quantile regression estimator does not have a closed form solution. However, we can obtain a closed form expression for the relevant factor estimates $\hat{\boldsymbol{f}}_t$ and their

[5] We noticed a decrease in performance for the forecasts obtained from the Qcov3PRF compared to the PQR when T is small (less than 100) in simulations, and obtaining multicollinearity issues for extreme quantile values (e.g. 1% and 99%). This is due to the additional estimation of the unconditional quantile for the target and the discrete nature of $I(z_{\ell t} - Q^\tau_{z_{\ell t}} > 0)$. This difference vanishes as T increases.

[6] For simplicity we have considered one-period ahead forecasts, however, the algorithm can be perfectly implemented for h-periods ahead forecasts. For instance, in our first empirical application with monthly data, we consider 3-periods and 12-periods ahead.

Algorithm 3 Quantile-covariance Three-Pass Regression Filter (Qcov3PRF)

Pass 0: Let $\boldsymbol{z}_\ell^* = I(\boldsymbol{z}_\ell - \hat{Q}^\tau_{z_\ell} > 0)$, where $\hat{Q}^\tau_{z_\ell}$ is an estimator for the unconditional quantile of z_ℓ for $\ell = 1, \ldots, K_f$.

Pass 1: Run a time series least squares regression of $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{iT})$ on $\boldsymbol{z}^*_\tau = (\boldsymbol{z}^{*'}_{1,\tau}, \ldots, \boldsymbol{z}^{*'}_{T,\tau})'$ ($\boldsymbol{z}^*_t = (z^*_{1t}, \ldots, z^*_{K_f t})$, $t = 1, \ldots, T$) for $i = 1, \ldots, N$,

$$x_{it} = \phi_{0,i} + \boldsymbol{\phi}'_{i,\tau}\boldsymbol{z}^*_{t,\tau} + \epsilon_{it},$$

and retain the estimate $\hat{\boldsymbol{\phi}}_{i,\tau}$.

Pass 2: Run a cross section least squares regression of $\boldsymbol{x}_t = (x_{1t}, \ldots, x_{Nt})$ on $\hat{\boldsymbol{\phi}}_\tau = (\hat{\boldsymbol{\phi}}'_{1,\tau}, \ldots, \hat{\boldsymbol{\phi}}'_{N,\tau})'$ ($\hat{\boldsymbol{\phi}}_{i,\tau} = (\hat{\phi}_{1i,\tau}, \ldots, \hat{\phi}_{K_f i,\tau})$, $i = 1, \ldots, N$) for $t = 1, \ldots, T$,

$$x_{it} = \phi_{0,t,\tau} + \boldsymbol{f}'_t\hat{\boldsymbol{\phi}}_{i,\tau} + \varepsilon_{it},$$

and retain the estimate $\hat{\boldsymbol{f}}_t$.

Pass 3: Run time series quantile regression of y_{t+1} on the estimated factors $\hat{\boldsymbol{f}} = (\hat{\boldsymbol{f}}'_1, \ldots, \hat{\boldsymbol{f}}'_T)'$ ($\hat{\boldsymbol{f}}_t = (\hat{f}_{1t}, \ldots, \hat{f}_{K_f t})$, $t = 1, \ldots, T$),

$$y_{t+1} = \beta_{0,\tau} + \boldsymbol{\beta}'_{f,\tau}\hat{\boldsymbol{f}}_t + u^\tau_{t+1}. \tag{15}$$

This gives the forecast for the conditional τ quantile of y_{t+1} denoted by $\hat{Q}^\tau_{y_{t+1}|\hat{\boldsymbol{f}}_t}$.

---

corresponding loadings by combining Pass 1 and Pass 2. These are given by:

$$\hat{\boldsymbol{\Phi}}_\tau = (\boldsymbol{W}_{zz})^{-1}\boldsymbol{W}'_{xz}, \tag{16}$$

$$\hat{\boldsymbol{F}}' = \boldsymbol{W}_{zz}(\boldsymbol{W}'_{xz}\boldsymbol{J}_N\boldsymbol{W}_{xz})^{-1}\boldsymbol{W}'_{xz}\boldsymbol{J}_N\boldsymbol{X}', \tag{17}$$

where $\hat{\boldsymbol{\Phi}}_\tau$, $\hat{\boldsymbol{F}}$ and $\boldsymbol{X}$ are the matrices resulting from stacking $\hat{\boldsymbol{\phi}}_{i,\tau}$ over i, and $\hat{\boldsymbol{F}}_t$ and $\boldsymbol{x}_t$ over t, respectively. Also, $\boldsymbol{J}_N = \boldsymbol{I}_N - \frac{1}{N}\boldsymbol{\iota}_N\boldsymbol{\iota}'_N$, $\boldsymbol{I}_N$ is the identity matrix of dimension N, $\boldsymbol{\iota}'_N$ is a vector of ones of dimension N, $\boldsymbol{W}_{zz} = \boldsymbol{Z}^{*'}_\tau\boldsymbol{J}_T\boldsymbol{Z}^*_\tau$, $\boldsymbol{W}_{xz} = \boldsymbol{X}'\boldsymbol{J}_T\boldsymbol{Z}^*_\tau$, and $\boldsymbol{W}_{zz} = \boldsymbol{Z}^{*'}_\tau\boldsymbol{J}_T\boldsymbol{Z}^*_\tau$ ($\boldsymbol{J}_T$ is analogous to $\boldsymbol{J}_N$). This closed form expression is particularly useful to show the consistency of the relevant factors and their corresponding loadings relying on the results of Kelly and Pruitt, 2015.

Now, we explore the implementation of automatic-proxies for Qcov3PRF generated when only $\boldsymbol{y}$ and $\boldsymbol{X}$ are available as it is common in practice that the researcher does not count with proxy variables. Compared to 3PRF, we cannot generate the proxies as residuals from a least squares regression. To see this, consider a scale model with two relevant factors $y_{t+1} = (f_{1t} + f_{2t} + \sigma_y)u_{t+1}$ where $f_{1t} \sim \mathcal{U}[0, 1]$, $f_{2t} \sim \mathcal{U}[0, 2]$, σ_y is a constant and $u_{t+1} \sim \mathcal{N}(0, 1)$. Hence, the conditional mean of y_{t+1} is equal to 0 but f_{1t} and f_{2t} are relevant for any quantile except the median. Consider Algorithm 2 applying Qcov3PRF instead of 3PRF, when $\ell = 1$, with the automatic proxy y_{t+1}, we get an estimate for the factor with the largest variance $\hat{f}_{1t}$. Next, when $\ell = 2$ the automatic proxies are y_{t+1} and the residual when running a least squares regression of y_{t+1} on f_{1t}, say, $\tilde{y}_{t+1}$. Then, Pass 1 in Qcov3PRF will potentially face a multicollinearity problem since the set of automatic proxies $\boldsymbol{r}_t = (y_{t+1}, \tilde{y}_{t+1})$ basically contains the target variable repeated twice as y_{t+1} is not affected by f_{1t} in the conditional mean. As a result, f_{2t} cannot be estimated.

As we mentioned earlier, from Eq.(4), the natural candidates for the automatic-proxies in the conditional quantile case are given by $I(\boldsymbol{y} - \hat{Q}^{\tau}_{y} > 0)$, $I(\boldsymbol{y} - \hat{Q}^{\tau}_{\boldsymbol{y}|\hat{\boldsymbol{f}}^{(1)}_1} > 0)$, $\ldots$, $I(\boldsymbol{y} - \hat{Q}^{\tau}_{\boldsymbol{y}|\hat{\boldsymbol{f}}^{(K_f-1)}_1,\ldots,\hat{\boldsymbol{f}}^{(K_f-1)}_{K_f-1}} > 0)$. Specifically, these automatic-proxies are obtained by applying Algorithm 4 in the same spirit as Algorithm 2. It can be shown that these proxies are linearly independent, but not necessarily uncorrelated of $\boldsymbol{g}_j$, for $j = K_f+1, \ldots, K$, if u^{τ} depends on $\boldsymbol{g}_j$. This is a consequence of the fact that $\boldsymbol{g}_j$ does not affect the conditional τ-quantile of $\boldsymbol{y}$ but it can affect some other part of the conditional distribution of $\boldsymbol{y}$. For example, consider $y_{t+1} = 1 + f_{1,t} + (0.5 f_{1,t} + f_{2,t}) u_{t+1}$, where $Q^{\tau}_{u} = 0$, and $f_{3,t}$ is some other factor independent of y_{t+1}. Then, $Q^{\tau}_{y_{t+1}|\boldsymbol{f}} = \beta_{0,\tau} + \beta_{1,\tau} f_{1,t} + \beta_{2,\tau} f_{2,t} + \beta_{3,\tau} f_{3,t}$ with $(\beta_{0,\tau}, \beta_{1,\tau}, \beta_{2,\tau}, \beta_{3,\tau}) = (1, 1, 0, 0)$. However, if we run a least squares regression of $I(y_{t+1} - Q^{\tau}_{y} > 0)$ on $\boldsymbol{f}_t = (f_{1,t}, f_{2,t}, f_{3,t})$ and a constant, we get the estimated coefficients $\boldsymbol{\alpha}^*_{\tau} = (\alpha^*_{0,\tau}, \alpha^*_{1,\tau}, \alpha^*_{2,\tau}, \alpha^*_{3,\tau})$ where $\alpha^*_{0,\tau} \neq 0$, $\alpha^*_{1,\tau} \neq 0$, $\alpha^*_{3,\tau} = 0$ but $\alpha^*_{2,\tau} \neq 0$.

Algorithm 4 Automatic proxy selection for Qcov3PRF

For a quantile value $\tau \in (0, 1)$ initialize $\boldsymbol{r}_{0,\tau} = \boldsymbol{y} - \hat{Q}^{\tau}_{y}$.

for $\ell = 1, \ldots, K_f$ **do**

Step 0: Let the ℓth automatic proxy be $\boldsymbol{r}_{\ell-1,\tau}$. Stop if $\ell = K_f$; otherwise proceed.

Step 1: Apply Qcov3PRF using the predictors $\boldsymbol{X}$, and the automatic proxies $\boldsymbol{r}_{0,\tau}, \ldots, \boldsymbol{r}_{\ell-1,\tau}$. Denote the forecast obtained as $\hat{Q}^{\tau}_{y|\hat{f}_1,\ldots,\hat{f}_{\ell-1}}$.

Step 2: Let the residual $\boldsymbol{r}_{\ell,\tau} = \boldsymbol{y} - \hat{Q}^{\tau}_{y|\hat{f}_1,\ldots,\hat{f}_{\ell-1}}$ be the $(\ell+1)$th automatic proxy.

Therefore, the following model is suggested:

$$r^*_{\ell t,\tau} = I(r_{\ell t,\tau} > 0) = I(y_{t+1} - \hat{Q}^{\tau}_{y_{t+1}|\hat{f}^{(\ell)}_{1,t},\ldots,\hat{f}^{(\ell)}_{\ell,t}} > 0) = \alpha_{0,\tau} + \boldsymbol{\alpha}^{(\ell)}_{f,\tau}\boldsymbol{f}_t + \boldsymbol{\alpha}^{(\ell)}_{g,\tau}\boldsymbol{g}_t + \xi_{t+1}, \qquad (18)$$

where ξ_{t+1} depends on u^{τ}_{t+1}, $\alpha^{(\ell)}_{j,f,\tau} \neq 0$ if $\beta_{j\tau} \neq 0$ for $j = 1, \ldots, K_f$, $\alpha^{(\ell)}_{k,g,\tau} = 0$ if g_{kt} is independent of y_{t+1} for $k = K_f+1, \ldots, K$. However, the value of $\alpha^{(\ell)}_{k,g,\tau}$ when g_k is not independent of y_{t+1}, specifically, if u^{τ}_{t+1} depends of g_k, is not guaranteed to be zero. Although our multiple simulation experiments always indicate that under this situation $\alpha^{(\ell)}_{k,g,\tau}$ is much lower than $\alpha^{(\ell)}_{j,f,\tau}$ for any $j = 1, \ldots, K_f$ and $k = K_f+1, \ldots, K$ and very close to zero for most cases it is not always the case that an irrelevant factor for the τ-conditional quantile of y_{t+1} would be irrelevant for the automatic-proxy $r^*_{\ell t,\tau}$ as well. In order to guarantee this is always the case we need to restrict the error term u^{τ}_{t+1} such that the automatic proxies are only spanned by the factors $\boldsymbol{f}$. That is stated in the following assumption.

Assumption 3. *The error term* u_{t+1}^{τ} *from Eq.(12) allows the following model:*

$$I(y_{t+1} - Q_{y_{t+1}}^{\tau} > 0) = \alpha_{0,\tau} + \boldsymbol{\alpha}'_{f,\tau}\boldsymbol{f}_t + \xi_{t+1}, \tag{19}$$

where the error term ξ_{t+1} *depends on* u_{t+1}^{τ} *and satisfies* $\mathbb{E}_t(\xi_{t+1}) = \mathbb{E}(\xi_{t+1}|y_t, \boldsymbol{F}_t, y_{t-1}, \boldsymbol{F}_{t-1}, \dots) = 0$, $\boldsymbol{E}(\xi_{t+1}^4) \leq M$ *and* $T^{-1/2}\sum_{t=1}^{T}\boldsymbol{F}_t\xi_{t+1} = \boldsymbol{O}_p(1)$ *for all* t *and* $M < \infty$*, and* $Q_{y_{t+1}}^{\tau}$ *denotes the unconditional* τ*-quantile of* y_{t+1}.

By Assumption 3 we can obtain K_f automatic-proxies that are uncorrelated with the irrelevant factors, this is,

$$\begin{aligned}
\text{q}\hat{\text{cov}}(\boldsymbol{y}, \boldsymbol{g}_j) &= \hat{\text{cov}}(I(\boldsymbol{y} - Q_y^{\tau} > 0), \boldsymbol{g}_j) = \hat{\text{cov}}(I(\boldsymbol{r}_{0,\tau} > 0), \boldsymbol{g}_j) = 0, \qquad (20)\\
\text{qp}\hat{\text{cov}}(\boldsymbol{y}, \boldsymbol{g}_j|\hat{\boldsymbol{f}}_1^{(1)}) &= \hat{\text{cov}}(I(\boldsymbol{y} - \hat{\beta}_{1f,\tau}^{(1)}\hat{\boldsymbol{f}}_1^{(1)} > 0), \boldsymbol{g}_j) = \hat{\text{cov}}(I(\boldsymbol{r}_{1,\tau} > 0), \boldsymbol{g}_j) = 0,\\
&\vdots\\
\text{qp}\hat{\text{cov}}(\boldsymbol{y}, \boldsymbol{g}_j|\hat{\boldsymbol{f}}_1^{(K_f-1)}, \dots, \hat{\boldsymbol{f}}_{K_f-1}^{(K_f-1)}) &= \hat{\text{cov}}(I(\boldsymbol{y} - \hat{\beta}_{1f,\tau}^{(K_f-1)}\hat{\boldsymbol{f}}_1^{(K_f-1)} - \dots - \hat{\beta}_{(K_f-1)f,\tau}^{(K_f-1)}\hat{\boldsymbol{f}}_{K_f-1}^{(K_f-1)} > 0), \boldsymbol{g}_j)\\
&= \hat{\text{cov}}(I(\boldsymbol{r}_{K_f-1,\tau} > 0), \boldsymbol{g}_j) = 0,
\end{aligned}$$

where q$\hat{\text{cov}}$ and qp$\hat{\text{cov}}$ denote the sample quantile covariance and the sample quantile partial covariance, respectively. It can also be shown that the proxies $\boldsymbol{r}_{t,\tau}^{*}$ are linearly independent. Specifically, the automatic proxies generated by Qcov3PRF are spanned precisely by the set of (relevant) factors that affect the conditional quantile of the target variable, as is stated in Section 4, Theorem 3. Therefore, Assumption 2 is satisfied by replacing $\boldsymbol{z}_{t,\tau}^{*}$ with $\boldsymbol{r}_{t,\tau}^{*} = (r_{0t,\tau}^{*}, \dots, r_{(K_f-1)t,\tau}^{*})$.

3.3 Quantile-covariance Three Pass Regression Filter for mixed frequency data

Similar to 3PRF, Qcov3PRF usually allows more tractable extensions since each pass involves least squares regressions. Now, we present a simple extension of Qcov3PRF for mixed frequency (MF) data inspired by Hepenstrick and Marcellino, 2019. Let us assume that the target variable and proxies are low frequency whereas the predictors are high frequency.[7] In terms of notation, assume that the covariates $\boldsymbol{x}_t$ can be observed for each t, whereas the target variable y_t and the proxies $\boldsymbol{z}_t$ can be observed only every h periods. For example, when $h = 3$ low frequency is quarterly (e.g. GDP growth), and high frequency is monthly (e.g. Industrial Production, prices, financial variables). The aggregate (low) frequency is indicated by s. An approach for Qcov3PRF that deals with this type of mixed frequency data is presented in Algorithm 5.

[7] Cases where the predictors are low frequency and the target and proxies are high frequency, and ragged edges presented in Hepenstrick and Marcellino, 2019 can also be used in Qcov3PRF.

Algorithm 5 Quantile Three-Pass Regression Filter with low frequency (quarterly) target variable and/or proxies, and high frequency (monthly) predictors (MF-Qcov3PRF).

Pass 0: Let $\boldsymbol{z}^*_\ell = I(\boldsymbol{z}_\ell - \hat{Q}^\tau_{z_\ell} > 0)$, where $\hat{Q}^\tau_{z_\ell}$ is an estimator for the unconditional quantile of $\boldsymbol{z}_\ell$ for $\ell = 1, \ldots, K_f$.

Pass 1: Run time series least squares regressions, in low frequency s, of $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{iT})$ on $\boldsymbol{z}^*_\tau = (\boldsymbol{z}^{*'}_{1,\tau}, \ldots, \boldsymbol{z}^{*'}_{T/3,\tau})'$ ($\boldsymbol{z}^*_{s,\tau} = (z^*_{1s,\tau}, \ldots, z^*_{K_f s,\tau})$, $s = 1, \ldots, \lfloor T/3 \rfloor$) for $i = 1, \ldots, N$,

$$x_{is} = \phi_{0,i} + \boldsymbol{\phi}'_i \boldsymbol{z}^*_{s,\tau} + \epsilon_{is}, \quad \text{for} \quad s = 1, \ldots, \lfloor T/3 \rfloor$$

where $x_{is} = \left(x_{i(3s-2)} + x_{i(3s-1)} + x_{i(3s)}\right)$, and retain the estimate $\hat{\boldsymbol{\phi}}_{i,\tau}$.

Pass 2: Run cross section least squares regressions of $\boldsymbol{x}_t = (x_{1t}, \ldots, x_{Nt})$ on $\hat{\boldsymbol{\phi}}_\tau = (\hat{\boldsymbol{\phi}}'_{1,\tau}, \ldots, \hat{\boldsymbol{\phi}}'_{N,\tau})'$ ($\hat{\boldsymbol{\phi}}_{i,\tau} = (\hat{\phi}_{1i,\tau}, \ldots, \hat{\phi}_{K_f i,\tau})$, $i = 1, \ldots, N$) for $t = 1, \ldots, T$,

$$x_{it} = \phi_{0,t,\tau} + \boldsymbol{f}'_t \hat{\boldsymbol{\phi}}_{i,\tau} + \varepsilon_{it},$$

and retain the estimate $\hat{\boldsymbol{f}}_t$ for each month $t = 1, \ldots, T$.

Pass 3: Split the estimated monthly factors $\hat{\boldsymbol{f}}_t = (\hat{f}_{1t}, \ldots, \hat{f}_{K_f t})$ into three quarterly factors $(\bar{\boldsymbol{f}}^1_s, \bar{\boldsymbol{f}}^2_s, \bar{\boldsymbol{f}}^3_s)$, where $\bar{\boldsymbol{f}}^j_s = \hat{\boldsymbol{f}}_{3(s-1)+j}$ and $\bar{\boldsymbol{f}}^j_s = (\bar{f}^j_{1s}, \ldots, \bar{f}^j_{K_f s})$ for $j = 1, 2, 3$. Then, run a time series quantile regression of y_{s+1} on $\left(\bar{\boldsymbol{f}}^1_s + \bar{\boldsymbol{f}}^2_s + \bar{\boldsymbol{f}}^3_s\right)$:

$$y_{s+1} = \beta_{0,\tau} + \boldsymbol{\beta}'_{f,\tau} \left(\bar{\boldsymbol{f}}^1_s + \bar{\boldsymbol{f}}^2_s + \bar{\boldsymbol{f}}^3_s\right) + u^\tau_{s+1}. \tag{21}$$

This gives the forecast for the conditional τ quantile of y_{t+1} denoted by $\hat{Q}^\tau_{y_{s+1}|\hat{\boldsymbol{f}}_t}$.

Compared to Algorithm 3, the passes that involve time series regressions are modified to allow mixed frequency (Pass 1 and Pass 3). Now, there are three sets of relevant factors $\bar{\boldsymbol{f}}_s^1$, $\bar{\boldsymbol{f}}_s^2$ and $\bar{\boldsymbol{f}}_s^3$ corresponding to the months in each quarter. We do not regress the target on $(\bar{\boldsymbol{f}}_s^1, \bar{\boldsymbol{f}}_s^2, \bar{\boldsymbol{f}}_s^3)$ separately since, by Assumptions 1 and 2, for Qcov3PRF (analogously to 3PRF) that regression will not result in consistent forecasts. This can be easily seen if we consider, say, the matrices $\boldsymbol{A}_1, \boldsymbol{A}_2$ and $\boldsymbol{A}_3$ where $\boldsymbol{A}_j$ is a matrix $\lfloor T/3 \rfloor \times T$, where $\lfloor \cdot \rfloor$ denotes the floor integer, with $\boldsymbol{A}_j = [\boldsymbol{\iota}_j; \boldsymbol{\iota}_{j+3}; \ldots; \boldsymbol{\iota}_{T-3+j}]$ for $j = 1, 2, 3$, where $\boldsymbol{\iota}_j$ is a $1 \times T$ vector equal to one in the jth entry and zero everywhere else. Then, from Eq.(11) we have:

$$\boldsymbol{X}\bar{\boldsymbol{A}}' = \boldsymbol{\phi}_0 \bar{\boldsymbol{A}}' + \boldsymbol{\Phi}' \boldsymbol{F} \bar{\boldsymbol{A}}' + \boldsymbol{\varepsilon} \bar{\boldsymbol{A}}',$$

where $\bar{\boldsymbol{A}} = \boldsymbol{A}_1 + \boldsymbol{A}_2 + \boldsymbol{A}_3$. Instead of Assumption 2, we have the following DGP:

$$\boldsymbol{Z}_\tau^* = \boldsymbol{\lambda}_{0,\tau} + \boldsymbol{\Lambda}_\tau' \boldsymbol{F} \bar{\boldsymbol{A}}' + \boldsymbol{\omega}.$$

These two DGP's are consistent with Pass 1 by running $\boldsymbol{X}_s$ on $\boldsymbol{z}_{s,\tau}^*$ since:

$$\begin{aligned} \boldsymbol{X}\bar{\boldsymbol{A}}' &= \boldsymbol{\psi}_0 + \boldsymbol{\Psi}' \boldsymbol{Z}_\tau^* + \boldsymbol{\epsilon} \\ &= \boldsymbol{\psi}_0 + \boldsymbol{\Psi}' \left[\boldsymbol{\lambda}_{0,\tau} + \boldsymbol{\Lambda}_\tau' \boldsymbol{F} \bar{\boldsymbol{A}}' + \boldsymbol{\omega}_\tau \right] + \boldsymbol{\epsilon} \\ &= \left(\boldsymbol{\psi}_0 + \boldsymbol{\Psi}' \boldsymbol{\lambda}_{0,\tau} \right) + \boldsymbol{\Psi}' \boldsymbol{\Lambda}_\tau' \boldsymbol{F} \bar{\boldsymbol{A}}' + \left(\boldsymbol{\Psi}' \boldsymbol{\omega}_\tau + \boldsymbol{\epsilon} \right), \end{aligned}$$

which is identifiable with Assumption 1 by making $\boldsymbol{\phi}_0 \bar{\boldsymbol{A}}' = \left(\boldsymbol{\psi}_0 + \boldsymbol{\Psi}' \boldsymbol{\lambda}_{0,\tau} \right)$, $\boldsymbol{\phi}_f \boldsymbol{f} = \boldsymbol{\Phi}' \boldsymbol{\Lambda}_\tau' \boldsymbol{F}$ where $\boldsymbol{\Phi} = [\boldsymbol{\phi}_f \;\; \boldsymbol{\phi}_g]$ and $\boldsymbol{\phi}_f$ ($\boldsymbol{\phi}_g$) denotes the corresponding loadings for $\boldsymbol{f}$ ($\boldsymbol{g}$), and $\boldsymbol{\varepsilon}\bar{\boldsymbol{A}}' + \boldsymbol{\phi}_g \boldsymbol{g} \bar{\boldsymbol{A}}' = \left(\boldsymbol{\Psi}' \boldsymbol{\omega}_\tau + \boldsymbol{\epsilon} \right)$. Matrices $\boldsymbol{A}_1$, $\boldsymbol{A}_2$ and $\boldsymbol{A}_3$ need to be considered jointly as a sum, otherwise, Pass 1 will not lead consistent forecasts. In other words, the aggregation scheme of the low frequency predictors x_{is} are generated and used in Pass 1 should be the same to the one we consider for the proxy $\boldsymbol{z}_s$ and $(\bar{\boldsymbol{f}}_s^1, \bar{\boldsymbol{f}}_s^2, \bar{\boldsymbol{f}}_s^3)$ in Pass 3 (since the set of factors becomes $\boldsymbol{F}\bar{\boldsymbol{A}}'$). In Algorithm 5 we consider the simple aggregation of the three months in each quarter. Here, we are assuming that all monthly predictors are flow variables so we transform them into quarterly frequency via aggregation. Pass 1 and Pass 3 should be modified accordingly being consistent with the same transformation. For example, with stock variables we should consider only the observation every three months.

Remark 1. *By considering that $\boldsymbol{X}$ follows a factor structure at s frequency, i.e., $\boldsymbol{x}_s = \boldsymbol{\phi}_0 + \boldsymbol{\Phi} \boldsymbol{F}_s + \boldsymbol{\varepsilon}_s$, we can replace Pass 3 with:*

$$y_{s+1} = \beta_{0,\tau} + \boldsymbol{\beta}_{1f,\tau}' \bar{\boldsymbol{f}}_s^1 + \boldsymbol{\beta}_{2f,\tau}' \bar{\boldsymbol{f}}_s^2 + \boldsymbol{\beta}_{3f,\tau}' \bar{\boldsymbol{f}}_s^3 + u_{s+1}^\tau,$$

as was considered in Hepenstrick and Marcellino, 2019. We prefer the initial Pass 3 in Algorithm 5 since *imposing factor structure at the frequency s for the predictors is more restrictive in practice.*

4 Assumptions and asymptotic results

In addition to Assumptions 1, 2, and 3, we state the following assumptions needed for the estimation of Qcov3PRF. These assumptions come from standard factor models estimation, 3PRF and quantile regression.

Assumption 4. *(Prediction Stage)*

1. *Let* $0 < m < M < \infty$. *The conditional density of* u_{t+1} *given* $\boldsymbol{F}_t$ *denoted as* $h_\tau(u_{t+1}|\boldsymbol{F}_t)$*:*
   (a) *is continuous.*
   (b) $m \leq h_\tau(u_{t+1}|\boldsymbol{F}_t) \leq M$ *for all* t.
   (c) $h_\tau(u_{t+1}|\boldsymbol{F}_t)$ *is Lipchitz continuous, i.e.,* $|h_\tau(u|\boldsymbol{F}_t) - h_\tau(u'|\boldsymbol{F}_t)| \leq M|u - u'|$ *for all* t.
2. $\mathbb{P}\left[u_{t+1} \leq 0|y_t, \boldsymbol{F}_t, y_{t-1}, \boldsymbol{F}_{t-1}, \ldots\right] = \tau$.
3. $T^{-1}\boldsymbol{F}'\boldsymbol{J}_T\boldsymbol{F}h_\tau(0|\boldsymbol{F}) \xrightarrow[T\to\infty]{p} \boldsymbol{\Delta}_{F,f}$.

Assumption 4.1 allows for a bias representation of the quantile regression in the prediction stage following Angrist et al., 2006. This permits us to show consistency of the conditional quantile forecasts. Assumption 4.2 corresponds to quantile regression in Eq.(12). Assumption 4.3 is required to show asymptotic normality in quantile regression.

Assumption 5. *(Factors, Loadings and Residuals). Let* $M < \infty$. *For any* i, s, t

1. $\mathbb{E}\|\boldsymbol{F}_t\|^4 < M, \quad T^{-1}\sum_{s=1}^T \boldsymbol{F}_s \xrightarrow[T\to\infty]{p} \boldsymbol{\mu}_F, \quad T^{-1}\boldsymbol{F}'\boldsymbol{J}_T\boldsymbol{F} \xrightarrow[T\to\infty]{p} \boldsymbol{\Delta}_F,$
2. $\mathbb{E}\|\boldsymbol{\phi}_i\|^4 \leq M, \quad N^{-1}\sum_{j=1}^N \boldsymbol{\phi}_j \xrightarrow[N\to\infty]{p} \boldsymbol{\mu}_\Phi, \quad N^{-1}\boldsymbol{\Phi}'\boldsymbol{J}_N\boldsymbol{\Phi} \xrightarrow[N\to\infty]{p} \boldsymbol{\Delta}_\Phi, \quad N^{-1}\boldsymbol{\Phi}'\boldsymbol{J}_N\boldsymbol{\phi}_0 \xrightarrow[N\to\infty]{p} \boldsymbol{\Delta}_{1,\Phi},$
3. $\mathbb{E}(\varepsilon_{it}) = 0, \qquad \mathbb{E}|\varepsilon_{it}|^8 \leq M,$
4. $\mathbb{E}\left(\boldsymbol{\omega}_t\right) = 0, \qquad \mathbb{E}||\boldsymbol{\omega}_t||^4 \leq M, \qquad T^{-1/2}\sum_{s=1}^T \boldsymbol{\omega}_s = \boldsymbol{O}_p(1), \qquad T^{-1}\boldsymbol{\omega}'\boldsymbol{J}_T\boldsymbol{\omega} \xrightarrow[T\to\infty]{p} \boldsymbol{\Delta}_\omega$
5. $\mathbb{E}(\psi_\tau(u_{t+1})^4) \leq M$, *and* u_{t+1} *is independent of* $\phi_i(m)$ *and* ε_{it}.

Assumptions 5.1-5.3 are the same considered by Bai and Ng, 2002, Bai, 2003 and Stock and Watson, 2002a. Assumption 5.4 is required in the 3PRF for conditional mean forecasts, and it is also required in the Qcov3PRF because this method uses proxies to extract the factors. The moments of the proxies noise $\boldsymbol{\omega}_t$ are bounded in the same manner as the bounds on factor moments. Assumption 5.5 is required to show consistency of estimators of factors and loadings, and asymptotic normality of the forecasts.

Assumption 6. *(Dependence) There exists a constant* $M < \infty$ *and for any* i, j, t, s, m_1, m_2

1. $\mathbb{E}(\varepsilon_{it}\varepsilon_{js}) = \sigma_{ij,ts}$, $|\sigma_{ij,ts}| \leq \bar{\sigma}_{ij}$, $|\sigma_{ij,ts}| \leq \bar{\sigma}_{ts}$, $N^{-1}\sum_{i,j=1}^{N} \bar{\sigma}_{ij} \leq M$, $T^{-1}\sum_{t,s=1}^{T} \bar{\sigma}_{ts} \leq M$, *and* $N^{-1}\sum_{i,s} |\sigma_{ij,ts}| \leq M$,

2. $\mathbb{E}|N^{-1/2}T^{-1/2}\sum_{s=1}^{T}\sum_{i=1}^{N} [\varepsilon_{is}\varepsilon_{it} - \mathbb{E}(\varepsilon_{is}\varepsilon_{it})]|^2 \leq M$,

3. $\mathbb{E}|T^{-1/2}\sum_{t=1}^{T} F_{m_1,t}\omega_{m_2,t}|^2 \leq M$,

4. $\mathbb{E}|T^{-1/2}\sum_{t=1}^{T} \omega_{m_1,t}\varepsilon_{it}|^2 \leq M$.

As in 3PRF, we allow some degree of cross sectional correlation and serial dependence in the factor structure through Assumptions 6.1-6.2. This follows the work of Chamberlain and Rothschild, 1982 and Stock and Watson, 2002a. In Assumptions 6.3-6.4 some proxy noise dependence with factors and idiosyncratic errors ε_{it} is allowed.

Assumption 7. *(Normalization and orthogonalization). For any* m_1, m_2*:*

1. $\boldsymbol{\Delta}_\Phi = I$, $\boldsymbol{\Delta}_{1,\Phi} = 0$,

2. $\boldsymbol{\Delta}_F$ *is diagonal, positive definite, and each diagonal element is unique.*

3. $\boldsymbol{f}_t$ *is independent of* $g_{j,t}$ *for* $j = K_f + 1, \ldots, K$.

Assumptions 7.1-7.2 give a unique representation of the latent factor and factor loadings in the same way as in Kelly and Pruitt, 2015. They select a normalization such that the covariance of predictor loadings is the identity matrix, and the factors are orthogonal to one another.

Assumption 8. *(Central Limit Theorems). For any* i, t

1. $\frac{1}{\sqrt{N}}\sum_{i=1}^{N} \boldsymbol{\phi}_i\varepsilon_{it} \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Sigma}_{\phi\varepsilon})$, *where* $\boldsymbol{\Sigma}_{\phi\varepsilon} = \text{plim}_{N\to\infty} \frac{1}{N}\sum_{i,j=1}^{N} \mathbb{E}\left[\varepsilon_{it}\varepsilon_{jt}\boldsymbol{\phi}_i\boldsymbol{\phi}_j'\right] > 0$,

2. $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} \boldsymbol{F}_t\varepsilon_{it} \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Sigma}_{F\varepsilon})$, *where* $\boldsymbol{\Sigma}_{F\varepsilon} = \text{plim}_{T\to\infty} \frac{1}{T}\sum_{t,s=1}^{T} \mathbb{E}\left[\varepsilon_{it}\varepsilon_{is}\boldsymbol{F}_t\boldsymbol{F}_s'\right] > 0$,

3. $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} \boldsymbol{F}_t\psi_\tau(u_{t+1}) \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Sigma}_{F\psi})$, *where* $\boldsymbol{\Sigma}_{F\psi} = \text{plim}_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}\left[\psi_\tau^2(u_{t+1})\boldsymbol{F}_t\boldsymbol{F}_t'\right] > 0$.

Assumption 8 states the central limit theorems needed to show asymptotic normality of the estimates. Next, we obtain that as N and T become large, the conditional quantile forecasts determined with Qcov3PRF converges to the infeasible best forecasts. This result is presented in Theorem 1, whose proof makes use of the convergence rates from Lemma 1, quasi-maximum likelihood results, equivariance properties of the quantile regression, and the weighted (quantile) regression from Angrist et al., 2006.

Theorem 1. *Let Assumptions 1-2 and 4-8 hold. Then, the forecast in the predictive stage Eq.(15) in Pass 3 of Qcov3PRF satisfies*

$$\hat{Q}^{\tau}_{y_{t+1}|\hat{\boldsymbol{F}}_t} = \hat{\beta}_{0,\tau} + \hat{\boldsymbol{\beta}}_\tau'\hat{\boldsymbol{F}}_t \xrightarrow[N,T\to\infty]{p} \boldsymbol{\beta}_{0,\tau} + \boldsymbol{\beta}_\tau'\boldsymbol{F}_t,$$

where $\hat{Q}^{\tau}_{y_{t+1}|\hat{\boldsymbol{F}}_t}$ *is the conditional* τ*-quantile forecast of* y_{t+1} *given the unobserved factors.*

Proof. See Appendix. □

To show the asymptotic normality of the infeasible quantile forecasts we need to make use of Lemma 2 (see Appendix 8.2), whose proof comes from Theorem 6 in Kelly and Pruitt, 2015 but now we consider the transformed proxies $\boldsymbol{z}^{*}_{t,\tau}$ instead of $\boldsymbol{z}_t$.

Theorem 2. *Let Assumptions 1-2 and 4-8 hold and* $\sqrt{N}/T \rightarrow 0$. *Then, the forecast in Pass 3 of Qcov3PRF satisfies,*

$$\hat{\boldsymbol{V}}_{\tau}^{-1/2}\left(\hat{\boldsymbol{\beta}}_{0,\tau} + \hat{\boldsymbol{\beta}}_{\tau}'\hat{\boldsymbol{F}}_t - \boldsymbol{\beta}_{0,\tau} - \boldsymbol{\beta}_{\tau}'\boldsymbol{F}_t\right) \xrightarrow[N,T\to\infty]{d} \mathcal{N}(0,1), \tag{22}$$

where $\hat{\boldsymbol{V}}_{\tau} = \frac{1}{T}\hat{\boldsymbol{\beta}}_{\tau}'(\boldsymbol{\Sigma}_{\hat{F}})\hat{\boldsymbol{\beta}}_{\tau} + \frac{1}{N}\hat{\boldsymbol{F}}_t'(\boldsymbol{\Sigma}_{\hat{\beta}_{\tau}})\hat{\boldsymbol{F}}_t$.

Proof. See Appendix. □

Theorem 2 guarantees that the estimates for the infeasible forecasts obtained from Qcov3PRF are asymptotically normal. However, this is more a theoretical result than practical since, similar to 3PRF, Qcov3PRF only recovers the estimates for the relevant factors so $\boldsymbol{\Sigma}_{\hat{F}}$ can not be estimated. To get an estimate for the variance we can implement an approximation of the finite distribution (or variance) using, for example, block bootstrap.

Next, we state formally that Algorithm 4 under Assumption 3 provides proxies consistent for the conditional quantile forecasts. Note that all of our simulations and the first two empirical applications described in the next sections consider that the only available information is the set of predictors $\boldsymbol{x}_t$ and the target y_{t+1}.

Theorem 3. *Let Assumptions 1 and 4-8 hold with the exception of Assumptions 5.4, 6.3 and 6.4. Then the* K_f *automatic-proxies obtained in Algorithm 4 under Assumption 3 satisfy Assumptions 2, 5.4, 6.3 and 6.4 when the number of relevant factors is* K_f. *As a result, the* K_f*-automatic-proxy forecast is consistent according to Theorem 1 and asymptotically normal according to Theorem 2.*

Proof. See Appendix. □

Lastly, we present Corollary 1 that presents the consistency for the infeasible forecast using the MF-Qcov3PRF covered in Section 3.3. In this case, Assumption 9 replaces Assumptions 1 and 2 which assumes the target and proxies are generated with low frequency.

Assumption 9. *(Data Generating Processes for Mixed Frequency data) The data for a fixed level* $\tau \in$

(0, 1) *is generated as follows:*

$$\begin{aligned} \boldsymbol{x}_t &= \boldsymbol{\phi}_0 + \boldsymbol{\Phi}\boldsymbol{F}_t + \boldsymbol{\varepsilon}_t, & (23)\\ y_{s+1} &= \beta_{0,\tau} + \boldsymbol{\beta}_\tau' \left(\boldsymbol{F}_{3(s-1)+1} + \boldsymbol{F}_{3(s-1)+2} + \boldsymbol{F}_{3(s-1)+3}\right) + u_{s+1}^\tau, & (24)\\ \boldsymbol{z}_{s,\tau}^* &= \boldsymbol{\lambda}_{\tau,0} + \boldsymbol{\Lambda}_\tau \left(\boldsymbol{F}_{3(s-1)+1} + \boldsymbol{F}_{3(s-1)+2} + \boldsymbol{F}_{3(s-1)+3}\right) + \boldsymbol{\omega}_s, & (25)\end{aligned}$$

where $\boldsymbol{y} = (y_2, \ldots, y_{S+1})$ *is an* $S \times 1$ *vector denoting the target variable time series. In addition,* $\boldsymbol{z}_{s,\tau}^* = (\boldsymbol{z}_{s,1}^*, \ldots, \boldsymbol{z}_{s,K_f}^*)$ *with* $\boldsymbol{z}_{s,\ell}^* = I(z_{s,\ell} - Q_{z_{s,\ell}}^\tau > 0)$*,* $\boldsymbol{z}_s$ *is a* $K_f \times 1$ *vector of proxy data. Here* s *represents low frequency and* t *high frequency. For example,* $S = \lfloor T/3 \rfloor$ *when* s *denotes quarterly data and* t *monthly data. All else elements have the same interpretation and dimensions from Assumptions 1 and 2.*

In Assumption 9 we are assuming that the predictors are flow variables so that the monthly aggregation makes sense.

Corollary 1. *Let Assumptions 4-9 hold. Then, the forecast in the prediction stage Eq.(24) in Pass 3 of Qcov3PRF with mixed frequency satisfies*

$$\hat{Q}^\tau_{y_{s+1}|\hat{\boldsymbol{F}}_t} \xrightarrow[N,T\to\infty]{p} \boldsymbol{\beta}_{0,\tau} + \boldsymbol{\beta}_\tau' \left(\boldsymbol{F}_{3(s-1)+1} + \boldsymbol{F}_{3(s-1)+2} + \boldsymbol{F}_{3(s-1)+3}\right),$$

where $\hat{Q}^\tau_{y_{s+1}|\hat{\boldsymbol{F}}_t}$ *is the conditional* τ*-quantile forecast of* y_{s+1} *given the unobserved (high frequency) factors.*

It is straightforward to show the asymptotic normality of the forecast estimates obtained from Algorithm 5, similar to what we do in the proof of Theorem 2.

5 Monte Carlo simulations

In this section we present the forecasting power of Qcov3PRF and the consistency of the forecasts in finite samples through Monte Carlo simulations. For the forecasting accuracy we examine three DGPs based on the simulations of Kelly and Pruitt, 2015, Giglio et al., 2016 and Chen et al., 2021. The first and third DGPs consist on location scale models with dependent idiosyncratic errors, while the second DGP is a location model with a focus on the estimation of the conditional median. We run 1000 simulations in each case when $N = 100$, $T = 200$ and $N = 200$, $T = 400$, respectively. The factor loadings follow a standard normal distribution and are i.i.d. The structure of the predictors is $\boldsymbol{x}_t = \boldsymbol{\Phi}\boldsymbol{F}_t + \boldsymbol{\varepsilon}_t$. The idiosyncratic errors of the predictors allow the presence of cross sectional and serial correlation following the expression:

$$\boldsymbol{\varepsilon} = \boldsymbol{\Sigma}_N \bar{\boldsymbol{\varepsilon}} \boldsymbol{\Sigma}_T,$$

where $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_T)$ with $\boldsymbol{\varepsilon}_t$, $t = 1, \ldots, T$ a column vector of size N. Also, $\bar{\boldsymbol{\varepsilon}}$ is a $N \times T$ matrix where each row $\bar{\varepsilon}_i$, $i = 1, \ldots, N$ is standard normally distributed, and $\boldsymbol{\Sigma}_N$ is an $N \times N$ matrix whose entries are given by $\sigma_N^{i,j} = \rho_N^{|i-j|/2}$, for all $i = 1, \ldots, N$ and $j = 1, \ldots, N$. We determine $\boldsymbol{\Sigma}_T$ similarly such that $\sigma_T^{t,s} = \rho_T^{|t-s|/2}$, for all $t = 1, \ldots, T$ and $s = 1, \ldots, T$. Then, higher values of ρ_N and ρ_T indicate high levels of cross sectional and serial correlation in the idiosyncratic component, respectively.

We report the out-of-sample R^2_τ defined in Koenker and Machado, 1999 evaluating forecasts one period ahead starting at the middle of the sample which is defined as:

$$R^2_\tau = 1 - \frac{\Sigma_{t=T/2}^{T} \left(y_{t+1} - \hat{Q}^\tau_{y_{t+1}|\hat{\boldsymbol{F}}_t} \right) \left(\tau - I(y_{t+1} - \hat{Q}^\tau_{y_{t+1}|\hat{\boldsymbol{F}}_t} < 0) \right)}{\Sigma_{t=T/2}^{T} \left(y_{t+1} - \hat{Q}^\tau_{y_{t+1}} \right) \left(\tau - I(y_{t+1} - \hat{Q}^\tau_{y_{t+1}} < 0) \right)}, \tag{26}$$

where $\hat{Q}^\tau_{y_{t+1}}$ is an estimate for the τ unconditional quantile of y_t considering $\{y_1, \ldots, y_t\}$, and $\hat{Q}^\tau_{y_{t+1}|\hat{\boldsymbol{F}}_t}$ is a conditional quantile estimate for y_{t+1} given $\hat{\boldsymbol{F}}_t$ using the DGPs presented below.

The first DGP consists on the following location scale model:

$$y_{t+1} = f_{1,t} + f_{2,t} + f_{3,t} + (\sigma_y + g_{1,t})\eta_{t+1}, \tag{27}$$

where $\eta_{t+1} = \bar{\eta}_{t+1} - Q^\tau_{\bar{\eta}_{t+1}}$, with $Q^\tau_{\bar{\eta}_{t+1}}$ the τ unconditional quantile of $\bar{\eta}_{t+1}$, $\sigma_y = 0.50$, and $\bar{\eta}_{t+1} \sim \mathcal{N}(0, 1/\sqrt{2})$. In this DGP, we assume the predictors are determined by six factors, three of which are relevant and the other three are irrelevant for the conditional τ-quantile of the target. However, for any other conditional quantile different from τ there are four relevant factors ($f_{1,t}$, $f_{2,t}$, $f_{3,t}$ and $g_{1,t}$). This DGP is particularly interesting since in real life applications the same factors are not necessarily relevant/irrelevant for all the quantiles across the target's conditional distribution. The relevant factors, $f_{1,t}$, $f_{2,t}$, $f_{3,t}$, are independent normally distributed with standard deviations $\{1, 1.5, 2\}$ and the irrelevant factors, $g_{1,t}$, $g_{2,t}$, $g_{3,t}$, are independent uniform distributed with the lower bound equal to zero and the upper bound equal to $\{3, 4, 5\}$, respectively. We report the R^2_τ using three methods: Qcov3PRF with one, two and three factors (denoted as Qcov3PRF1, Qcov3PRF2, Qcov3PRF3, respectively), PQR (Giglio et al., 2016) (that only estimates one relevant factor), and PCQR (Giglio et al., 2016) with six PCA factors.

Table 1 shows that Qcov3PRF3 provides the highest predictive power over many cases. We can see that Qcov3PRF1 and PQR report similar R^2_τ values to each other and always lower than those of Qcov3PRF2 and Qcov3PRF3. These results suggest that both methods, PQR and Qcov3PRF1, estimate the same relevant factor, and the clear advantage of Qcov3PRF over PQR would appear when more than one relevant factor are estimated. As expected, Qcov3PRF1 provides slightly lower performance than PQR as Pass 1 requires the dummies (transformed proxies) generated in Pass 0, whereas PQR does not consider a transformation of the proxies. In addition, as a result of these transformations, PCQR6 performs slightly better for the cases where ρ_N and ρ_T are low. All these differences between methods decrease as N and T augment. The advantage of using Qcov3PRF is more evident when ρ_N and ρ_T are high. The reason is that as serial and cross-sectional correlations increase, the idiosyncratic component is wrongly seen as an additional factor(s), so considering only six factors for PCQR is not enough to capture $f_{1,t}$, $f_{2,t}$, $f_{3,t}$ (and $g_{1,t}$, $g_{2,t}$, $g_{3,t}$). In contrast, Qcov3PRF performs well even when ρ_N and ρ_T are high as it focuses on only estimating the relevant factors.

In general, the out-of-sample R^2_τ obtained with Qcov3PRF3 is only marginally different compared to Qcov3PRF1. The reason is possibly related to the results obtained in Ahn and Bae, 2021. These authors found that there is only one relevant factor using PLS in cases where the variance for the relevant factors is the same, and even in cases where the relevant factors variances are different, additional relevant factors provide only marginal increments in forecasting power.

[Table 1 about here.]

The second DGP considers the following location model:

$$y_{t+1} = f_{1,t} + f_{2,t} + \sigma_y \eta_{t+1}, \tag{28}$$

where $\eta_t = B_t \mathcal{N}(0,1) + (1 - B_t)\text{Cauchy}(0,1)$, $B_t \sim \text{Bernoulli}(0.80)$. The value of σ_y is equal to 0.75. In this case the expected value of the target does not exist, since part of η_{t+1} consists of a Cauchy distribution, but the median does. Hence, the use of methods that focus on the estimation of the conditional median becomes crucial as methods including PCR and 3PRF are not valid. In DGP (28), we assume the predictors are determined by four factors $\boldsymbol{F}_t = (\boldsymbol{f}_{1,t}, \boldsymbol{f}_{2,t}, \boldsymbol{g}_{1,t}, \boldsymbol{g}_{2,t})$, two of which are relevant ($\boldsymbol{f}_{1,t}$ and $\boldsymbol{f}_{2,t}$) and the other two are irrelevant ($\boldsymbol{g}_{1,t}$ and $\boldsymbol{g}_{2,t}$). We use the same methods used in DGP (27) plus PCR with four factors (denoted as PCR4) and 3PRF with two relevant factors (denoted as 3PRF2). We also vary the level of cross-sectional and serial correlations, and we additionally report the cases where the factors are AR(1) processes with standard normal distributed i.i.d. errors, with standard deviations $\{1, 1.5, 2.0, 2.5\}$, and autorregressive parameters α^f, $\alpha^g \in \{0, 0.3, 0.6\}$ for relevant and irrelevant factors, respectively.

In Table 2 we see that Qcov3PRF2 gives the best forecasts on average. Similar to the previous DGP, Qcov3PRF1 and PQR report similar values to each other and always below Qcov3PRF2, with Qcov3PRF1 having slightly lower performance than PQR. Also, PCQR4 performs slightly better for the cases where ρ_N and ρ_T are low. These differences drop as N and T increase. Qcov3PRF performs substantially better when ρ_N and ρ_T are high. When α^f is high (and α^g is low), it is easier to distinguish the relevant factors from the irrelevant ones resulting in greater R^2_τ, the opposite happens when α^g is high (and α^f is low) as the irrelevant factors are more dominant. PCR and 3PRF do not give the best performance as the unconditional mean does not exist in DGP (28).

[Table 2 about here.]

For the third DGP, we refer to as a heteroskedastic location scale model. Specifically,

$$y_{t+1} = 2f_{1,t} + (\sigma_y + |f_{2,t}|)\eta_{t+1}, \tag{29}$$

where $\eta_t \sim \mathcal{N}(0,1)$ i.i.d. and $\sigma_y = 1.0$. In DGP (29), we also assume the predictors are determined by four factors as in the previous DGP, two of which are relevant and the other two are irrelevant, with standard deviations equal to $\{1.5, 1.0, 2.0, 2.5\}$, respectively. Simulation results are reported in Table 3. We can see that the highest R^2_τ comes from using Qcov3PRF with two relevant factors (Qcov3PRF2), showing similar patterns as in DGPs (27) and (28).

[Table 3 about here.]

To check the consistency of the infeasible forecast $\hat{\beta}_{0,\tau} + \hat{\beta}'_\tau \hat{\boldsymbol{F}}_t$ in finite samples, we report the Mean Absolute Error (MAE), the Mean Squared Error (MSE) and standard correlation between the true conditional quantile of $Q^\tau_{y_{t+1}|\boldsymbol{F}_t}$ and the estimated conditional quantile $\hat{Q}^\tau_{y_{t+1}|\hat{\boldsymbol{F}}_t}$. We examine the following location-scale model:

$$\begin{aligned} y_{t+1} &= -f_{1,t} - 0.5f_{2,t} + (0.5f_{2,t} + f_{3,t})\eta_{t+1}, && (30)\\ \boldsymbol{x}_t &= \boldsymbol{\phi}'\boldsymbol{f}_t + \boldsymbol{\psi}'\boldsymbol{g}_t + \boldsymbol{e}_t, && (31)\end{aligned}$$

where $e_{it} \sim \mathcal{N}(0,1)$, $\phi_i \sim \mathcal{N}(0,1)$, $\psi_i \sim \mathcal{N}(0,1)$, $\eta_t \sim \mathcal{N}(0,1)$ i.i.d. We consider two data generating processes. The first alternative (31.a), where $f_{1,t} \sim U[0,1]$, $f_{2,t} \sim U[0,2]$, $f_{3,t} \sim U[0,3]$, and $g_{i,t} \sim N(0,1)$, $i = 1,2,3$. And the second alternative (31.b), where $f_{1,t} \sim U[0,1]$, $f_{2,t} \sim U[0,2]$, $f_{3,t} \sim U[0,3]$ and $g_{i,t}$ is normally skewed distributed i.i.d. with location, scale and skewness parameters equal to 0, σ_i and 100, respectively, with $\sigma_i \in \{1.25, 1.5, 1.75\}$ for $i = 1,2,3$. Hence, the conditional quantile of y_{t+1} is given by $Q^\tau_{y_{t+1}|\boldsymbol{F}_t} = \beta_1 f_{1,t} + \beta_2 f_{2,t} + \beta_3 f_{3,t}$, where $\beta_1 = -1$, $\beta_2 = -0.5 + 0.5 \times Q^\tau_{\eta_{t+1}}$ and $\beta_3 = Q^\tau_{\eta_{t+1}}$.

In Table 4 we see that the MAE and the MSE approach to 0, and the correlation approaches to 1 as T and N increase for both alternatives in DGP (30)-(31), regardless of which quantile we focus on. Although the MAE and MSE when using Qcov3PRF for three relevant factors (denoted as Qcov3PRF3) are not the highest, they are very close to the highest ones which come from using PCQR with six factors (denoted as PCQR6). PCQR6 serves as a benchmark since we know by Ando and Tsay, 2011 and Giglio et al., 2016 that the forecasts obtained with PCQR are consistent. By looking at the results obtained from PQR and Qcov3PRF1 in Table 4, we see that the MAE and the MSE converge to zero as N and T increase. Similarly, we notice a convergence of the correlation towards one as N and T increase. However, as the true number of relevant factors is equal to three, the MAE, MSE and the correlation provide better estimates using Qcov3PRF3 and PCQR6, which indicates the inconsistency of PQR and Qcov3PRF1.[8]

[Table 4 about here.]

6 Applications to forecasting macroeconomic vulnerability and climate change

In this section we show that Qcov3PRF outperforms competitive alternatives in three different applications. The first two applications take into account the relationship between financial indicators (e.g., the variables that construct the National Finance Condition Index (NFCI, Federal Reserve Bank of Chicago)) and the conditional distribution of the economic activity (e.g., GDP growth). The connection between the financial and economic sector focused on the left tail conditional quantiles of the latter is often referred to as Growth at Risk (GaR, Adrian et al., 2019). The third application studies the relationship between the carbon dioxide emissions and the conditional distribution of the global temperature changes, with a particular focus on the right tail quantiles. We called such effect Climate at Risk (CaR).

6.1 Forecasting Monthly GaR Using the NFCI Components

For the first empirical application we consider the approach of Adrian et al., 2019, Adams et al., 2021 and Adrian et al., 2022. In these works the authors obtained significant relationships between macroeconomic indicators (Real GDP growth, the consensus forecasts of Real GDP growth, Unemployment rate and Inflation) and the financial sector. Specifically, there is a significant effect of the financial variables represented by the NFCI on predicting macroeconomic variables in the left tail of the distribution. The NFCI is a weighted index constructed with 107 financial variables and is published monthly by the Federal Reserve Bank of Chicago. Instead of considering a simple quantile regression of the Real GDP annual growth on its lags and the NFCI (as done in Adrian et al., 2019), we run a quantile regression with the relevant factors obtained from the set of indicators that construct the NFCI. This is motivated by the fact that the relevant factor(s) required to predict the left tail of the distribution for the macroeconomic indicator can be different than the ones that best predict the right tail. The forecast horizons considered are three months and twelve months ahead.

The NFCI index and the indicators used to construct it can be downloaded from the webpage https://www.chicagofed.org/publications/nfci/index. We also make use of the Risk, Credit, Leverage and Nonfinancial Leverage subindexes of the NFCI, which consider subsets of the 107 variables. We use

[8] Indeed, we know from Kelly and Pruitt, 2015 that the forecast estimates with one relevant factor when $K_f \geq 2$ are inconsistent albeit they result in in-sample and out-of-sample R^2_τ close to the measures obtained with the true value of K_f. The only case where one relevant factor forecasts are consistent is under the knife-edge case. This case is where the variance of the relevant factor and factor loadings are the same for all the factors.

the Index for Industrial Production (IP) activity, which can be downloaded from the FRED database of the Federal Reserve Bank of St. Louis, as a monthly approximation of the Real GDP. The IP growth is computed as the natural logarithm of the IP in the current period minus the natural logarithm of the IP h months before for $h \in \{3, 12\}$. Our sample period is from 2007M06 to 2023M12 resulting in 199 months.

6.1.1 Out-of-sample evaluation

Tables 5 and 6 contain the out-of-sample R^2_τ with IP growth as the dependant variable for three and twelve months ahead, respectively. We present the results for Qcov3PRF and PCQR for up to three factors, PQR and PCQR with LASSO implemented in a second step to select the latent factors (out of 40 factors). We also report the R^2_τ when the predictors are the NFCI, the Risk, Credit, Leverage and Nonfinancial Leverage, respectively. We consider both rolling and expanding window schemes starting in August 2011 and October 2015. The rolling window scheme consists of 50 observations. Quantiles in the tails are considered with $\tau \in \{0.05, 0.10, 0.90, 0.95\}$ as well as quantiles near the center with $\tau \in \{0.25, 0.50, 0.75\}$ following Adrian et al., 2019.

Overall, we can see greater forecasting power of Qcov3PRF compared to other alternative methods, particularly for IP growth 3-month ahead. The predetermined indexes perform poorly, with significant power concentrated in the left tail using the expanding window scheme. Qcov3PRF exhibits a robust performance across different quantiles, forecast horizons, and estimation window schemes. In contrast, PCQR results in several negative R^2_τ in the center and right tail of the distribution.

[Table 5 about here.]

[Table 6 about here.]

6.1.2 In-sample prediction of the conditional distribution

The use of Qcov3PRF can be relevant not only for forecasting but for in-sample prediction as the estimated relevant factors (being referred as indices) may condense important information of the economic/financial activity for an specific forecast target (e.g., Real GDP growth).

In our GaR application, Figure 1 illustrates the in-sample prediction for various conditional quantiles of IP growth three and twelve months ahead. The estimates obtained after applying Qcov3PRF (with 3 factors) are consistent with Adrian et al., 2019 findings. The indicators of financial conditions have a stronger effect on predicting the left tail of the economic activity conditional distribution. Specifically, 5% quantile prediction exhibits more variability compared to 95% quantile estimation in both forecast horizons. This pattern is also found using the qunatile regression with NFCI 5% quantile and 95% quantile predictions, but the red and blue lines only show significant variability in the financial crisis of 2008 and COVID-pandemic episodes. Regarding the methods PCQR3 and PQR, we can not notice significant differences in variability between the 5% quantile and 95% quantile predictions, the estimates only mirror the blue and red lines except for the case with PCQR3 and $h = 3$. Table 7 indicates the greater accuracy using our method compared to the alternatives giving the lowest value in the check loss, most apparently in both tails.

[Figure 1 about here.]

[Table 7 about here.]

Once we have estimated the conditional quantiles of IP growth, we follow the two-step procedure proposed by Adrian et al., 2019 to estimate the conditional parametric distribution of the target. In the first step, the conditional quantiles of the target variable are estimated, and in the second step the parameters that describe the conditional distribution are estimated by fitting the conditional quantiles according to a quadratic loss function. This step requires to assume an specific parametric form of the conditional distribution for the target variable. Specifically, this method considers the following optimization problem:

$$\min_{\boldsymbol{\theta}} \sum_{\tau} \left(\hat{Q}^{\tau}_{y_{t+1}|\hat{\boldsymbol{F}}_t} - G^{-1}(\tau|\boldsymbol{F}_t;\boldsymbol{\theta}) \right)^2, \tag{32}$$

where $G^{-1}(\tau|\boldsymbol{F}_t)$ are the quantiles corresponding to the the conditional distribution of y_{t+1} given $\boldsymbol{F}_t$, denoted as $G(y_{t+1}|\boldsymbol{F}_t)$. Adrian et al., 2019 assumed that the conditional density of y_{t+1} has the form of a skewed t-distribution :

$$g(y;\mu,\sigma,\alpha,\nu) = \frac{2}{\sigma} t\left(\frac{y-\mu}{\sigma};\nu\right) T\left(\alpha \frac{y-\mu}{\sigma}\sqrt{\frac{\nu+1}{\nu+\left(\frac{y-\mu}{\sigma}\right)^2}};\nu+1\right), \tag{33}$$

where $t(\cdot)$ and $T(\cdot)$ denote the PDF and CDF of the Student t-distribution, respectively. Then, the estimates $\hat{\boldsymbol{\theta}} = (\hat{\mu}_{t+1}, \hat{\sigma}_{t+1}, \hat{\alpha}_{t+1}, \hat{\nu}_{t+1})$, where α is the skewness parameter and ν is the degrees of freedom parameter determining the kurtosis, are obtained by solving the problem (32) with the distribution (33), where the subscripts in the estimated parameters indicate dependence on $\boldsymbol{F}_t$, and $\hat{\mu}_{t+1} \in \mathbb{R}$, $\hat{\sigma}_{t+1} \in \mathbb{R}^+$, $\hat{\alpha}_{t+1} \in \mathbb{R}$, and $\hat{\nu}_{t+1} \in \mathbb{Z}^+$. The predictions correspond to the quantiles $\tau \in \{0.05, 0.25, 0.75, 0.95\}$ to exactly identify $\hat{\boldsymbol{\theta}}$.

Figure 2 shows the estimated in-sample conditional distribution with the conditional quantiles obtained with NFCI (quantile regression), PCQR, PQR and Qcov3PRF for three and twelve months ahead. Overall, there is more heterogeneity in the conditional distribution estimated with Qcov3PRF over time. The plot confirms the findings from Qcov3PRF estimation shown in Figure 1, which exhibits more variability in the left tail of the distribution over time.

[Figure 2 about here.]

Next, for robustness in our findings we calculate the conditional expected shortfall (ES), a popular risk measure, and the expected longrise (EL) as in Adrian et al., 2019. Both measures can be calculated from fitting the conditional quantiles into a skew-t distribution as described above in Eq.(33). Specifically, for a chosen τ probability the ES and EL are defined as:

$$\begin{aligned} \mathrm{ES}_{t+1}^{\tau}(\boldsymbol{F}_t) &= \mathbb{E}[y_{t+1}|y_{t+1} \leq \mathrm{GaR}_{\tau}(y_{t+1}|\boldsymbol{F}_t)], \\ \mathrm{EL}_{t+1}^{1-\tau}(\boldsymbol{F}_t) &= \mathbb{E}[y_{t+1}|y_{t+1} \geq \mathrm{GaR}_{1-\tau}(y_{t+1}|\boldsymbol{F}_t)], \end{aligned}$$

where $\mathrm{GaR}_{\tau}(y_{t+1}|\boldsymbol{F}_t)$ denotes the conditional τ-quantile of y_{t+1} given $\boldsymbol{F}_t$. Figure 3 presents the in-sample IP growth ES at $\tau = 5\%$ level and EL at $1-\tau = 95\%$ level, estimated with NFCI, PCQR, PCQR and Qcov3PRF. We see that the results from using Qcov3PRF, again, show the highest heterogeneity between the expected shortfall and the expected longrise. In particular, the expected longrise is flatter and shows less variation than the expected shortfall.

[Figure 3 about here.]

For the in-sample evaluation of the expected shortfall and longrise we make use of the FZ loss function proposed by Fissler and Ziegel, 2016 for the case where the loss function is homogeneous of degree zero (denoted as FZ0) as in Patton et al., 2019. This is defined as:

$$L_{FZ0}(y_{t+1}, Q^{\tau}_{y_{t+1}|\boldsymbol{F}_t}, \mathrm{ES}^{\tau}_{t+1}; \tau) = -\frac{1}{\tau \mathrm{ES}^{\tau}_{t+1}} I(y_{t+1} \leq Q^{\tau}_{y_{t+1}})(Q^{\tau}_{y_{t+1}|\boldsymbol{F}_t} - y_{t+1}) + \frac{Q^{\tau}_{y_{t+1}|\boldsymbol{F}_t}}{\mathrm{ES}^{\tau}_{t+1}} + \log(-\mathrm{ES}^{\tau}_{t+1}) - 1, \tag{34}$$

where ES^{τ}_{t+1} is the τ-conditional expected shortfall. Since this loss function is only valid for positive values of the expected shortfall, we rescale y_{t+1}, $Q^{\tau}_{y_{t+1}|\boldsymbol{F}_t}$ and ES^{τ}_{t+1} by a constant.[9] Table 8 reports the in-sample mean of the loss given in Eq.(34) across alternatives. The results indicate better prediction in both, the expected shortfall and longrise, when using the conditional quantiles obtained by Qcov3PRF. This suggests that the larger effects of the financial indicators towards downside economic risks, as can be seen in Figure 3, are consistent with the financial conditions indices (factors) that best predict the conditional distribution of the real activity growth (i.e., the correlation between ES_{τ}, and the relevant factors and the NFCI is higher in absolute terms compared to the correlation between EL_{τ} and the indices/factors).

[Table 8 about here.]

[9]For the case of expected longrise we use Eq.(34) evaluated in $-y_{t+1}$, $-Q^{\tau}_{y_{t+1}|\boldsymbol{F}_t}$ and $-\mathrm{ES}^{\tau}_{t+1}$, i.e., we mirror the conditional distribution of y_{t+1} such that the $1-\tau$ expected longrise of y_{t+1} is the τ expected shortfall of $-y_{t+1}$.

6.2 Forecasting Quarterly GaR Using Financial Risk Measures through mixed frequency data

As a second empirical application we study the case where the conditional distribution of the Real GDP growth is affected by financial variables. We make use of the financial risk measures (16 variables) considered in Giglio et al., 2016 to illustrate that GDP growth is more affected by an index of these measures instead of adding all of them explicitly. This index is the most relevant factor of the financial variables on the annual Real GDP growth one-quarter ahead, estimated by applying Qcov3PRF. In addition, we use the Chicago Fed National Activity Index (CFNAI) as a proxy variable. This follows from the evidence found in Giglio et al., 2016 regarding the significance of the relevant factor obtained through PQR using these 16 financial risk measures as predictors on the (quarterly) shocks from the CFNAI. In particular, the significant effect is on the left tail of the distribution, i.e. there is presence of GaR. The financial risk measures can be downloaded from www.sethpruitt.net/GKPwebdata.zip, and are based on data for financial institutions identified by two-digit SIC codes 60 through 67 (finance, insurance, and real estate). The CFNAI is constructed by the Federal Reserve Bank of Chicago. The risk measures are transformed to quarterly frequency by taking simple average from monthly data. Current Real GDP growth is considered as an additional predictor. The sample data is from the 1971Q1 to 2011Q4.

Table 9 contains the out-of-sample R^2_τ through different methods focusing on the left tail of the target distribution, specifically, $\tau \in \{0.05, 0.10, 0.25\}$. These methods are: quantile regression (QR) with the 16 risk measures, QR with CFNAI, QR with LASSO, PCQR with one factor, PQR, Qcov3PRF and MF-Qcov3PRF (monthly predictors, and quarterly target and proxy). We use these last three methods with 1 automatic proxy and the CFNAI as a proxy, respectively. For the forecast evaluation, we consider a rolling window scheme of length 40, 80, and 120 with starting dates at 1981Q1, 1991Q1 and 2001Q1, respectively.

We see that Qcov3PRF is more effective capturing the predictive power of the financial variables compared to the quantile regression. Moreover, the R^2_τ is greater overall when CFNAI is considered as a proxy using quarterly data. With mixed frequency data, we do not see a significant difference between using either an automatic proxy or the CFNAI. This result suggests that both CFNAI and Real GDP growth depend on the same relevant factor contained in the risk measures. Markedly, Qcov3PRF improves PQR for the case of one relevant factor. One possible reason is that Giglio et al., 2016 assume that PQR provides consistent forecasts when the idiosyncratic component in the factor structure of the predictors are i.i.d. and the distribution for the irrelevant factors is symmetric. These assumptions are not required in our method (see the online Appendix of Giglio et al., 2016).

[Table 9 about here.]

6.3 Climate at Risk (CaR)

The next empirical application is related to the work of Rivas and Gonzalo, 2020 and the empirical application for the climate change in Chen et al., 2021. These papers show there is heterogeneity in the global temperature changes (through different kinds of trends and latent factors, respectively) depending on what side (or characteristics) of the distribution we focus. In our application we want to see whether Qcov3PRF improves forecasting power for the annual global changes in temperature through the carbon dioxide (CO_2) emissions of several countries, we call this effect of CO_2 on the right tail quantiles of the global temperature changes as *Climate at Risk (CaR)*.

For the changes in global temperature we use the Global Land-Ocean Temperature index which can be downloaded from the NASA website https://climate.nasa.gov/vital-signs/global-temperature/. For the set of predictors we consider the CO_2 annual emissions of 86 countries. The data can be downloaded in https://ourworldindata.org/co2-emissions. The sample period is 1930-2021 resulting in 92 years. We consider the annual growth rate of the CO_2 emissions by taking the log difference. The forecast horizon is one year ahead.

We compare the same forecasting methods included in Section 6.1. For Qcov3PRF and PCQR we report the results up to five factors. In all the methods we include the current value of the Global Land-Ocean Temperature index as an additional predictor. We forecast the Global Land-Ocean temperature index one year ahead with expanding and rolling windows. The rolling window considers the 70 previous observations. Table 10 reports the R^2_τ for several values of τ with the validation period starting in 2001. We can see that Qcov3PRF generally shows greater predictive power. In particular, the results obtained by using PCQR do not show good performance even when five factors are contemplated. This may suggest that the relevant factors for predicting the changes in temperature are not the eigenvectors associated with the largest eigenvalues in the covariance matrix of the predictors. We can see a higher forecasting power in the right tail when we use Qcov3PRF, i.e., there is forecasting evidence towards CaR. This is consistent with the fact that over time the CO_2 emissions have been increased in many countries affecting more extreme increments in temperature.

[Table 10 about here.]

7 Conclusions

We have proposed a new method called Qcov3PRF that estimates the conditional quantile of a target variable with a large set of predictors by incorporating a quantile-covariance concept. As a result and in contrast to other existing supervised methods, our method successfully extracts more than one relevant factor of the predictors. Qcov3PRF exploits the quantile-covariance (*qcov*) between the target and the predictors in a similar way as 3PRF (or PLS) exploits the covariance between the target and predictors to obtain the conditional mean forecast of the target. Qcov3PRF demonstrates strong forecasting performance, often superior to alternatives, across simulation specifications and in empirical applications consistent with the the recent literature on GaR (Adrian et al., 2019 and Giglio et al., 2016). We also showed asymptotic properties of the resulting forecasts. Extensions for Qcov3PRF can easily be implemented as in 3PRF such as allowing Markov-Switching regimes (Guérin et al., 2020), time-varying (Su and Wang, 2017) or state-varying factor loadings (Pelger and Xiong, 2022). We leave them for further research.

8 Appendix

In Section 8.1 we first go into detail about how Qcov3PRF and PLS are related. Just like Kelly and Pruitt, 2015 showed that PLS is a particular case of 3PRF, we justify that Qcov3PRF contains a suitable extension of PLS designed for conditional quantile forecasting which maximizes the *qcov* between the target and the predictors. Section 8.2 contains the mathematical proofs of the main results and auxiliary lemmas.

8.1 PLS for conditional quantile prediction is a particular case of Qcov3PRF

As with the Three-Pass Regression Filter and Principal Components, Partial Least Squares (PLS) constructs forecasting indices, or latent factors, as linear combinations of the underlying predictors. These predictive indices are referred as "directions". The PLS forecast, based on the first j directions, denoted as $\hat{\boldsymbol{y}}^{(j)}$, aims to solve the following optimization problem:

$$\boldsymbol{\phi}^{(j)} = \underset{\boldsymbol{\phi}, \|\boldsymbol{\phi}\|=1}{\arg\max}\{\hat{\text{cov}}(\boldsymbol{X}^{(j-1)}\boldsymbol{\phi}, \boldsymbol{y}^{(j-1)})'\hat{\text{cov}}(\boldsymbol{X}^{(j-1)}\boldsymbol{\phi}, \boldsymbol{y}^{(j-1)})\}, \tag{35}$$

where $\boldsymbol{y} = (y_2, \dots, y_{T+1})'$, $\boldsymbol{X} = (\boldsymbol{x}_1^{'}, \dots, \boldsymbol{x}_T^{'})$, $\boldsymbol{\phi}$ a $N \times 1$ vector, and $\hat{\text{cov}}(\boldsymbol{X}^{(j-1)}\boldsymbol{\phi}, \boldsymbol{X}^{(j-1)}\boldsymbol{\phi}_l) = 0$ for $l = 1, \dots, j$. This restriction is equivalent to normalize the latent factors to be orthogonal following the PCA literature. $\boldsymbol{X}^{(j-1)}$ denotes the deflated underlying set of predictors, and $\boldsymbol{y}^{(j-1)}$ is the deflated target variable. Compared to 3PRF framework, $\boldsymbol{\phi}$ are the factor loadings and $\boldsymbol{X}^{(j-1)}\boldsymbol{\phi}$ can be seen as the residual of $\boldsymbol{X}$ containing the remaining $j-1$ relevant factors, the K_g irrelevant factors and the idiosyncratic component. This becomes clear as in each jth iteration in PLS the predictors $\boldsymbol{X}^{(j)}$ are deflated by the $\boldsymbol{f}_j$ factor. The PLS procedure is presented in Algorithm 6.

Comparing 3PRF with PLS, it is clear that the former is equivalent to the latter when three things happen: the underlying predictors are standarized, the proxies for 3PRF are given by $\boldsymbol{y}^{(j)} = \boldsymbol{y} - q_j\boldsymbol{f}_j$, and the resulting predictors are not deflated, i.e., $\boldsymbol{X}^{(j)} = \boldsymbol{X}^{(j-1)}$ with $j = 1, \dots, K_f$. In other words, 3PRF shows that it is only required to substract the effect of the relevant factor j on $\boldsymbol{y}$ in order to

Algorithm 6 Partial Least Squares (PLS)

Let $\boldsymbol{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N)$. Standardize each $\boldsymbol{x}_i$, $i = 1, \ldots, N$ to have mean zero and variance one.
Set $\boldsymbol{X}^{(0)} = \boldsymbol{X}$ and $\boldsymbol{y}^{(0)} = \boldsymbol{y}$.
for $j = 1, \ldots, K_f$ **do**

1. Compute $\phi_i^{(j-1)} = \hat{\text{cov}}(\boldsymbol{x}_i^{(j-1)}, \boldsymbol{y}^{(j-1)})$.
   Then, $\boldsymbol{\phi}^{(j-1)} = (\phi_1^{(j-1)}, \ldots, \phi_N^{(j-1)})'$.
2. Calculate the score vector (latent factor) as $\boldsymbol{f}_j = \boldsymbol{X}^{(j-1)}\boldsymbol{\phi}^{(j-1)}$, the loading vector (factor loading) of $\boldsymbol{X}$ as $\boldsymbol{p}_j = \frac{\boldsymbol{X}^{(j-1)\prime}\boldsymbol{f}_j}{\boldsymbol{f}_j'\boldsymbol{f}_j}$, and the loading of $\boldsymbol{y}$ as $q_j = \frac{\boldsymbol{y}^{(j-1)\prime}\boldsymbol{f}_j}{\boldsymbol{f}_j'\boldsymbol{f}_j}$.
3. Deflat $\boldsymbol{X}^{(j-1)}$ and $\boldsymbol{y}^{(j-1)}$ such that $\boldsymbol{X}^{(j)} = \boldsymbol{X}^{(j-1)} - \boldsymbol{f}_j\boldsymbol{p}_j'$ and $\boldsymbol{y}^{(j)} = \boldsymbol{y}^{(j-1)} - q_j\boldsymbol{f}_j$.

get the forecasts. This is because only the condition $\text{cov}(\boldsymbol{y}^{(j)}, \boldsymbol{f}_j) = 0$ is required to avoid possible multicollinearity problems in Pass 1 of 3PRF.

The particular way of *deflating* and the use of mean covariance in Step 1 in Algorithm 6 is what makes possible the implementation of Pass 1 and Pass 2 through linear regressions in 3PRF. Deflating $\boldsymbol{y}$ but not $\boldsymbol{X}$ implies that the relevant factor estimates are not equal in both methods. Specifically, the relevant factors estimated in 3PRF are not orthogonal[10], in contrast to the factors estimated via PLS.

With respect to conditional quantile prediction, we can see that Qcov3PRF extends the fast Partial Quantile Regression (fPQR) of Méndez-Civieta et al., 2022 with one modification described below. The algorithm for fPQR is obtained by adapting the objective function (35) using the quantile-covariance defined in Section 3.1. The new objective function is given by:

$$\begin{aligned}
\boldsymbol{\phi}^{(j)} &= \underset{\boldsymbol{\phi}, \|\boldsymbol{\phi}\|=1}{\arg\max}\{\text{q}\hat{\text{cov}}_\tau(\boldsymbol{X}^{(j-1)}\boldsymbol{\phi}, \boldsymbol{y}^{(j-1)})'\text{q}\hat{\text{cov}}_\tau(\boldsymbol{X}^{(j-1)}\boldsymbol{\phi}, \boldsymbol{y}^{(j-1)})\} \\
&= \underset{\boldsymbol{\phi}, \|\boldsymbol{\phi}\|=1}{\arg\max}\{\hat{\text{cov}}(\boldsymbol{X}^{(j-1)}\boldsymbol{\phi}, \psi_\tau(\boldsymbol{y}^{(j-1)} - Q^\tau_{y^{(j-1)}}))'\hat{\text{cov}}(\boldsymbol{X}^{(j-1)}\boldsymbol{\phi}, \psi_\tau(\boldsymbol{y}^{(j-1)} - Q^\tau_{y^{(j-1)}}))\}
\end{aligned}$$

where $\hat{\text{cov}}(\boldsymbol{X}^{(j-1)}\boldsymbol{\phi}, \boldsymbol{X}^{(j-1)}\boldsymbol{\phi}_l) = 0$ for $l = 1, \ldots, j$. Méndez-Civieta et al., 2022 adapt Algorithm 6 by changing the mean covariance with quantile-covariance in Step 1. It is important to note that this procedure can not be extended to Qcov3PRF if the automatic proxies procedure for 3PRF is kept. To see this, consider the fact that Step 3 in iteration $j-1$ and Step 1 in iteration j in PLS are linked through the following partial mean covariance:

$$\begin{aligned}
\hat{\text{cov}}(\boldsymbol{y}^{(j)}, \boldsymbol{x}_i^{(j)}) &= \mathbb{E}\left[(\boldsymbol{y}^{(j-1)} - \mathbb{E}(\boldsymbol{y}^{(j-1)}|\boldsymbol{f}_{j-1}))(\boldsymbol{x}_i^{(j-1)} - \mathbb{E}(\boldsymbol{x}_i^{(j-1)}|\boldsymbol{f}_{j-1}))\right] \\
&= \text{p}\hat{\text{cov}}(\boldsymbol{y}^{(j-1)}, \boldsymbol{x}_i^{(j-1)}|\boldsymbol{f}_j), \quad \text{for} \quad i = 1, \ldots, N,
\end{aligned}$$

where $\boldsymbol{x}_i$ is a $T \times 1$ vector. Regarding fPQR, the objective in each iteration j can be seen as to deflat $\boldsymbol{X}^{(j)}$ and $\boldsymbol{y}^{(j)}$ such that $\hat{\text{cov}}(\boldsymbol{X}^{(j)}, \boldsymbol{f}_j) = 0$ and $\text{q}\hat{\text{cov}}_\tau(\boldsymbol{y}^{(j)}, \boldsymbol{f}_j) = 0$. The first covariance is still with respect to the mean since the restriction $\hat{\text{cov}}(\boldsymbol{X}^{(j-1)}\boldsymbol{\phi}, \boldsymbol{X}^{(j-1)}\boldsymbol{\phi}_l) = 0$ does not change. However, with

[10]Regarding Qcov3PRF, the non-orthogonality of the estimated factors is clear by looking at Lemma 1.

respect to the variable $\boldsymbol{y}^{(j)}$ we need to consider a partial quantile-covariance. From Ma et al., 2017 we consider the partial quantile-covariance defined as follows:

$$\begin{aligned} \mathrm{q\hat{c}ov}_\tau(\boldsymbol{y}^{(j)}, \boldsymbol{x}_i^{(j)}) &= \mathbb{E}\left[(\psi_\tau(\boldsymbol{y}^{(j-1)} - Q^\tau_{\boldsymbol{y}^{(j-1)}|\boldsymbol{f}_{j-1}})(\boldsymbol{x}_i^{(j-1)} - \mathbb{E}(\boldsymbol{x}_i^{(j-1)}|\boldsymbol{f}_{j-1}))\right] \\ &= \mathrm{qp\hat{c}ov}_\tau(\boldsymbol{y}^{(j)}, \boldsymbol{x}_i^{(j)}|\boldsymbol{f}_j) \quad \text{for} \quad i = 1, \ldots, N. \end{aligned}$$

To obtain $\mathrm{c\hat{o}v}(\boldsymbol{X}^{(j)}, \boldsymbol{f}_j) = 0$, we deflat $\boldsymbol{X}^{(j-1)}$ in the same way as in Step 3. This is setting $\boldsymbol{X}^{(j)}$ equal to the residuals obtained by regressing each $\boldsymbol{x}_i^{(j-1)}$ on $\boldsymbol{f}_j$ using least squares. For $\mathrm{q\hat{c}ov}_\tau(\boldsymbol{y}^{(j)}, \boldsymbol{f}_j) = 0$, we set $\boldsymbol{y}^{(j)}$ equal to the residuals obtained by regressing $\boldsymbol{y}^{(j-1)}$ on $\boldsymbol{f}_j$ using quantile regression. Therefore, a PLS extension for quantile regression that incorporates *qcov* need to satisfy $\mathrm{c\hat{o}v}(\boldsymbol{X}^{(j)}, \boldsymbol{f}_j) = 0$ and $\mathrm{q\hat{c}ov}_\tau(\boldsymbol{y}^{(j)}, \boldsymbol{f}_j) = 0$.[11] Deflating $\boldsymbol{y}^{(j-1)}$ with quantile regression motivates the automatic proxies procedure for Qcov3PRF. This is the modification in Step 3 of fPQR which makes it a special case in Qcov3PRF. We then require that:

$$\mathrm{q\hat{c}ov}_\tau(\boldsymbol{y}^{(j)}, \boldsymbol{f}_j) = \mathrm{c\hat{o}v}\left(I\left(\boldsymbol{y}^{(j)} - Q^\tau_{\boldsymbol{y}^{(j)}} > 0\right), \boldsymbol{f}_j\right) = 0, \tag{36}$$

which is satisfied with $\boldsymbol{y}^{(j)} = \boldsymbol{y}^{(j-1)} - \hat{q}_j \boldsymbol{f}_j$ where $\hat{q}_j$ is the estimate of running a quantile regression of $\boldsymbol{y}^{(j-1)}$ on $\boldsymbol{f}_j$.

To illustrate the relationship between Qcov3PRF and fPQR with the correct deflat in $\boldsymbol{y}$, we consider the case where a single predictive index is constructed. Applying fPQR we have:

1. Set $\phi_i^\tau = \boldsymbol{x}_i' \boldsymbol{y}^\tau$, and $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_N)'$, where $\boldsymbol{y}^\tau = I(\boldsymbol{y} - Q^\tau_y > 0)$.

2. Set $f_t = \boldsymbol{x}_t' \boldsymbol{\phi}^\tau$, and $\boldsymbol{f} = (f_1, \ldots, f_T)'$.

3. The forecast for the conditional τ quantile of y_{t+1} is obtained from running a quantile regression of y_{t+1} on f_t.

This forecast is the same as what we obtained from Qcov3PRF with 1 automatic-proxy.

8.2 Proofs and additional lemmas

In this section we provide the proofs of the main results of the paper and auxiliary lemmas. Lemma 1 and Lemma 3 are very similar to Lemma 3 and Lemma 4, respectively (with the corresponding normalization assumption), in Kelly and Pruitt, 2015. The proof of Theorem 3 is similar to Theorem 7 in Kelly and Pruitt, 2015. Theorem 1 uses the same initial steps considered in the proof of Theorem 1 in Giglio et al., 2016. Theorem 2 is new.

[11] As we noted in Section 3.2, we also need Assumption 3 to guarantee the transformation of deflated residuals $\psi_\tau(\boldsymbol{y}^{(j-1)} - Q^\tau_{y^{(j-1)}})$ is not affected by the irrelevant factors.

Lemma 1 provides asymptotic limits of the estimates for the relevant factors and their corresponding factor loadings. In general, the estimates $\hat{\boldsymbol{f}}_t$ are not orthogonal between each other since their covariance matrix does not converge to a diagonal matrix, in contrast to the estimated factors resulting from principal components.

Lemma 1. *Let Assumptions 1-2 and 4-8 hold. Then, the probability limits of $\hat{\boldsymbol{\Phi}}_\tau$ and $\hat{\boldsymbol{F}}_t$ are:*

$$\begin{aligned}\hat{\boldsymbol{\Phi}}_\tau &\xrightarrow[T,N\to\infty]{p} \left[(\boldsymbol{\Lambda}_{f,\tau}\boldsymbol{\Lambda}'_{f,\tau} + \boldsymbol{\Delta}_f^{-1}\boldsymbol{\Delta}_\omega)^{-1}\boldsymbol{\Lambda}_{f,\tau}\boldsymbol{\Phi}'_f \quad \mathbf{0}\right] \quad \text{and} \\ \hat{\boldsymbol{F}}_t &\xrightarrow[T,N\to\infty]{p} \left[\left(\boldsymbol{\Lambda}_{f,\tau}\boldsymbol{\Lambda}'_{f,\tau} + \boldsymbol{\Delta}_f^{-1}\boldsymbol{\Delta}_\omega\right)\boldsymbol{\Lambda}'^{-1}_{f,\tau}\boldsymbol{f}_t \quad \mathbf{0}\right],\end{aligned}$$

where $\boldsymbol{f}_t$ is the part of $\boldsymbol{F}_t$ corresponding to $\boldsymbol{\Lambda}_{f,\tau}$ such that $\boldsymbol{\Lambda}_\tau\boldsymbol{F}_t = [\boldsymbol{\Lambda}_{f,\tau} \quad \mathbf{0}]\boldsymbol{F}_t = [\boldsymbol{\Lambda}_{f,\tau}\boldsymbol{f}_t \quad \mathbf{0}]$, and $\boldsymbol{\Delta}_f$ is the covariance matrix of $\boldsymbol{f}_t$.

Proof. The proof makes use of the closed form expression of $\hat{\boldsymbol{\Phi}}_\tau$ and $\hat{\boldsymbol{F}}$ (Eq.(17)), and Assumptions 5, 6 and 8. It is covered in detail in Kelly and Pruitt, 2015 in their Lemma 2, and it is further simplified by Assumptions 2 and 7. □

Proof of Theorem 1. Following Lemma 1, the relevant factors estimated in Pass 1 and Pass 2 are asymptotically relevant for the quantile linear model in Pass 3. Let us now consider the prediction stage of Qcov3PRF such that the optimal coefficients are given by:

$$(\hat{\beta}_{0.\tau}, \hat{\boldsymbol{\beta}}_\tau) = \underset{\beta_{0,\tau},\boldsymbol{\beta}_\tau}{\arg\min} \frac{1}{T}\sum_{t=1}^{T} \rho_\tau(y_{t+1} - \beta_{0,\tau} - \boldsymbol{\beta}'_\tau\hat{\boldsymbol{F}}_t).$$

Since $\boldsymbol{F}_t$ linearly depends on the vector $(\hat{\boldsymbol{F}}_t, \hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau\boldsymbol{F}_t)$, a regression that considers this vector nests the correctly specified quantile forecast regression. Using Corollary 5.12 in White, 1994 and the equivariance properties of quantile regression we have that the regression coefficients which solve:

$$(\dot{\beta}_{0,\tau}, \dot{\boldsymbol{\beta}}_\tau, \dot{\boldsymbol{\beta}}_{1,\tau}) = \underset{\beta_0,\boldsymbol{\beta},\boldsymbol{\beta}_1}{\arg\min} \frac{1}{T}\sum_{t=1}^{T} \rho_\tau\left(y_{t+1} - \beta_{0,\tau} - \boldsymbol{\beta}'_\tau\hat{\boldsymbol{F}}_t - \boldsymbol{\beta}'_{1,\tau}(\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau\boldsymbol{F}_t)\right),$$

are such that,

$$\sqrt{T}(\dot{\boldsymbol{\beta}}_\tau - \boldsymbol{\beta}'_\tau\boldsymbol{H}_\tau^{-1}) \xrightarrow[N,T\to\infty]{p} \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\dot{\boldsymbol{\beta}}}).$$

Now, from Theorem 1 in Angrist et al., 2006 we have that,

$$\hat{\boldsymbol{\beta}}_\tau = \dot{\boldsymbol{\beta}}_\tau + \left(\sum_{t=1}^{T} c_t\hat{\boldsymbol{F}}_t\hat{\boldsymbol{F}}'_t\right)^{-1}\left(\sum_{t=1}^{T} c_t\hat{\boldsymbol{F}}_t\dot{\boldsymbol{\beta}}'_{1,\tau}(\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau\boldsymbol{F}_t)\right), \tag{37}$$

where $c_t = \frac{1}{2}\int_0^1 h_\tau \left(u(\dot{\boldsymbol{\beta}}_\tau \hat{\boldsymbol{F}}_t - \boldsymbol{\beta}_\tau' \boldsymbol{F}_t)|\boldsymbol{F}_t\right) du$. Also, we can rewrite the forecast error as follows:

$$\hat{\boldsymbol{\beta}}_\tau' \hat{\boldsymbol{F}}_t - \boldsymbol{\beta}_\tau' \boldsymbol{F}_t = \hat{\boldsymbol{\beta}}_\tau'(\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau \boldsymbol{F}_t) + (\hat{\boldsymbol{\beta}}_\tau' - \boldsymbol{\beta}_\tau' \boldsymbol{H}_\tau^{-1})\boldsymbol{H}_\tau \boldsymbol{F}_t.$$

What remains to find is the convergence order of $(\hat{\boldsymbol{\beta}}_\tau' - \boldsymbol{\beta}_\tau' \boldsymbol{H}_\tau^{-1})$. From Eq.(37) we get:

$$(\hat{\boldsymbol{\beta}}_\tau' - \boldsymbol{\beta}_\tau' \boldsymbol{H}_\tau^{-1}) = (\dot{\boldsymbol{\beta}}_\tau' - \boldsymbol{\beta}_\tau' \boldsymbol{H}_\tau^{-1}) + \left(\frac{1}{T}\sum_{t=1}^T c_t \hat{\boldsymbol{F}}_t \hat{\boldsymbol{F}}_t'\right)^{-1} \left(\frac{1}{T}\sum_{t=1}^T c_t \hat{\boldsymbol{F}}_t \dot{\boldsymbol{\beta}}_{1,\tau}'(\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau \boldsymbol{F}_t)\right).$$

Now, let us focus on the numerator of the second term in the previous expression and use $\hat{\boldsymbol{F}}_t \equiv \hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau \boldsymbol{F}_t + \boldsymbol{H}_\tau \boldsymbol{F}_t$, then we get:

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^T c_t \hat{\boldsymbol{F}}_t \dot{\boldsymbol{\beta}}_{1,\tau}'(\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau \boldsymbol{F}_t) &= \delta_{NT}^{-2}\frac{1}{T}\sum_{t=1}^T c_t \delta_{NT}(\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau \boldsymbol{F}_t)\dot{\boldsymbol{\beta}}_{1,\tau}' \delta_{NT}(\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau \boldsymbol{F}_t) \\
&\quad + \delta_{NT}^{-1}\frac{1}{T}\sum_{t=1}^T c_t(\boldsymbol{H}_\tau \boldsymbol{F}_t)\dot{\boldsymbol{\beta}}_{1,\tau}' \delta_{NT}(\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau \boldsymbol{F}_t) \\
&= \delta_{NT}^{-2}\frac{1}{\sqrt{T}}\frac{1}{T}\sum_{t=1}^T c_t \delta_{NT}(\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau \boldsymbol{F}_t)\sqrt{T}(\dot{\boldsymbol{\beta}}_{1,\tau}' - \boldsymbol{\beta}_{1,\tau}')\delta_{NT}(\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau \boldsymbol{F}_t) \\
&\quad + \delta_{NT}^{-2}\frac{1}{T}\sum_{t=1}^T c_t \delta_{NT}(\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau \boldsymbol{F}_t)\boldsymbol{\beta}_{1,\tau}' \delta_{NT}(\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau \boldsymbol{F}_t) \\
&\quad + \delta_{NT}^{-1}\frac{1}{\sqrt{T}}\frac{1}{T}\sum_{t=1}^T c_t(\boldsymbol{H}_\tau \boldsymbol{F}_t)\sqrt{T}(\dot{\boldsymbol{\beta}}_{1,\tau}' - \boldsymbol{\beta}_{1,\tau}')\delta_{NT}(\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau \boldsymbol{F}_t) \\
&\quad + \delta_{NT}^{-1}\frac{1}{T}\sum_{t=1}^T c_t(\boldsymbol{H}_\tau \boldsymbol{F}_t)\boldsymbol{\beta}_{1,\tau}' \delta_{NT}(\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau \boldsymbol{F}_t) \\
&= \delta_{NT}^{-2}\boldsymbol{O}_p(T^{-1/2}) + \delta_{NT}^{-2}\boldsymbol{O}_p(1) + \delta_{NT}^{-1}\boldsymbol{O}_p(T^{-1/2}) + \delta_{NT}^{-1}\boldsymbol{O}_p(1) \\
&= \boldsymbol{O}_p(\delta_{NT}^{-1}),
\end{aligned}$$

where $\delta_{NT} = \min(\sqrt{N}, T)$. The above result implies that $(\hat{\boldsymbol{\beta}}_\tau' - \boldsymbol{\beta}_\tau' \boldsymbol{H}_\tau^{-1}) = \boldsymbol{O}_p(T^{-1/2}) + \boldsymbol{O}_p(1)\boldsymbol{O}_p(\delta_{NT}^{-1}) = \boldsymbol{O}_p(\delta_{NT}^{-1})$. Therefore, $\hat{\boldsymbol{\beta}}_\tau' \hat{\boldsymbol{F}}_t - \boldsymbol{\beta}_\tau' \boldsymbol{F}_t = \boldsymbol{O}_p(1)\boldsymbol{O}_p(\delta_{NT}^{-1}) + \boldsymbol{O}_p(\delta_{NT}^{-1})\boldsymbol{O}_p(1) = \boldsymbol{O}_p(\delta_{NT}^{-1})$. □

To show the asymptotic normality of the infeasible quantile foecasts we need to make use of the following lemma, whose proof comes from Theorem 6 in Kelly and Pruitt, 2015, but we consider the transformed proxies $\boldsymbol{z}_{t,\tau}^*$ instead of $\boldsymbol{z}_t$.

Lemma 2. *Let Assumptions 1-2 and 4-8 hold. We have for all* t,

(i). If $\sqrt{N} = o(T)$ *or* $N = O(T)$ *then:*

$$\sqrt{N}\left[\hat{\boldsymbol{F}}_t - \tilde{\boldsymbol{H}}_\tau \boldsymbol{F}_t\right] \xrightarrow[N,T\to\infty]{d} \mathcal{N}(\boldsymbol{0}, \operatorname{plim} \boldsymbol{\Sigma}_{\hat{F}_t}).$$

(ii). If $T = o(\sqrt{N})$ *then:*

$$T\left[\hat{\boldsymbol{F}}_t - \tilde{\boldsymbol{H}}_\tau \boldsymbol{F}_t\right] = \boldsymbol{O}_p(1).$$

where $\operatorname{plim} \boldsymbol{\Sigma}_{\hat{F}_t} = \left(\boldsymbol{\Lambda}_\tau \boldsymbol{\Delta}_F \boldsymbol{\Lambda}_\tau' + \boldsymbol{\Delta}_\omega\right)\left(\boldsymbol{\Lambda}_\tau \boldsymbol{\Delta}_F^2 \boldsymbol{\Lambda}_\tau'\right)^{-1} \boldsymbol{\Lambda}_\tau \boldsymbol{\Delta}_F \boldsymbol{\Sigma}_{\phi\varepsilon} \boldsymbol{\Delta}_F \boldsymbol{\Lambda}_\tau' \left(\boldsymbol{\Lambda}_\tau \boldsymbol{\Delta}_F^2 \boldsymbol{\Lambda}_\tau'\right)^{-1} \left(\boldsymbol{\Lambda}_\tau \boldsymbol{\Delta}_F \boldsymbol{\Lambda}_\tau' + \boldsymbol{\Delta}_\omega\right)'$, *and* $\tilde{\boldsymbol{H}}_\tau = \hat{\boldsymbol{F}}_A \hat{\boldsymbol{F}}_B^{-1} N^{-1} T^{-1} \boldsymbol{Z}_\tau^{*'} \boldsymbol{J}_T \boldsymbol{X} \boldsymbol{J}_N \boldsymbol{\Phi}$, *with* $\hat{\boldsymbol{F}}_A = T^{-1} \boldsymbol{Z}_\tau^{*'} \boldsymbol{J}_T \boldsymbol{Z}_\tau^*$ *and* $\hat{\boldsymbol{F}}_B = N^{-1} T^{-2} \boldsymbol{Z}_\tau^{*'} \boldsymbol{J}_T \boldsymbol{X} \boldsymbol{J}_N \boldsymbol{X}' \boldsymbol{J}_T \boldsymbol{Z}_\tau^*$.

Proof of Theorem 2. Without loss of generality, let us assume that the target variable does not have an intercept. We can rewrite the forecast error as follows:

$$\begin{aligned} \hat{\boldsymbol{\beta}}_\tau' \hat{\boldsymbol{F}}_t - \boldsymbol{\beta}_\tau' \boldsymbol{F}_t &= \hat{\boldsymbol{\beta}}_\tau' \hat{\boldsymbol{F}}_t - \boldsymbol{\beta}_\tau' \boldsymbol{H}_\tau^{-1} \hat{\boldsymbol{F}}_t + \boldsymbol{\beta}_\tau' \boldsymbol{H}_\tau^{-1} \hat{\boldsymbol{F}}_t - \boldsymbol{\beta}_\tau' \boldsymbol{F}_t \\ &= \sqrt{T}(\hat{\boldsymbol{\beta}}_\tau' - \boldsymbol{\beta}_\tau' \boldsymbol{H}_\tau^{-1}) \frac{\hat{\boldsymbol{F}}_t}{\sqrt{T}} + \frac{\boldsymbol{\beta}_\tau' \boldsymbol{H}_\tau^{-1}}{\sqrt{N}} \sqrt{N}(\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau \boldsymbol{F}_t). \end{aligned}$$

Since,

$$\begin{aligned} (\hat{\boldsymbol{\beta}}_\tau' - \boldsymbol{\beta}_\tau' \boldsymbol{H}_\tau^{-1}) &= (\dot{\boldsymbol{\beta}}_\tau' - \boldsymbol{\beta}_\tau' \boldsymbol{H}_\tau^{-1}) + \left(\frac{1}{T} \sum_{t=1}^T c_t \hat{\boldsymbol{F}}_t \hat{\boldsymbol{F}}_t' \right)^{-1} \left(\frac{1}{T} \sum_{t=1}^T c_t \hat{\boldsymbol{F}}_t \dot{\boldsymbol{\beta}}_{1,\tau}' (\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau \boldsymbol{F}_t) \right) \\ &= (\dot{\boldsymbol{\beta}}_\tau' - \boldsymbol{\beta}_\tau' \boldsymbol{H}_\tau^{-1}) + \mathrm{I}^{-1} \cdot \mathrm{II}. \end{aligned}$$

We have that:

$$\begin{aligned} |\mathrm{I}| &\leq \left(\frac{1}{T} \sum_{t=1}^T c_t^2 \right)^{1/2} \left(\frac{1}{T^2} \sum_{t=1}^T \|\hat{\boldsymbol{F}}_t\|^4 \right)^{1/2} \\ &\leq \boldsymbol{O}_p(1) \cdot \boldsymbol{O}_p(1) \\ &= \boldsymbol{O}_p(1), \end{aligned}$$

where the second inequality is due to Cauchy-Schwartz inequality and the last inequality holds by Assumption 4.1 and Lemma 1. For the numerator term we get that:

$$\begin{aligned} |\mathrm{II}| &\leq \frac{1}{T^2} \sum_{t=1}^T \frac{c_t \hat{\boldsymbol{F}}_t}{\sqrt{NT}} \sqrt{T}(\dot{\boldsymbol{\beta}}_{1,\tau}' - \boldsymbol{\beta}_{1,\tau}') \sqrt{N}(\hat{\boldsymbol{F}}_t \boldsymbol{H}_\tau - \boldsymbol{F}_t) + \frac{1}{T^2} \boldsymbol{\beta}_{1,\tau}' \sum_{t=1}^T \frac{c_t \hat{\boldsymbol{F}}_t}{\sqrt{N}} \sqrt{N}(\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau \boldsymbol{F}_t) \\ &\leq \frac{1}{N^{1/2} T} \|\sqrt{T}(\hat{\boldsymbol{\beta}}_{1,\tau} - \boldsymbol{\beta}_{1,\tau})\| \left(\frac{1}{T} \sum_{t=1}^T c_t^2 \right)^{1/2} \left(\frac{1}{T} \sum_{t=1}^T \|\hat{\boldsymbol{F}}_t\|^2 \right)^{1/2} \left(\frac{1}{T} \sum_{t=1}^T (\sqrt{N}(\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau \boldsymbol{F}_t))^2 \right)^{1/2} \\ &\quad + \frac{1}{N^{1/2} T^{1/2}} \|\boldsymbol{\beta}_{1,\tau}\| \left(\frac{1}{T} \sum_{t=1}^T c_t^2 \right)^{1/2} \left(\frac{1}{T} \sum_{t=1}^T \|\hat{\boldsymbol{F}}_t\|^2 \right)^{1/2} \left(\frac{1}{T} \sum_{t=1}^T (\sqrt{N}(\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau \boldsymbol{F}_t))^2 \right)^{1/2} \\ &= \boldsymbol{O}_p(N^{-1/2} T^{-1/2}) \\ &= \boldsymbol{o}_p(1), \end{aligned}$$

by using Cauchy-Schwartz inequality, Assumption 4.1, Assumption 5.1, Lemma 2 and the asymptotic normality of linear estimates in quantile regression. Then, by the continuous mapping theorem, we have:

$$\hat{\boldsymbol{\beta}}_\tau' \hat{\boldsymbol{F}}_t - \boldsymbol{\beta}_\tau' \boldsymbol{F}_t = \sqrt{T}(\dot{\boldsymbol{\beta}}_\tau' - \boldsymbol{\beta}_\tau' \boldsymbol{H}_\tau^{-1}) \frac{\hat{\boldsymbol{F}}_t}{\sqrt{T}} + \frac{\boldsymbol{\beta}_\tau' \boldsymbol{H}^{-1}}{\sqrt{N}} \sqrt{N}(\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau \boldsymbol{F}_\tau) + o_p(1).$$

Since

$$\sqrt{N}(\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau \boldsymbol{F}_t) \xrightarrow[N,T\to\infty]{d} \mathcal{N}(\mathbf{0}, \text{plim}\, \boldsymbol{\Sigma}_{\hat{F}_t}),$$
$$\sqrt{T}(\hat{\boldsymbol{\beta}}_\tau - \boldsymbol{\beta}_\tau) \xrightarrow[N,T\to\infty]{d} \mathcal{N}(\mathbf{0}, \text{plim}\, \boldsymbol{\Sigma}_{\hat{\beta}}),$$

where $\boldsymbol{\Sigma}_{\hat{F}_t}$ is defined in Lemma 2 and $\text{plim}\, \boldsymbol{\Sigma}_{\hat{\beta}} = \mathbb{E}[h_\tau(0|\boldsymbol{F}_t)\boldsymbol{F}_t\boldsymbol{F}_t']^{-1}\mathbb{E}[\psi_\tau(u_{t+1}^2)\boldsymbol{F}_t\boldsymbol{F}_t']\mathbb{E}[h_\tau(0|\boldsymbol{F}_t)\boldsymbol{F}_t\boldsymbol{F}_t']^{-1}$ which is feasible by Assumption 4.3, 5.5 and 8.2. Since $\sqrt{N}(\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau \boldsymbol{F}_t)$ depend on $\boldsymbol{\varepsilon}_{i,t}$, and $\sqrt{T}(\hat{\boldsymbol{\beta}}_\tau - \boldsymbol{\beta}_\tau)$ depend on u_{t+1}, both terms are independent. It follows that $\hat{\boldsymbol{\beta}}_\tau'\hat{\boldsymbol{F}}_t - \boldsymbol{\beta}_\tau'\boldsymbol{F}_t$ is normal with estimated variance $\hat{\boldsymbol{V}}_\tau = \frac{1}{T}\hat{\boldsymbol{\beta}}_\tau'(\boldsymbol{\Sigma}_{\hat{F}})\hat{\boldsymbol{\beta}}_\tau + \frac{1}{N}\hat{\boldsymbol{F}}_t'(\boldsymbol{\Sigma}_{\hat{\beta}_\tau})\hat{\boldsymbol{F}}_t$ given that $\hat{\boldsymbol{\beta}}_\tau = \boldsymbol{\beta}_\tau + o_p(1)$. □

Lemma 3 shows the consistency of the linear estimates from Eq.(19). We need it in order to show that the automatic proxies resulting from Algorithm 4 are linearly independent such that Assumption 2 is satisfied.

Lemma 3. *Let Assumptions 1-2 and 4-8 hold. Then, under Assumption 3 the probability limit of the estimated coefficients* $\hat{\boldsymbol{\alpha}}_\tau$ *obtained by running a least squares regression of* $I(y_{t+1} - Q^\tau_{y_{t+1}} > 0)$ *on* $\hat{\boldsymbol{F}}_t$ *is:*

$$\hat{\boldsymbol{\alpha}}_\tau \xrightarrow[T,N\to\infty]{p} \left[\left(\boldsymbol{\Lambda}_{f,\tau}\boldsymbol{\Lambda}_{f,\tau}' + \boldsymbol{\Delta}_f^{-1}\boldsymbol{\Delta}_\omega\right)^{-1} \boldsymbol{\Lambda}_{f,\tau}\boldsymbol{\alpha}_{f,\tau} \quad \mathbf{0} \right],$$

where $\boldsymbol{\alpha}_\tau = (\boldsymbol{\alpha}_{f,\tau}' \quad \mathbf{0})'$, *with* $\boldsymbol{\alpha}_{f,\tau}$ *a vector of size* $K_f \times 1$.

Proof. The estimation of $\boldsymbol{\alpha}_\tau$ using least squares regression results in:

$$\begin{aligned}
\hat{\boldsymbol{\alpha}}_\tau &= \left(\hat{\boldsymbol{F}}'\boldsymbol{J}_T\hat{\boldsymbol{F}}\right)^{-1}\hat{\boldsymbol{F}}'\boldsymbol{J}_T I(\boldsymbol{y} - Q^\tau_{y+1}) \\
&= (T^{-1}\boldsymbol{Z}_\tau^{*'}\boldsymbol{J}_T\boldsymbol{Z}_\tau^*)^{-1}N^{-1}T^{-2}\boldsymbol{Z}_\tau^{*'}\boldsymbol{J}_T\boldsymbol{X}\boldsymbol{J}_N\boldsymbol{X}'\boldsymbol{J}_T\boldsymbol{Z}_\tau^* \\
&\quad (N^{-2}T^{-3}\boldsymbol{Z}_\tau^{*'}\boldsymbol{J}_T\boldsymbol{X}\boldsymbol{J}_N\boldsymbol{X}'\boldsymbol{J}_T\boldsymbol{X}\boldsymbol{J}_N\boldsymbol{X}'\boldsymbol{J}_T\boldsymbol{Z}_\tau^*)^{-1}N^{-1}T^{-2}\boldsymbol{Z}_\tau^{*'}\boldsymbol{J}_T\boldsymbol{X}\boldsymbol{J}_N\boldsymbol{X}'\boldsymbol{J}_T I(\boldsymbol{y} - Q^\tau_{y+1}) \\
&= \hat{\boldsymbol{\alpha}}_{1,\tau}^{-1}\hat{\boldsymbol{\alpha}}_{2,\tau}\hat{\boldsymbol{\alpha}}_{3,\tau}^{-1}\hat{\boldsymbol{\alpha}}_{4,\tau},
\end{aligned}$$

where the second equality follows from Eq.(17). From Lemma 3 in Kelly and Pruitt, 2015, by replacing the term $\boldsymbol{Z}$ with $\boldsymbol{Z}_\tau^*$ and further simplifying by Assumptions 2 and 7 we have the following probability limits:

$$\hat{\boldsymbol{\alpha}}_{1,\tau} \xrightarrow[T,N\to\infty]{p} \boldsymbol{\Lambda}_\tau\boldsymbol{\Delta}_F\boldsymbol{\Lambda}_\tau' + \boldsymbol{\Delta}_\omega, \tag{38}$$
$$\hat{\boldsymbol{\alpha}}_{2,\tau} \xrightarrow[T,N\to\infty]{p} \boldsymbol{\Lambda}_\tau\boldsymbol{\Delta}_F^2\boldsymbol{\Lambda}_\tau'. \tag{39}$$

Regarding the terms $\hat{\boldsymbol{\alpha}}_{3,\tau}$ and $\hat{\boldsymbol{\alpha}}_{4,\tau}$, we rely on the web appendix from Kelly and Pruitt, 2015. Then, by simply replacing their term $\boldsymbol{y}$ with $I(\boldsymbol{y} - Q^{\tau}_{y+1})$, we obtain that

$$\hat{\boldsymbol{\alpha}}_{3,\tau} \xrightarrow[T,N\to\infty]{p} \boldsymbol{\Lambda}_\tau \boldsymbol{\Delta}^3_F \boldsymbol{\Lambda}'_\tau \tag{40}$$

$$\hat{\boldsymbol{\alpha}}_{4,\tau} \xrightarrow[T,N\to\infty]{p} \boldsymbol{\Lambda}_\tau \boldsymbol{\Delta}^2_F \boldsymbol{\alpha}_\tau. \tag{41}$$

By Eqs.(38)-(41), Assumptions 2 and 7, and the continuous mapping theorem, the main result is obtained.

□

Proof of Theorem 3. We begin by showing that Assumption 2 is generally satisfied, this is that the loadings of the automatic proxies are linear independent (full rank matrix) and that they are uncorrelated with irrelevant factors. If $K_f = 1$, it is clearly seen that $\boldsymbol{r}_{0,\tau} = \boldsymbol{y} - \hat{Q}^{\tau}_y$ is not correlated with irrelevant factors by Assumption 3. If $K_f = 2$, to see that $\boldsymbol{r}_{1,\tau} = \boldsymbol{y} - \hat{Q}^{\tau}_{y|\hat{f}_1,\ldots,\hat{f}_{\ell-1}}$ is uncorrelated with irrelevant factors we have that:

$$\begin{aligned}
\boldsymbol{r}_{1,\tau} &= \boldsymbol{y} - \hat{Q}^{\tau}_{y|\hat{\boldsymbol{f}}^{(1)}_1} \\
&= \boldsymbol{F}\boldsymbol{\beta}_\tau + \boldsymbol{u}^\tau - \hat{\boldsymbol{f}}^{(1)}_1 \hat{\beta}^{(1)}_{1,\tau} \\
&= \boldsymbol{F}\boldsymbol{\beta}_\tau + \boldsymbol{u}^\tau - \boldsymbol{X}\boldsymbol{\Omega}^{(1)}_\tau \hat{\beta}^{(1)}_{1,\tau} \\
&= \boldsymbol{F}\boldsymbol{\beta}_\tau + \boldsymbol{u}^\tau - \boldsymbol{X}\boldsymbol{\Omega}^{(1)}_\tau (\hat{\beta}^{(1)}_{1,\tau} - \tilde{\boldsymbol{H}}^{(1)-1}_\tau \beta^{(1)}_{1,\tau}) - \boldsymbol{f}_1 \beta^{(1)}_{1,\tau} \\
&= \boldsymbol{F}\boldsymbol{\beta}_\tau + \boldsymbol{u}^\tau - \boldsymbol{X}\boldsymbol{\Omega}^{(1)}_\tau (\hat{\boldsymbol{\beta}}^{(1)}_{1,\tau} - \tilde{\boldsymbol{H}}^{(1)-1}_\tau \beta^{(1)}_{1,\tau}) - \boldsymbol{f}_1 \beta_{1,\tau} - \boldsymbol{f}_1 (\mathbb{E}[\tilde{w}^{(1)}_\tau(\boldsymbol{F})\boldsymbol{f}^2_1])^{-1} \mathbb{E}[\tilde{w}^{(1)}_\tau(\boldsymbol{F})\boldsymbol{f}_1 \boldsymbol{f}'_{-1}\boldsymbol{\beta}_{-1,\tau}] \\
&= -\boldsymbol{f}_1 b^{(1)}_{1,\tau} + \boldsymbol{f}_{-1}\boldsymbol{\beta}_{-1,\tau} - \boldsymbol{F}\tilde{\boldsymbol{H}}^{(1)}_\tau (\hat{\beta}^{(1)}_{1,\tau} - \tilde{\boldsymbol{H}}^{(1)-1}_\tau \beta^{(1)}_{1,\tau}) + \boldsymbol{u}^\tau - \boldsymbol{\varepsilon}\boldsymbol{\Omega}^{(1)}_\tau (\hat{\beta}^{(1)}_{1,\tau} - \tilde{\boldsymbol{H}}^{(1)-1}_\tau \beta^{(1)}_{1,\tau}),
\end{aligned}$$

where $\boldsymbol{\Omega}^{(1)}_\tau = \boldsymbol{J}_N \boldsymbol{W}^{(1)}_{xr} (\boldsymbol{W}^{(1)'}_{xr} \boldsymbol{J}_N \boldsymbol{W}^{(1)}_{xr})^{-1}$, $\boldsymbol{W}^{(1)}_{xr} = \boldsymbol{X}'\boldsymbol{J}_T \boldsymbol{r}^*_{1,\tau}$, $\tilde{\boldsymbol{H}}^{(1)}_\tau = \boldsymbol{\Phi}'\boldsymbol{\Omega}^{(1)}_\tau$,
$b^{(1)}_{1,\tau} = (\mathbb{E}[\tilde{w}^{(1)}_\tau(\boldsymbol{F})\boldsymbol{f}^2_1])^{-1}\mathbb{E}[\tilde{w}^{(1)}_\tau(\boldsymbol{F})\boldsymbol{f}_1\boldsymbol{f}'_{-1}\boldsymbol{\beta}_{-1,\tau}]$, and $\tilde{w}^{(1)}_\tau(\boldsymbol{F}) = \frac{1}{2}\int_0^1 h_\tau(\boldsymbol{u}\Delta_\tau(\boldsymbol{F}, \beta^{(1)}_{1,\tau})|\boldsymbol{F})du$ with $\Delta_\tau(\boldsymbol{F}, \beta^{(1)}_{1,\tau}) = \boldsymbol{f}_1\beta^{(1)}_{1,\tau} - Q_\tau(\boldsymbol{y}|\boldsymbol{F})$. In the fifth inequality we use the partial quantile regression and omitted variable bias formulation of Angrist et al., 2006. Recalling that $\boldsymbol{F}\tilde{\boldsymbol{H}}^{(1)}_\tau = [\boldsymbol{f}\tilde{\boldsymbol{H}}^{(1)}_{f,\tau} \quad \boldsymbol{0}]$, and by Assumptions 4.2 and 8.2 the conditional quantile of $\boldsymbol{r}_{1,\tau}$ is uncorrelated with irrelevant factors $\boldsymbol{g}$. Then, by Assumption 3 and Lemma 3 we can guarantee that $\boldsymbol{r}^*_{1,\tau} = I(\boldsymbol{r}_{1,\tau} > 0)$ contains loadings equal to zero corresponding to irrelevant factors, and therefore Assumption 2 is satisfied.

For $K_f > 2$, we proceed by induction to show that Algorithm 4 constructs a set of proxies that satisfies Assumption 2. In particular, we want to show that the automatic proxies have a loading matrix on relevant factors $\boldsymbol{\Lambda}_{f,\tau}$ that is full rank, and that $\boldsymbol{\Lambda}_{g,\tau} = 0$. Suppose by hypothesis that we have $k < K_f$ automatically generated proxies, the loadings corresponding to irrelevant factors are equal to zero by the same argument when $K_f = 2$, i.e., $\mathbf{\Lambda}_\tau^k = [\mathbf{\Lambda}_{f,\tau}^k \quad \mathbf{0}]$, where $\mathbf{\Lambda}_{f,\tau}^k$ is a $k \times K_f$ matrix, specifically,

$$\begin{aligned}\boldsymbol{r}_{k,\tau} \quad &= \quad -\boldsymbol{f}_{1:k}\boldsymbol{b}_{1:k,\tau}^{(k)} + \boldsymbol{f}_{k+1:K_f}\boldsymbol{\beta}_{k+1:K_f,\tau} - \boldsymbol{F}\tilde{\boldsymbol{H}}_\tau^{(k)}(\hat{\boldsymbol{\beta}}_{1:k,\tau}^{(k)} - \tilde{\boldsymbol{H}}_\tau^{(k)-1}\boldsymbol{\beta}_{1:k,\tau}^{(k)}) \\ &\quad +\boldsymbol{u}^\tau - \boldsymbol{\varepsilon}\mathbf{\Omega}_\tau^{(k)}(\hat{\boldsymbol{\beta}}_{1:k,\tau}^{(k)} - \tilde{\boldsymbol{H}}_\tau^{(k)-1}\boldsymbol{\beta}_{1:k,\tau}),\end{aligned} \tag{42}$$

where the notation $\boldsymbol{v}_{k_1:k_2}$ indicates the elements from k_1 to k_2 of the $\boldsymbol{v}$ vector. Also, the terms $\mathbf{\Omega}_\tau^{(k)}$, $\tilde{\boldsymbol{H}}_\tau^{(k)}$ and $\boldsymbol{b}_{1:k,\tau}^{(k)}$ are defined similarly as in the $K_f = 2$ case but now with the first $k+1$ automatically generated proxies after the kth iteration following Algorithm 4. The equality above implies that $\boldsymbol{r}_{k,\tau}^* = I(\boldsymbol{r}_{k,\tau} > 0)$ has loadings equal to zero corresponding to irrelevant factors by Assumption 3 and Lemma 3.

Now, we will show that the $\boldsymbol{r}_{k,\tau}$'s loading on relevant factors is linearly independent of the rows of $\mathbf{\Lambda}_{f,\tau}^k$. This follows from a similar argument done in the proof for Theorem 7 in the Appendix in Kelly and Pruitt, 2015. First, note that the relevant-factor loadings are equal to $\Big((-\boldsymbol{b}_{1:k,\tau}^{(k)'}, \boldsymbol{\beta}'_{k+1:K_f,\tau})' - \tilde{\boldsymbol{H}}_{f,\tau}^{(k)}(\hat{\boldsymbol{\beta}}_{1:k,\tau}^{(k)} - \tilde{\boldsymbol{H}}_\tau^{(k)-1}\boldsymbol{\beta}_{1:k,\tau}^{(k)})\Big)$ by Eq.(42), where the matrix $\tilde{\boldsymbol{H}}_{f,\tau}^{(k)}$ is constructed based on the proxies $(\boldsymbol{r}_{0,\tau}^*, \ldots, \boldsymbol{r}_{k-1,\tau}^*)$. Next, project $\boldsymbol{r}_{k,\tau}$'s relevant-factor loadings onto the column space of $\mathbf{\Lambda}_{f,\tau}^{k'}$. Then, the residual's loading vector is linearly independent of $\mathbf{\Lambda}_{f,\tau}^{k'}$ if the difference between it and its projection on $\mathbf{\Lambda}_{f,\tau}^{k'}$ is non-zero. Specifically, this difference is presented as:

$$\Big(\boldsymbol{I} - \mathbf{\Lambda}_{f,\tau}^{k'}(\mathbf{\Lambda}_{f,\tau}^k\mathbf{\Lambda}_{f,\tau}^{k'})^{-1}\mathbf{\Lambda}_{f,\tau}^k\Big)\Big((-\boldsymbol{b}_{1:k,\tau}^{(k)'}, \boldsymbol{\beta}'_{k+1:K_f,\tau})' - \tilde{\boldsymbol{H}}_{f,\tau}^{(k)}(\hat{\boldsymbol{\beta}}_{1:k,\tau}^{(k)} - \tilde{\boldsymbol{H}}_\tau^{(k)-1}\boldsymbol{\beta}_{1:k,\tau}^{(k)})\Big).$$

As it is clear that the relevant-factor loadings are not equal to zero with probability zero, this difference is equal to zero only when $\mathbf{\Lambda}_{f,\tau}^{k'}(\mathbf{\Lambda}_{f,\tau}^k\mathbf{\Lambda}_{f,\tau}^{k'})^{-1}\mathbf{\Lambda}_{f,\tau}^k = \boldsymbol{I}$. The induction hypothesis ensures that this is not the case as long as $k < K_f$. Thus, the difference between the $\boldsymbol{r}_{1,\tau}$'s loading vector and its projection onto the column space of $\mathbf{\Lambda}_{f,\tau}^{k'}$ is nonzero, implying that the loading vector is linearly independent of the rows of $\mathbf{\Lambda}_{f,\tau}^{k'}$. Lastly, by Assumption 3, the relevant-factor loadings of the transformed proxy $\boldsymbol{r}_{k,\tau}^*$ is also linearly independent of the rows of $\mathbf{\Lambda}_{f,\tau}^{k'}$. Therefore, the transformed proxies $(\boldsymbol{r}_{1,\tau}^*, \ldots, \boldsymbol{r}_{K_f,\tau}^*)$ satisfy Assumption 2.

Finally, it is left to check that the automatic proxies satisfy Assumptions 5.4, 6.3 and 6.4 when the remaining parts of Assumptions 1–8 hold. To see this, by Assumption 3 and Eq.(42) we can rewrite an automatic proxy $r_{kt,\tau}$ (suppressing constants) as $r_{kt,\tau} = \tilde{\mathbf{\Lambda}}_{f,\tau}^{k+1'}\boldsymbol{f}_t + \tilde{\omega}_t^k$ with $\tilde{\omega}_t^k = u_{t+1} + N^{-1}\tilde{\boldsymbol{a}}'_{f,\tau}\boldsymbol{\epsilon}_t$ as we can show it with the following:

$$
\begin{aligned}
r_{kt,\tau} &= \boldsymbol{y} - \hat{Q}^{\tau}_{y|\hat{f}_1^{(k)},\ldots,\hat{f}_k^{(k)}} \\
&= \boldsymbol{\beta}'_{f,\tau}\boldsymbol{f}_t + u^{\tau}_{t+1} - \hat{\beta}^{(k)'}_{1:k,\tau}\boldsymbol{\Omega}^{(k)}_{\tau}\boldsymbol{x}_t \\
&= \boldsymbol{\beta}'_{f,\tau}\boldsymbol{f}_t + u^{\tau}_{t+1} - N^{-1}\hat{\beta}^{(k)'}_{1:k,\tau}T^{-1}\boldsymbol{W}^{(k)'}_{rr}(N^{-1}T^{-2}\boldsymbol{W}^{(k)'}_{xr}\boldsymbol{J}_N\boldsymbol{W}^{(k)}_{xr})^{-1}T^{-1}\boldsymbol{W}^{(k)}_{xr}\boldsymbol{J}_N\boldsymbol{x}_t \\
&= \boldsymbol{\beta}'_{f,\tau}\boldsymbol{f}_t + u^{\tau}_{t+1} - N^{-1}\tilde{\boldsymbol{a}}'_{f,\tau}\boldsymbol{x}_t \\
&= \tilde{\boldsymbol{\Lambda}}^{k+1'}_{f,\tau}\boldsymbol{f}_t + u^{\tau}_{t+1} - \tilde{\boldsymbol{a}}^{k'}_{f,\tau}N^{-1}\boldsymbol{\varepsilon}_t,
\end{aligned}
$$

where $\tilde{\boldsymbol{a}}^{k}_{f,\tau} = \boldsymbol{O}_p(1)$ by the proof of Lemma 1 in Kelly and Pruitt, 2015 (replacing $\boldsymbol{Z}$ with $(\boldsymbol{r}^*_{0,\tau},\ldots,\boldsymbol{r}^*_{k,\tau})$). Then, by Assumption 3, we can rewrite $r^*_{kt,\tau} = \boldsymbol{\Lambda}^{k+1'}_{f,\tau}\boldsymbol{f}_t + \xi_{t+1} - \boldsymbol{a}^{k'}_{f,\tau}N^{-1}\boldsymbol{\varepsilon}_t = \boldsymbol{\Lambda}^{k+1'}_{f,\tau}\boldsymbol{f}_t + \omega^k_t$, where ξ_{t+1} depends on u^{τ}_{t+1} and $\boldsymbol{a}^k_{f,\tau} = \boldsymbol{O}_p(1)$.

By Assumption 5.5, the ξ_{t+1} and $\boldsymbol{a}'_{f,\tau}N^{-1}\boldsymbol{\varepsilon}_t$ components of ω_t are independent (since ξ_{t+1} is a function of u^{τ}_{t+1}) so they can be treated separately. By Assumptions 3 and 5.5, the ξ_{t+1} component satisfies Assumptions 5.4, 6.3 and 6.4. For the second component, given that $\boldsymbol{a}^k_{f,\tau}$ is a vector of constants, Assumption 5.4 is satisfied by Assumption 5.3 and 6.2, Assumption 6.3 is satisfied by Assumption 8.2. Finally, Assumption 6.4 holds by Assumption 6.2. □

References

Adams, P. A., Adrian, T., Boyarchenko, N., & Giannone, D. (2021). Forecasting macroeconomic risks. *International Journal of Forecasting*, *37*(3), 1173–1191.

Adrian, T., Boyarchenko, N., & Giannone, D. (2019). Vulnerable growth. *American Economic Review*, *109*(4), 1263–89.

Adrian, T., Grinberg, F., Liang, N., Malik, S., & Yu, J. (2022). The term structure of growth-at-risk. *American Economic Journal: Macroeconomics*, *14*(3), 283–323.

Ahn, S. C., & Bae, J. (2021). *Forecasting with partial least squares when a large number of predictors are available* (tech. rep.). Working paper.

Ando, T., & Tsay, R. S. (2011). Quantile regression models with factor-augmented predictors and information criterion. *The Econometrics Journal*, *14*(1), 1–24.

Angrist, J., Chernozhukov, V., & Fernández-Val, I. (2006). Quantile regression under misspecification, with an application to the us wage structure. *Econometrica*, *74*(2), 539–563.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, *71*(1), 135–171.

Bai, J., & Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, *70*(1), 191–221.

Bai, J., & Ng, S. (2006). Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions. *Econometrica*, *74*(4), 1133–1150.

Bai, J., & Ng, S. (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics*, *146*(2), 304–317.

Chamberlain, G., & Rothschild, M. (1982). Arbitrage, factor structure, and mean-variance analysis on large asset markets.

Chen, L., Dolado, J. J., & Gonzalo, J. (2021). Quantile factor models. *Econometrica*, *89*(2), 875–910.

Dodge, Y., & Whittaker, J. (2009). Partial quantile regression. *Metrika*, *70*(1), 35–57.

Fissler, T., & Ziegel, J. F. (2016). Higher order elicitability and osand's principle. *The Annals of Statistics*, *44*(4), 1680–1707.

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

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, *33*(5), 2223–2273.

Guérin, P., Leiva-Leon, D., & Marcellino, M. (2020). Markov-switching three-pass regression filter. *Journal of Business & Economic Statistics*, *38*(2), 285–302.

Hepenstrick, C., & Marcellino, M. (2019). Forecasting gross domestic product growth with large unbalanced data sets: The mixed frequency three-pass regression filter. *Journal of the Royal Statistical Society: Series A*, *182*(1), 69–99.

Huang, D., Jiang, F., Li, K., Tong, G., & Zhou, G. (2022). Scaled PCA: A new approach to dimension reduction. *Management Science*, *68*(3), 1678–1695.

Huang, D., Jiang, F., Tu, J., & Zhou, G. (2015). Investor sentiment aligned: A powerful predictor of stock returns. *The Review of Financial Studies*, *28*(3), 791–837.

Huang, D., Li, J., & Wang, L. (2021). Are disagreements agreeable? evidence from information aggregation. *Journal of Financial Economics*, *141*(1), 83–101.

Kelly, B., & Pruitt, S. (2013). Market expectations in the cross-section of present values. *The Journal of Finance*, *68*(5), 1721–1756.

Kelly, B., & Pruitt, S. (2015). The three-pass regression filter: A new approach to forecasting using many predictors. *Journal of Econometrics*, *186*(2), 294–316.

Koenker, R., & Bassett Jr, G. (1978). Regression quantiles. *Econometrica*, 33–50.

Koenker, R., & Machado, J. A. (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association*, *94*(448), 1296–1310.

Li, G., Li, Y., & Tsai, C.-L. (2015). Quantile correlations and quantile autoregressive modeling. *Journal of the American Statistical Association*, *110*(509), 246–261.

Light, N., Maslov, D., & Rytchkov, O. (2017). Aggregation of information about the cross section of stock returns: A latent variable approach. *The Review of Financial Studies*, *30*(4), 1339–1381.

Lyle, M. R., & Wang, C. C. (2015). The cross section of expected holding period returns and their dynamics: A present value approach. *Journal of Financial Economics*, *116*(3), 505–525.

Ma, S., Li, R., & Tsai, C.-L. (2017). Variable screening via quantile partial correlation. *Journal of the American Statistical Association*, *112*(518), 650–663.

Méndez-Civieta, Á., Aguilera-Morillo, M. C., & Lillo, R. E. (2022). Fast partial quantile regression. *Chemometrics and Intelligent Laboratory Systems*, *223*, 104533.

Patton, A. J., Ziegel, J. F., & Chen, R. (2019). Dynamic semiparametric models for expected shortfall (and value-at-risk). *Journal of Econometrics*, *211*(2), 388–413.

Pelger, M., & Xiong, R. (2022). State-varying factor models of large dimensions. *Journal of Business & Economic Statistics*, *40*(3), 1315–1333.

Rivas, M. D. G., & Gonzalo, J. (2020). Trends in distributional characteristics: Existence of global warming. *Journal of Econometrics*, *214*(1), 153–174.

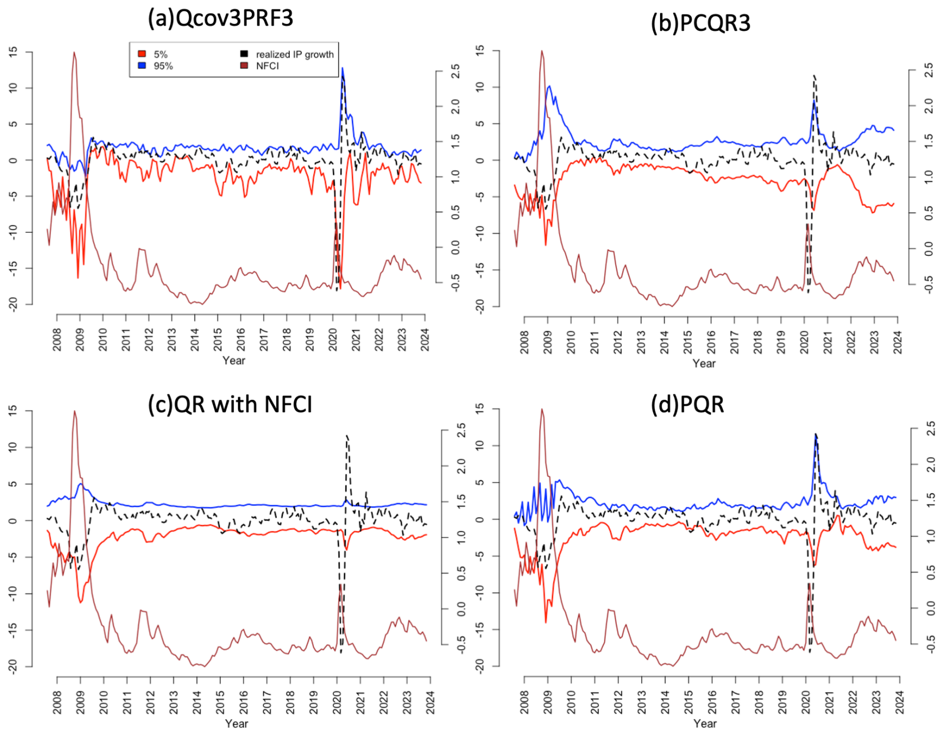
Stock, J. H., & Watson, M. W. (2002a). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, *97*(460), 1167–1179.

Stock, J. H., & Watson, M. W. (2002b). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, *20*(2), 147–162.

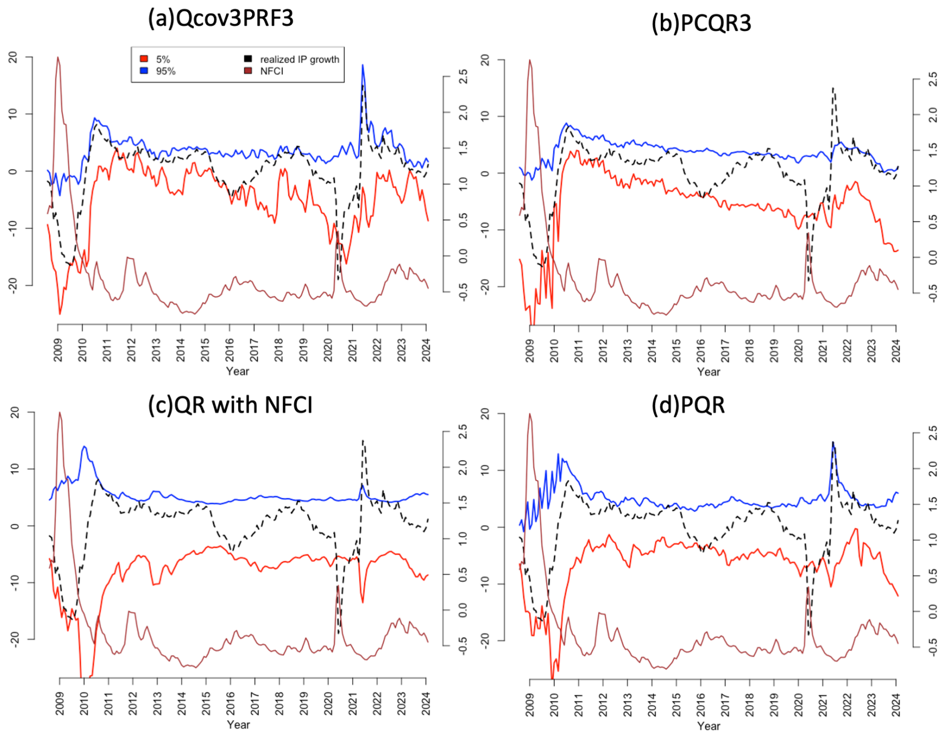
Su, L., & Wang, X. (2017). On time-varying factor models: Estimation and testing. *Journal of Econometrics*, *198*(1), 84–101.

White, H. (1994). *Estimation, inference and specification analysis*. Cambridge university press.

Wold, H. (1966). Estimation of principal components and related models by iterative least squares. *Multivariate Analysis*, 391–420.

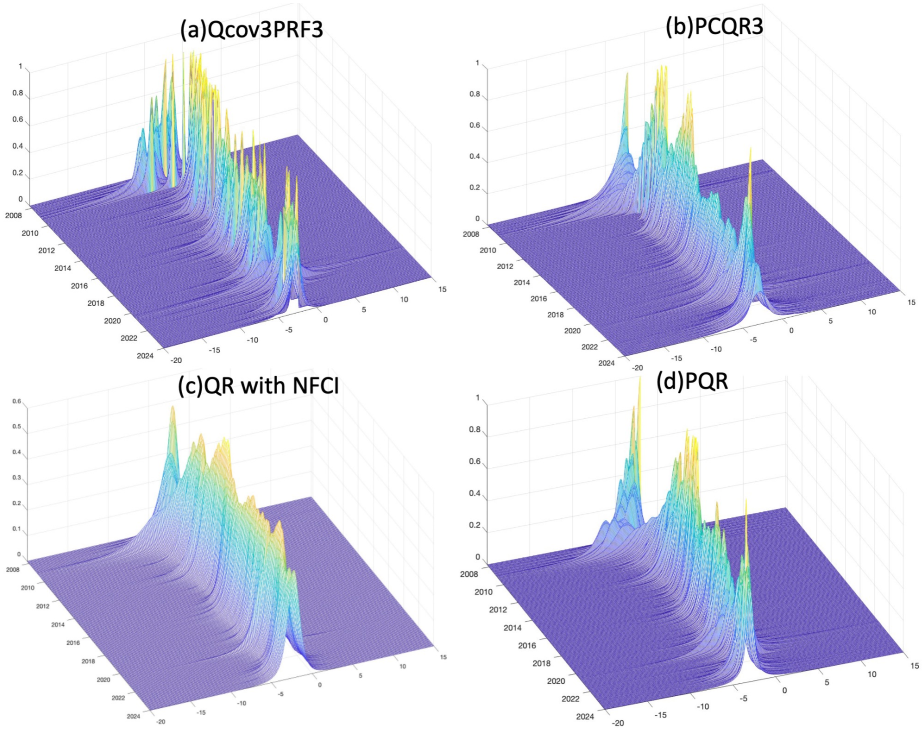


(i) 3-month ahead

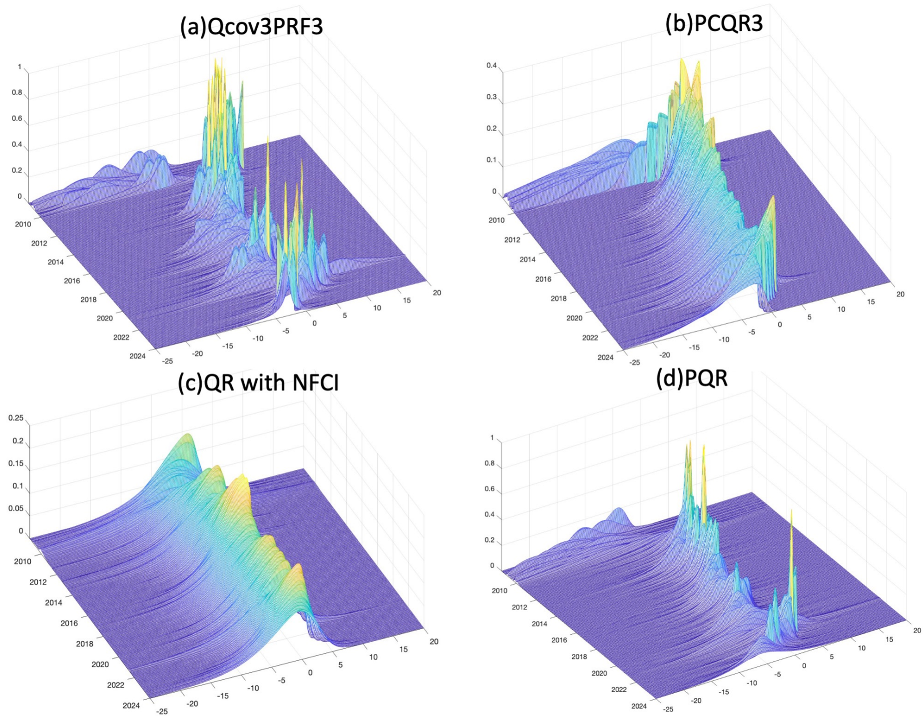


(ii) 12-month ahead

Figure 1: Conditional quantiles and realized IP growth, corresponding to the vertical left axis, estimated with (a) Qcov3PRF with 3 factors, (b) PCQR with 3 factors, (c) Quantile Regression with NFCI as a regressor, and (d) PQR. The vertical right axis corresponds to NFCI.

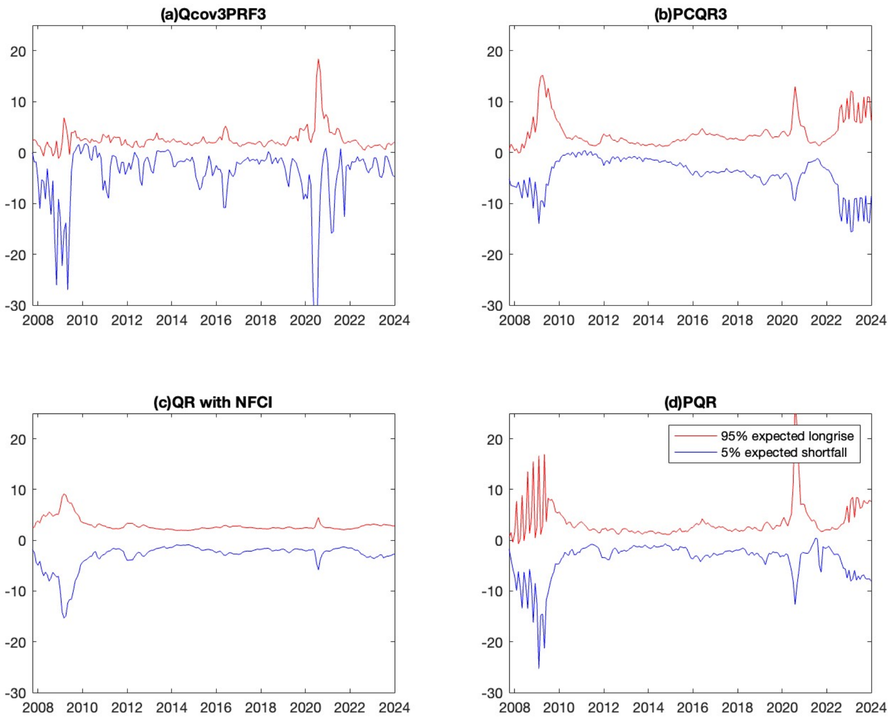


(i) 3-month ahead

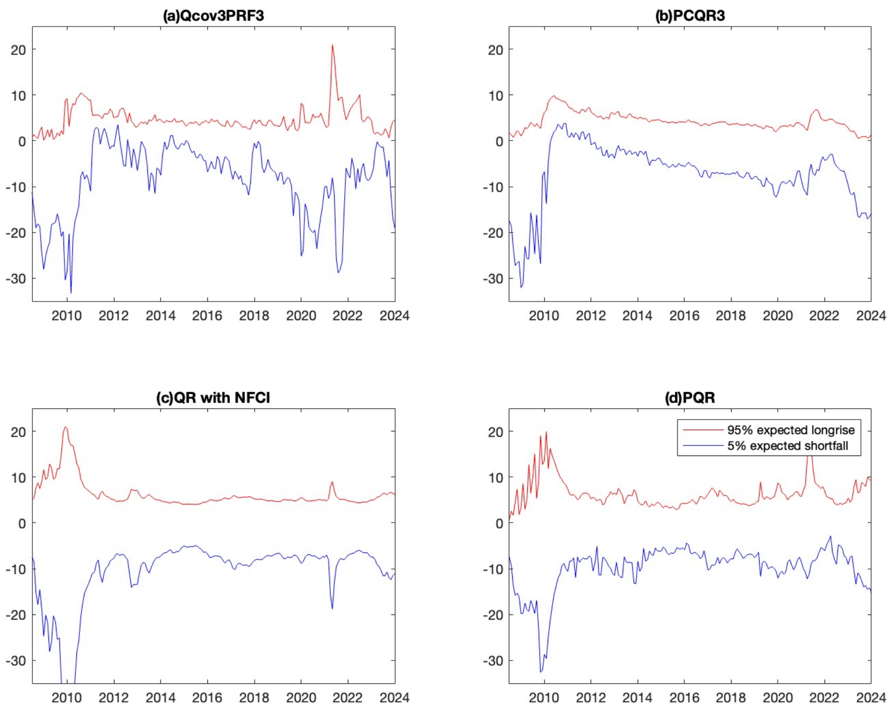


(ii) 12-month ahead

Figure 2: Conditional parametric distribution of the IP growth based on conditional quantiles after being estimated with (a) Qcov3PRF with 3 factors, (b) PCQR with 3 factors, (c) Quantile Regression with NFCI as a regressor, and (d) PQR.



(i) 3-month ahead



(ii) 12-month ahead

Figure 3: Conditional expected shortfall (longrise) of the IP growth after the conditional parametric distribution is estimated with the conditional quantiles etimated by: (a) Qcov3PRF with 3 factors, (b) PCQR with 3 factors, (c) Quantile Regression with NFCI as a regressor, and (d) PQR.

Table 1: Out-of-sample $R^2_\tau(\%)$ for DGP (27) with $h = 1$.

		$N = 100$, $T = 200$					$N = 200$, $T = 400$				
τ	$\rho_T = \rho_N$	Qcov3PRF3	Qcov3PRF2	Qcov3PRF1	PQR	PCQR6	Qcov3PRF3	Qcov3PRF2	Qcov3PRF1	PQR	PCQR6
0.10	0	50.9	49.6	42.6	44.5	51.5	52.9	51.8	45.3	45.9	53.1
	0.2	50.6	49.4	42.4	43.7	50.9	52.5	51.5	44.8	45.7	52.8
	0.4	48.9	47.7	40.9	42.3	46.4	51.6	50.7	44.1	44.8	51.6
	0.6	46.4	44.0	36.1	38.1	33.5	49.8	48.3	41.3	42.2	38.2
	0.8	41.0	33.9	17.7	20.5	22.4	43.4	37.3	23.6	26.3	9.3
0.25	0	51.9	50.7	43.7	44.4	52.0	53.4	52.3	45.4	45.5	53.4
	0.2	51.4	50.3	43.3	44.1	51.5	53.0	51.9	45.2	45.4	53.1
	0.4	50.2	49.1	42.1	42.9	47.3	52.0	50.9	44.0	44.5	51.7
	0.6	48.3	46.1	38.3	39.4	34.9	50.8	49.3	42.0	42.3	38.9
	0.8	43.6	37.2	21.7	23.6	25.3	45.3	39.7	26.5	28.6	9.8
0.50	0	51.9	50.7	44.1	44.5	52.0	53.2	52.2	45.4	45.6	53.3
	0.2	51.5	50.4	43.8	44.0	51.4	53.0	52.0	45.3	45.4	53.0
	0.4	50.2	49.0	42.1	42.6	47.2	52.4	51.2	44.5	44.5	51.7
	0.6	48.8	46.3	38.6	39.2	34.8	51.0	49.3	42.4	42.7	38.9
	0.8	44.1	38.1	22.7	24.9	24.6	45.5	40.2	27.7	29.2	10.4
0.75	0	51.7	50.5	43.7	44.1	52.0	53.2	52.0	45.5	45.9	53.3
	0.2	51.2	50.0	43.5	44.0	51.2	52.9	51.9	45.1	45.3	52.9
	0.4	49.9	48.6	41.6	42.4	46.8	52.2	51.2	44.4	44.8	51.9
	0.6	48.3	45.8	38.0	39.0	34.4	50.6	49.1	41.7	42.1	38.8
	0.8	42.8	36.5	20.7	23.2	24.3	45.4	39.9	27.0	29.0	10.6
0.90	0	51.5	50.1	43.3	45.0	51.7	53.1	51.9	45.0	45.9	53.3
	0.2	50.3	49.4	41.9	43.6	50.5	52.5	51.5	44.8	45.5	52.7
	0.4	48.6	47.6	40.4	41.9	46.1	51.4	50.6	43.6	44.3	51.5
	0.6	46.2	43.4	35.8	37.9	32.7	49.6	48.3	41.4	42.5	38.7
	0.8	40.6	34.2	17.5	20.5	22.5	43.3	37.1	22.7	25.8	8.4

Note: The median of R^2_τ out of 1000 simulations in each case is reported. R^2_τ evaluates the forecasts of the second half of the sample (100 and 200 observations, respectively). We set $\sigma_y = 0.5$. ρ_T represents the level of serial correlation, and ρ_N is the level of cross-section correlation. The model consists on six latent factors, three are relevant. Qcov3PRF# denotes Qcov3PRF implementation with # number of factors, similarly for PCQR#.

Table 2: Out-of-sample $R^2_\tau(\%)$ for DGP (28) with $h = 1$ and $\tau = 0.50$.

$\rho_T = \rho_N$	α^f	α^g	Qcov3PRF2	Qcov3PRF1	PQR	PCQR4	3PRF2	PCR4
			$N = 100$, $T = 200$					
0	0	0	45.2	36.8	37.9	47.4	40.1	41.3
0.2	0	0	45.0	36.0	37.7	47.0	40.0	41.3
0.4	0	0	43.1	34.2	35.4	37.5	37.0	32.2
0.6	0	0	39.7	28.9	30.2	5.1	31.9	2.6
0.8	0	0	28.4	12.6	13.6	-1.4	19.5	-3.4
0	0.6	0.3	52.8	42.9	43.9	54.9	47.6	49.0
0.2	0.6	0.3	53.7	43.2	44.8	55.8	49.0	50.6
0.4	0.6	0.3	51.3	41.2	42.5	50.2	45.9	45.0
0.6	0.6	0.3	47.3	36.1	37.3	20.5	41.2	16.2
0.8	0.6	0.3	35.4	16.9	17.8	0.7	29.1	-1.6
0	0.3	0.6	46.2	34.2	36.0	49.7	41.7	43.9
0.2	0.3	0.6	45.1	33.8	35.1	48.3	38.8	41.6
0.4	0.3	0.6	43.2	31.6	32.9	40.5	37.5	34.7
0.6	0.3	0.6	39.0	26.7	28.1	8.3	33.3	4.9
0.8	0.3	0.6	28.0	11.9	12.9	-1.1	20.6	-3.1
			$N = 200$, $T = 400$					
0	0	0	45.8	39.2	40.0	47.8	41.2	42.2
0.2	0	0	44.7	38.1	39.0	46.8	39.3	40.1
0.4	0	0	44.1	37.5	38.3	45.5	37.2	39.5
0.6	0	0	41.6	33.6	34.7	17.2	32.5	13.3
0.8	0	0	31.0	16.3	17.5	-0.6	19.9	-2.1
0	0.6	0.3	53.7	46.0	46.9	55.4	49.3	49.9
0.2	0.6	0.3	52.9	45.1	45.9	54.6	48.3	49.3
0.4	0.6	0.3	52.6	44.3	45.0	53.7	46.9	48.6
0.6	0.6	0.3	49.4	41.1	41.8	30.7	42.5	26.7
0.8	0.6	0.3	39.0	22.2	23.2	0.5	31.4	-1.3
0	0.3	0.6	46.2	37.9	39.1	49.1	40.5	42.3
0.2	0.3	0.6	46.3	37.8	39.1	48.9	41.2	42.8
0.4	0.3	0.6	44.5	35.9	37.2	46.9	39.1	41.2
0.6	0.3	0.6	42.0	32.9	34.2	21.2	33.3	16.5
0.8	0.3	0.6	31.7	16.0	17.0	-0.5	19.3	-2.3

Note: The median of R^2_τ out of 1000 simulations in each case is reported. R^2_τ evaluates the forecasts of the second half of the sample (100 and 200 observations, respectively). We set $\sigma_y = 0.75$. The predictors contain four latent factors, two are relevant. All the factors follow independent AR(1) processes. α^f (α^g_t) denotes the value of the coefficient for the autorregresive term of the relevant (irrelevant) factors. ρ_T represents the level of serial correlation, and ρ_N is the level of cross-section correlation. Qcov3PRF# denotes Qcov3PRF implementation with # number of factors, similarly for PCQR#, PCR# and 3PRF#.

Table 3: Out-of-sample $R^2_\tau(\%)$ for DGP (29) with $h = 1$.

		$N = 100, T = 200$				$N = 200, T = 400$			
τ	$\rho_T = \rho_N$	Qcov3PRF2	Qcov3PRF1	PQR	PCQR4	Qcov3PRF2	Qcov3PRF1	PQR	PCQR4
0.10	0	42.2	35.4	38.0	42.4	43.7	39.6	41.5	43.6
	0.2	41.7	34.9	37.6	41.6	43.7	39.9	41.3	43.7
	0.4	39.9	32.6	35.4	39.4	43.4	39.0	40.9	43.2
	0.6	36.0	27.3	30.5	7.7	40.3	35.1	37.4	28.8
	0.8	25.4	9.6	11.2	-4.1	30.6	15.5	17.9	-1.7
0.25	0	46.1	40.6	42.4	46.2	47.0	43.9	44.8	47.0
	0.2	46.0	40.0	41.6	46.0	46.9	43.7	44.7	46.9
	0.4	44.3	38.2	39.7	42.8	46.2	42.8	43.7	46.1
	0.6	41.8	33.1	35.3	8.9	44.5	39.6	40.8	30.4
	0.8	31.7	14.4	16.3	-2.1	36.4	20.5	22.5	-0.8
0.50	0	47.2	42.1	43.4	47.2	48.0	45.1	46.0	48.1
	0.2	46.8	41.7	42.7	46.8	47.9	45.1	45.6	47.9
	0.4	45.8	39.9	41.2	44.2	47.3	44.4	45.1	47.1
	0.6	43.5	34.9	36.6	9.2	45.6	41.3	42.1	32.1
	0.8	34.8	16.5	18.6	-1.4	38.6	22.9	24.5	-0.7
0.75	0	46.2	40.9	42.6	46.2	47.2	43.9	44.8	47.2
	0.2	45.8	40.2	42.1	45.8	46.7	43.2	44.3	46.6
	0.4	44.8	39.0	40.9	43.6	46.6	42.9	44.0	46.3
	0.6	41.4	32.9	34.8	8.9	44.6	39.9	41.0	31.1
	0.8	32.7	14.8	16.8	-2.0	36.6	20.8	23.1	-0.9
0.90	0	42.2	35.5	38.8	42.4	43.5	39.4	41.3	43.7
	0.2	41.9	35.0	38.0	42.2	43.9	39.6	41.5	43.9
	0.4	39.8	32.7	35.2	38.9	42.8	38.6	40.5	42.7
	0.6	37.0	27.6	31.1	6.7	40.8	35.3	37.2	29.2
	0.8	25.7	10.2	12.2	-4.3	30.4	15.1	17.3	-1.8

Note: The median of R^2_τ out of 1000 simulations in each case is reported. R^2_τ evaluates the forecasts of the second half of the sample (100 and 200 observations, respectively). We set $\sigma_y = 1.0$. ρ_T represents the degree of serial correlation, and ρ_N is the degree of cross-section correlation. The model consists on four latent factors, two are relevant. Qcov3PRF# denotes Qcov3PRF implementation with # number of factors, similarly for PCQR#.

Table 4: In-sample MAE, MSE and correlations with DGP (30)-(31).

			DGP (31.a)				DGP (31.b)			
	$T = N$	τ	Qcov3PRF3	Qcov3PRF1	PQR	PCQR6	Qcov3PRF3	Qcov3PRF1	PQR	PCQR6
MAE	100	0.05	0.843	0.880	0.805	0.857	0.459	0.528	0.509	0.394
MSE			1.187	1.285	1.080	1.257	0.336	0.439	0.410	0.257
Correlation			0.844	0.794	0.840	0.819	0.637	0.406	0.460	0.755
MAE	200	0.05	0.593	0.873	0.796	0.568	0.340	0.476	0.457	0.269
MSE			0.584	1.223	1.024	0.543	0.184	0.359	0.331	0.118
Correlation			0.914	0.770	0.815	0.923	0.824	0.533	0.576	0.892
MAE	1000	0.05	0.272	0.561	0.515	0.240	0.132	0.328	0.303	0.111
MSE			0.120	0.502	0.418	0.095	0.028	0.179	0.154	0.020
Correlation			0.984	0.912	0.929	0.989	0.976	0.789	0.821	0.983
MAE	100	0.10	0.685	0.814	0.769	0.667	0.372	0.428	0.419	0.311
MSE			0.786	1.067	0.960	0.756	0.222	0.290	0.281	0.159
Correlation			0.821	0.693	0.736	0.835	0.701	0.473	0.512	0.792
MAE	200	0.10	0.483	0.667	0.626	0.451	0.265	0.378	0.365	0.212
MSE			0.388	0.716	0.632	0.339	0.113	0.227	0.213	0.073
Correlation			0.915	0.798	0.829	0.928	0.856	0.593	0.631	0.910
MAE	1000	0.10	0.232	0.453	0.426	0.192	0.108	0.242	0.229	0.090
MSE			0.087	0.322	0.283	0.061	0.019	0.099	0.088	0.013
Correlation			0.981	0.912	0.924	0.989	0.976	0.839	0.858	0.985
MAE	100	0.25	0.589	0.580	0.573	0.506	0.295	0.317	0.313	0.231
MSE			0.587	0.544	0.534	0.431	0.142	0.160	0.157	0.088
Correlation			0.739	0.629	0.654	0.789	0.726	0.524	0.553	0.801
MAE	200	0.25	0.409	0.472	0.461	0.348	0.210	0.269	0.262	0.161
MSE			0.279	0.358	0.342	0.202	0.071	0.116	0.111	0.042
Correlation			0.862	0.743	0.764	0.896	0.859	0.650	0.676	0.909
MAE	1000	0.25	0.197	0.331	0.321	0.150	0.095	0.163	0.155	0.068
MSE			0.063	0.170	0.160	0.037	0.015	0.045	0.041	0.007
Correlation			0.963	0.873	0.882	0.982	0.966	0.870	0.884	0.984
MAE	100	0.75	0.629	0.504	0.504	0.506	0.290	0.281	0.281	0.218
MSE			0.664	0.418	0.420	0.438	0.137	0.126	0.127	0.079
Correlation			0.632	0.632	0.647	0.708	0.684	0.504	0.529	0.767
MAE	200	0.75	0.449	0.397	0.391	0.341	0.209	0.237	0.234	0.151
MSE			0.335	0.253	0.246	0.193	0.071	0.090	0.088	0.037
Correlation			0.786	0.752	0.768	0.864	0.826	0.626	0.647	0.889
MAE	1000	0.75	0.206	0.270	0.263	0.145	0.093	0.145	0.139	0.064
MSE			0.069	0.110	0.105	0.035	0.015	0.035	0.033	0.007
Correlation			0.946	0.882	0.889	0.976	0.956	0.854	0.867	0.981
MAE	100	0.90	0.694	0.718	0.686	0.675	0.344	0.372	0.369	0.291
MSE			0.809	0.834	0.768	0.774	0.190	0.218	0.217	0.139
Correlation			0.777	0.708	0.746	0.785	0.661	0.453	0.485	0.752
MAE	200	0.90	0.490	0.565	0.530	0.440	0.252	0.325	0.319	0.199
MSE			0.403	0.519	0.458	0.327	0.102	0.168	0.162	0.065
Correlation			0.887	0.822	0.851	0.913	0.824	0.566	0.597	0.886
MAE	1000	0.90	0.245	0.336	0.316	0.188	0.103	0.210	0.199	0.083
MSE			0.098	0.179	0.157	0.059	0.018	0.074	0.066	0.011
Correlation			0.972	0.939	0.948	0.986	0.969	0.822	0.842	0.981
MAE	100	0.95	0.837	0.951	0.890	0.857	0.418	0.459	0.448	0.368
MSE			1.171	1.437	1.281	1.251	0.279	0.327	0.315	0.225
Correlation			0.785	0.670	0.726	0.778	0.596	0.387	0.432	0.714
MAE	200	0.95	0.578	0.764	0.687	0.569	0.318	0.414	0.400	0.251
MSE			0.558	0.943	0.771	0.547	0.161	0.267	0.251	0.103
Correlation			0.902	0.790	0.839	0.907	0.790	0.498	0.544	0.867
MAE	1000	0.95	0.283	0.439	0.394	0.237	0.125	0.282	0.263	0.102
MSE			0.132	0.313	0.250	0.094	0.026	0.130	0.114	0.017
Correlation			0.977	0.934	0.949	0.987	0.969	0.777	0.808	0.980

Note: MAE and MSE denote the mean absolute error and the mean squared error, respectively. These measures consider the difference between the true conditional quantile of the target variable and the estimated forecast. Correlation reports the standard linear correlation and considers the same two time series. We run 1000 simulations. In both DGPs we consider six factors, three are relevant. Qcov3PRF# denotes Qcov3PRF implementation with # number of factors, similarly for PCQR#.

Table 5: Out-of-sample $R^2_\tau(\%)$ IP growth with $h = 3$.

	τ=0.05	τ=0.1	τ=0.25	τ=0.5	τ=0.75	τ=0.9	τ=0.95
Rolling window. Starting in 2011M08.							
NFCI	-0.5	1.5	-0.3	-8.1	-15.1	-12.8	-8.7
Risk	-39.1	-1.6	-3.0	-6.1	-15.2	-19.6	-7.1
Credit	7.1	2.1	1.9	-5.1	-16.8	-8.2	0.2
Leverage	3.0	9.4	5.7	0.9	-7.2	-19.8	14.1
Nonfin. Lev.	-31.8	-10.7	-5.0	-8.0	-3.9	-15.7	-51.7
Qcov3PRF1	34.7	13.9	6.8	4.2	1.4	15.4	16.1
Qcov3PRF2	39.8	26.4	13.2	3.6	7.0	9.4	10.9
Qcov3PRF3	39.0	24.6	2.7	1.9	12.1	9.3	11.2
PQR	40.5	16.5	-0.5	-1.3	7.1	9.0	10.6
PCQR1	-6.2	-8.6	-4.1	-10.7	-13.4	-22.9	-7.7
PCQR2	6.8	11.6	2.4	-1.7	-10.1	-5.0	-0.3
PCQR3	-7.6	15.5	4.1	-3.4	0.5	-10.1	3.9
PCQR-LASSO	25.0	14.0	9.0	9.0	3.0	-11.0	16.0
Rolling window. Starting in 2015M10.							
NFCI	-7.6	-8.6	-4.7	-7.9	-17.0	-13.8	-8.0
Risk	-54.4	-15.2	-6.6	-3.2	-16.6	-22.2	-5.8
Credit	0.1	-11.4	-4.8	-6.0	-20.1	-9.3	1.0
Leverage	-3.8	0.5	4.0	3.3	-5.5	-22.7	16.3
Nonfin. Lev.	-45.9	-26.5	-12.9	-10.0	-5.2	-18.4	-58.2
Qcov3PRF1	31.7	0.0	-1.5	2.0	-1.8	14.5	16.0
Qcov3PRF2	34.7	15.7	4.8	2.0	6.0	10.8	12.7
Qcov3PRF3	34.3	14.0	-7.7	0.9	11.8	9.5	13.0
PQR	36.1	3.4	-13.0	-5.6	5.5	6.9	9.6
PCQR1	-15.3	-21.4	-8.2	-9.4	-13.6	-24.3	-5.6
PCQR2	-1.6	-1.1	-7.4	-4.9	-14.5	-8.0	0.4
PCQR3	-20.8	3.1	-4.5	-5.3	-0.5	-11.4	5.4
PCQR-LASSO	26.0	7.0	2.0	9.0	0.0	-18.0	20.0
Expanding window. Starting in 2011M08.							
NFCI	15.8	12.2	0.3	-4.9	-5.2	0.5	1.0
Risk	13.2	10.1	-6.4	-11.1	-7.0	0.1	-0.5
Credit	18.9	14.5	6.0	-0.3	-3.0	1.1	3.0
Leverage	18.1	14.2	1.3	-4.5	-8.2	-5.4	-0.8
Nonfin. Lev.	17.2	11.2	4.6	0.7	-1.7	-2.2	-2.1
Qcov3PRF1	19.1	14.9	8.2	2.9	4.1	6.4	10.4
Qcov3PRF2	28.9	23.9	13.6	6.4	2.1	10.2	19.9
Qcov3PRF3	27.2	26.8	14.2	8.8	6.1	14.7	20.0
PQR	19.9	14.9	7.8	0.2	2.0	6.6	14.3
PCQR1	17.0	12.8	2.0	-4.8	-7.8	-3.5	-2.4
PCQR2	20.6	14.0	4.0	-3.9	-2.0	1.8	0.4
PCQR3	20.8	14.0	1.8	-10.6	-14.8	-6.1	9.3
PCQR-LASSO	24.0	24.0	6.0	-2.0	-1.0	3.0	4.0
Expanding window. Starting in 2015M10.							
NFCI	7.5	6.1	-2.1	-0.1	0.1	0.1	0.4
Risk	6.4	5.8	-6.7	-3.0	-0.2	-0.3	-1.4
Credit	7.5	4.6	1.4	1.9	0.1	0.5	2.7
Leverage	9.8	10.1	0.4	1.2	-0.1	-1.7	0.9
Nonfin. Lev.	7.0	1.3	0.7	2.6	1.0	-1.6	-2.7
Qcov3PRF1	11.3	8.2	3.8	4.2	4.9	3.0	8.7
Qcov3PRF2	18.0	11.7	5.8	7.0	3.1	9.7	20.0
Qcov3PRF3	14.4	15.0	6.3	8.3	6.8	13.6	21.6
PQR	13.1	8.5	3.3	3.6	3.9	4.4	14.1
PCQR1	8.1	6.6	-0.5	0.5	-0.2	-0.3	-0.8
PCQR2	9.0	2.7	-1.3	0.8	2.7	0.5	-1.2
PCQR3	9.2	2.4	-3.6	-8.1	-10.4	-6.7	11.2
PCQR-LASSO	29.0	21.0	-1.0	1.0	4.0	9.0	17.0

Note: The out-of-sample R^2_τ is reported. It evaluates the out-of-sample performance of the samples 2011M08-2023M12 and 2015M08-2023M12, respectively. August 2011 represents the observation 51 and October 2015 the observation 101. The size of the rolling window is 50. The variables Risk, Credit, Leverage and Nonfi. Lev. consider a subgroup of indicators that construct the NFCI that do not overlap. The IP growth is equal to the log difference of the current IP and the IP 3 months before. Qcov3PRF# and PCQR# denote Qcov3PRF and PCQR implementation with # number of factors, respectively. PCQR-LASSO selects predictors from the 40 factors corresponding to the highest eigenvalues. The predictors are standardized.

Table 6: Out-of-sample $R^2_\tau(\%)$ for IP growth with $h = 12$.

	τ=0.05	τ=0.1	τ=0.25	τ=0.5	τ=0.75	τ=0.9	τ=0.95
Rolling window. Starting in 2011M08.							
NFCI	1.0	11.2	5.4	-5.8	-10.6	-6.5	1.1
Risk	-5.2	11.3	5.2	-2.1	-11.9	-8.4	-4.4
Credit	6.8	10.2	4.4	-4.5	-10.4	1.4	7.3
Leverage	5.6	11.5	7.2	-5.1	-2.3	-10.0	12.8
Nonfin. Lev.	-13.2	10.7	-8.1	-28.6	-17.7	-14.1	-30.5
Qcov3PRF1	32.9	20.0	20.4	5.8	3.7	15.4	21.7
Qcov3PRF2	39.2	34.6	34.3	10.7	13.4	21.4	13.9
PQR	36.2	35.8	25.5	6.0	9.7	10.1	17.1
PCQR1	21.4	19.5	5.6	-7.4	-7.4	-9.8	-9.5
PCQR2	8.3	15.0	-4.8	-19.7	-1.4	6.0	-0.9
PCQR3	19.8	30.2	18.1	-0.4	10.5	14.1	1.0
PCQR-LASSO	14.0	34.0	35.0	18.0	11.0	12.0	31.0
Rolling window. Starting in 2015M10.							
NFCI	-13.3	0.9	3.7	-4.5	-14.8	-13.5	-5.9
Risk	-22.5	-1.3	1.2	0.5	-16.5	-15.1	-12.0
Credit	-1.4	-0.9	3.1	-3.7	-15.6	-4.3	1.3
Leverage	-8.9	-3.3	1.4	-2.4	-1.4	-13.3	13.4
Nonfin. Lev.	-36.0	-7.1	-25.4	-40.9	-30.1	-26.9	-48.2
Qcov3PRF1	22.3	0.1	9.2	3.0	-1.9	8.6	16.9
Qcov3PRF2	30.8	19.8	28.4	2.6	4.7	16.4	5.6
PQR	27.8	24.1	16.3	2.5	4.5	1.2	10.3
PCQR1	13.9	9.5	0.0	-4.4	-10.3	-16.4	-17.1
PCQR2	-7.4	-2.3	-22.3	-32.6	-11.9	-2.6	-11.4
PCQR3	9.2	19.5	10.4	-5.6	3.1	6.4	-7.3
PCQR-LASSO	2.0	28.0	20.0	16.0	4.0	1.0	27.0
Expanding window. Starting in 2011M08.							
NFCI	24.8	20.4	-4.1	-2.6	0.1	7.1	13.3
Risk	17.2	13.7	-1.9	-6.1	0.1	3.9	8.8
Credit	30.9	24.5	-4.5	0.9	0.3	8.3	15.8
Leverage	26.6	25.7	2.5	-2.6	-1.4	-0.2	10.4
Nonfin. Lev.	33.8	33.0	13.8	3.1	-1.2	2.3	7.3
Qcov3PRF1	30.7	24.9	12.7	7.7	13.0	17.4	20.4
Qcov3PRF2	40.1	40.8	21.9	14.4	23.6	23.7	28.7
Qcov3PRF3	37.5	46.1	26.3	21.6	24.0	28.3	28.8
PQR	35.5	29.8	6.6	5.3	10.1	15.3	17.6
PCQR1	24.1	20.9	-7.3	0.8	-0.1	0.1	8.7
PCQR2	36.6	31.9	5.3	4.0	4.9	9.2	3.2
PCQR3	38.4	32.8	8.4	7.5	11.0	26.0	27.7
PCQR-LASSO	30.0	20.0	26.0	8.0	10.0	15.0	15.0
Expanding window. Starting in 2015M10.							
NFCI	7.9	-2.7	-4.9	1.1	0.1	1.0	5.4
Risk	-0.3	-9.5	-2.1	-1.0	-0.1	-4.7	-1.6
Credit	14.8	2.2	-10.6	3.3	0.5	3.4	10.1
Leverage	14.8	10.2	9.8	0.2	-0.3	-0.1	7.3
Nonfin. Lev.	24.9	19.7	18.3	11.8	2.9	-6.5	-3.7
Qcov3PRF1	17.1	7.3	7.5	3.9	10.2	9.4	12.1
Qcov3PRF2	18.9	18.4	6.9	10.2	21.8	14.8	20.6
Qcov3PRF3	19.6	23.9	14.7	18.3	20.9	19.9	20.5
PQR	18.4	8.7	3.5	5.7	7.0	8.3	9.3
PCQR1	10.0	0.7	-2.8	6.6	1.1	-0.4	6.1
PCQR2	19.2	6.5	-9.2	-0.2	0.7	-2.4	-10.9
PCQR3	21.9	8.3	-4.5	5.0	9.0	21.4	21.6
PCQR-LASSO	29.0	19.0	17.0	13.0	12.0	9.0	6.0

Note: The out-of-sample R^2_τ is reported. It evaluates the out-of-sample performance of the samples 2011M08-2023M12 and 2015M08-2023M12, respectively. August 2011 represents the observation 51 and October 2015 the observation 101. The size of the rolling window is 50. The variables Risk, Credit, Leverage and Nonfi. Lev. consider a subgroup of indicators that construct the NFCI that do not overlap. The IP growth is equal to the log difference of the current IP and the IP 12 months before. Qcov3PRF# and PCQR# denote Qcov3PRF and PCQR implementation with # number of factors, respectively. PCQR-LASSO selects predictors from the 40 factors corresponding to the highest eigenvalues. The predictors are standardized.

Table 7: Check Loss for IP growth in-sample prediction.

Method	τ=0.05	τ=0.25	τ=0.50	τ=0.75	τ=0.95
			3-month ahead		
NFCI	0.319	0.552	0.611	0.506	0.226
Qcov3PRF1	0.291	0.494	0.557	0.436	0.137
Qcov3PRF2	0.206	0.430	0.474	0.376	0.113
Qcov3PRF3	0.172	0.399	0.452	0.355	0.094
PQR	0.287	0.496	0.550	0.428	0.135
PCQR1	0.310	0.541	0.600	0.497	0.227
PCQR2	0.295	0.528	0.596	0.477	0.216
PCQR3	0.292	0.528	0.595	0.474	0.191
			12-month ahead		
NFCI	0.611	1.695	1.760	1.209	0.383
Qcov3PRF1	0.518	1.170	1.345	0.845	0.274
Qcov3PRF2	0.377	0.860	1.033	0.750	0.191
Qcov3PRF3	0.326	0.752	0.952	0.631	0.178
PQR	0.503	1.199	1.303	0.799	0.267
PCQR1	0.591	1.557	1.660	1.195	0.393
PCQR2	0.461	1.227	1.330	0.935	0.342
PCQR3	0.444	1.182	1.278	0.918	0.340

Note: IP growth is equal to the log difference of the current IP and IP h months before. Qcov3PRF# and PCQR# denote Qcov3PRF and PCQR implementation with # number of factors, respectively.

Table 8: Out-of-sample average FZ0 loss for IP growth prediction.

	QR with NFCI	Qcov3PRF3	PCQR3	PQR
		3-month ahead		
expected shortfall (5%)	0.96	0.43	0.82	0.84
expected longrise (95%)	0.57	0.07	0.36	0.17
		12-month ahead		
expected shortfall (5%)	1.15	0.53	0.95	1.09
expected longrise (95%)	0.68	0.06	0.80	0.15

Note: For the expected longrise we consider the FZ0 loss function but evaluated with the negative quantile and expected longrise of the target. The values are divided by 100 for ease of presentation. Qcov3PRF# and PCQR# denote Qcov3PRF and PCQR implementation with # number of factors, respectively. The values are divided by 100 for ease of presentation.

Table 9: Out-of-sample R^2_τ for one-quarter ahead forecast of Real GDP growth.

Starting Validation Date	1981Q1			1991Q1			2001Q1		
	τ=0.05	τ=0.1	τ=0.25	τ=0.05	τ=0.1	τ=0.25	τ=0.05	τ=0.1	τ=0.25
QR	-111.7	-29.5	15.0	46.2	49.8	56.9	48.9	66.9	66.2
QR with $CFNAI_t$	-117.4	-31.5	16.5	44.8	51.7	55.1	49.0	57.9	63.9
QR-LASSO	-8.7	2.6	-3.6	34.3	16.3	-5.0	35.2	20.3	11.9
PCQR1	14.9	18.9	5.5	46.6	33.2	15.5	12.1	29.2	38.9
PQR-$z_t = y_{t+1}$	18.6	28.1	15.7	54.9	45.3	34.8	58.9	53.6	58.8
PQR-$z_t = CFNAI_t$	19.3	30.4	23.3	58.5	46.6	44.3	58.6	57.7	55.4
Qcov3PRF1-$z_t = y_{t+1}$	21.3	27.8	24.8	57.9	48.5	38.1	58.5	65.6	56.5
Qcov3PRF1-$z_t = CFNAI_t$	22.1	32.5	25.3	58.3	56.6	44.4	61.3	53.8	60.0
MF-Qcov3PRF1-$z_t = y_{t+1}$	55.0	60.8	51.8	64.6	56.7	50.5	61.0	59.5	61.0
MF-Qcov3PRF1-$z_t = CFNAI_t$	54.2	61.6	51.6	61.5	61.0	50.2	61.4	60.0	60.7

Note: The forecasts are evaluated using a rolling window scheme of length 40, 80, and 120 with validation period starting in 1981Q1, 1991Q1 and 2001Q1, respectively. QR denotes quantile regression. QR-LASSO quantile regression with LASSO. Qcov3PRF1-$z_t = y_{t+1}$, Qcov3PRF with 1 automatic proxy. Qcov3PRF1-$z_t = CFNAI_t$, Qcov3PRF with $CFNAI_t$ as a proxy. Similarly for PQR and MF-Qcov3PRF1. MF-Qcov3PRF1 works with monthly predictors, and quarterly proxy and target.

Table 10: Out-of-sample R^2_τ for one-year ahead forecast of the change in Global Land-Ocean Temperature Index.

Method	τ=0.05	τ=0.1	τ=0.25	τ=0.5	τ=0.75	τ=0.9	τ=0.95
Rolling window.							
Qcov3PRF1	4.5	6.8	18.6	25.4	38.0	37.5	18.8
Qcov3PRF2	8.5	14.1	35.6	43.3	52.1	42.5	17.4
Qcov3PRF3	11.4	20.9	43.4	54.9	56.2	44.9	25.6
Qcov3PRF4	16.7	27.2	48.6	58.3	60.5	46.6	18.7
Qcov3PRF5	22.7	36.7	52.9	60.2	63.5	46.7	17.8
PQR	13.3	14.2	18.2	18.8	39.2	32.6	20.0
PCQR1	8.2	8.3	13.6	18.1	15.3	16.2	15.3
PCQR2	11.5	10.8	13.6	19.8	16.2	23.8	23.6
PCQR3	10.9	10.6	13.5	19.2	15.1	23.9	25.5
PCQR4	11.8	12.1	18.5	18.6	12.5	26.6	22.0
PCQR5	12.0	11.7	19.4	19.5	14.4	23.8	26.2
PCQR-LASSO	16.4	11.7	10.0	14.1	41.7	25.3	0.3
Expanding window.							
Qcov3PRF1	6.9	3.3	11.1	25.8	45.7	39.8	30.8
Qcov3PRF2	8.7	13.1	27.0	48.7	61.2	57.9	34.7
Qcov3PRF3	11.9	16.4	36.5	61.2	66.7	57.3	27.2
Qcov3PRF4	16.5	22.1	46.1	63.6	69.6	57.1	39.0
Qcov3PRF5	18.7	29.7	51.0	64.7	70.4	58.0	34.0
PQR	13.3	14.2	18.2	18.8	39.2	32.6	20.0
PCQR1	5.2	3.0	3.8	9.5	11.4	13.0	21.2
PCQR2	3.4	2.7	3.8	11.0	15.0	15.6	23.8
PCQR3	4.5	2.1	5.7	12.9	15.7	14.8	25.2
PCQR4	6.3	5.7	9.7	11.0	13.4	13.3	21.2
PCQR5	5.8	5.7	9.8	9.2	14.8	16.7	12.0
PCQR-LASSO	17.6	23.3	25.0	30.5	25.9	6.6	-18.9

Note: The forecasts are evaluated starting in 2001. The size of the rolling window is 70. We take the first difference of the Global Land-Ocean Temperature index and the annual growth of the CO_2 emissions (difference in logs). In all the specifications we include the current value of the Global land-ocean temperature index as a predictor. Qcov3PRF# and PCQR# denote Qcov3PRF and PCQR implementation with # number of factors, respectively.