# Quantile-Covariance Three-Pass Regression Filter

Pedro Isaac Chavez-Lopez* Tae-Hwy Lee†

October 13, 2025

### Summary

We propose a factor model for quantile regression using *quantile-covariance*($qcov$), called the Quantile-Covariance Three-Pass Regression Filter (Qcov3PRF). This method estimates the supervised factors from a set of predictors to forecast the conditional quantile of a target. Our approach differs from the Partial Quantile Regression (PQR) as Qcov3PRF successfully allows the estimation of more than one relevant factor by virtue of using *qcov*. By estimating the true number of relevant factors, Qcov3PRF forecasts are consistent and asymptotically normal when both time and cross sectional dimensions become large. Simulations confirms these asymptotic results, showing Qcov3PRF exhibits good finite sample properties. Empirical applications to forecasting Growth-at-Risk highlights merits of Qcov3PRF over PQR. R codes to replicate the results are available.

***Keywords***: *Factor models, quantile-covariance, quantile regression, Qcov3PRF, PQR, Growth-at-Risk.*

***JEL Classification***: C13; C22; C53

---

*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 approach has usually been applied in a setup where in the first step a set of $N$ covariates is decomposed 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, in finite samples and under high levels of serial and cross sectional correlation in the idiosyncratic component of the predictors, 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, the Partial Least Squares (PLS, as a special case of 3PRF) of Wold, 1966, 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. 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, show that 3PRF is successful at extracting factors for predicting stock returns and economic activities using 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 conditional quantile prediction of the target some examples are the Principal Components Quantile Regression (PCQR, Ando and Tsay, 2011 and Giglio et al., 2016) and the Partial Quantile Regression (PQR, 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, as PQR only deals with one proxy, it only obtains one relevant factor for each quantile. This limitation happens as extending 3PRF to the PQR approach, in the step that concerns the estimation of relevant factors loadings, 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. Because of this issue, PQR runs instead a quantile regression of the target/proxy on each predictor but one can only include one proxy in that step by construction. Then, as a consequence, 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 these methods exploit the covariance between the target variable and the predictors when forecasting the conditional mean of a target variable. However, when the goal is to obtain the conditional $\tau$-quantile forecasts of the target variable, a linear relationship between the target variable and the predictors indexed by the $\tau$-quantile has to be considered. Therefore, in Qcov3PRF we instead incorporate the $\tau$-*quantile-covariance* (*qcov*) definition of Li et al., 2015. This measure is implied by the first order condition from a univariate linear quantile regression. In contrast to PQR, our method can successfully estimate more than one relevant factor. Our algorithm is easy to implement and only requires least square regressions to extract the relevant factors.

We are not the first to incorporate *quantile-covariance* measure 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 different definitions of quantile-covariance. 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 justify 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 only show the effectiveness of their method in the median. Neither of both methods provide asymptotic properties of the forecasts generated.

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 and provide a feasible estimate for their asymptotic variance. In particular, the consistency and asymptotic normality hold under the use of *automatic-proxies*, which are proxies generated by Qcov3PRF that are always available and constructed only with the target and the set of predictors. 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 an empirical application that involves the real growth vulnerability from a set of financial variables related to forecasting Growth-at-Risk (Adrian et al., 2019). Our results suggest that our method surpasses PCQR, PQR and other similar alternatives in in-sample prediction and forecasting accuracy. In addition, similar conclusions are found when comparing the performance using functionals constructed with the predicted conditional quantiles such as the target's expected shortfall.

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 an empirical application study, 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 $y_{t+1}$. The error components $\boldsymbol{\varepsilon}_t$, $u_{t+1}$ and $\boldsymbol{\omega}_t$ follow standard assumptions in the factor models literature, with $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\omega}_t$ allowing some degree of cross-sectional and serial correlation, and $u_{t+1}$ being uncorrelated over time.

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 $\boldsymbol{F}_t = (\boldsymbol{f}'_t, \boldsymbol{g}'_t)'$ with their corresponding loadings $\boldsymbol{\Phi} = (\boldsymbol{\Phi}_f, \boldsymbol{\Phi}_g)$. 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 $\boldsymbol{\Lambda} = [\boldsymbol{\Lambda}_f \quad \boldsymbol{0}]$. In 3PRF, relevant factors are recovered through Eq.(3) 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 proxies such that after applying 3PRF Eq.(1) can be written as $\boldsymbol{x}_t = \hat{\boldsymbol{\phi}}_0 + \hat{\boldsymbol{\Phi}}(\boldsymbol{z}_t)\hat{\boldsymbol{f}}_t + \hat{\boldsymbol{\varepsilon}}_t$ where the factor loadings $\hat{\boldsymbol{\Phi}}(\boldsymbol{z}_t) = \hat{\boldsymbol{\Phi}}_f$ are identified by Eq.(3) where $\hat{\boldsymbol{\Phi}}_f \hat{\boldsymbol{f}}_t \xrightarrow[N,T\to\infty]{p} \boldsymbol{\Phi}_f \boldsymbol{f}_t$ and $\hat{\boldsymbol{\varepsilon}}_t \xrightarrow[N,T\to\infty]{p} \boldsymbol{\Phi}_g \boldsymbol{g}_t + \boldsymbol{\varepsilon}_t$. The method 3PRF is presented in Algorithm 1.

Pass 1 and Pass 2, through time series and cross sectional linear regressions, respectively, determine the relevant factors with $\boldsymbol{z}_t$[1]. Once the relevant factors are estimated, in Pass 3 the conditional mean target's forecast is determined. 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

---

[1] The idiosyncratic component $\boldsymbol{\omega}_t$ in Eq.(3) is crucial for 3PRF in the sense that, without it, the proxies are perfect determinants of 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.

```
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 $y_{t+1}$.
```

and Pruitt, 2015 proposed an algorithm (Table 2 in Kelly and Pruitt, 2015) to determine automatic-proxies by only using the data for $\boldsymbol{x}_t$ and $y_{t+1}$.[2] The authors showed that these proxies are linearly independent and uncorrelated with the irrelevant factors.

## 2.2 Related methods that forecast the conditional quantile of the target

Giglio et al., 2016 extend 3PRF to predict conditional quantiles of the target variable developing method known as Partial Quantile Regression (PQR). However, PQR only extracts one relevant factor. Specifically, when 3PRF is compared to PQR, the latter method replaces the prediction stage (Pass 3 in 3PRF) with $\tau$-quantile regression. More importantly, PQR replaces Pass 1 in 3PRF by running a time series $\tau$-quantile regression of $y_{t+1}$ on $x_{it}$ and a constant, and keep the factor loadings estimates, say, $\hat{\phi}_{i\tau}$, for $i =$

---

[2] Then, after applying 3PRF Eq.(1) can be written as $\boldsymbol{x}_t = \hat{\boldsymbol{\phi}}_0 + \hat{\boldsymbol{\Phi}}(y_{t+1})\hat{\boldsymbol{f}}_t + \hat{\varepsilon}_t$, which highlights the role of the target when estimating relevant factors. 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.

$1, \ldots, N$ for the relevant factor estimation in the next Pass.[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. 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 $\tau$-quantile of $\boldsymbol{z}$.

Moreover, Giglio et al., 2016 present the Principal Components Quantile Regression (PCQR). This method uses *all* the factors estimated by PCA to predict the conditional quantile of a target, i.e., it simply substitutes the prediction stage of PCR with quantile regression. The authors show that both PQR (for the one relevant factor case) 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, PCQR shares the same limitation as PCR given that both methods use *all* factors. As a consequence, this method may deliver poor forecasts in the presence of high serial and cross sectional correlations in finite samples.

Another supervised method for quantile prediction involving factor models is the fast

---

[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.

PQR (fPQR) developed by Méndez-Civieta et al., 2022. In this model the authors follow a similar approach to 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 for the median case. Méndez-Civieta et al., 2022 do not provide asymptotic properties for fPQR. Refinements and extensions for PQR and fPQR are presented in Beyaztas et al., 2024 and Mutis et al., 2025 without showing asymptotic properties. The method fPQR is similar in nature to ours but Qcov3PRF is applied for *any* conditional quantile prediction of the target and we provide its asymptotic properties. Table 1 shows a comparison between our method and the factor models listed above. Additional methods that involve the use of *qcov* when forecasting the conditional variable of a target with a large a set of predictors include the variable screening approaches of Ma et al., 2017 and Chen and Lee, 2024.

Lastly, it is worth mentioning that the components estimated by Qcov3PRF are consistent to the corresponding relevant components for *any* quantile, this is $\hat{\boldsymbol{\Phi}}_{f,\tau}^{\prime Qcov3PRF} \hat{\boldsymbol{f}}_t^{Qcov3PRF} \xrightarrow[N,T\to\infty]{p} \boldsymbol{\Phi}_f \boldsymbol{f}_t$. Then, the components estimated via Qcov3PRF have the same asymptotic limit as the ones corresponding to the same factors $\boldsymbol{f}_t$ estimated by PCA ($\hat{\boldsymbol{f}}_t^{PCA}$). Therefore, the factors estimated by Qcov3PRF are different from the *quantile* factors estimated by Quantile Factor Model (QFM, Chen et al., 2021) or similar (Ando and Bai, 2020, Clark et al., 2024 and Korobilis and Schröder, 2025), where the corresponding estimated components are such that $\hat{\boldsymbol{\Phi}}_\tau^{\prime} \hat{\boldsymbol{F}}_{\tau,t} \xrightarrow[N,T\to\infty]{p} \boldsymbol{\Phi}_\tau \boldsymbol{F}_{\tau,t}$. These factor models take into account the quantile component of predictors $\boldsymbol{x}_t$, whereas in Qcov3PRF we only work on the quantile component in $y_{t+1}$ (which is the same quantile in $\boldsymbol{z}_t$).

# 3 Quantile-covariance Three Pass Regression Filter

In this Section we introduce our estimation method to forecast the $\tau$-conditional quantile of a target variable. We first present the $\tau$-quantile covariance used in our method, and

Table 1: Comparison between factor models that forecast conditional quantiles.

| | Qcov3PRF | PCQR | PQR | fPQR |
|---|---|---|---|---|
| Supervised for target prediction | yes | no | yes | yes |
| Can $K_f^*$ be larger than 1? | yes | yes | no | yes |
| Consistency is shown | yes | yes | yes | no |
| Asymptotic normality is shown | yes | yes | no | no |
| $\tau$ | any | any | any | 0.5 |

Note: $K_f^*$ denotes the number of relevant factors that predict target's $\tau$-conditional quantile.

then go over the general factor structure in our framework. We approach that factor structure by estimating target's relevant factors through proxies generated using the $\tau$-quantile covariance with information of the target and predictors.

## 3.1 Quantile-covariance (*qcov*)

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 $Q_y^\tau$ be the $\tau$th unconditional quantile of $y$ and $Q_{y|f}^\tau$ be the $\tau$th 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:

$$\begin{aligned} \text{qcov}_\tau(y, f) &= \text{cov}(I(y - Q_y^\tau > 0), f) \\ &= \mathbb{E}[\kappa_\tau(y - Q_y^\tau)(f - \mathbb{E}(f))], \end{aligned} \tag{4}$$

where $\kappa_\tau(v) = \tau - I(v < 0)$. It is worth to mention that *qcov* does not possess a symmetry property, i.e., $\text{qcov}_\tau(y, f) \neq \text{qcov}_\tau(f, y)$. However, $\text{cov}(I(y - Q_y^\tau > 0), f) = \text{cov}(f, I(y - Q_y^\tau > 0))$. 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 $j = 1, \ldots, K_f$

---

[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.

and look for the following minimizers:

$$(b_{0\tau}^{*}, b_{j\tau}^{*}) = \underset{b_{0\tau}, b_{j\tau}}{\arg\min}\, \mathbb{E}\left[\rho_\tau\left(y - b_{0\tau} - b_{j\tau} f_j\right)\right] \tag{5}$$

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 $\mathrm{qcov}_\tau(y, f_j) = \nu(b_{j\tau}^{*})$ where $\nu(\cdot)$ is a continuous and increasing function, and, more importantly, $\nu(b_{j\tau}^{*}) = 0$ if and only if $b_{j\tau}^{*} = 0$. In our framework we also need to invoke the result for the quantile-covariance between $y$ and $u^\tau$. We show in Lemma 1 in the online Appendix that $\mathrm{qcov}_\tau(y, u) = \nu_u(b_{j\tau}^{*})$ where $\nu_u(\cdot)$ is a continuous function and $\nu_u(b_{j\tau}^{*}) \neq 0$. The previous results only hold for the case of one regressor at a time. However, in order to recover more than one relevant factor, Qcov3PRF exploits the linear relationship in the $\tau$-quantile of the target between the target and the set of multiple predictors (not only one) in its first stage. We will see in the next section that, indeed, $\tau$-*qcov* is enough to generate automatic-proxies in Qcov3PRF that span the set of relevant factors and these are linearly independent, allowing a correct estimation of the conditional $\tau$-quantile forecasts.

## 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 factor model called Qcov3PRF. This forecasting method relies on the quantile-covariance *qcov* defined in the previous subsection. The crucial difference between Qcov3PRF and PQR is the implementation of *qcov* via multivariate least squares regressions associated with a specific $\tau$-quantile in order to estimate the factor loadings in a setup with automatic-proxies, compared to the univariate $\tau$-quantile regressions implemented in the first step of PQR. This change allows us to recover more than one relevant factor in Qcov3PRF, in contrast to PQR.

For more detail, we first establish the environment considered for this method. There is a target variable that we wish to forecast its conditional $\tau$-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, but now the difference is that the proxies are driven by the relevant factors that affect the target variable in the $\tau$th quantile *and* by the factors relevant factors to any other $\tilde{\tau}$th quantile. This is a consequence of the target's DGP (Eqs.(8) and (9)). 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** *(Data Generating Process)*

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

$$\begin{aligned} \boldsymbol{x}_t &= \boldsymbol{\phi}_0 + \boldsymbol{\Phi}\boldsymbol{F}_t + \boldsymbol{\varepsilon}_t, && (6)\\ \boldsymbol{z}_{t,\tau} &= \boldsymbol{\lambda}_{0,\tau} + \boldsymbol{\Psi}_t\boldsymbol{\Lambda}_\tau\boldsymbol{F}_t + \boldsymbol{\omega}_t, && (7)\\ y_{t+1} &= \beta_{0,\tau} + \boldsymbol{\beta}_\tau'\boldsymbol{F}_t + u_{t+1}^\tau, && (8)\\ u_{t+1}^\tau &= (\sigma_u + \boldsymbol{\gamma}'\boldsymbol{F}_t)\xi_{t+1}^\tau && (9) \end{aligned}$$

*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 $\tau$ is because the conditional quantile of $y_{t+1}$ is $Q_{y_{t+1}|\boldsymbol{F}_t}^\tau = \beta_{\tau,0} + \boldsymbol{\beta}_\tau'\boldsymbol{F}_t$ where $\mathbb{E}[\tau - I(u_{t+1}^\tau < 0)|y_t, \boldsymbol{F}_t, y_{t-1}, \boldsymbol{F}_{t-1}, \ldots] = 0$. Also,*

*$\boldsymbol{\beta}_\tau = (\boldsymbol{\beta}_{f,\tau}^{'}, \mathbf{0}')'$ with $\boldsymbol{\beta}_{f,\tau} \neq \mathbf{0}$ is a vector of dimension $K_f \times 1$, and $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_f^{'}, \mathbf{0}')'$ with $\boldsymbol{\gamma}_f$ is a vector of dimension $\tilde{K}_f \times 1$ and $\boldsymbol{\gamma}_f \neq \mathbf{0}$. It is also assumed that $\sigma_u + \boldsymbol{\gamma}'\boldsymbol{F}_t \geq 0$. Let $K_f^*$ equal to the cardinality of the union set of those factors corresponding to $\boldsymbol{\beta}_{f,\tau}$ and $\boldsymbol{\gamma}_f$. The term $\boldsymbol{z}_{t,\tau}$ is a $K_f^* \times 1$ vector of proxy data, where $\boldsymbol{F}_t = (\boldsymbol{f}_t', \boldsymbol{g}_t')'$, and $\boldsymbol{\Lambda}_\tau = [\boldsymbol{\Lambda}_{f,\tau} \quad \boldsymbol{\Lambda}_{g,\tau}]$ 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 $K_f^* + K_g = K$. The term $\boldsymbol{\Psi}_t$ is a $K_f^* \times K_f^*$ random matrix with elements $\boldsymbol{\psi}_{jkt}$, $j,k = 1,...,K_f^*$.*

2. *The vector of proxies $\boldsymbol{z}_{t,\tau}$ is only spanned by the set of relevant factors, i.e., $\boldsymbol{\Lambda}_\tau = [\boldsymbol{\Lambda}_{f,\tau} \quad \mathbf{0}]$, and $\boldsymbol{\Lambda}_{f,\tau}$ is nonsingular.*

3. *$\mathbb{E}(\boldsymbol{\psi}_{jkt}) \neq 0$, $\mathbb{E}(\boldsymbol{\psi}_{jkt}^2) < \infty$ and independent of $\boldsymbol{F}_t$.*

Eq.(6) assumes a standard factor structure for a large set of predictors (Bai and Ng, 2002), Eq.(8) 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 $\varepsilon_t$ and the existence of several irrelevant factors with high variance, the estimation can result in forecasts with poor performance.

Based on Assumption 1, 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 $K_f^* = K_f$ and Eq.(7) with $\boldsymbol{\Psi}_t$ equal to identity matrix gives a link to extract the relevant factors. However, for the conditional quantile case, the implementation needs to account additional $K_f^* - K_f$ factors which are relevant for some target's conditional $\tilde{\tau}$-quantile where $\tilde{\tau} \neq \tau$. The reason for this is that the automatic-proxies generated with the Qcov3PRF are spanned by $K_f^*$ factors. Additionally, in Eq.(7) the factors may be multiplied by a random variable with mean different from zero and independent of the factors. Compared to Eq.(3), Eq.(7) is a generalization with further relatively weak assumptions. This DGP is required as the automatic proxies obtained with Qcov3PRF follow this structure.

As we have noticed in the methods described in Section 2, a key point to successfully extend 3PRF to conditional $\tau$-quantile prediction relies on obtaining suitable proxies that are spanned only by relevant factors of the target's conditional $\tau$-quantile. In Qcov3PRF we generate those proxies with only data from $y_{t+1}$ and $\boldsymbol{x}_t$ by applying *qcov* relationship between the target and the predictors. This quantile covariance is used to allow a kind of "inverse" form for the quantile regression, such that Pass 1 in Algorithm 1 is implemented to estimate the relevant factor loadings indexed to the $\tau$-quantile.

Qcov3PRF is implemented through Algorithm 2. In this algorithm the variables $\hat{\boldsymbol{z}}_{\ell,\tau}$, $\ell = 1, ..., K_f^*$ are called automatic-proxies and act as the proxies $\boldsymbol{z}_\tau$ in Eq.(7). As an initial step, the unconditional quantile for the target is estimated and we set the first proxy as $I(\boldsymbol{y} - \hat{Q}_y^\tau > 0)$, which is the same transformation used in *qcov* (Eq.(4)). Then, if $K_f^* = 1$ there is only one relevant factor and Eq.(8) is a univariate model implying that $\hat{\boldsymbol{z}}_{1,\tau} = I(\boldsymbol{y} - \hat{Q}_y^\tau > 0)$ is only required for the estimation of the relevant factor loadings. Otherwise, if $K_f^* > 1$ we need to construct $K_f^*$ proxies. Following step 1(a) in Algorithm 2 we use the first two passes of 3PRF to obtain $\ell$ relevant factor estimates in the $\ell$th iteration and construct the corresponding $(\ell+1)$th proxy for $\ell = 2, ..., K_f^*$. We repeat step 1(a) until we obtain $K_f^*$ proxies. Once all the $K_f^*$ factors are estimated, we simply run a quantile regression of the target on these factors (and a constant) in Step 2. In that way we construct the forecasts one period ahead[5] for the observation $y_{T+2}$ with the updated set of predictors $\boldsymbol{x}_t$, $t = 1, \ldots, T+1$ to obtain $\hat{\boldsymbol{f}}_{T+1}^{(K_f^*)}$ and compute $\hat{y}_{T+2} = \hat{\beta}_{0,\tau} + \hat{\boldsymbol{\beta}}'_{f,\tau} \hat{\boldsymbol{f}}_{T+1}^{(K_f^*)}$.

Only for the case $K_f^* = 1$ we explicitly implement *qcov* in the estimation of the cor-

---

[5] 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 2** Quantile-Covariance Three Pass Regression Filter(Qcov3PRF)

```
Step 0: For a quantile value $\tau \in (0,1)$ initialize $\hat{\boldsymbol{z}}_{1,\tau} = I(\boldsymbol{y} - \hat{Q}_y^{\tau} > 0)$, where $\hat{Q}_y^{\tau}$ is the estimate after running a time series quantile regression of $\boldsymbol{y}$ on a constant, and apply the first two passes of 3PRF with $\{\boldsymbol{y}, \boldsymbol{X}, \hat{\boldsymbol{z}}_{1,\tau}\}$ to get $\hat{\boldsymbol{f}}_1^{(1)}$.
Step 1: If $K_f^* = 1$ then go to Step 2. Otherwise, for $\ell = 2, \dots, K_f^*$:
    (a) Let $\hat{\boldsymbol{z}}_{\ell,\tau} = I(\boldsymbol{y} - \tilde{\beta}_{0,\tau}^{(\ell-1)} - \tilde{\beta}_{1,\tau}^{(\ell-1)}\hat{\boldsymbol{f}}_1^{(\ell-1)} - \cdots - -\tilde{\beta}_{\ell-1,\tau}^{(\ell-1)}\hat{\boldsymbol{f}}_{\ell-1}^{(\ell-1)} > 0)$, where $(\tilde{\beta}_{0,\tau}^{(\ell-1)}, \tilde{\beta}_{1,\tau}^{(\ell-1)}, \dots, \tilde{\beta}_{\ell-1,\tau}^{(\ell-1)})$ are the linear estimates after running a time series quantile regression of $\boldsymbol{y}$ on a constant and $(\hat{\boldsymbol{f}}_1^{(\ell-1)}, \dots, \hat{\boldsymbol{f}}_{\ell-1}^{(\ell-1)})$, and apply the first two passes of 3PRF with $\{\boldsymbol{y}, \boldsymbol{X}, \hat{\boldsymbol{z}}_{1,\tau}, \dots, \hat{\boldsymbol{z}}_{\ell-1,\tau}, \hat{\boldsymbol{z}}_{\ell,\tau}\}$ to get $(\hat{\boldsymbol{f}}_1^{(\ell)}, \dots, \hat{\boldsymbol{f}}_\ell^{(\ell)})$.
Step 2: Run a time series $\tau$-quantile regression of $\boldsymbol{y}$ on $(\hat{\boldsymbol{f}}_1^{(K_f^*)}, \dots, \hat{\boldsymbol{f}}_{K_f^*}^{(K_f^*)})$.
```

responding factor loadings in the Pass 1 of 3PRF and then estimate the relevant factor $\hat{\boldsymbol{f}}_1^{(1)}$ in Pass 2 of 3PRF. For the case where $K_f^* > 1$, we are based on the transformation $I(\cdot > 0)$ to sequentially generate the subsequent automatic-proxies but now we substract the conditional $\tau$-quantile of $\boldsymbol{y}$ considering the factors $(\hat{\boldsymbol{f}}_1^{(\ell)}, \dots, \hat{\boldsymbol{f}}_\ell^{(K_f^*)})$ generated in the $\ell$th iteration such that the $\ell$th proxy in the $\ell$th iteration is equal to $I(\boldsymbol{y} - \tilde{\beta}_{1,\tau}^{(\ell-1)}\hat{\boldsymbol{f}}_1^{(\ell-1)} - \cdots - \tilde{\beta}_{\ell-1,\tau}^{(\ell-1)}\hat{\boldsymbol{f}}_{\ell-1}^{(\ell-1)} > 0)$. The resulting proxies are associated to the corresponding $\tau$-quantile through the estimated $\tilde{\beta}_{j,\tau}^{(\ell-1)}$ for $j = 2, ..., \ell - 1$ while we can still run OLS regressions. In particular, the first two passes of 3PRF can be implemented to estimate $(\hat{\boldsymbol{f}}_1^{(\ell)}, \dots, \hat{\boldsymbol{f}}_\ell^{(\ell)})$. This is not possible if we instead use quantile regressions, because regressing each predictor on the proxies, which is Pass 1 in 3PRF, would provide the predictor's conditional quantile estimate and not the target's conditional quantile. It can be shown that the $K_f^*$ automatic-proxies generated are linearly independent, spanned by the $K_f^*$ factors and uncorrelated with the irrelevant factors $\boldsymbol{g}$. This is shown in Theorem 3 in next section. The similarity of our algorithm to the automatic-proxy procedure implemented in 3PRF is very clear as in that method the proxies are equal to the subsequent residuals of regressing a time series OLS regression of $\boldsymbol{y}$ on $(\hat{\boldsymbol{f}}_1^{(\ell)}, \dots, \hat{\boldsymbol{f}}_\ell^{(\ell)})$ (see Table 2 in Kelly and Pruitt, 2015), while in our method the proxies are equal to the subsequent residuals of regressing a time series quantile regression of $\boldsymbol{y}$ on $\hat{\boldsymbol{f}}_1^{(\ell)}$ and additionally these residuals are transformed through the function $I(\cdot > 0)$.[6]

In general $K_f^*$ proxies have to be considered and not only $K_f$ proxies using Algorithm 2 as it can be shown that the first $K_f$ proxies are linearly independent, but not necessarily uncorrelated with the remaining factors $\boldsymbol{f}_j$, for $j = K_f + 1, \dots, K_f^*$. This is a consequence of the fact that $\boldsymbol{f}_j$ does not affect the conditional $\tau$-quantile of $\boldsymbol{y}$ but it can affect some other part of the conditional distribution of $\boldsymbol{y}$. For example, consider $y_{t+1} = 1 + f_{1,t} + (0.5 f_{1,t} + f_{2,t})\xi_{t+1}$, where $Q_\xi^\tau = 0$, and $g_{3,t}$ is some other factor independent of $y_{t+1}$. Then, $Q^\tau_{y_{t+1}|\boldsymbol{f}} = \beta_{0,\tau} + \beta_{1,\tau} f_{1,t} + \beta_{2,\tau} f_{2,t} + \beta_{3,\tau} g_{3,t}$ with $(\beta_{0,\tau}, \beta_{1,\tau}, \beta_{2,\tau}, \beta_{3,\tau}) = (1, 1, 0, 0)$. However, if we run, say, a least squares regression of the first automatic proxy generated $I(\boldsymbol{y} - \hat{Q}^\tau_{\boldsymbol{y}} > 0)$ on $\boldsymbol{f}_t = (f_{1,t}, f_{2,t}, g_{3,t})$ and a constant, we get the estimated coefficients $\hat{\boldsymbol{\alpha}}_\tau = (\hat{\alpha}_{0,\tau}, \hat{\alpha}_{1,\tau}, \hat{\alpha}_{2,\tau}, \hat{\alpha}_{3,\tau})$ where $\hat{\alpha}_{0,\tau} \neq 0$, $\hat{\alpha}_{1,\tau} \neq 0$, $\hat{\alpha}_{3,\tau} = 0$, but $\hat{\alpha}_{2,\tau} \neq 0$, statistically.

To see more clearly why our setup focuses on estimating $K_f^*$ relevant factors instead of $K_f$ factors we provide a sketch of the proof which shows that the resulting automatic-proxies in Algorithm 2 are only spawn by the $K_f^*$ relevant factors making use of quantile-covariance. The formal proof is in the online Appendix as part of the proof of Theorem 3. For ease of presentation let $K_f^* = 3$ and $y_{t+1} = \beta_{0,\tau} + \beta_{1,\tau} f_{1,t} + \beta_{2,\tau} f_{2,t} + u^\tau_{t+1}$, with $u^\tau_{t+1} = (\sigma_u + \gamma_2 f_{2,t} + \gamma_3 f_{3,t})\xi^\tau_{t+1}$. This prediction equation is *not* a univariate regression, however, after some rearrangements we obtain:

$$y_{t+1} = \beta_{0,\tau} + \tilde{\beta}_\tau \tilde{f}_t + u^\tau_{t+1},$$

where $\tilde{\beta}_\tau \neq 0$, $\tilde{f}_t = \frac{\beta_{1,\tau}}{\tilde{\beta}_\tau} f_{1,t} + \frac{\beta_{2,\tau}}{\tilde{\beta}_\tau} f_{2,t}$. Since $\mathrm{qcov}_\tau(y_{t+1}, \tilde{f}_t) = \nu(\tilde{\beta}_\tau) \neq 0$ by Lemma 1 in Li et al., 2015, and $\mathrm{qcov}_\tau(y_{t+1}, u^\tau_{t+1}) = \nu_u(\tilde{\beta}_\tau) \neq 0$ by Lemma 4 in the online Appendix we

---

[6] Similar to the fact that 3PRF is a generalization of PLS, in the online Appendix we explore how PLS for conditional quantile prediction and Qcov3PRF are related.

have:

$$\begin{aligned}\hat{z}_{1,t}^{\tau} &= I(y_{t+1} > \hat{Q}_y^{\tau}) = \alpha_{0,\tau} + \bar{\nu}(\tilde{\beta}_{\tau})\tilde{f}_t + \bar{\nu}_u(\tilde{\beta}_{\tau})u_{t+1}^{\tau} + \bar{u}_{t+1}, \\ &= \alpha_{0,\tau} + \bar{\beta}_{1,\tau}f_{1,t} + (\bar{\beta}_{2,\tau} + \bar{\gamma}_{2,\tau}\xi_{t+1}^{\tau})f_{2,t} + \bar{\gamma}_{3,\tau}\xi_{t+1}^{\tau}f_{3,t} + \bar{\omega}_{t+1}\end{aligned}$$

where $\bar{\beta}_{j,\tau} = \bar{\nu}(\tilde{\beta}_{\tau})\frac{\beta_{j,\tau}}{\tilde{\beta}_{\tau}}$ for $j = 1, 2$, $\bar{\gamma}_k = \bar{\nu}_u(\tilde{\beta}_{\tau})\gamma_k$ for $k = 2, 3$, and $\bar{u}_{t+1}$ is uncorrelated with $\tilde{f}_t$ and $u_{t+1}^{\tau}$. Hence, we have: $\text{cov}(\hat{z}_{1,t}^{\tau}, f_{1,t}|f_{2,t}, f_{3,t}) = \bar{\beta}_{1,\tau}$, $\text{cov}(\hat{z}_{1,t}^{\tau}, f_{2,t}|f_{1,t}, f_{3,t}) = \bar{\beta}_{2,\tau} + \bar{\gamma}_{2,\tau}\mathbb{E}(\xi_{t+1}^{\tau})$, $\text{cov}(\hat{z}_{1,t}^{\tau}, f_{3,t}|f_{1,t}, f_{2,t}) = \bar{\gamma}_{3,\tau}\mathbb{E}(\xi_{t+1}^{\tau})$ and $\hat{z}_{1,t}^{\tau}$ is uncorrelated with $f_{j,t}$, $j = K_f^* + 1, ..., K$. We can show that the proxies $\hat{z}_{2,t}^{\tau}$ and $\hat{z}_{3,t}^{\tau}$ are also only spanned by ($f_{1,t}$, $f_{2,t}$, and $f_{3,t}$) using a similar approach. Moreover, it is not difficult to show that these proxies are linearly independent.

In contrast to 3PRF, the 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 corresponding loadings by combining Pass 1 and Pass 2. These are given by:

$$\hat{\boldsymbol{\Phi}}_{\tau} = \left(\boldsymbol{W}_{zz}\right)^{-1}\boldsymbol{W}'_{xz}, \tag{10}$$

$$\hat{\boldsymbol{f}}' = \boldsymbol{W}_{zz}(\boldsymbol{W}'_{xz}\boldsymbol{J}_N\boldsymbol{W}_{xz})^{-1}\boldsymbol{W}'_{xz}\boldsymbol{J}_N\boldsymbol{X}', \tag{11}$$

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 and asymptotic normality of the relevant factors and their corresponding loadings (as it is shown in Lemmas 2 and 3 in the online Appendix).

**Remark 1** *So far, we have assumed that the number of relevant factors $K_f^*$ is known. In practice, an information criterion (IC) can be used to select the correct number of*

*Qcov3PRF factors. If the researcher counts with theoretically-motivated proxies, then the appropriate number of factors should be justified, in principle, by some economic/finance theory. Otherwise, the use of an IC is important when the researcher needs to produce automatic-proxies. For 3PRF, Kelly and Pruitt, 2015 apply a modified BIC with the degrees of freedom for PLS proposed by Krämer and Sugiyama, 2011. That approach does not extend directly to our method, so we consider Cross-Validation (CV) (Ahn and Bae, 2021). The procedure is as follows: we divide the entire available data (with T observations) into two parts, training and test data. The training data covers the first 70% of the sample, that is, the observations from* $t = 2$ *to* $t = T^* + 1$*, and the remaining observations are the test data, i.e., the observations from* $t = T^* + 2$ *to* $t = T + 1$*. Then, for a given time* $s \in [T^* + 2, T + 1]$*, we forecast* $y_s$ *using a given number of relevant factors and the latent factors estimated from the training data from* $t = 2$ *to* $t = s - 1$*. Let MCLE(*$K_f^*$*) be the Mean Check Loss Error of the forecasts for* $y_s$ *obtained using* $K_f^*$ *automatic-proxies. Then, the cross-validation estimate of the optimal number of Qcov3PRF factors, denoted by* $RCV$*, is the value of* $K_f^*$ *that minimizes MCLE(*$K_f^*$*).*

# 4 Asymptotic results

This section describes the main asymptotic properties of Qcov3PRF forecasts. The remaining required assumptions (Assumptions 2-5) are stated in the online Appendix and are based on those required in Kelly and Pruitt, 2015 and Giglio et al., 2016 following standard factor models and quantile regression. First, we obtain that as $N$ and $T$ become large, the conditional quantile forecasts converge to the infeasible best forecasts. This result is presented in Theorem 1, whose proof makes use of the convergence rates from Lemma 2 (see online Appendix), quasi-maximum likelihood results, the equivariance properties of the quantile regression, and the weighted (quantile) regression of Angrist et al., 2006.

**Theorem 1** *Let Assumptions 1-5 hold. Then, the forecast using Pass 1 and Pass 2 from Algorithm 1 and runnig a $\tau$-quantile regression of $y_{t+1}$ on $\hat{\boldsymbol{f}}_t$ satisfies*

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

*where $\hat{Q}^{\tau}_{y_{t+1}|\hat{\boldsymbol{f}}_t}$ is the conditional $\tau$-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 3 (see online Appendix).

**Theorem 2** *Let Assumptions 1-5 hold and $\sqrt{N}/T \to 0$. Then, the forecast using Pass 1 and Pass 2 from Algorithm 1 and runnig a $\tau$-quantile regression of $y_{t+1}$ on $\hat{\boldsymbol{f}}_t$ satisfies*

$$\hat{\boldsymbol{V}}_{\tau}^{-1/2}\left(\hat{\beta}_{0,\tau} + \hat{\boldsymbol{\beta}}'_{\tau}\hat{\boldsymbol{f}}_t - \beta_{0,\tau} - \boldsymbol{\beta}'_{\tau}\boldsymbol{F}_t\right) \xrightarrow[N,T\to\infty]{d} \mathcal{N}(0,1), \tag{12}$$

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

**Proof:** See Appendix. ☐

Theorem 2 guarantees that the estimates for the infeasible forecasts are asymptotically normal. Under the independence between $\boldsymbol{f}_t$ and $g_{j,t}$ for $j = K^*_f + 1, \ldots, K$ (see Assumption 5.3 in the online Appendix) an estimate for the asymptotic covariance of $\hat{\boldsymbol{f}}_t$ is given by:

$$\hat{\boldsymbol{\Sigma}}_{\hat{f}} = \hat{\boldsymbol{F}}_A\hat{\boldsymbol{F}}_C^{-1}\hat{\boldsymbol{F}}_D\hat{\boldsymbol{\Sigma}}_{\Phi\varepsilon}\hat{\boldsymbol{F}}'_D\hat{\boldsymbol{F}}_C^{'-1}\hat{\boldsymbol{F}}'_A, \tag{13}$$

where $\hat{\boldsymbol{F}}_A = T^{-1}\boldsymbol{Z}'_{\tau}\boldsymbol{J}_T\boldsymbol{Z}_{\tau}$, $\hat{\boldsymbol{F}}_C = N^{-1}T^{-2}\boldsymbol{Z}'_{\tau}\boldsymbol{J}_T\hat{\boldsymbol{f}}'\hat{\boldsymbol{\Phi}}_{\tau}\boldsymbol{J}_N\hat{\boldsymbol{\Phi}}'_{\tau}\hat{\boldsymbol{f}}\boldsymbol{J}_T\boldsymbol{Z}_{\tau}$, and $\hat{\boldsymbol{F}}_D = T^{-1}\boldsymbol{Z}'_{\tau}\boldsymbol{J}_T\hat{\boldsymbol{f}}'$. Also $\hat{\boldsymbol{\Sigma}}_{\Phi\varepsilon}$ can be calculated by Heteroskedasticity and Autocorrelation Consistent (HAC) covariance matrix estimators given the vector series $\hat{\boldsymbol{\Phi}}_i\hat{\boldsymbol{\varepsilon}}_i$, where $\hat{\boldsymbol{\varepsilon}}_i = \boldsymbol{X}_i - \hat{\boldsymbol{\Phi}}'_i\hat{\boldsymbol{f}}$.

For a feasible estimator of $\hat{\boldsymbol{\Sigma}}_{\hat{\beta}_{\tau}}$ we consider the more general estimator that allows misspecification:

$$\hat{\boldsymbol{\Sigma}}_{\hat{\beta}_\tau} = \left(\frac{1}{2T\delta}\sum_{t=1}^{T}\hat{\boldsymbol{f}}_t\hat{\boldsymbol{f}}_t^{'}\mathbf{1}(|\hat{u}^{\tau}_{t+1}|<\delta)\right)^{-1}\left(\frac{1}{\delta}\sum_{t=1}^{T}\hat{\boldsymbol{f}}_t\hat{\boldsymbol{f}}_t^{'}\hat{\kappa}^2_{t\tau}\right)\left(\frac{1}{2T\delta}\sum_{t=1}^{T}\hat{\boldsymbol{f}}_t\hat{\boldsymbol{f}}_t^{'}\mathbf{1}(|\hat{u}^{\tau}_{t+1}|<\delta)\right)^{-1}, \tag{14}$$

where $\hat{\kappa}_{t\tau} = \tau - \mathbf{1}(y_{t+1} < \hat{\beta}_{0,\tau} + \hat{\boldsymbol{\beta}}_\tau^{'}\hat{\boldsymbol{f}}_t)$, $\hat{u}^{\tau}_{t+1} = y_{t+1} - \hat{\beta}_{0,\tau} + \hat{\boldsymbol{\beta}}_\tau^{'}\hat{\boldsymbol{f}}_t$, and $\delta$ is a bandwidth hyperparameter whose optimal value can be determined by cross-validation.

Next, we state formally that Algorithm 2 provides automatic-proxies consistent for the conditional quantile forecasts. To show Theorem 3 we make use of Lemmas 4 and 5, which are stated in the online Appendix. Note that all of our simulations and the empirical application described in the next sections consider that the only available information is the set of predictors $\boldsymbol{x}_t$ and the target $y_{t+1}$, then we can only generate forecasts with automatic-proxies.

**Theorem 3** *Let Assumptions 1-5 hold with the exception of Assumptions 1.2,1.3, 3.4, 4.3, 4.4 and 4.5. Then, the $K_f^*$ automatic-proxies obtained in Algorithm 2 satisfy Assumptions 1.2, 1.3, 3.4, 4.3, 4.4 and 4.5 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. □

# 5 Monte Carlo simulations

In this section we present the forecasting power of Qcov3PRF and the asymptotic normality of the forecasts in finite samples through Monte Carlo simulations. Additional simulations to show the in-sample consistency of the predictions in finite samples are included in the online Appendix.

The general prediction equation of the DGP is as follows:

$$y_{t+1} = \beta_{1,\tau} f_{1,t} + \beta_{2,\tau} f_{2,t} + (\sigma_u + \gamma_1 f_{1,t} + \gamma_3 f_{3,t})\xi_{t+1}^{\tau}, \tag{15}$$

such that $\xi_{t+1}^{\tau} = \bar{\xi}_{t+1} - Q_{\bar{\xi}_{t+1}}^{\tau}$, with $\bar{\xi}_{t+1} \sim \mathcal{N}(0,1)$. If $\beta_{1,\tau} + \gamma_1 Q_{\bar{\xi}_{t+1}}^{\tilde{\tau}} \neq 0$, $\beta_{2,\tau} \neq 0$ and $\gamma_3 Q_{\bar{\xi}_{t+1}}^{\tilde{\tau}} \neq 0$ for some $\tilde{\tau} \neq \tau$ there are three relevant factors, i.e., $K_f^* = 3$, but for a given $\tau$ there are two relevant factors, i.e., $K_f = 2$. 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.

We work with Eq.(15) in the simulations presented below with different values on the parameters $\{\beta_{1,\tau}, \beta_{2,\tau}, \sigma_u, \gamma_1, \gamma_3\}$ and a different number of irrelevant factors, i.e., different values for $K_g$. The structure of the predictors is $\boldsymbol{x}_t = \boldsymbol{\Phi} \boldsymbol{F}_t + \boldsymbol{\varepsilon}_t$, with all factor loadings standard normally distributed and the noise term may contain some degree of serial or cross-sectional correlation. The relevant factors are distributed as follows, $f_{1,t} \sim \mathcal{U}[-1/5, 1.5]$, $f_{2,t} \sim \mathcal{N}(0,1)$, and $f_{3,t} \sim \mathcal{U}[-2,2]$. We focus on the quantile $\tau = 0.1$ as with other values of $\tau$ we obtain similar results and conclusions in the experiments studied in this section.

## 5.1 Asymptotic normality in finite samples

We evaluate the accuracy of finite sample approximations based on the asymptotic distribution we have derived in Theorem 2. We examine the distributions of the conditional quantile forecasts. For each Monte Carlo draw, we first compute the estimates $\hat{Q}_{y_{t+1}|\hat{f}_t}^{\tau}$. Then, we standardize each estimate based on Theorem 2 by subtracting the true conditional quantile of the target and dividing by the respective asymptotic standard error estimate. We consider the predictions obtained by Qcov3PRF which should follow a standard normal distribution for large $N$ and $T$ based on Theorem 3. We plot the distribution of standardized forecasts across simulations (histogram part) versus the standard normal pdf (solid line).

We work on Eq.(15) with two different groups of settings and present three panels for each illustrating asymptotic convergence towards a standard normal distribution as $N$ and $T$ increase. In particular, the first DGP(A) consists on two relevant factors and there are no irrelevant factors so $K = K_f^* = 2$, with $\beta_{1,\tau} = 2$, $\beta_{2,\tau} = 0$, $\gamma_1 = \gamma_2 = 0.5$, $\sigma_u = 2.0$, Also, the noise term $\varepsilon_{i,t}$ are i.i.d. standard normal. The second DGP(B) consists of three relevant factors and there are no irrelevant factors, i.e. $K = K_f^* = 3$ with the same parameters and distributions as in DGP(A) except $\beta_{2,\tau} = 2$. We simulate 5000 samples in each experiment. The figures show that the standard normal distribution successfully describes the finite sample behavior of these predictions, consistent with Theorem 3.

[Figure 1 about here.]

Regarding the first model (first row), in all cases we fail to reject the standard normal null hypothesis for standardized estimates. For the second model (second row), we reject the standard normal null hypothesis for the standardized predictions in all cases but when $N = 2500$ and $T = 5000$. There is a clear presence of significant bias for $N = 100, T = 200$ and $N = 500, T = 1000$, but this bias vanishes as $N$ and $T$ increase. Altogether, simulations provide evidence that the Qcov3PRF accurately estimates the infeasible best quantile forecasts, and that its theoretical asymptotic distributions successfully approximate the finite sample distributions for Qcov3PRF predictions.

## 5.2 Out of sample forecasting performance

For the forecasting accuracy we consider Eq.(15) and 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 $\rho_N$ and $\rho_T$ indicate high levels of cross sectional and serial correlation in the idiosyncratic component, respectively.

We report the out-of-sample $R_\tau^2$ (Koenker and Machado, 1999) evaluating forecasts one period ahead for the next 50 observations which is defined as:

$$R_\tau^2 = 1 - \frac{\sum_{t=T+1}^{T+50} \left(y_{t+1} - \hat{Q}^\tau_{y_{t+1|\hat{\boldsymbol{F}}_t}}\right)\left(\tau - I(y_{t+1} - \hat{Q}^\tau_{y_{t+1|\hat{\boldsymbol{F}}_t}} < 0)\right)}{\sum_{t=T+1}^{T+50} \left(y_{t+1} - \hat{Q}^\tau_{y_{t+1}}\right)\left(\tau - I(y_{t+1} - \hat{Q}^\tau_{y_{t+1}} < 0)\right)}, \tag{16}$$

where $\hat{Q}^\tau_{y_{t+1}}$ is an estimate for the $\tau$ 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.

Specifically, DGP(C) consists on Eq.(15) with $\beta_{1,\tau} = 2$, $\beta_{2,\tau} = 2$, $\gamma_1 = \gamma_2 = 0.5$, $\sigma_u = 2.0$. In DGP(C) we assume the predictors are determined by seven factors, three of which are relevant ($f_{1,t}$, $f_{2,t}$ and $f_{3,t}$) and the other four are irrelevant ($g_{1,t}$, $g_{2,t}$, $g_{3,t}$ and $g_{4,t}$). Then, $K_f^* = 3$ and $K_g = 4$. The irrelevant factors, $g_{1,t}$, $g_{2,t}$, $g_{3,t}$ and $g_{4,t}$ are independent distributed as $\mathcal{U}[-2.5, 2.5]$, $\mathcal{N}(0, 1.5)$, $\mathcal{N}(0, 1.75)$ and $\mathcal{N}(0, 2)$, respectively. We report the $R_\tau^2$ measuring the out-of-sample performance of the next 50 observations 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 seven PCA factors. We run 1000 simulations in each case when $(N, T) = \{(100, 100), (100, 200), (200, 100), (200, 200), (200, 500), (500, 200)\}$.

Table 2 shows that Qcov3PRF3 provides the highest predictive power over many cases. The advantage of using Qcov3PRF is more evident when $\rho_N$ and $\rho_T$ are high. The reason is that as serial and cross-sectional correlations increase, the idiosyncratic component is mistakenly seen as an additional factor(s), so considering only seven factors for PCQR is not enough to capture $f_{1,t}$, $f_{2,t}$, $f_{3,t}$ (and $g_{1,t}$, $g_{2,t}$, $g_{3,t}$, $g_{4,t}$). In contrast, Qcov3PRF performs well even when $\rho_N$ and $\rho_T$ are high as it focuses on only estimating the relevant factors.

We can see that Qcov3PRF1 and PQR report similar $R^2_\tau$ 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 is estimated. As expected, Qcov3PRF1 provides slightly lower performance than PQR as Step 1(a) in Algorithm 2 requires the dummy corresponding to the first automatic-proxy, whereas PQR does not consider an indicator function. In addition, as a result of the use of the indicator function, PCQR7 performs slightly better for the cases where $\rho_N$ and $\rho_T$ are low. All these differences between methods decrease as $N$ and $T$ augment.

In general, the out-of-sample $R^2_\tau$ obtained with Qcov3PRF3 is only marginally different compared to Qcov3PRF2. 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 variances of the relevant factors are different, the additional relevant factors provide only marginal increments in the forecasting power.

# 6 Application to forecasting monthly Growth-at-Risk using the NFCI components

In this section we show that Qcov3PRF outperforms competitive alternatives in taking into account the relationship between financial indicators (e.g., the variables that construct the National Finance Condition Index (NFCI, Federal Reserve Bank of Chicago))

Table 2: Out-of-sample $R^2_\tau(\%)$ for DGP(C) with $\tau = 0.10$ and $h = 1$.

| $\rho_T = \rho_N$ | Qcov3PRF3 | Qcov3PRF2 | Qcov3PRF1 | PQR | PCQR7 | Qcov3PRF3 | Qcov3PRF2 | Qcov3PRF1 | PQR | PCQR7 |
|---|---|---|---|---|---|---|---|---|---|---|
| | $N = 100$, $T = 100$ | | | | | $N = 200$, $T = 200$ | | | | |
| 0.0 | 48.0 | 44.4 | 33.2 | 36.1 | 48.8 | 50.3 | 48.4 | 38.4 | 40.7 | 51.0 |
| 0.2 | 46.4 | 43.2 | 33.2 | 35.2 | 47.2 | 49.4 | 47.7 | 38.6 | 40.3 | 50.6 |
| 0.4 | 44.4 | 40.8 | 31.7 | 33.8 | 32.7 | 48.7 | 46.2 | 37.3 | 39.2 | 47.0 |
| 0.6 | 39.0 | 34.4 | 27.4 | 28.6 | 16.1 | 44.8 | 41.6 | 32.8 | 35.0 | 19.4 |
| 0.8 | 30.2 | 24.5 | 19.5 | 19.4 | 11.6 | 33.7 | 27.3 | 20.9 | 21.7 | 15.3 |
| | $N = 200$, $T = 100$ | | | | | $N = 100$, $T = 200$ | | | | |
| 0.0 | 47.8 | 44.0 | 33.7 | 36.9 | 48.9 | 50.6 | 48.3 | 38.0 | 40.6 | 51.3 |
| 0.2 | 47.5 | 44.0 | 33.8 | 36.3 | 48.6 | 49.7 | 47.5 | 37.2 | 39.6 | 49.7 |
| 0.4 | 44.6 | 41.8 | 31.5 | 33.8 | 40.9 | 47.5 | 45.5 | 36.2 | 37.9 | 38.8 |
| 0.6 | 39.0 | 34.7 | 27.4 | 29.3 | 15.7 | 45.1 | 40.3 | 31.5 | 33.5 | 18.8 |
| 0.8 | 26.9 | 21.3 | 17.7 | 18.3 | 9.4 | 36.5 | 28.4 | 21.4 | 21.6 | 15.5 |
| | $N = 500$, $T = 200$ | | | | | $N = 200$, $T = 500$ | | | | |
| 0.0 | 51.4 | 49.8 | 39.5 | 42.3 | 51.7 | 52.8 | 51.6 | 44.6 | 46.0 | 53.1 |
| 0.2 | 50.5 | 48.4 | 39.5 | 41.6 | 51.0 | 53.2 | 51.9 | 44.5 | 45.9 | 53.5 |
| 0.4 | 49.6 | 47.5 | 38.5 | 40.2 | 49.8 | 52.3 | 50.5 | 43.2 | 44.6 | 51.3 |
| 0.6 | 46.3 | 43.2 | 35.0 | 36.5 | 20.3 | 51.0 | 48.0 | 39.6 | 41.1 | 21.5 |
| 0.8 | 32.5 | 27.0 | 21.6 | 21.8 | 14.5 | 42.9 | 34.9 | 24.8 | 25.6 | 18.4 |

Note: The median of $R^2_\tau$ out of 1000 simulations in each case is reported. $R^2_\tau$ evaluates the forecasts for the next 50 observations. $\rho_T$ represents the level of serial correlation, and $\rho_N$ is the level of cross-section correlation. The model consists on seven latent factors, three are relevant. Qcov3PRF# denotes Qcov3PRF implementation with # number of factors, similarly for PCQR#.

and the conditional distribution of the economic activity (e.g., Gross Domestic Product 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).

For the 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. In this application we study in-sample and out-of-sample prediction. The forecast horizons considered are three months and twelve months ahead. Additional results regarding forecasting prediction using rolling windows are reported in the online Appendix.

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 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 In-sample prediction of the conditional quantiles (GaR) and risk measures (Expected Shortfalls)

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 2 illustrates the in-sample prediction for various conditional quantiles of IP growth three and twelve months ahead. From the methods considered, Qcov3PRF (with the number of factors selected by cross-validation whose setup is described in Section 3.2) shows better performance as the region covered by 5% and

95% quantile predictions captures the downfall and then the immediate peak in the following months of the realized IP growth during the COVID period in both horizons, the other three methods perform very poorly in capturing the extreme values in the COVID period while all the methods are successful in capturing the 2008 financial crisis event.

Additionally, the estimates obtained after applying Qcov3PRF 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 quantile 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 PCQR-IC and PQR, we can also notice significant differences in variability between the 5% quantile and 95% quantile predictions, but lower compared to those from Qcov3PRF-CV. Table 3 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 2 about here.]

Next, we calculate the conditional expected shortfall, a popular risk measure, and the expected longrise as in Adrian et al., 2019. Both measures can be calculated with the conditional quantiles from the estimated conditional parametric distribution of the target determined by the two-step procedure proposed by Adrian et al., 2019. 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.

Table 3: Check Loss for IP growth in-sample prediction.

| Method | $\tau$=0.05 | $\tau$=0.25 | $\tau$=0.50 | $\tau$=0.75 | $\tau$=0.95 |
|---|---|---|---|---|---|
| $h = 3$ months | | | | | |
| QR with NFCI | 0.32 | 0.55 | 0.61 | 0.51 | 0.23 |
| Qcov3PRF1 | 0.29 | 0.49 | 0.56 | 0.44 | 0.14 |
| Qcov3PRF2 | 0.21 | 0.43 | 0.47 | 0.38 | 0.11 |
| Qcov3PRF3 | 0.17 | 0.40 | 0.45 | 0.36 | 0.09 |
| Qcov3PRF-CV | 0.16(4) | 0.40(3) | 0.45(3) | 0.36(3) | 0.14(1) |
| PQR | 0.29 | 0.50 | 0.55 | 0.43 | 0.14 |
| PCQR1 | 0.31 | 0.54 | 0.60 | 0.50 | 0.23 |
| PCQR2 | 0.30 | 0.53 | 0.60 | 0.48 | 0.22 |
| PCQR3 | 0.29 | 0.53 | 0.60 | 0.47 | 0.19 |
| PCQR-IC | 0.29(3) | 0.49(5) | 0.60(3) | 0.46(5) | 0.19(4) |
| $h = 12$ months | | | | | |
| QR with NFCI | 0.61 | 1.70 | 1.76 | 1.21 | 0.38 |
| Qcov3PRF1 | 0.52 | 1.17 | 1.35 | 0.85 | 0.27 |
| Qcov3PRF2 | 0.38 | 0.86 | 1.03 | 0.75 | 0.19 |
| Qcov3PRF3 | 0.33 | 0.75 | 0.95 | 0.63 | 0.18 |
| Qcov3PRF-CV | 0.33(3) | 0.75(3) | 0.85(4) | 0.63(3) | 0.18(3) |
| PQR | 0.50 | 1.20 | 1.30 | 0.80 | 0.27 |
| PCQR1 | 0.59 | 1.56 | 1.66 | 1.20 | 0.40 |
| PCQR2 | 0.46 | 1.23 | 1.33 | 0.94 | 0.34 |
| PCQR3 | 0.44 | 1.18 | 1.28 | 0.92 | 0.34 |
| PCQR-IC | 0.43(5) | 1.13(5) | 1.25(4) | 0.83(4) | 0.20(5) |

Note: IP growth is equal to the log difference of the current IP and IP $h$ months before. The sample period is from 2007M06 to 2023M12. Qcov3PRF# and PCQR# denote Qcov3PRF and PCQR implementation with # number of factors, respectively. Qcov3PRF-CV and PCQR-IC takes into account a maximum of 5 factors selecting the number of factors using cross-validation and Akaike information criteria, respectively. The number in parenthesis shows the estimated number of factors determined by these criteria. The predictors are standardized.

Specifically, this method considers the following optimization problem:

$$\min_{\boldsymbol{\theta}} \sum_{\tau} \left( \hat{Q}^{\tau}_{y_{t+h}|\boldsymbol{F}_t} - G^{-1}(\tau|\boldsymbol{F}_t; \boldsymbol{\theta}) \right)^2, \tag{17}$$

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

$$g(y; \mu, \sigma, \alpha, \nu) = \frac{2}{\sigma} t\left(\frac{y-\mu}{\sigma}; \nu\right) T\left(\alpha \frac{y-\mu}{\sigma} \sqrt{\frac{\nu+1}{\nu + \left(\frac{y-\mu}{\sigma}\right)^2}}; \nu+1\right), \tag{18}$$

where $t(\cdot)$ and $T(\cdot)$ denote the PDF and CDF of the Student $t$-distribution, respectively.

Then, the estimates $\hat{\boldsymbol{\theta}} = (\hat{\mu}_{t+h}, \hat{\sigma}_{t+h}, \hat{\alpha}_{t+h}, \hat{\nu}_{t+h})$, where $\alpha$ is the skewness parameter and $\nu$ is the degrees of freedom parameter determining the kurtosis, are obtained by solving the problem (17) with the distribution (18), where the subscripts in the estimated parameters indicate dependence on $\boldsymbol{F}_t$, and $\hat{\mu}_{t+h} \in \mathbb{R}$, $\hat{\sigma}_{t+h} \in \mathbb{R}^+$, $\hat{\alpha}_{t+h} \in \mathbb{R}$, and $\hat{\nu}_{t+h} \in \mathbb{Z}^+$. The predictions correspond to the quantiles $\tau \in \{0.05, 0.25, 0.75, 0.95\}$ to exactly identify $\hat{\boldsymbol{\theta}}$.

Once we have estimated a skew-$t$ conditional distribution for the IP growth for each period, we calculate the expected shortfall (ES) and expected longrise (EL). Specifically, for a chosen $\tau$ probability the ES and EL are defined as:

$$\begin{aligned}
\mathrm{ES}^{\tau}_{t+h}(\boldsymbol{F}_t) &= \mathbb{E}[y_{t+h} | y_{t+h} \leq Q^{\tau}_{y_{t+h}|\boldsymbol{F}_t})] = \frac{1}{\tau} \int_0^{\tau} G^{-1}(\alpha | \boldsymbol{F}_t; \boldsymbol{\theta}) d\alpha, \\
\mathrm{EL}^{1-\tau}_{t+h}(\boldsymbol{F}_t) &= \mathbb{E}[y_{t+h} | y_{t+h} \geq Q^{1-\tau}_{y_{t+h}|\boldsymbol{F}_t})] = \frac{1}{\tau} \int_{1-\tau}^{1} G^{-1}(\alpha | \boldsymbol{F}_t; \boldsymbol{\theta}) d\alpha.
\end{aligned}$$

We estimate the ES and the EL via the numerical integrals of the estimated skewed-$t$ distributions over $\alpha \in [0, \tau]$ and $\alpha \in [1 - \tau, 1]$, respectively. Figure 3 presents the in-sample IP growth ES at $\tau = 5\%$ level and the EL at $1 - \tau = 95\%$ level, estimated with QR 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:

$$L_{FZ0}(y_{t+h}, Q^{\tau}_{y_{t+h}|\boldsymbol{F}_t}, \mathrm{ES}^{\tau}_{t+h}; \tau) = -\frac{1}{\tau \mathrm{ES}^{\tau}_{t+h}} I(y_{t+h} \leq Q^{\tau}_{y_{t+h}|\boldsymbol{F}_t})(Q^{\tau}_{y_{t+h}|\boldsymbol{F}_t} - y_{t+h}) + \frac{Q^{\tau}_{y_{t+h}|\boldsymbol{F}_t}}{\mathrm{ES}^{\tau}_{t+h}} + \log(-\mathrm{ES}^{\tau}_{t+h}) - 1, \tag{19}$$

where $\mathrm{ES}^{\tau}_{t+h}$ is the $\tau$-conditional expected shortfall. Since this loss function is only valid for positive values of the expected shortfall, we rescale $y_{t+h}$, $Q^{\tau}_{y_{t+h}|\boldsymbol{F}_t}$ and $\mathrm{ES}^{\tau}_{t+h}$ by a constant.[7] Table 4 reports the in-sample mean of the loss given in Eq.(19) 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 $\mathrm{ES}_{\tau}$, and the relevant factors and the NFCI is higher in absolute terms compared to the correlation between $\mathrm{EL}_{\tau}$ and the indices/factors).

Table 4: In-sample average FZ0 loss for IP growth prediction.

| | QR with NFCI | Qcov3PRF3 | Qcov3PRF-CV | PCQR3 | PCQR-IC | PQR |
|---|---|---|---|---|---|---|
| | | | $h = 3$ months | | | |
| expected shortfall (5%) | 2.43 | 0.95 | 1.00 | 1.68 | 1.63 | 1.75 |
| expected longrise (95%) | 1.70 | 0.99 | 1.32 | 1.46 | 1.36 | 1.53 |
| | | | $h = 12$ months | | | |
| expected shortfall (5%) | 2.61 | 1.56 | 1.56 | 1.96 | 1.99 | 2.31 |
| expected longrise (95%) | 2.13 | 1.40 | 1.40 | 2.04 | 1.75 | 1.81 |

Note: For the expected longrise we consider the FZ0 loss function but evaluated with the negative quantile and expected longrise of the target. Qcov3PRF# and PCQR# denote Qcov3PRF and PCQR implementation with # number of factors, respectively. Qcov3PRF-CV and PCQR-IC takes into account a maximum of 5 factors selecting the number of factors using cross-validation and Akaike information criteria, respectively.

## 6.2 Out-of-sample prediction of the conditional quantiles (GaR) and risk measures (Expected Shortfalls)

Table 5 contains the out-of-sample $R^2_{\tau}$ with IP growth as the dependent variable for three months ahead and twelve months ahead, respectively. We present the results for Qcov3PRF and PCQR for up to three factors, Qcov3PRF with the number of factors determined by cross-validation for each forecast, PCQR with the number of factors determined by the Akaike information criteria (Bai and Ng, 2002) for each forecast and

---

[7] For the case of expected longrise we use Eq.(19) evaluated in $-y_{t+h}$, $-Q^{\tau}_{y_{t+h}|\boldsymbol{F}_t}$ and $-\mathrm{ES}^{\tau}_{t+h}$, i.e., we mirror the conditional distribution of $y_{t+h}$ such that the $1-\tau$ expected longrise of $y_{t+h}$ is the $\tau$ expected shortfall of $-y_{t+h}$.

PQR. We also report the $R^2_\tau$ from using QR when the predictors are the NFCI, the Risk, Credit, Leverage and Nonfinancial Leverage, respectively. We consider an expanding window scheme evaluating the out-of-sample samples 2011M08-2023M12 and 2015M10-2023M12. Quantiles in the tails are considered with $\tau \in \{0.05, 0.10, 0.90, 0.95\}$ as well as quantiles near the center of the distribution 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. The predetermined indexes perform poorly, with significant power concentrated in the left tail. Qcov3PRF exhibits a robust performance across different quantiles, forecast horizons, and estimation window schemes. In contrast, PCQR results in several negative $R^2_\tau$ in the center and right tail of the distribution. This comparison between Qcov3PRF and PCQR holds in general although we see a drop in performance in the right tail using cross-validation in Qcov3PRF compared to PCQR using the Akaike information criteria. This finding can be attributed to the focus of cross-validation in in-sample estimation rather than out-of-sample prediction. In general, the values obtained by using PQR are lower suggesting the existence of more than one relevant factor to forecast the target over different horizons and quantiles. All these findings also hold using a rolling window scheme as can be seen in an additional table shown in the online Appendix.

Lastly, Table 6 presents results for the out-of-sample mean of the loss given in Eq.(19) across alternatives. While the values presented indicate mixed results and none method gives the lowest prediction in all cases, Qcov3PRF3 and Qcov3PRF-CV result in more stable and relative low values in both expected shortfall and expected longrise, with different forecast horizons and different validating windows. For example, PCQR3 and PCQR-IC provide good prediction when $h = 12$, but show very poor performance for the expected longrise when $h = 3$.

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

| | $\tau$=0.05 | $\tau$=0.1 | $\tau$=0.25 | $\tau$=0.5 | $\tau$=0.75 | $\tau$=0.9 | $\tau$=0.95 |
|---|---|---|---|---|---|---|---|
| $h = 3$ months and evaluating over the OOS sample 2011M08-2023M12. | | | | | | | |
| 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 |
| Qcov3PRF-CV | 23.9 | 17.3 | 15.3 | 7.5 | 7.0 | 4.4 | 10.6 |
| 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-IC | 20.0 | 14.9 | 4.2 | -8.5 | -6.1 | 6.7 | 5.9 |
| $h = 3$ months and evaluating over the OOS sample 2015M10-2023M12. | | | | | | | |
| 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 |
| Qcov3PRF-CV | 10.4 | 3.4 | 6.2 | 4.7 | 7.1 | 1.9 | 8.5 |
| 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-IC | 8.8 | 2.3 | -3.4 | -5.6 | -3.0 | 7.9 | 9.6 |
| $h = 12$ months and evaluating over the OOS sample 2011M08-2023M12. | | | | | | | |
| 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 |
| Qcov3PRF-CV | 41.2 | 39.9 | 21.4 | 17.5 | 17.0 | 17.4 | 23.5 |
| 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-IC | 37.4 | 32.2 | 6.7 | 4.8 | 14.8 | 29.5 | 31.0 |
| $h = 12$ months and evaluating over the OOS sample 2015M10-2023M12. | | | | | | | |
| 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 |
| Qcov3PRF-CV | 20.8 | 16.8 | 8.0 | 12.1 | 10.5 | 4.6 | 13.7 |
| 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-IC | 21.8 | 9.8 | -4.3 | 4.9 | 13.7 | 25.4 | 29.1 |

Note: The out-of-sample $R^2_\tau$ is reported. It evaluates the performance using the samples 2011M08-2023M12 and 2015M08-2023M12. 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 $h$ months before. Qcov3PRF# and PCQR# denote Qcov3PRF and PCQR implementation with # number of factors, respectively. Qcov3PRF-CV and PCQR-IC takes into account a maximum of 5 factors selecting the number of factors using cross-validation and Akaike information criteria, respectively. The predictors are standardized.

Table 6: Out-of-sample average FZ0 loss for IP growth prediction.

| | QR with NFCI | Qcov3PRF3 | Qcov3PRF-CV | PCQR3 | PCQR-IC | PQR |
|---|---|---|---|---|---|---|
| $h = 3$ months and evaluating over the OOS sample 2011M08-2023M12. | | | | | | |
| expected shortfall (5%) | 2.71 | 2.48 | 2.53 | 2.44 | 2.43 | 2.53 |
| expected longrise (95%) | 2.36 | 2.31 | 2.29 | 2.94 | 4.10 | 2.23 |
| $h = 3$ months and evaluating over the OOS sample 2015M10-2023M12. | | | | | | |
| expected shortfall (5%) | 3.04 | 2.83 | 2.90 | 2.73 | 2.69 | 2.75 |
| expected longrise (95%) | 2.58 | 2.51 | 2.50 | 3.45 | 5.18 | 2.34 |
| $h = 12$ months and evaluating over the OOS sample 2011M08-2023M12. | | | | | | |
| expected shortfall (5%) | 3.00 | 2.64 | 2.60 | 2.68 | 2.68 | 2.73 |
| expected longrise (95%) | 2.56 | 2.49 | 2.49 | 2.49 | 2.45 | 2.73 |
| $h = 12$ months and evaluating over the OOS sample 2015M10-2023M12. | | | | | | |
| expected shortfall (5%) | 3.27 | 2.85 | 2.93 | 2.91 | 2.86 | 2.99 |
| expected longrise (95%) | 2.74 | 2.70 | 2.69 | 2.66 | 2.55 | 3.03 |

Note: For the expected longrise we consider the FZ0 loss function but evaluated with the negative quantile and expected longrise of the target. It evaluates the performance using the samples 2011M08-2023M12 and 2015M08-2023M12. Qcov3PRF# and PCQR# denote Qcov3PRF and PCQR implementation with # number of factors, respectively. Qcov3PRF-CV and PCQR-IC takes into account a maximum of 5 factors selecting the number of factors using cross-validation and Akaike information criteria, respectively.

# 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. 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. In particular, *qcov* permits us to run time series least squares regressions of each regressor on a set of proxies indexed to a specific quantile of the target that only depend on the target's relevant factors. These proxies are automatic in the sense that always can be constructed with only the target and predictors. Without *qcov*, running quantile regressions of each predictor on proxies is not appropriate because regressing each predictor on the proxies would give the predictor's conditional quantile and not the target's quantiles. Then, as a consequence of running a quantile regression of the target on each predictor, only one factor is recovered with PQR. Our approach differs from PQR as Qcov3PRF successfully allows the estimation of more than one relevant factor.

Qcov3PRF demonstrates strong forecasting performance, often superior to alternatives, across simulation specifications. Moreover, Qcov3PRF results in superior performance in an empirical application consistent with the the recent literature on GaR (Adrian et al., 2019 and Giglio et al., 2016) when exploring in-sample and out-of-sample evaluation in different forecast horizons using the target's predicted conditional quantiles and some risk tail measures constructed with these conditional quantiles. We also showed asymptotic properties of the resulting forecasts.

# 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., & Bai, J. (2020). Quantile co-movement in financial markets: A panel quantile model with unobserved heterogeneity. *Journal of the American Statistical Association*, *115*(529), 266–279.

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.

Beyaztas, U., Tez, M., & Lin Shang, H. (2024). Robust scalar-on-function partial quantile regression. *Journal of Applied Statistics*, *51*(7), 1359–1377.

Chen, H., & Lee, J. H. (2024). Predictive quantile regression with high-dimensional predictors: The variable screening approach. *arXiv preprint arXiv:2410.15097*.

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

Clark, T. E., Huber, F., Koop, G., Marcellino, M., & Pfarrhofer, M. (2024). Investigating growth-at-risk using a multicountry nonparametric quantile factor model. *Journal of Business & Economic Statistics*, *42*(4), 1302–1317.

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.

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.

Korobilis, D., & Schröder, M. (2025). Monitoring multi-country macroeconomic risk: A quantile factor-augmented vector autoregressive (qfavar) approach. *Journal of Econometrics*, *249*, 105730.

Krämer, N., & Sugiyama, M. (2011). The degrees of freedom of partial least squares regression. *Journal of the American Statistical Association*, *106*(494), 697–705.

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.

Mutis, M., Beyaztas, U., Karaman, F., & Lin Shang, H. (2025). On function-on-function linear quantile regression. *Journal of Applied Statistics*, *52*(4), 814–840.

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.

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.

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

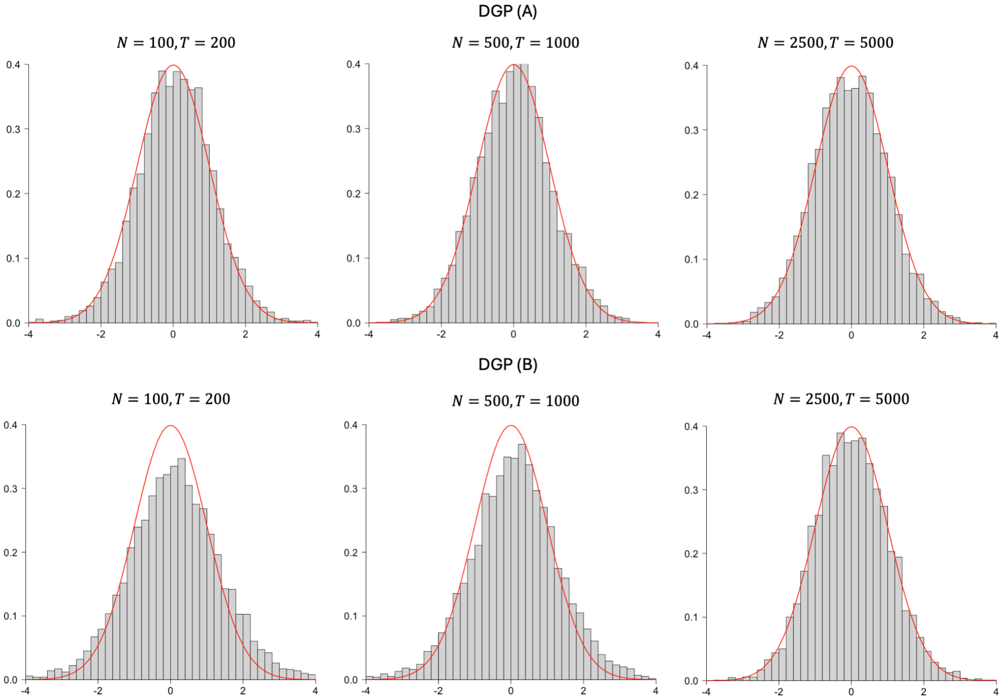
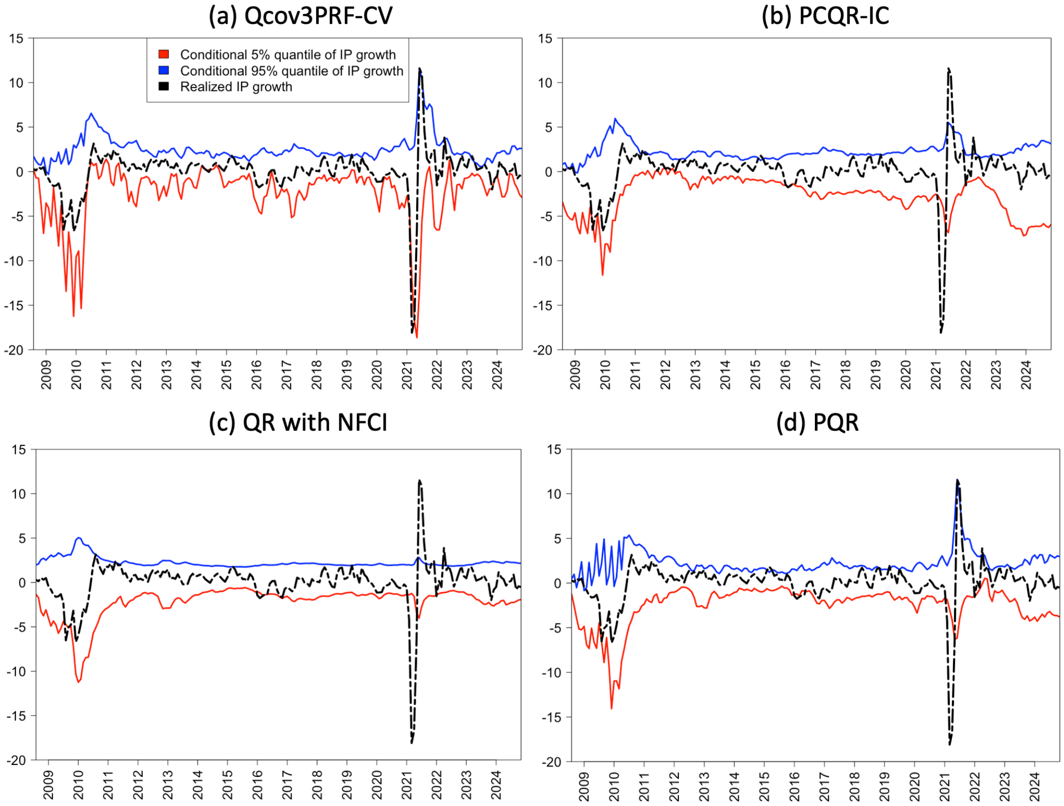
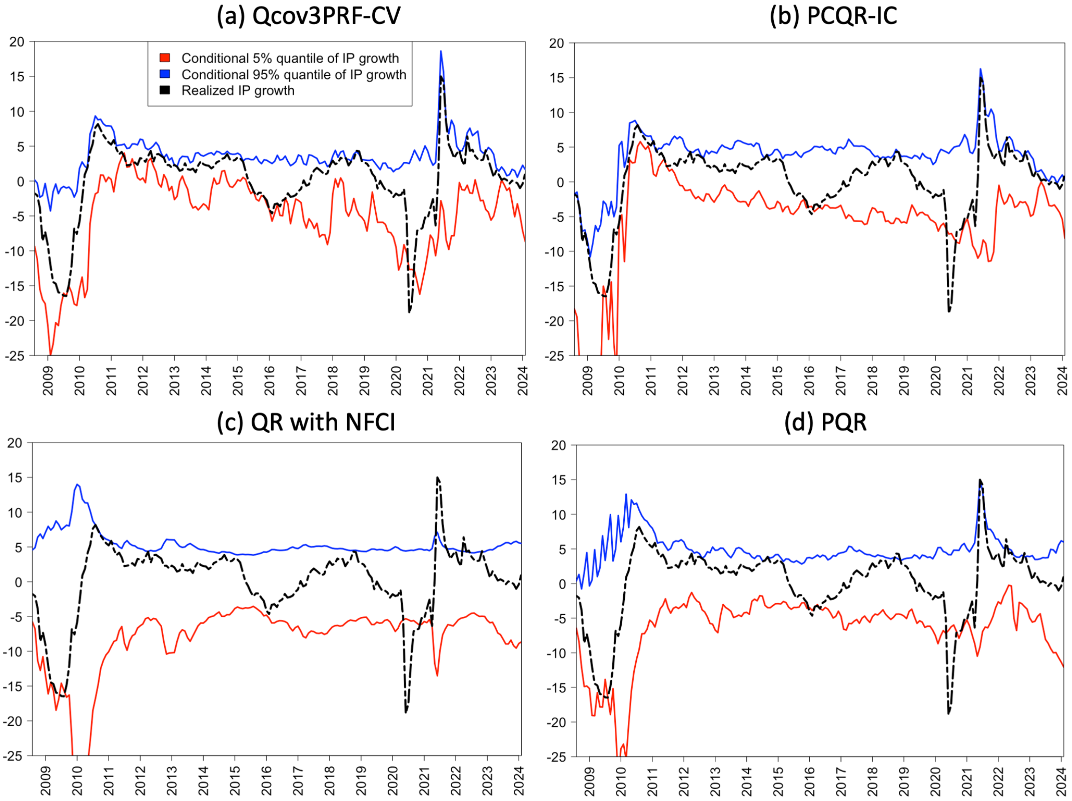


Figure 1: Simulated distribution of $\hat{Q}^{\tau}_{y_{t+h}|\hat{f}_t}$ with $h = 1$ and $\tau = 0.1$.

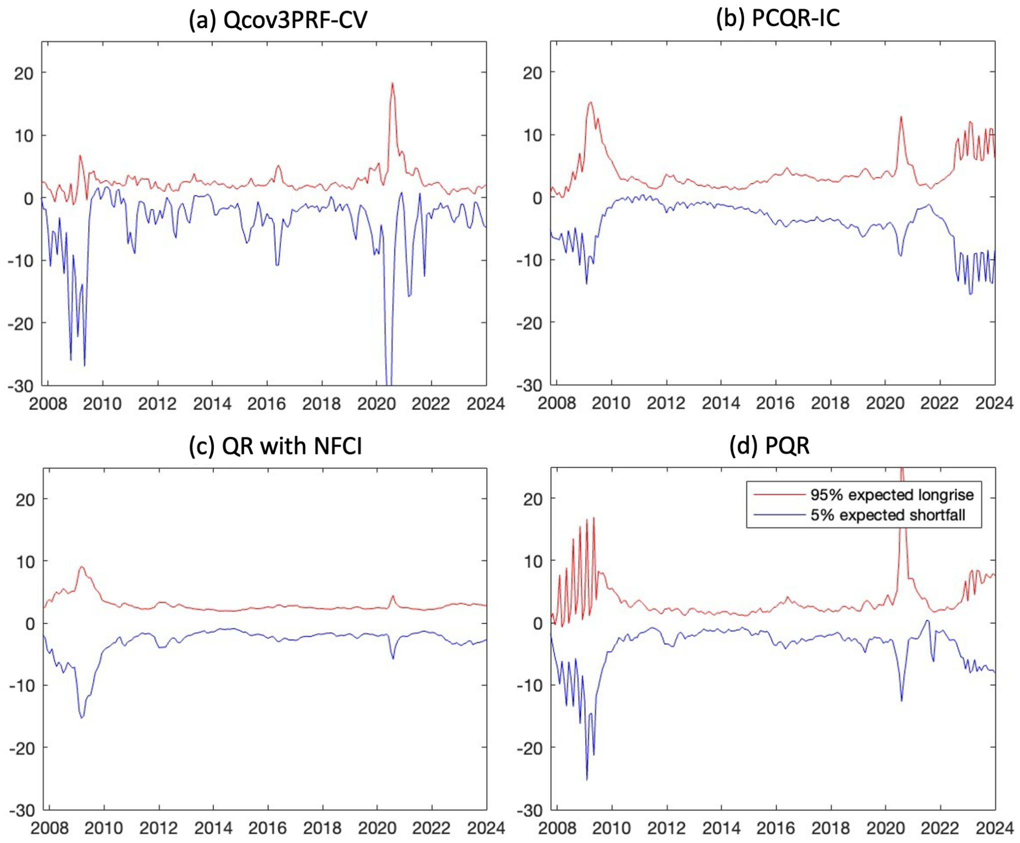


(i) $h$=3 months

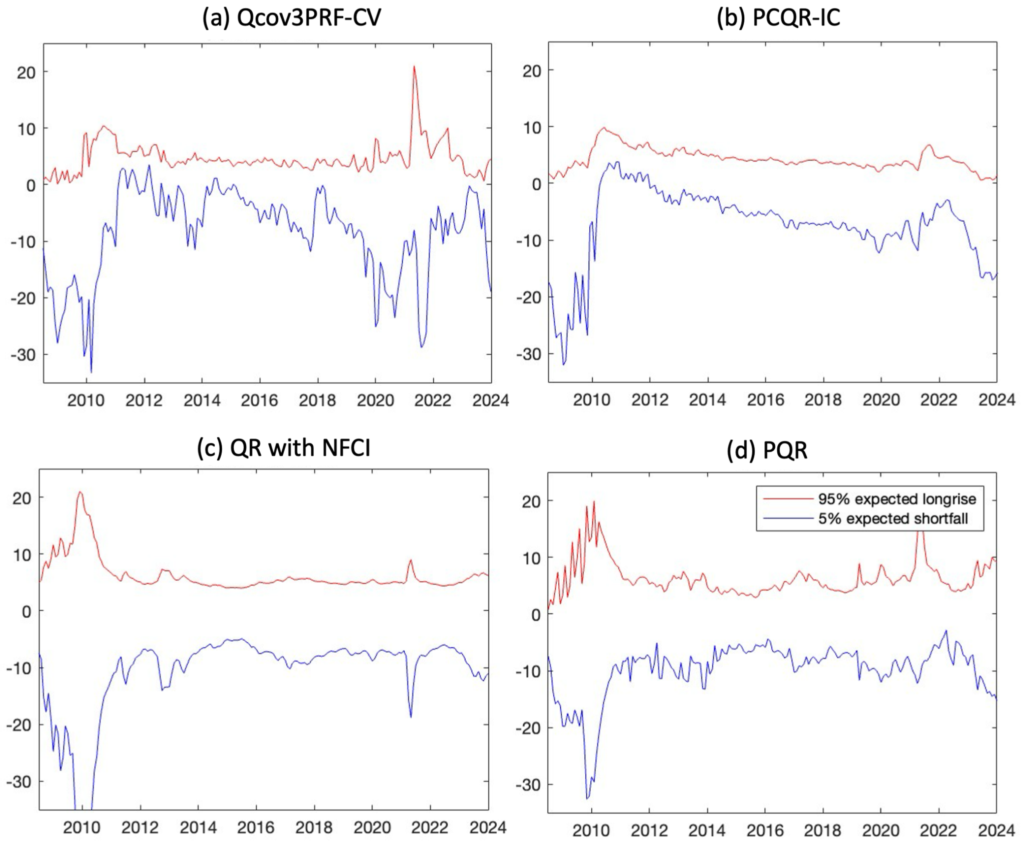


(ii) $h$=12 months

Figure 2: Conditional quantiles and realized IP growth, corresponding to the vertical left axis, estimated with (a) Qcov3PRF with the number of factors determined by cross-validation, (b) PCQR with the number of factors determined by AIC, (c) Quantile Regression with NFCI as a regressor, and (d) PQR.



(i) $h$=3 months



(ii) $h$=12 months

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

# Online Appendix of Quantile-Covariance Three-Pass Regression Filter

Pedro Isaac Chavez-Lopez* Tae-Hwy Lee†

October 13, 2025

In this appendix in Section A.1 we list the required Assumptions for Qcov3PRF not presented in the main text. Section A.2 contains the mathematical proofs of the main results and auxiliary lemmas. In Section A.3 we 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 A.4 shows additional simulations concerning the consistency of the forecasts using Qcov3PRF, and Section A.5 shows further results from the the empirical application presented in the main text.

## A.1 Assumptions

In addition to Assumption 1, in this section we state the following Assumptions needed for the estimation of Qcov3PRF. These Assumptions come from standard factor models estimation, the method 3PRF and quantile regression.

**Assumption 2.** *(Prediction Stage)*

1. *Let $0 < m < M < \infty$. The conditional density of $u_{t+1}$ given $\boldsymbol{F}_t$ denoted as $h_\tau(u_{t+1}|\boldsymbol{F}_t)$:*

---

*Economics Research Division, Bank of Mexico, pedro.chavez@banxico.org.mx

†Department of Economics, University of California-Riverside, taelee@ucr.edu

*(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.,*

$$|h_\tau(u_{t+1}|\boldsymbol{F}_t) - h_\tau(u^{'}_{t+1}|\boldsymbol{F}_t)| \leq M|u_{t+1} - u^{'}_{t+1}| \text{ for all } t.$$

2. $\mathbb{E}(|u_{t+1}|) \leq M$ *and* $T^{-1}\boldsymbol{F'}\boldsymbol{J}_T\boldsymbol{F}h_\tau(0|\boldsymbol{F}) \xrightarrow[T\to\infty]{p} \boldsymbol{\Delta}_{f,h} > 0.$

Assumption 2.1 allows for a bias representation of the quantile regression in the prediction stage following Angrist et al., 2006, along with Assumption 2.2, we can show the consistency of the conditional quantile forecasts if $\boldsymbol{F}_t$ were known.

**Assumption 3.** *(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. $\mathbb{E}\|\boldsymbol{\phi}_i\|^4 \leq M, \quad N^{-1}\sum_{j=1}^{N}\boldsymbol{\phi}_j \xrightarrow[N\to\infty]{p} \boldsymbol{\mu}_\Phi, \quad N^{-1}\boldsymbol{\Phi'}\boldsymbol{J}_N\boldsymbol{\Phi} \xrightarrow[N\to\infty]{p} \boldsymbol{\Delta}_\Phi,$
   $N^{-1}\boldsymbol{\Phi'}\boldsymbol{J}_N\boldsymbol{\phi}_0 \xrightarrow[N\to\infty]{p} \boldsymbol{\Delta}_{1,\Phi},$
3. $\mathbb{E}(\varepsilon_{it}) = 0, \qquad \mathbb{E}|\varepsilon_{it}|^8 \leq M$ *, and* $\varepsilon_{it}$ *is independent of* $F_t(m)$.
4. $\mathbb{E}||\boldsymbol{\omega}_t||^4 \leq M, \quad T^{-1/2}\sum_{s=1}^{T}\boldsymbol{\omega}_s = \boldsymbol{O}_p(1), \quad T^{-1}\boldsymbol{\omega'}\boldsymbol{J}_T\boldsymbol{\omega} \xrightarrow[T\to\infty]{p} \boldsymbol{\Delta}_\omega,$
   $T^{-1}\sum_{s=1}^{T}\boldsymbol{\Psi}_s \xrightarrow[T\to\infty]{p} \quad \boldsymbol{\mu}_\psi \neq \boldsymbol{0}, \quad T^{-1}\boldsymbol{\Psi'}\boldsymbol{\Psi} \xrightarrow[T\to\infty]{p} \boldsymbol{\Delta}_\psi,$
5. $\xi_{t+1}$ *is independent of* $\phi_i(m)$, $F_t(m)$ *and* $\varepsilon_{it}$.
6. $(y_{t+1}, \boldsymbol{\varepsilon}_t, \boldsymbol{F}_t)$ *are strictly stationary, ergodic, for some* $r > 4$, $\mathbb{E}|y_{t+1}|^r < \infty$, $\mathbb{E}\|\boldsymbol{F}_t\|^r < \infty$, *and* $\mathbb{E}\|\boldsymbol{\varepsilon}_t\|^r < \infty$, *and the mixing coefficients for* $(y_{t+1}, \boldsymbol{\varepsilon}_t, \boldsymbol{F}_t)$ *satisfy* $\sum_{\ell=1}^{\infty}\alpha(\ell)^{1-4/r} < \infty$.

Assumptions 3.1-3.3 are the same considered by Bai and Ng, 2002, Bai, 2003 and Stock and Watson, 2002. Assumption 3.4 is required in 3PRF for conditional mean forecasts, and it is also required in Qcov3PRF because this method uses proxies to extract the factors following a similar equation but now assuming additionally finite first and second moments for the random matrix $\boldsymbol{\Psi}_t$. The moments of the proxies noise $\boldsymbol{\omega}_t$ are bounded in the same manner as the bounds in factor moments. Assumption 3.5 is required to show consistency of estimators of factors and loadings, and asymptotic normality of the forecasts. Assumption 3.6 is required for the central limit theorems of the vector series $\boldsymbol{F}_t \varepsilon_{it}$ and $\boldsymbol{F}_t \psi_\tau(u_{t+1})$.

**Assumption 4.** *(Dependence) There exists a constant* $M < \infty$ *and for any* $i, j, t, s, m_1, m_2, m_3$

1. $\mathbb{E}(\varepsilon_{it}\varepsilon_{js}) = \sigma_{ij,ts}, \quad |\sigma_{ij,ts}| \leq \bar{\sigma}_{ij}, \quad |\sigma_{ij,ts}| \leq \bar{\sigma}_{ts}, \quad N^{-1}\sum_{i,j=1}^{N} \bar{\sigma}_{ij} \leq M,$
   $T^{-1}\sum_{t,s=1}^{T} \bar{\sigma}_{ts} \leq M$, *and* $\quad N^{-1}\sum_{i,s} |\sigma_{ij,ts}| \leq M$,
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. $\boldsymbol{\omega}_t$ *is independent*[1] *of* $F_t(m_1)$,
4. $\mathbb{E}|T^{-1/2}\sum_{t=1}^{T} \omega_t(m_1)\varepsilon_{it}|^2 \leq M$,
5. $\mathbb{E}|T^{-1}\sum_{t=1}^{T} \psi_t^2(m_1, m_2)\omega_t^2(m_3)| \leq M$ *and* $\mathbb{E}|T^{-1}\sum_{t=1}^{T} \psi_t^2(m_1, m_2)\xi_{t+1}^2| \leq M$.

As in 3PRF, we allow some degree of cross sectional correlation and serial dependence in the factor structure through Assumptions 4.1-4.2. This follows the work of Chamberlain and Rothschild, 1982 and Stock and Watson, 2002. Assumption 4.3 is stronger than what is required in Kelly and Pruitt, 2015, we can relax it with the assupmtion $\mathbb{E}|F_t^2(m_1)\omega_t^2(m_2)| = \mathbb{E}|F_t^2(m_1)|\mathbb{E}|\omega_t^2(m_2)|$. In Assumptions 4.4-4.5 some proxy noise dependence with idiosyncratic errors and the error term is allowed.

**Assumption 5.** *(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* $j = K_f^* + 1, \ldots, K$.

---

[1] It can be shown that $\mathbb{E}|T^{-1/2}\sum_{t=1}^{T} F_t(m_1)\omega_t(m_2)|^2 \leq M$ for all $m_1$ and $m_2$, by this independence assumption and by Assumptions 3.1 and 3.4.

Assumptions 5.1 and 5.2 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. In addition, we require independence between relevant and irrelevant factors as it allows the estimation of feasible covariance matrix estimate regarding asymptotic normality of the forecasts.

It is worth to mention that Assumption 4 in Kelly and Pruitt, 2015, which contains the central limit theorems, is not explicitly assumed in this article although it is required in almost all proofs for the main results in 3PRF. The reason is that that Assumptions 2, 3.1 and 3.6 imply Assumptions 4.2 and 4.3 in Kelly and Pruitt, 2015 under prediction equation (8). Assumption 4.1 in Kelly and Pruitt, 2015 is implied by Assumptions 3.2, 3.3 and 5.1.

## A.2 Proofs and additional lemmas

In this section we provide the proofs of the main results of the paper and auxiliary lemmas. Lemma 1 provides additional results to Lemma 2 in Kelly and Pruitt, 2015 required to show consistency and asymptotic normality of the estimates by considering the proxy equation (7). Lemma 2 and Lemma 3 are similar to Lemma 3 and Theorem 6 (with the corresponding normalization Assumption) in Kelly and Pruitt, 2015. The proof of Theorem 1 is very similar to Theorem 1 in Giglio et al., 2016. Lemma 4, Lemma 5, Theorem 2 and Theorem 3 are new.

The following lemma gives asymptotic orders, additional to Lemmas 1 and 2 in Kelly and Pruitt, 2015, used to show the asymptotic limits of the factors and factor loadings. They are a generalization of some results in Lemma 2 of Kelly and Pruitt, 2015 using $\boldsymbol{F}_{\psi^*}$ instead of $\boldsymbol{F}$ under Eq.(7).

**Lemma 1.** *Let Assumptions 1, 3 and 4 hold. Then, for any $\bar{m}$, $\bar{m}_1$, $\bar{m}_2$ as $N, T \to \infty$*

1. $T^{-1/2}\boldsymbol{F}'_{\psi^*}(\bar{m})\boldsymbol{J}_T\boldsymbol{\omega} = \boldsymbol{O}_p(1)$

2. $T^{-1/2}\boldsymbol{F}'_{\psi^*}(\bar{m})\boldsymbol{J}_T\boldsymbol{\xi} = \boldsymbol{O}_p(1)$

3. $N^{-1}T^{-1}\boldsymbol{\Phi}'\boldsymbol{J}_N\boldsymbol{\varepsilon}'\boldsymbol{J}_T\boldsymbol{F}_{\psi^*}(\bar{m}) = \boldsymbol{O}_p(\delta_{NT}^{-1})$

4. $N^{-1}T^{-3/2}\boldsymbol{F}'_{\psi^*}(\bar{m}_1)\boldsymbol{J}_T\boldsymbol{\varepsilon}\boldsymbol{J}_N\boldsymbol{\varepsilon}'\boldsymbol{J}_T\boldsymbol{F}_{\psi^*}(\bar{m}_2) = \boldsymbol{O}_p(\delta_{NT}^{-1})$

5. $N^{-1}T^{-3/2}\boldsymbol{\omega}'\boldsymbol{J}_T\boldsymbol{\varepsilon}\boldsymbol{J}_N\boldsymbol{\varepsilon}'\boldsymbol{J}_T\boldsymbol{F}_{\psi^*}(\bar{m}) = \boldsymbol{O}_p(\delta_{NT}^{-1})$

6. $N^{-1}T^{-1/2}\boldsymbol{F}'_{\psi^*}(\bar{m})\boldsymbol{J}_T\boldsymbol{\varepsilon}\boldsymbol{J}_N\boldsymbol{\varepsilon}_t = \boldsymbol{O}_p(\delta_{NT}^{-1})$

7. $N^{-1}T^{-3/2}\boldsymbol{\xi}'\boldsymbol{J}_T\boldsymbol{\varepsilon}\boldsymbol{J}_N\boldsymbol{\varepsilon}'\boldsymbol{J}_T\boldsymbol{F}_{\psi^*}(\bar{m}) = \boldsymbol{O}_p(\delta_{NT}^{-1})$

*where* $\delta_{NT} = \min(\sqrt{N}, \sqrt{T})$ *and* $\boldsymbol{F}_{\psi^*}(\bar{m}) = [\boldsymbol{F}'_1\boldsymbol{\psi}^*_1(\bar{m}), \ldots, \boldsymbol{F}'_T\boldsymbol{\psi}^*_T(\bar{m})]'$, $\boldsymbol{\psi}^*_t(\bar{m})$ *denotes the* $\bar{m}$*th row of matrix* $\boldsymbol{\psi}^*_t$,

$\boldsymbol{\psi}^*_t = \begin{bmatrix}\boldsymbol{\psi}_t & \mathbf{0}_{K^*_f\times(K-K^*_f)}; \mathbf{0}_{(K-K^*_f)\times K^*_f} & \mathbf{0}_{(K-K^*_f)\times(K-K^*_f)}\end{bmatrix}$, $t = 1, \ldots, T$.

*Proof.* <u>Item 1</u>: The object of interest is a $K \times K^*_f$ matrix, for each entry $(m_1, m_2)$ we have:

$$\begin{aligned}
\mathbb{E}\left[|T^{-1/2}\boldsymbol{F}'_{\psi^*}(\bar{m}, m_1)\boldsymbol{J}_T\boldsymbol{\omega}(m_2)|^2\right] &= T^{-1}\sum_t \mathbb{E}\left[\psi^{*2}_t(\bar{m}, m_1)F^2_t(m_1)\omega^2_t(m_2)\right] \\
&- \left[T^{-1}\sum_t \mathbb{E}[F_t(m_1)^2\psi^{*2}_t(\bar{m}, m_1)]\right]\left[T^{-1}\sum_t \mathbb{E}[\omega^2_t(m_2)]\right] \\
&= T^{-1}\sum_t \mathbb{E}[F^2_t(m_1)]\mathbb{E}[\psi^{*2}_t(\bar{m}, m_1)\omega^2_t(m_2)] \\
&- \left[T^{-1}\sum_t \mathbb{E}[F^2_t(m_1)]\mathbb{E}[\psi^{*2}_t(\bar{m}, m_1)]\right]\left[T^{-1}\sum_t \mathbb{E}[\omega^2_t(m_2)]\right] \\
&= O_p(1),
\end{aligned}$$

by Assumptions 3.1, 3.4, 4.3, 4.5 and since the second moment of $\boldsymbol{\psi}_t$ is bounded from above.

Item 2: The object of interest is a vector of size $K$, for each element $m_1$ we have:

$$
\begin{aligned}
\mathbb{E}\left[|T^{-1/2}\boldsymbol{F}'_{\psi^*}(\bar{m},m_1)\boldsymbol{J}_T\boldsymbol{\xi}|^2\right] &= T^{-1}\sum_t \mathbb{E}\left[\psi_t^{*2}(\bar{m},m_1)F_t^2(m_1)\xi_{t+1}^2\right] \\
&- \left[T^{-1}\sum_t \mathbb{E}[F_t(m_1)^2\psi_t^{*2}(\bar{m},m_1)]\right]\left[T^{-1}\sum_t \mathbb{E}[\xi_{t+1}^2]\right] \\
&= T^{-1}\sum_t \mathbb{E}[F_t^2(m_1)]\mathbb{E}[\psi_t^{*2}(\bar{m},m_1)\xi_{t+1}^2] \\
&- \left[T^{-1}\sum_t \mathbb{E}[F_t^2(m_1)]\mathbb{E}[\psi_t^{*2}(\bar{m},m_1)]\right]\left[T^{-1}\sum_t \mathbb{E}[\xi_{t+1}^2]\right] \\
&= O_p(1),
\end{aligned}
$$

by Assumptions 3.1, 3.5, 3.6, 4.5 and since the second moment of $\boldsymbol{\psi}_t$ is bounded from above.

Item 3: The object of interest is a $K \times K_f^*$ matrix, for each entry $(m_1, m_2)$ we have:

$$
\begin{aligned}
&N^{-1}T^{-1}\sum_{i,t}\phi_i(m_1)F_{\psi^*,t}(\bar{m},m_2)\varepsilon_{it} - N^{-2}T^{-1}\sum_{i,j,t}\phi_i(m_1)F_{\psi^*,t}(\bar{m},m_2)\varepsilon_{jt} \\
&-N^{-1}T^{-2}\sum_{j,s,t}\phi_j(m_1)F_{\psi^*,s}(\bar{m},m_2)\varepsilon_{jt} + N^{-2}T^{-2}\sum_{i,j,s,t}\phi_i(m_1)F_{\psi^*,s}(\bar{m},m_2)\varepsilon_{jt} \\
&\qquad = 3.I + 3.II + 3.III + 3.IV,
\end{aligned}
$$

$$
\begin{aligned}
3.I &= N^{-1}T^{-1}\sum_{i,t}\phi_i(m_1)F_t(m_2)\psi_t^*(\bar{m},m_2)\varepsilon_{it} \\
&\leq T^{-1/2}\left(N^{-1}\sum_i \phi_i(m_1)^2\right)^{1/2}\left(N^{-1}\sum_i\left[T^{-1/2}\sum_t \varepsilon_{it}F_t(m_2)\psi_t^*(\bar{m},m_2)\right]^2\right)^{1/2} \\
&= O_p(T^{-1/2}),
\end{aligned}
$$

where the inequality holds by the Cauchy-Schwartz inequality, and the second equality holds by Assumption 3.2 and by invoking the central limit theorem between $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{F}_{\psi^*}$

given Assumptions 3.1, 3.3 and 3.6.

$$\begin{aligned} 3.II &= T^{-1/2}\left(N^{-1}\sum_i \phi_i(m_1)\right)\left(N^{-1}\sum_j T^{-1/2}\sum_t \varepsilon_{it}F_t(m_2)\psi_t^*(\bar{m},m_2)\right) \\ &= O_p(T^{-1/2}), \end{aligned}$$

by Assumption 3.2 and by invoking the central limit theorem between $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{F}_{\psi^*}(\bar{m})$.

$$\begin{aligned} 3.III &\leq N^{-1/2}\left(T^{-1}\sum_s \psi_s^*(\bar{m},m_2)F_s(m_2)\right)\left(T^{-1}\sum_t N^{-1/2}\sum_j \varepsilon_{jt}\phi_i(m_1)\right) \\ &= O_p(N^{-1/2}), \end{aligned}$$

by the Cauchy-Schwartz inequality, Assumptions 1 and 3.1 and by invoking the central limit theorem between $\boldsymbol{\varepsilon}_t$ and $\boldsymbol{\phi}$ given Assumptions 3.2 and 3.3.

$$\begin{aligned} 3.IV &\leq T^{-1/2}N^{-1/2}\left(N^{-1}\sum_i \phi_i(m_1)\right)\left(T^{-1}\sum_s \psi_s^*(\bar{m},m_2)F_s(m_2)\right)\left(T^{-1/2}N^{-1/2}\sum_{j,t}\varepsilon_{jt}\right) \\ &= O_p(T^{-1/2}N^{-1/2}), \end{aligned}$$

by the Cauchy-Schwartz inequality, Assumptions 1, 3.1 and 3.2 and Lemma 1.3 in Kelly and Pruitt, 2015. Summing these terms we get that Item 3 is $\boldsymbol{O}_p(\delta_{NT}^{-1})$.

Item 4: The object of interest is a $K \times K$ matrix, for each entry $(m_1, m_2, m_3, m_4)$ we have:

$$
\begin{aligned}
N^{-1}T^{-3/2}\sum_{i,s,t} F_{\psi^*,s}(m_1)\varepsilon_{i,s}\varepsilon_{i,t}F_{\psi^*,t}(\bar{m}_1,m_2) \\
-N^{-1}T^{-5/2}2\sum_{i,s,t,u} F_{\psi^*,s}(\bar{m}_2,m_1)\varepsilon_{i,s}\varepsilon_{i,t}F_{\psi^*,u}(\bar{m},m_2) \\
+N^{-1}T^{-7/2}\sum_{i,s,t,u,v} F_{\psi^*,s}(\bar{m}_1,m_1)\varepsilon_{i,t}\varepsilon_{i,u}F_{\psi^*,v}(\bar{m}_2,m_2) \\
+N^{-2}T^{-3/2}\sum_{i,j,s,t} F_{\psi^*,s}(\bar{m}_1,m_1)\varepsilon_{i,s}\varepsilon_{j,t}F_{\psi^*,t}(\bar{m}_2,m_2) \\
-N^{-2}T^{-5/2}2\sum_{i,j,s,t,u} F_{\psi^*,s}(\bar{m}_1,m_1)\varepsilon_{i,s}\varepsilon_{j,t}F_{\psi^*,u}(\bar{m}_2,m_2) \\
-N^{-2}T^{-7/2}\sum_{i,j,s,t,u,v} F_{\psi^*,s}(\bar{m}_1,m_1)\varepsilon_{i,t}\varepsilon_{j,u}F_{\psi^*,v}(\bar{m}_2,m_2) \\
= 4.I - \cdots - 4.VI,
\end{aligned}
$$

$$
\begin{aligned}
4.I &= T^{-1/2}\left[N^{-1}\sum_i\left(T^{-1/2}\sum_s \psi_s^*(\bar{m}_1,m_1)F_s(m_1)\varepsilon_{i,s}\right)\left(T^{-1/2}\sum_t \psi_t^*(\bar{m}_2,m_2)F_t(m_2)\varepsilon_{i,t}\right)\right] \\
&= O_p(T^{-1/2}),
\end{aligned}
$$

by invoking the central limit theorem between $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{F}_{\psi^*}(\bar{m})$.

$$
\begin{aligned}
4.II &= \left(T^{-1}\sum_u \psi_u^*(\bar{m}_2,m_2)F_u(m_2)\right)T^{-1}\sum_t\left(N^{-1}T^{-1/2}\sum_{i,s}\psi_s^*(\bar{m}_1,m_1)F_s(m_1)\varepsilon_{i,s}\varepsilon_{i,t}\right) \\
&= O_p(\delta_{NT}^{-1}),
\end{aligned}
$$

by Assumptions 1, 3.1 and a slight modification of the result of Lemma 1.12 in Kelly and Pruitt, 2015. This modification involves the use of the Cauchy-Schwartz inequality and the fact that the first moments of the variable $\boldsymbol{\psi}_t$ are bounded and it is independent of

$\boldsymbol{F}_t$, which leads that the term $N^{-1}T^{-1/2}\sum_{i,s}\psi_s^*(\bar{m}_1, m_1)F_s(m_1)\varepsilon_{i,s}\varepsilon_{i,t}$ is $O_p(\delta_{NT}^{-1})$.

$$\begin{aligned}
4.III &= \left(T^{-1}\sum_s \psi_s^*(\bar{m}_1, m_1)F_s(m_1)\right)\left(T^{-1}\sum_u N^{-1}T^{-1/2}\sum_{i,t}\varepsilon_{i,t}\varepsilon_{i,u}\right) \\
&\quad \left(T^{-1}\sum_v \psi_v^*(\bar{m}_2, m_2)F_v(m_2)\right) \\
&= O_p(\delta_{NT}^{-1}),
\end{aligned}$$

by Assumptions 1, 3.1 and Lemma 1.10 in Kelly and Pruitt, 2015.

$$\begin{aligned}
4.IV &= T^{-1/2}\left[N^{-1}\sum_i\left(T^{-1/2}\sum_s \psi_s^*(\bar{m}_1, m_1)F_s(m_1)\varepsilon_{i,s}\right)\right. \\
&\quad \left. N^{-1}\sum_j\left(T^{-1/2}\sum_t \psi_t^*(\bar{m}_2, m_2)F_t(m_2)\varepsilon_{j,t}\right)\right] \\
&= O_p(T^{-1/2}),
\end{aligned}$$

by invoking the central limit theorem between $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{F}_{\psi^*}(\bar{m})$.

$$\begin{aligned}
4.V &= N^{-1/2}T^{-1/2}\left[N^{-1}\sum_i\left(T^{-1/2}\sum_s \psi_s^*(\bar{m}_1, m_1)F_s(m_1)\varepsilon_{i,s}\right)\left(N^{-1/2}T^{-1/2}\sum_{j,t}\varepsilon_{j,t}\right)\right. \\
&\quad \left.\left(T^{-1}\sum_u \psi_u^*(\bar{m}_2, m_2)F_u(m_2)\right)\right] \\
&= O_p(N^{-1/2}T^{-1/2}),
\end{aligned}$$

by invoking the central limit theorem between $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{F}_{\psi^*}(\bar{m})$, Assumptions 1 and 3.1, and Lemma 1.3 in Kelly and Pruitt, 2015.

$$\begin{aligned}
4.VI &= N^{-1}T^{-1/2}\left[\left(T^{-1}\sum_s \psi_s^*(\bar{m}_1, m_1)F_s(m_1)\right)\left(N^{-1/2}T^{-1/2}\sum_{i,t}\varepsilon_{i,t}\right)\right. \\
&\quad \left.\left(N^{-1/2}T^{-1/2}\sum_{j,t}\varepsilon_{j,u}\right)\left(T^{-1}\sum_v \psi_v^*(\bar{m}_2, m_2)F_v(m_2)\right)\right] \\
&= O_p(N^{-1}T^{-1/2}),
\end{aligned}$$

by invoking the central limit theorem between $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{F}_{\psi^*}(\bar{m})$, and Lemma 1.3 in Kelly and Pruitt, 2015. Summing these terms we get that Item 4 is $\boldsymbol{O}_p(\delta_{NT}^{-1})$.

The proof of Item 5 follows the same argument as Item 4 but replace where appropriate $\omega_s(m)$ for $F_{\psi^*,s}(\bar{m}, m)$, Lemma 1.13 in Kelly and Pruitt, 2015 for Lemma 1.12 in Kelly and Pruitt, 2015, and changing the use of central limit theorem between $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{F}_{\psi^*}(\bar{m})$ by Assumption 4.4. Then, Item 5 is $\boldsymbol{O}_p(\delta_{NT}^{-1})$.

Item 6: The object of interest is a vector of size $K$, for each entry $m$th we have:

$$
\begin{aligned}
N^{-1}T^{-1/2}\sum_{i,s}\psi_s^*(\bar{m},m)F_s(m)\varepsilon_{i,s}\varepsilon_{i,t} - N^{-2}T^{-1/2}\sum_{i,j,s}\psi_s^*(\bar{m},m)F_s(m)\varepsilon_{i,s}\varepsilon_{j,t} \\
-N^{-1}T^{-3/2}\sum_{i,s,u}\psi_s^*(\bar{m},m)F_s(m)\varepsilon_{i,u}\varepsilon_{i,t} + N^{-2}T^{-3/2}\sum_{i,j,s,u}\psi_s^*(\bar{m},m)F_s(m)\varepsilon_{i,u}\varepsilon_{j,t} \\
= 6.I + 6.II + 6.III + 6.IV,
\end{aligned}
$$

$6.I = O_p(\delta_{NT}^{-1})$ by Assumptions 1, 3.1 and a slight modification of the result of Lemma 1.12 in Kelly and Pruitt, 2015.

$$
\begin{aligned}
6.II &= N^{-1/2}\left[N^{-1}\sum_i\left(T^{-1/2}\sum_s\psi_s^*(\bar{m},m)F_s(m)\varepsilon_{i,s}\right)\right]\left(N^{-1/2}\sum_j\varepsilon_{j,t}\right) \\
&= O_p(N^{-1/2}),
\end{aligned}
$$

by invoking the central limit theorem between $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{F}_{\psi^*}$ and Lemma 1.3 in Kelly and Pruitt, 2013.

$$
\begin{aligned}
6.III &= \left(T^{-1}\sum_s\psi_s^*(\bar{m},m)F_s(m)\right)\left(N^{-1}T^{-1/2}\sum_{i,u}\varepsilon_{i,u}\varepsilon_{i,t}\right) \\
&= O_p(\delta_{NT}^{-1}),
\end{aligned}
$$

by Assumption 1 and 3.1, and Lemma 1.10 in Kelly and Pruitt, 2015.

$$\begin{aligned} 6.IV &= N^{-1}\left(T^{-1}\sum_{s}\psi_s^*(\bar{m},m)F_s(m)\right)\left(N^{-1/2}T^{-1/2}\sum_{i,u}\varepsilon_{i,u}\right)\left(N^{-1/2}\sum_{j}\varepsilon_{j,t}\right)\\ &= O_p(N^{-1}), \end{aligned}$$

by Assumption 1 and 3.1, and Lemma 1.3 in Kelly and Pruitt, 2015. Summing all the terms we obtain that Item 6 is $\boldsymbol{O}_p(\delta_{NT}^{-1})$.

Item 7: The object of interest is a vector of size $K$, for each entry $m$th we have:

$$\begin{aligned} &N^{-1}T^{-3/2}\sum_{i,s,t}F_{\psi^*,s}(\bar{m},m)\varepsilon_{i,s}\varepsilon_{i,t}\xi_{t+1} - N^{-1}T^{-3/2}\sum_{i,s,t,u}F_{\psi^*,s}(\bar{m},m)\varepsilon_{i,s}\varepsilon_{i,t}\xi_{u+1}\\ &+N^{-1}T^{-5/2}\sum_{i,s,t,u}F_{\psi^*,s}(\bar{m},m)\varepsilon_{i,t}\varepsilon_{i,u}\xi_{u+1} + N^{-1}T^{-7/2}\sum_{i,s,t,u,v}F_{\psi^*,s}(\bar{m},m)\varepsilon_{i,t}\varepsilon_{i,u}\xi_{v+1}\\ &-N^{-2}T^{-3/2}\sum_{i,j,s,t}F_{\psi^*,s}(\bar{m},m)\varepsilon_{i,s}\varepsilon_{j,t}\xi_{t+1} + N^{-2}T^{-5/2}\sum_{i,j,s,t,u}F_{\psi^*,s}(\bar{m},m)\varepsilon_{i,s}\varepsilon_{j,t}\xi_{u+1}\\ &+N^{-2}T^{-5/2}\sum_{i,j,s,t,u}F_{\psi^*,s}(\bar{m},m)\varepsilon_{i,t}\varepsilon_{j,u}\xi_{u+1} - N^{-2}T^{-7/2}\sum_{i,j,s,t,u,v}F_{\psi^*,s}(\bar{m},m)\varepsilon_{i,t}\varepsilon_{j,u}\xi_{v+1}\\ &= 7.I - \cdots - 7.VIII, \end{aligned}$$

$$\begin{aligned} 7.I &= T^{-1/2}\left[N^{-1}\sum_{i}\left(T^{-1/2}\sum_{s}\psi_s^*(\bar{m},m)F_s(m)\varepsilon_{i,s}\right)\left(T^{-1/2}\sum_{t}\xi_{t+1}\varepsilon_{i,t}\right)\right]\\ &= O_p(T^{-1/2}), \end{aligned}$$

by invoking the central limit theorem between $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{F}_{\psi^*}$, and Lemma 1.5 in Kelly and Pruitt, 2015.

$$\begin{aligned} 7.II &= T^{-1/2}\left(N^{-1}T^{-1/2}\sum_{i,s}\psi_s^*(\bar{m},m)F_s(m)\varepsilon_{i,s}\varepsilon_{i,t}\right)\left(T^{-1/2}\sum_{u}\xi_{u+1}\right)\\ &= O_p(T^{-1/2}\delta_{NT}^{-1}), \end{aligned}$$

by a slight modification of the result of Lemma 1.12 in Kelly and Pruitt, 2015 and Lemma 1.4 in Kelly and Pruitt, 2015.

$$\begin{aligned} 7.III &= \left(T^{-1}\sum_s \psi_s^*(\bar{m},m)F_s(m)\right)\left(N^{-1}T^{-3/2}\sum_{i,u,t}\xi_{u+1}\varepsilon_{i,u}\varepsilon_{i,t}\right) \\ &= O_p(\delta_{NT}^{-1}), \end{aligned}$$

by Assumption 1 and 3.1, and Lemma 1.11 in Kelly and Pruitt, 2015.

$$\begin{aligned} 7.IV &= T^{-1/2}\left(T^{-1}\sum_s \psi_s^*(\bar{m},m)F_s(m)\right)\left[T^{-1}\sum_t\left(N^{-1}T^{-1/2}\sum_{i,u}\varepsilon_{i,t}\varepsilon_{i,u}\right)\right]\left(T^{-1/2}\sum_v \xi_{v+1}\right) \\ &= O_p(T^{-1/2}\delta_{NT}^{-1}), \end{aligned}$$

by Assumption 1, 3.1, Lemma 1.10 and Lemma 1.4 in Kelly and Pruitt, 2015.

$$\begin{aligned} 7.V &= N^{-1/2}T^{-1/2}\left[N^{-1}\sum_i\left(T^{-1/2}\sum_s \psi_s^*(\bar{m},m)F_s(m)\varepsilon_{i,s}\right)\right]\left(N^{-1/2}T^{-1/2}\sum_{j,t}\varepsilon_{j,t}\xi_{t+1}\right) \\ &= O_p(N^{-1/2}T^{-1/2}), \end{aligned}$$

by invoking the central limit theorem between $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{F}_{\psi^*}(\bar{m})$, and Lemma 1.6 in Kelly and Pruitt, 2015.

$$\begin{aligned} 7.VI &= N^{-1/2}T^{-1}\left[N^{-1}\sum_i\left(T^{-1/2}\sum_s \psi_s^*(\bar{m},m)F_s(m)\varepsilon_{i,s}\right)\right] \\ &\quad \left(N^{-1/2}T^{-1/2}\sum_{j,t}\varepsilon_{j,t}\right)\left(T^{-1/2}\sum_v \xi_{v+1}\right) \\ &= O_p(N^{-1/2}T^{-1}), \end{aligned}$$

by invoking the central limit theorem between $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{F}_{\psi^*}(\bar{m})$, Lemma 1.3 and Lemma

1.4 in Kelly and Pruitt, 2015.

$$\begin{aligned}
7.VII &= N^{-1}T^{-1/2}\left(T^{-1}\sum_{s}\psi_s^*(\bar{m},m)F_s(m)\right)\left(N^{-1/2}T^{-1/2}\sum_{i,t}\varepsilon_{i,t}\right) \\
&\quad \left(N^{-1/2}T^{-1/2}\sum_{j,u}\xi_{u+1}\varepsilon_{j,u}\right) \\
&= O_p(N^{-1}T^{-1/2}),
\end{aligned}$$

by Assumption 1, 3.1, Lemma 1.3 and Lemma 1.6 in Kelly and Pruitt, 2015.

$$\begin{aligned}
7.VIII &= N^{-1}T^{-1}\left(T^{-1}\sum_{s}\psi_s^*(\bar{m},m)F_s(m)\right)\left(N^{-1/2}T^{-1/2}\sum_{i,t}\varepsilon_{i,t}\right) \\
&\quad \left(N^{-1/2}T^{-1/2}\sum_{j,u}\varepsilon_{j,u}\right)\left(T^{-1/2}\sum_{v}\xi_{v+1}\right) \\
&= O_p(N^{-1}T^{-1}),
\end{aligned}$$

by Assumption 1, 3.1, Lemma 1.3 and Lemma 1.4 in Kelly and Pruitt, 2015. Summing all the terms we obtain that Item 7 is $\boldsymbol{O}_p(\delta_{NT}^{-1})$. □

Lemma 2 provides asymptotic limits of the estimates for the relevant factors and their corresponding factor loadings.

**Lemma 2.** *Let Assumptions 1-5 hold. Then, the probability limits of* $\hat{\boldsymbol{\Phi}}_\tau$ *and* $\hat{\boldsymbol{F}}_t$ *are:*

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

*where* $\boldsymbol{f}_t$ *is the part of* $\boldsymbol{F}_t$ *corresponding to* $\boldsymbol{\Lambda}_{f,\tau}$ *such that* $\boldsymbol{\Lambda}_\tau\boldsymbol{F}_t = [\boldsymbol{\Lambda}_{f,\tau}\quad \boldsymbol{0}]\boldsymbol{F}_t = [\boldsymbol{\Lambda}_{f,\tau}\boldsymbol{f}_t\quad \boldsymbol{0}]$, $\hat{\boldsymbol{\Phi}}_{f,\tau}$ *are the corresponding loadings of* $\boldsymbol{f}_t$, *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.(16)), and

Assumptions 4 and 5. It is similar to Lemma 3 in Kelly and Pruitt, 2015, and it is further simplified by Assumptions 1 and 5. Specifically,

$$
\begin{aligned}
\hat{\boldsymbol{F}}_t &= T^{-1}\boldsymbol{Z}_\tau'\boldsymbol{J}_T\boldsymbol{Z}_\tau\left(N^{-1}T^{-2}\boldsymbol{Z}_\tau'\boldsymbol{J}_T\boldsymbol{X}\boldsymbol{J}_N\boldsymbol{X}'\boldsymbol{J}_T\boldsymbol{Z}_\tau\right)^{-1}N^{-1}T^{-1}\boldsymbol{Z}_\tau'\boldsymbol{J}_T\boldsymbol{X}\boldsymbol{J}_N\boldsymbol{x}_t \\
&= \hat{\boldsymbol{F}}_A\boldsymbol{F}_B^{-1}\hat{\boldsymbol{F}}_{C,t}.
\end{aligned}
$$

Handling each of the three terms separately we obtain the following.

$$
\begin{aligned}
\hat{\boldsymbol{F}}_A &= T^{-1}\boldsymbol{Z}_\tau'\boldsymbol{J}_T\boldsymbol{Z}_\tau \\
&= \boldsymbol{\Lambda}_\tau\left(T^{-1}\boldsymbol{F}_{\psi^*}'\boldsymbol{J}_T\boldsymbol{F}_{\psi^*}\right)\boldsymbol{\Lambda}_\tau' + \boldsymbol{\Lambda}_\tau\left(T^{-1}\boldsymbol{F}_{\psi^*}'\boldsymbol{J}_T\boldsymbol{\omega}\right) + \left(T^{-1}\boldsymbol{\omega}'\boldsymbol{J}_T\boldsymbol{F}_{\psi^*}\right)\boldsymbol{\Lambda}_\tau' + T^{-1}\boldsymbol{\omega}'\boldsymbol{J}_T\boldsymbol{\omega} \\
&\xrightarrow[T,N\to\infty]{p} \boldsymbol{\Lambda}_\tau\mathrm{Var}(\boldsymbol{F}_{\psi^*})\boldsymbol{\Lambda}_\tau' + \boldsymbol{\Delta}_\omega = \boldsymbol{\Lambda}_\tau\boldsymbol{\Delta}_F\boldsymbol{\Delta}_{\psi^*}\boldsymbol{\Lambda}_\tau' + \boldsymbol{\Delta}_\omega = \begin{bmatrix}\boldsymbol{\Lambda}_{f,\tau}\boldsymbol{\Delta}_{f,\tau}\boldsymbol{\Delta}_\psi\boldsymbol{\Lambda}_{f,\tau}' + \boldsymbol{\Delta}_\omega & \boldsymbol{0}\end{bmatrix}.
\end{aligned}
$$

$$
\begin{aligned}
\hat{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 \\
&= \boldsymbol{\Lambda}_\tau\left(T^{-1}\boldsymbol{F}'_{\psi^*}\boldsymbol{J}_T\boldsymbol{F}\right)\left(N^{-1}\boldsymbol{\Phi}'\boldsymbol{J}_N\boldsymbol{\Phi}\right)\left(T^{-1}\boldsymbol{F}'\boldsymbol{J}_T\boldsymbol{F}_{\psi^*}\right)\boldsymbol{\Lambda}'_\tau \\
&\quad +\boldsymbol{\Lambda}_\tau\left(T^{-1}\boldsymbol{F}'_{\psi^*}\boldsymbol{J}_T\boldsymbol{F}\right)\left(N^{-1}\boldsymbol{\Phi}'\boldsymbol{J}_N\boldsymbol{\Phi}\right)\left(T^{-1}\boldsymbol{F}'\boldsymbol{J}_T\boldsymbol{\omega}\right) \\
&\quad +\boldsymbol{\Lambda}_\tau\left(T^{-1}\boldsymbol{F}'_{\psi^*}\boldsymbol{J}_T\boldsymbol{F}\right)\left(N^{-1}T^{-1}\boldsymbol{\Phi}'\boldsymbol{J}_N\boldsymbol{\varepsilon}\boldsymbol{J}_T\boldsymbol{F}_{\psi^*}\right)\boldsymbol{\Lambda}'_\tau \\
&\quad +\boldsymbol{\Lambda}_\tau\left(T^{-1}\boldsymbol{F}'_{\psi^*}\boldsymbol{J}_T\boldsymbol{F}\right)\left(N^{-1}T^{-1}\boldsymbol{\Phi}'\boldsymbol{J}_N\boldsymbol{\varepsilon}\boldsymbol{J}_T\boldsymbol{\omega}\right) \\
&\quad +\boldsymbol{\Lambda}_\tau\left(N^{-1}T^{-1}\boldsymbol{F}'_{\psi^*}\boldsymbol{J}_T\boldsymbol{\varepsilon}\boldsymbol{J}_N\boldsymbol{\Phi}\right)\left(T^{-1}\boldsymbol{F}'\boldsymbol{J}_T\boldsymbol{F}_{\psi^*}\right)\boldsymbol{\Lambda}'_\tau \\
&\quad +\boldsymbol{\Lambda}_\tau\left(N^{-1}T^{-1}\boldsymbol{F}'_{\psi^*}\boldsymbol{J}_T\boldsymbol{\varepsilon}\boldsymbol{J}_N\boldsymbol{\Phi}\right)\left(T^{-1}\boldsymbol{F}'\boldsymbol{J}_T\boldsymbol{\omega}\right) \\
&\quad +\boldsymbol{\Lambda}_\tau\left(N^{-1}T^{-2}\boldsymbol{F}'_{\psi^*}\boldsymbol{J}_T\boldsymbol{\varepsilon}\boldsymbol{J}_N\boldsymbol{\varepsilon}'\boldsymbol{J}_T\boldsymbol{F}_{\psi^*}\right)\boldsymbol{\Lambda}'_\tau \\
&\quad +\boldsymbol{\Lambda}_\tau\left(N^{-1}T^{-2}\boldsymbol{F}'_{\psi^*}\boldsymbol{J}_T\boldsymbol{\varepsilon}\boldsymbol{J}_N\boldsymbol{\varepsilon}'\boldsymbol{J}_T\boldsymbol{\omega}\right) \\
&\quad +\left(T^{-1}\boldsymbol{\omega}'\boldsymbol{J}_T\boldsymbol{F}\right)\left(N^{-1}\boldsymbol{\Phi}'\boldsymbol{J}_N\boldsymbol{\Phi}\right)\left(T^{-1}\boldsymbol{F}'\boldsymbol{J}_T\boldsymbol{F}_{\psi^*}\right)\boldsymbol{\Lambda}'_\tau \\
&\quad +\left(T^{-1}\boldsymbol{\omega}'\boldsymbol{J}_T\boldsymbol{F}\right)\left(N^{-1}\boldsymbol{\Phi}'\boldsymbol{J}_N\boldsymbol{\Phi}\right)\left(T^{-1}\boldsymbol{F}'\boldsymbol{J}_T\boldsymbol{\omega}\right) \\
&\quad +\left(T^{-1}\boldsymbol{\omega}'\boldsymbol{J}_T\boldsymbol{F}\right)\left(N^{-1}T^{-1}\boldsymbol{\Phi}'\boldsymbol{J}_N\boldsymbol{\varepsilon}\boldsymbol{J}_T\boldsymbol{F}_{\psi^*}\right)\boldsymbol{\Lambda}'_\tau \\
&\quad +\left(T^{-1}\boldsymbol{\omega}'\boldsymbol{J}_T\boldsymbol{F}\right)\left(N^{-1}T^{-1}\boldsymbol{\Phi}'\boldsymbol{J}_N\boldsymbol{\varepsilon}\boldsymbol{J}_T\boldsymbol{\omega}\right) \\
&\quad +\left(N^{-1}T^{-1}\boldsymbol{\omega}'\boldsymbol{J}_T\boldsymbol{\varepsilon}\boldsymbol{J}_N\boldsymbol{\Phi}\right)\left(T^{-1}\boldsymbol{F}'\boldsymbol{J}_T\boldsymbol{F}_{\psi^*}\right)\boldsymbol{\Lambda}'_\tau \\
&\quad +\left(N^{-1}T^{-1}\boldsymbol{\omega}'\boldsymbol{J}_T\boldsymbol{\varepsilon}\boldsymbol{J}_N\boldsymbol{\Phi}\right)\left(T^{-1}\boldsymbol{F}'\boldsymbol{J}_T\boldsymbol{\omega}\right) \\
&\quad +\left(N^{-1}T^{-2}\boldsymbol{\omega}'\boldsymbol{J}_T\boldsymbol{\varepsilon}\boldsymbol{J}_N\boldsymbol{\varepsilon}\boldsymbol{J}_T\boldsymbol{F}_{\psi^*}\right)\boldsymbol{\Lambda}'_\tau \\
&\quad +\left(N^{-1}T^{-2}\boldsymbol{\omega}'\boldsymbol{J}_T\boldsymbol{\varepsilon}\boldsymbol{J}_N\boldsymbol{\varepsilon}\boldsymbol{J}_T\boldsymbol{\omega}\right) \\
&\quad \xrightarrow[T,N\to\infty]{p} \boldsymbol{\Lambda}_\tau\mathbb{E}(\boldsymbol{F}'_{\psi^*}\boldsymbol{J}_T\boldsymbol{F})^2\boldsymbol{\Lambda}'_\tau = \boldsymbol{\Lambda}_\tau\boldsymbol{\Delta}_F^2\boldsymbol{\mu}_{\psi^*}^2\boldsymbol{\Lambda}'_\tau = \begin{bmatrix}\boldsymbol{\Lambda}_{f,\tau}\boldsymbol{\Delta}_{f,\tau}^2\boldsymbol{\mu}_\psi^2\boldsymbol{\Lambda}'_{f,\tau} & \boldsymbol{0}\end{bmatrix}.
\end{aligned}
$$

$$\begin{aligned}
\hat{\boldsymbol{F}}_{C,t} &= N^{-1}T^{-1}\boldsymbol{Z}'_{\tau}\boldsymbol{J}_T\boldsymbol{X}\boldsymbol{J}_N\boldsymbol{x}_t \\
&= \boldsymbol{\Lambda}_{\tau}\left(T^{-1}\boldsymbol{F}'_{\psi^*}\boldsymbol{J}_T\boldsymbol{F}\right)\left(N^{-1}\boldsymbol{\Phi}'\boldsymbol{J}_N\boldsymbol{\phi}_0\right) + \boldsymbol{\Lambda}_{\tau}\left(T^{-1}\boldsymbol{F}'_{\psi^*}\boldsymbol{J}_T\boldsymbol{F}\right)\left(N^{-1}\boldsymbol{\Phi}'\boldsymbol{J}_N\boldsymbol{\Phi}\right)\boldsymbol{F}_t \\
&\quad + \boldsymbol{\Lambda}_{\tau}\left(T^{-1}\boldsymbol{F}'_{\psi^*}\boldsymbol{J}_T\boldsymbol{F}\right)\left(N^{-1}\boldsymbol{\Phi}'\boldsymbol{J}_N\boldsymbol{\varepsilon}_t\right) \\
&\quad + \boldsymbol{\Lambda}_{\tau}\left(N^{-1}T^{-1}\boldsymbol{F}'_{\psi^*}\boldsymbol{J}_T\boldsymbol{\varepsilon}\boldsymbol{J}_N\boldsymbol{\phi}_0\right) + \boldsymbol{\Lambda}_{\tau}\left(N^{-1}T^{-1}\boldsymbol{F}'_{\psi^*}\boldsymbol{J}_T\boldsymbol{\varepsilon}\boldsymbol{J}_N\boldsymbol{\phi}\right)\boldsymbol{F}_t \\
&\quad + \boldsymbol{\Lambda}_{\tau}\left(N^{-1}T^{-1}\boldsymbol{F}'_{\psi^*}\boldsymbol{J}_T\boldsymbol{\varepsilon}\boldsymbol{J}_N\boldsymbol{\varepsilon}_t\right) \\
&\quad + \left(T^{-1}\boldsymbol{\omega}'\boldsymbol{J}_T\boldsymbol{F}\right)\left(N^{-1}\boldsymbol{\Phi}'\boldsymbol{J}_N\boldsymbol{\phi}_0\right) + \left(T^{-1}\boldsymbol{\omega}'\boldsymbol{J}_T\boldsymbol{F}\right)\left(N^{-1}\boldsymbol{\Phi}'\boldsymbol{J}_N\boldsymbol{\Phi}\right)\boldsymbol{F}_t \\
&\quad + \left(T^{-1}\boldsymbol{\omega}'\boldsymbol{J}_T\boldsymbol{F}\right)\left(N^{-1}\boldsymbol{\Phi}'\boldsymbol{J}_N\boldsymbol{\varepsilon}_t\right) \\
&\quad + \left(N^{-1}T^{-1}\boldsymbol{\omega}'\boldsymbol{J}_T\boldsymbol{\varepsilon}\boldsymbol{J}_N\boldsymbol{\phi}_0\right) + \left(N^{-1}T^{-1}\boldsymbol{\omega}'\boldsymbol{J}_T\boldsymbol{\varepsilon}\boldsymbol{J}_N\boldsymbol{\Phi}\right)\boldsymbol{F}_t \\
&\quad + \left(N^{-1}T^{-1}\boldsymbol{\omega}'\boldsymbol{J}_T\boldsymbol{\varepsilon}\boldsymbol{J}_N\boldsymbol{\varepsilon}_t\right) \\
&\xrightarrow[T,N\to\infty]{p} \boldsymbol{\Lambda}_{\tau}\mathbb{E}\left(\boldsymbol{F}'_{\psi^*}\boldsymbol{J}_T\boldsymbol{F}\right)\boldsymbol{F}_t = \boldsymbol{\Lambda}_{\tau}\boldsymbol{\mu}_{\psi^*}\boldsymbol{\Delta}_F\boldsymbol{F}_t = \left[\boldsymbol{\Lambda}_{f,\tau}\boldsymbol{\mu}_{\psi}\boldsymbol{\Delta}_{f,\tau}\boldsymbol{f}_t \quad \mathbf{0}\right].
\end{aligned}$$

Each convergence result follows from Lemma 1, Lemma 2 in Kelly and Pruitt, 2015 and Assumptions 1-5. The convergence of $\hat{\boldsymbol{F}}_t$ is obtained after applying the continuous mapping theorem. As $\hat{\boldsymbol{\Phi}}_{\tau} = \hat{\boldsymbol{F}}_A\left(T^{-1}\boldsymbol{Z}'_{\tau}\boldsymbol{J}_T\boldsymbol{X}\right)$ and since $\text{plim}_{N,T\to\infty} T^{-1}\boldsymbol{Z}'_{\tau}\boldsymbol{J}_T\boldsymbol{X} = \boldsymbol{\Lambda}_{\tau}\boldsymbol{\mu}_{\psi^*}\boldsymbol{\Delta}_F\boldsymbol{\Phi}' = \left[\boldsymbol{\Lambda}_{f,\tau}\boldsymbol{\mu}_{\psi}\boldsymbol{\Delta}_{f,\tau}\boldsymbol{\Phi}'_f \quad \mathbf{0}\right]$, we obtain the convergence result for $\hat{\boldsymbol{\Phi}}_{f,\tau}$ by the above results and lemmas mentioned, and the continuous mapping theorem. □

***Proof of Theorem 1.*** Following Lemma 2, 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:

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

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 solves:

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

are such that,

$$\sqrt{T}(\dot{\boldsymbol{\beta}}_{\tau} - \boldsymbol{\beta}'_{f,\tau} \boldsymbol{H}_{\tau}^{-1}) \xrightarrow[N,T\to\infty]{p} \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\dot{\boldsymbol{\beta}}}). \tag{A.21}$$

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

$$\hat{\boldsymbol{\beta}}_{\tau} = \dot{\boldsymbol{\beta}}_{\tau} + \left( \sum_{t=1}^{T} c_t \hat{\boldsymbol{f}}_t \hat{\boldsymbol{f}}'_t \right)^{-1} \left( \sum_{t=1}^{T} c_t \hat{\boldsymbol{f}}_t \dot{\boldsymbol{\beta}}'_{1,\tau} (\hat{\boldsymbol{f}}_t - \boldsymbol{H}_{\tau} \boldsymbol{f}_t) \right), \tag{A.22}$$

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

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

What remains to find is the convergence order of $(\hat{\boldsymbol{\beta}}'_{\tau} - \boldsymbol{\beta}'_{f,\tau} \boldsymbol{H}_{\tau}^{-1})$. From Eq.(A.22) we get:

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

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 $(\hat{\boldsymbol{\beta}}'_\tau - \boldsymbol{\beta}'_{f,\tau}\boldsymbol{H}_\tau^{-1}) = \boldsymbol{O}_p(T^{-1/2}) + \boldsymbol{O}_p(1)\boldsymbol{O}_p(\delta_{NT}^{-1}) = \boldsymbol{O}_p(\delta_{NT}^{-1})$. Therefore, $\hat{\boldsymbol{\beta}}'_\tau \hat{\boldsymbol{f}}_t - \boldsymbol{\beta}'_{f,\tau}\boldsymbol{f}_t = \boldsymbol{O}_p(1)\boldsymbol{O}_p(\delta_{NT}^{-1}) + \boldsymbol{O}_p(\delta_{NT}^{-1})\boldsymbol{O}_p(1) = \boldsymbol{O}_p(\delta_{NT}^{-1})$. □

To show the asymptotic normality of the infeasible quantile forecasts we need to make use of the following lemma, whose proof is similar to Theorem 6 in Kelly and Pruitt, 2015.

**Lemma 3.** *Let Assumptions 1-5 hold. We have for all t,*

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

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

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

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

*Where*

$\operatorname{plim}\boldsymbol{\Sigma}_{\hat{f}} = (\boldsymbol{\Delta}_\psi \boldsymbol{\Delta}_f \boldsymbol{\Lambda}_{f,\tau}\boldsymbol{\Lambda}'_{f,\tau} + \boldsymbol{\Delta}_\omega)\left(\boldsymbol{\mu}_\psi^2 \boldsymbol{\Lambda}_\tau \boldsymbol{\Delta}_f^2 \boldsymbol{\Lambda}'_\tau\right)^{-1} \boldsymbol{\mu}_\psi \boldsymbol{\Lambda}_\tau \boldsymbol{\Delta}_f \boldsymbol{\Sigma}_{\Phi\varepsilon}\boldsymbol{\Delta}_f \boldsymbol{\Lambda}'_\tau \boldsymbol{\mu}_\psi$

*$\left(\boldsymbol{\mu}_{\psi}^2\boldsymbol{\Lambda}_\tau\boldsymbol{\Delta}_f^2\boldsymbol{\Lambda}_\tau'\right)^{-1}(\boldsymbol{\Delta}_\psi\boldsymbol{\Lambda}_\tau\boldsymbol{\Delta}_f\boldsymbol{\Lambda}_\tau'+\boldsymbol{\Delta}_\omega)'$, and $\hat{\boldsymbol{H}}_\tau = \hat{\boldsymbol{F}}_A\hat{\boldsymbol{F}}_B^{-1}N^{-1}T^{-1}\boldsymbol{Z}_\tau'\boldsymbol{J}_T\boldsymbol{X}\boldsymbol{J}_N\boldsymbol{\Phi}$, with $\hat{\boldsymbol{F}}_A = T^{-1}\boldsymbol{Z}_\tau'\boldsymbol{J}_T\boldsymbol{Z}_\tau$ and $\hat{\boldsymbol{F}}_B = N^{-1}T^{-2}\boldsymbol{Z}_\tau'\boldsymbol{J}_T\boldsymbol{X}\boldsymbol{J}_N\boldsymbol{X}'\boldsymbol{J}_T\boldsymbol{Z}_\tau$, and a feasible estimator for the covariance matrix $\boldsymbol{\Sigma}_{\hat{f}}$ is given by:*

$$
\begin{aligned}
\hat{\Sigma}_f &= (T^{-1}\boldsymbol{Z}_\tau'\boldsymbol{J}_T\boldsymbol{Z}_\tau)\,(N^{-1}T^{-2}\boldsymbol{Z}_\tau'\boldsymbol{J}_T\boldsymbol{X}\boldsymbol{J}_N\boldsymbol{X}'\boldsymbol{J}_T\boldsymbol{Z}_\tau)\left(N^{-2}T^2\boldsymbol{Z}_\tau'\boldsymbol{J}_T\hat{\boldsymbol{f}}\sum_i\hat{\boldsymbol{\phi}}_{f,\tau,i}^2\hat{\varepsilon}_{it}^2\hat{\boldsymbol{f}}'\boldsymbol{J}_T\boldsymbol{Z}_\tau\right)\\
&\quad (N^{-1}T^{-2}\boldsymbol{Z}_\tau'\boldsymbol{J}_T\boldsymbol{X}\boldsymbol{J}_N\boldsymbol{X}'\boldsymbol{J}_T\boldsymbol{Z}_\tau)\,(T^{-1}\boldsymbol{Z}_\tau'\boldsymbol{J}_T\boldsymbol{Z}_\tau)\,,
\end{aligned}
$$

*where $\hat{\boldsymbol{f}}$ and $\hat{\boldsymbol{\phi}}_{f,\tau}$ are the relevant factors and loadings estimates obtained by Qcov3PPRF and $\hat{\varepsilon}_t = \boldsymbol{x}_t - \hat{\boldsymbol{\phi}}_{f,\tau}\hat{\boldsymbol{f}}_t$. .*

*Proof.* By Eq.(16), we obtain that (omitting the constant term):

$$
\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau\boldsymbol{F}_t = \hat{\boldsymbol{F}}_A\hat{\boldsymbol{F}}_B^{-1}N^{-1}T^{-1}\boldsymbol{Z}_\tau'\boldsymbol{J}_T\boldsymbol{X}\boldsymbol{J}_N\boldsymbol{\varepsilon}_t.
$$

Then, to find the asymptotic distribution of the term $\hat{\boldsymbol{F}}_t - \boldsymbol{H}_\tau\boldsymbol{F}_t$ we can focus on the term $N^{-1}T^{-1}\boldsymbol{Z}_\tau'\boldsymbol{J}_T\boldsymbol{X}\boldsymbol{J}_N\boldsymbol{\varepsilon}_t$ which results in the following:

$$
\begin{aligned}
N^{-1}T^{-1}\boldsymbol{Z}_\tau'\boldsymbol{J}_T\boldsymbol{X}\boldsymbol{J}_N\boldsymbol{\varepsilon}_t &= \boldsymbol{\Lambda}_\tau(T^{-1}\boldsymbol{F}_{\psi^*}'\boldsymbol{J}_T\boldsymbol{F})(N^{-1}\boldsymbol{\Phi}'\boldsymbol{J}_N\boldsymbol{\varepsilon}_t)\\
&+ \boldsymbol{\Lambda}_\tau(N^{-1}T^{-1}\boldsymbol{F}_{\psi^*}'\boldsymbol{J}_T\boldsymbol{\varepsilon}\boldsymbol{J}_N\boldsymbol{\varepsilon}_t)\\
&+ (T^{-1}\boldsymbol{\omega}'\boldsymbol{J}_T\boldsymbol{F})(N^{-1}\boldsymbol{\Phi}'\boldsymbol{J}_N\boldsymbol{\varepsilon}_t)\\
&+ (N^{-1}T^{-1}\boldsymbol{\omega}'\boldsymbol{J}_T\boldsymbol{\varepsilon}\boldsymbol{J}_N\boldsymbol{\varepsilon}_t)\\
&= \boldsymbol{O}_p(N^{-1/2}) + \boldsymbol{O}_p(\delta_{NT}^{-1}T^{-1/2}) + \boldsymbol{O}_p(N^{-1/2}T^{-1/2}) + \boldsymbol{O}_p(\delta_{NT}^{-1}T^{-1/2})\\
&= \boldsymbol{O}_p(N^{-1/2}) + \boldsymbol{O}_p(\delta_{NT}^{-1}T^{-1/2}),
\end{aligned}
$$

where the third equality follows from Lemma 1 and Lemma 2 in Kelly and Pruitt, 2015. Case (i) follows if the first term is the dominant, i.e., $\sqrt{N}/T \to 0$ and applying the continuous mapping theorem. When $\liminf \sqrt{N}/T \geq \kappa > 0$, we have $N^{-1}T^{-1}\boldsymbol{Z}_\tau'\boldsymbol{J}_T\boldsymbol{X}\boldsymbol{J}_N\boldsymbol{\varepsilon}_t = \boldsymbol{O}_p(1)$ by the continuous mapping theorem. The expression for the asymptotic covariance matrix plim $\boldsymbol{\Sigma}_{\hat{f}}$ is the asymptotic limit of the covariance estimator $\hat{\Sigma}_f$ whose consistency can be verified by Lemma 1, Lemma 2 in Kelly and Pruitt, 2015 and the asymptotic limits derived in Lemma 2. $\square$

***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:

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

Since,

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

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 3.1 and Lemma 2. 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 Cauchy-Schwartz inequality, Assumption 2, Assumption 3.1, Lemma 3 and the asymptotic normality of linear estimates in quantile regression. Then, by the continuous mapping theorem, we have:

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

Since

$$\begin{aligned}
\sqrt{N}(\hat{\boldsymbol{f}}_t-\boldsymbol{H}_\tau\boldsymbol{f}_t) &\xrightarrow[N,T\to\infty]{d} \mathcal{N}(\mathbf{0},\operatorname{plim}\boldsymbol{\Sigma}_{\hat{f}}),\\
\sqrt{T}(\hat{\boldsymbol{\beta}}_\tau-\boldsymbol{\beta}'_{f,\tau}\boldsymbol{H}_\tau^{-1}) &\xrightarrow[N,T\to\infty]{d} \mathcal{N}(\mathbf{0},\operatorname{plim}\boldsymbol{\Sigma}_{\hat{\beta}}),
\end{aligned}$$

by Lemma 3.i and Eq.(A.21), where $\boldsymbol{\Sigma}_{\hat{f}_t}$ is defined in Lemma 3 and $\operatorname{plim}\boldsymbol{\Sigma}_{\hat{\beta}}=\mathbb{E}[h_\tau(0|\boldsymbol{f}_t)\boldsymbol{f}_t\boldsymbol{f}'_t]^{-1}\mathbb{E}[\kappa_\tau(u_{t+1}^2)\boldsymbol{f}_t\boldsymbol{f}'_t]\mathbb{E}[h_\tau(0|\boldsymbol{f}_t)\boldsymbol{f}_t\boldsymbol{f}'_t]^{-1}=\boldsymbol{\Delta}_{f,h}^{-1}\mathbb{E}[\kappa_\tau(u_{t+1}^2)\boldsymbol{f}_t\boldsymbol{f}'_t]\boldsymbol{\Delta}_{f,h}^{-1}$ which is feasible by Assumptions 2, 3.1, 3.5 and 3.6. Since $\sqrt{N}(\hat{\boldsymbol{f}}_t-\boldsymbol{H}_\tau\boldsymbol{f}_t)$ depend on $\boldsymbol{\varepsilon}_{i,t}$, and $\sqrt{T}(\hat{\boldsymbol{\beta}}_\tau-\boldsymbol{\beta}_{f,\tau})$ depend on $u_{t+1}$, both terms are independent. It follows that $\hat{\boldsymbol{\beta}}'_\tau\hat{\boldsymbol{f}}_t-\boldsymbol{\beta}'_{f,\tau}\boldsymbol{f}_t$ is asymptotically 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 $\hat{\boldsymbol{\beta}}_\tau=\boldsymbol{\beta}_{f,\tau}+o_p(1)$. □

The following lemmma states that, in a univariate quantile regression, the *qcov* between the target variable and the error term is never zero and it is a function of the optimal linear coefficient. This lemma is a key result to justify the necessity of generating exactly $K_f^*$ automatic-proxies in Qcov3PRF instead of only $K_f$ proxies as we show in the proof of Theorem 3.

**Lemma 4.** *Consider the quantile regression $y = b_{0\tau} + b_\tau f + u^\tau$ and let $(b_{0\tau}^*, b_\tau^*) = \arg\min_{(b_{0\tau},b_\tau)} \mathbb{E}\left[\rho_\tau(y - b_{0\tau} - b_\tau f)\right]$. Suppose that $f$ and $u$ have a joint density and $\mathbb{E}[f^2] < \infty$. Then, $qcov_\tau(y,u) = \nu_u(b_\tau^*)$ where $\nu_u(\cdot)$ is a continuous function and $\nu_u(b_\tau^*) \neq 0$. If $b_{0\tau} = 0$ and $b_\tau > 0$, then $\nu_u(\cdot)$ is a decreasing function. If $b_{0\tau} = 0$ and $b_\tau < 0$, then $\nu_u(\cdot)$ is an increasing function.*

*Proof.* For $a, b, c \in \mathbb{R}$, denote the function $h(a,b,c)$ as follows:

$$\begin{aligned}
h(a,b,c) &= \mathbb{E}[\rho_\tau(cu - b_{0\tau} - b_\tau f)] \\
&= \mathbb{E}\left[(cu - b_{0\tau} - b_\tau f)\left(\tau - I(cu < b_{0\tau} + b_\tau f)\right)\right] \\
&= \tau \int_{-\infty}^{\infty} (cu - b_{0\tau} - b_\tau f) h(u) du - \int_{-\infty}^{b_{0\tau}+b_\tau f} (cu - b_{0\tau} - b_\tau f) h(u) du \\
&= (\tau - 1) \int_{-\infty}^{b_{0\tau}+b_\tau f} (cu - b_{0\tau} - b_\tau f) h(u) du + \tau \int_{b_{0\tau}+b_\tau f}^{\infty} (cu - b_{0\tau} - b_\tau f) h(u) du.
\end{aligned}$$

It is known that $h(b_{0\tau}, b_\tau, c)$ is a convex function in $b_{0\tau}$ and $b$ with $\lim_{b_{0\tau}^2 + b^2 \to \infty} h(b_{0\tau}, b, c) = +\infty$. Applying the Leibniz rule we obtain the first order derivatives of $h(b_{0\tau}, b_\tau, c)$ with respect to $b_{0\tau}$, $b_\tau$ and $c$, respectively as follows:

$$\begin{aligned}
\frac{\partial h(b_{0\tau}, b_\tau, c)}{\partial b_{0\tau}} &= -(\tau - 1) \int_{-\infty}^{b_{0\tau}+b_\tau f} h(cu) du - \tau \int_{b_{0\tau}+b_\tau f}^{\infty} h(cu) du \\
&= -\mathbb{E}[\tau - I(cu < b_{0\tau} + b_\tau f)] \\
&= -\mathbb{E}[\kappa_\tau(c - b_{0\tau} - b_\tau f)], \qquad \text{(A.23)}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial h(b_{0\tau}, b_\tau, c)}{\partial b_\tau} &= -(\tau-1)\int_{-\infty}^{b_{0\tau}+b_\tau f} f\cdot h(cu)du - \tau\int_{b_{0\tau}+b_\tau f}^{\infty} f\cdot h(cu)du \\
&= -\mathbb{E}[(\tau - I(cu < b_{0\tau}+b_\tau f))\, f] \\
&= -\mathbb{E}[\kappa_\tau(c - b_{0\tau} - b_\tau f)f], \qquad \text{(A.24)}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial h(b_{0\tau}, b_\tau, c)}{\partial c} &= (\tau-1)\int_{-\infty}^{b_{0\tau}+b_\tau f} u\cdot h(cu)du + \tau\int_{b_{0\tau}+b_\tau f}^{\infty} u\cdot h(cu)du \\
&= \mathbb{E}[(\tau - I(cu < b_{0\tau}+b_\tau f))\, u] \\
&= \mathbb{E}[\kappa_\tau(cu - b_{0\tau} - b_\tau f)u] = h_c(b_{0\tau}, b_\tau, c). \qquad \text{(A.25)}
\end{aligned}$$

All these expressions are continuous functions by the Cauchy-Schwartz inequality and the continuity of $f$ and $u$.Let $c = 1$. By setting Eqs.(A.23) and (A.24) to zero we obtain the optimal values of $b_{0\tau}$ and $b_\tau$ denoted by $Q^\tau_{u-b^*_\tau f}$ and $b^*_\tau$, respectively. Then, we have:

$$h_c(Q^\tau_{u-b^*_\tau f}, b^*_\tau, 1) = \mathbb{E}[\kappa_\tau(u - Q^\tau_{u-b^*_\tau f} - b^*_\tau f)u] = \nu_u(b^*_\tau),$$

which is a continuous function by the convexity of $h(Q^\tau_{u-b^*_\tau f}, b^*_\tau, 1)$ (as $h(b_{0\tau}, b_\tau, 1)$ is a convex function with respect to $b_{0\tau}$ and , $b_\tau$), and it is never zero unless $\tau\int_{Q^\tau_{u-b^*_\tau f}+b^*_\tau f}^{\infty} u\cdot h(u)du = (1-\tau)\int_{-\infty}^{Q^\tau_{u-b^*_\tau f}+b^*_\tau f} u\cdot h(u)du$, which is impossible by the continuity of $u$ and Assumption 2. On the other hand, taking the derivative of $h_c$ with respect to $b_\tau$ we get:

$$\frac{\partial h_c(b_{0\tau}, b_\tau, 1)}{\partial b_\tau} = -f(b_{0\tau}+b_\tau f)h(b_{0\tau}+b_\tau f),$$

this derivative is negative when $b_{0\tau} = 0$ and $b_\tau < 0$, implying that $\nu_u(b_\tau)$ is a decreasing function, and it is positive when $b_{0\tau} = 0$ and $b_\tau < 0$, implying that $\nu_u(b_\tau)$ is an increasing function. □

The next lemma is useful to show that the automatic-proxies obtained by Qcov3PRF do not depend on irrelevant factors and, therefore, these proxies are only spanned by the set of $K_f^*$ relevant factors.

**Lemma 5.** *Consider the quantile model* $y = \beta_{0\tau}^* + \beta_{1,\tau}^* f_1 + \cdots + \beta_{K,\tau}^* f_K + u_\beta^\tau$, *and* $\mathbb{E}[\tau - I(u^\tau < 0)|y, f_1, ..., f_K] = 0$. *Assume that the conditional density* $h_\tau(u|\boldsymbol{f})$ *exists. Also, consider a quantile regression of* $y$ *on* $f_{k_1}, \cdots, f_{k_f}$, *where* $k_1, \cdots, k_f \in \{1, \ldots, K\}$ *and* $f \leq K$, *i.e.,*

*let* $(b_{0\tau}^*, b_{k_1,\tau}^*, \ldots, b_{k_f,\tau}^*) = \arg\min_{(b_{0\tau}, b_{k_1,\tau}, \ldots, b_{k_f,\tau})} \mathbb{E}\left[\rho_\tau(y - b_{0\tau} - b_{k_1,\tau} f_{k_1} - \cdots - b_{k_f,\tau} f_{k_f})\right]$. *If* $\beta_{j,\tau}^* = 0$ *and* $f_j$ *is independent of* $\boldsymbol{f}_{-j} = (f_{k_1}, \ldots, f_{j-1}, f_{j+1}, \ldots, f_{k_f})$ , *then* $b_{j,\tau}^* = 0$. *Moreover, if* $\beta_{j,\tau}^* \neq 0$ *then* $b_{j,\tau}^* = \beta_{j,\tau}^* + d_{j,\tau}$, *where* $d_{j,\tau}$ *does not depend on those predictors* $f_k$*'s such that* $\beta_{k,\tau}^* = 0$.

*Proof.* Without loss of generality, we assume that $\mathbb{E}(f_j) = 0$ and $\mathrm{Var}(f_j) = 1$ for $j = 1, \cdots, K$. Following the procedure presented in Theorem 2 of Angrist et al., 2006, we obtain the coefficients $(b_{0\tau}^*, b_{k_1,\tau}^*, \ldots, b_{k_f,\tau}^*)$ which solve the following optimization problem:

$$(b_{0\tau}^*, b_{k_1,\tau}^*, \ldots, b_{k_f,\tau}^*) = \underset{(b_{0\tau}^*, b_{k_1,\tau}^*, \ldots, b_{k_f,\tau}^*)}{\arg\min} \mathbb{E}\left[c(\boldsymbol{f}) \cdot \left(Q_{y|\boldsymbol{f}}^\tau - b_{0\tau} - b_{k_1,\tau} f_{k_1} - \cdots - b_{k_f,\tau} f_{k_f}\right)^2\right],$$

where $Q_{y|\boldsymbol{f}}^\tau = \beta_{0\tau}^* + \beta_{1,\tau}^* f_1 + \cdots + \beta_{K,\tau}^* f_K$, $c(\boldsymbol{f}) = \frac{1}{2}\int_0^1 h_\tau(\varepsilon \cdot \boldsymbol{\Delta}_\tau|\boldsymbol{f}) d\varepsilon$, with $\boldsymbol{\Delta}_\tau = b_{0\tau} + b_{k_1,\tau} f_{k_1} + \cdots + b_{k_f,\tau} f_{k_f} - Q_{y|\boldsymbol{f}}^\tau$. As a result, for $j = k_1, \ldots, k_f$ we have:

$$b_{j,\tau}^* = \{\mathbb{E}(c(\boldsymbol{f}) f_j \boldsymbol{M}_{-j} f_j)\}^{-1} \mathbb{E}(c(\boldsymbol{f}) f_j \boldsymbol{M}_{-j} Q_{y|\boldsymbol{f}}^\tau) = \beta_{j,\tau}^* + d_{j,\tau}, \tag{A.26}$$

where $\boldsymbol{M}_{-j} = 1 - \boldsymbol{f}'_{-j}(\boldsymbol{f}_{-j}\boldsymbol{f}'_{-j})\boldsymbol{f}_{-j}$, and

$$d_{j,\tau} = \sum_{k \neq j} \beta_{k,\tau}^* \{\mathbb{E}(c(\boldsymbol{f}) f_j \boldsymbol{M}_{-j} f_j)\}^{-1} \mathbb{E}(c(\boldsymbol{f}) f_j \boldsymbol{M}_{-j} f_k). \tag{A.27}$$

If $\beta^*_{j,\tau} = 0$ for some $f_j$ independent of $\boldsymbol{f}_{-j}$ we have that:

$$b^*_{j,\tau} = \sum_{k \neq j} \beta^*_{k,\tau} \{\mathbb{E}(c(\boldsymbol{f}) f_j \boldsymbol{M}_{-j} f_j)\}^{-1} \mathbb{E}(c(\boldsymbol{f}) f_j \boldsymbol{M}_{-j} f_k),$$

by Eqs.(A.26) and (A.27). The solution of this equation is $b^*_{j,\tau} = 0$ as in that case the term $c(\boldsymbol{f}) = c(\boldsymbol{f}_{-j})$ and since $f_j$ is independent of $\boldsymbol{f}_{-j}$ each term $\mathbb{E}(c(\boldsymbol{f}_j) f_j \boldsymbol{M}_{-j} f_k) = \mathbb{E}(f_j)\mathbb{E}(c(\boldsymbol{f}_{-j})\boldsymbol{M}_{-j} f_k)$ is equal to zero. Moreover, for any $\beta^*_{k,\tau} \neq 0$ we have $b^*_{k,\tau} = \beta^*_{k,\tau} + d_{k,\tau}$ with $d_{k,\tau}$ independent of those predictors $f_j's$ such that $\beta^*_{j,\tau} = 0$ given that $c(\boldsymbol{f}) = c(\boldsymbol{f}_{-j})$. ☐

***Proof of Theorem 3.*** We begin by showing that Assumption 1.2 is generally satisfied, this is that the loadings of the automatic proxies are uncorrelated with irrelevant factors and are linearly independent (full rank matrix). If $K^*_f = 1$, it is clearly seen that $\hat{\boldsymbol{z}}_{1,\tau} = I(\boldsymbol{y} > \hat{Q}^\tau_y)$ is not correlated with irrelevant factors as $\text{qcov}(\boldsymbol{y}, \boldsymbol{f}_1) = \nu(\beta_{1,\tau}) \neq 0$ and $\text{qcov}(\boldsymbol{y}, \boldsymbol{g}_j) = \nu(0) = 0$ by Lemma 1 in Li et al., 2015. For $K^*_f > 1$, let $\boldsymbol{y} = \beta_{0,\tau}\boldsymbol{\iota} + \beta_{1,\tau}\boldsymbol{f}_1 + \cdots + \beta_{K^*_f,\tau}\boldsymbol{f}_{K^*_f} + \boldsymbol{u}^\tau$, with $\boldsymbol{u}^\tau = \text{diag}(\boldsymbol{\xi}^\tau)(\sigma_u\boldsymbol{\iota} + \gamma_1\boldsymbol{f}_1 + \cdots + \gamma_{K^*_f}\boldsymbol{f}_{K^*_f})$. Without loss of generality we rule out the case where $\beta_{j,\tau} = 0$ but $\gamma_j \neq 0$ for some $j = 1, \ldots, K^*_f$ as we can always sum and extract some term $\beta_{0j,\tau}\boldsymbol{f}_j$ with $\beta_{0j,\tau} \neq 0$ and redefine the corresponding random term as $\boldsymbol{\xi}^\tau_j = \boldsymbol{\xi}^\tau - \frac{\beta_{0j,\tau}}{\gamma_j}\boldsymbol{\iota}$. Nevertheless, we allow the case where $\beta_{j,\tau} \neq 0$ but $\gamma_j = 0$. The prediction equation presented above is *not* a univariate regression, however, after some rearrangements we obtain:

$$\boldsymbol{y} = \beta_{0,\tau}\boldsymbol{\iota} + \check{\beta}_\tau \check{\boldsymbol{f}} + \boldsymbol{u}^\tau, \tag{A.28}$$

where $\check{\beta}_\tau \neq 0$, $\check{\boldsymbol{f}} = \frac{\beta_{1,\tau}}{\check{\beta}_\tau}\boldsymbol{f}_1 + \cdots + \frac{\beta_{K^*_f,\tau}}{\check{\beta}_\tau}\boldsymbol{f}_{K^*_f}$. For the first automatic-proxy, $k = 1$, since $\text{qcov}_\tau(\boldsymbol{y}, \check{\boldsymbol{f}}) = \nu_1(\check{\beta}_\tau) \neq 0$ by Lemma 1 in Li et al., 2015, and $\text{qcov}_\tau(\boldsymbol{y}, \boldsymbol{u}^\tau) = \nu_{1,u}(\check{\beta}_\tau) \neq 0$

by Lemma 4, we have:

$$\begin{aligned}
\hat{\boldsymbol{z}}_{1,\tau} &= I(\boldsymbol{y} > \hat{Q}^{\tau}_{y}) = \alpha_{0,\tau}\boldsymbol{\iota} + \bar{\nu}_1(\check{\beta}_\tau)\check{\boldsymbol{f}} + \bar{\nu}_{1,u}(\check{\beta}_\tau)\boldsymbol{u}^{\tau} + \bar{\boldsymbol{u}}^{(1)}, \\
&= \alpha_{0,\tau}\boldsymbol{\iota} + (\bar{\beta}_{1,\tau}\boldsymbol{\iota} + \bar{\gamma}_{1,\tau}\mathrm{diag}(\boldsymbol{\xi}^{\tau}))\boldsymbol{f}_1 + \cdots + (\bar{\beta}_{K_f^*,\tau}\boldsymbol{\iota} + \bar{\gamma}_{K_f^*,\tau}\mathrm{diag}(\boldsymbol{\xi}^{\tau}))\boldsymbol{f}_{K_f^*} + \bar{\boldsymbol{\omega}}^{(1)}, \quad \text{(A.29)}
\end{aligned}$$

where $\bar{\beta}_{j,\tau} = \bar{\nu}_1(\check{\beta}_\tau)\frac{\beta_{j,\tau}}{\check{\beta}_\tau}$, $\bar{\gamma}_{j,\tau} = \bar{\nu}_{1,u}(\check{\beta}_\tau)\gamma_j$ for $j = 1, \ldots, K_f^*$, and $\bar{\boldsymbol{u}}^{(1)}$ is uncorrelated with $\tilde{\boldsymbol{f}}$ and $\boldsymbol{u}$. Hence, we have: $\mathrm{cov}(\hat{\boldsymbol{z}}_{1,\tau}, \boldsymbol{f}_j | \boldsymbol{f}_{-j}) = \bar{\beta}_{j,\tau} + \bar{\gamma}_{j,\tau}\mathbb{E}(\boldsymbol{\xi}^{\tau})$ and $\hat{\boldsymbol{z}}_{1,\tau}$ is uncorrelated with $\boldsymbol{f}_k$, $k = K_f^* + 1, ..., K$. Let the superscript $(k)$ denotes the forecast using $k$ proxies. When $k = 2$ we have that:

$$\begin{aligned}
\hat{\boldsymbol{z}}_{2,\tau} &= I(\boldsymbol{y} > \hat{Q}^{\tau}_{y|\hat{\boldsymbol{f}}_1^{(1)}}) \\
&= I(\beta_{0,\tau}\boldsymbol{\iota} + \boldsymbol{F}\boldsymbol{\beta}_\tau + \boldsymbol{u}^{\tau} > \tilde{\beta}^{(1)}_{0\tau}\boldsymbol{\iota} + \tilde{\beta}^{(1)}_{1,\tau}\hat{\boldsymbol{f}}_1^{(1)}) \\
&= I(\beta_{0,\tau}\boldsymbol{\iota} + \boldsymbol{F}\boldsymbol{\beta}_\tau + \boldsymbol{u}^{\tau} > \tilde{\beta}^{(1)}_{0\tau}\boldsymbol{\iota} + \boldsymbol{X}\boldsymbol{\Omega}^{(1)}_\tau\tilde{\beta}^{(1)}_{1,\tau}) \\
&= I(\beta_{0,\tau}\boldsymbol{\iota} + \boldsymbol{F}\boldsymbol{\beta}_\tau + \boldsymbol{u}^{\tau} > \tilde{\beta}^{(1)}_{0\tau}\boldsymbol{\iota} + \boldsymbol{X}\boldsymbol{\Omega}^{(1)}_\tau(\tilde{\boldsymbol{H}}_\tau^{(1)-1}\beta^{(1)}_{1,\tau} + o_p(1)) \\
&= I\left(\beta_{0,\tau}\boldsymbol{\iota} + \boldsymbol{F}\boldsymbol{\beta}_\tau + \boldsymbol{u}^{\tau} > \tilde{\beta}^{(1)}_{0\tau}\boldsymbol{\iota} + \boldsymbol{X}\boldsymbol{\Omega}^{(1)}_\tau\left((\boldsymbol{H}_\tau^{-1}\boldsymbol{\beta}_\tau)^{(1)} + d_\tau^{(1)} + o_p(1)\right)\right) \\
&= I\Bigg(\beta_{0,\tau}\boldsymbol{\iota} + \boldsymbol{F}\boldsymbol{\beta}_\tau + \boldsymbol{u}^{\tau} > \tilde{\beta}^{(1)}_{0\tau}\boldsymbol{\iota} + \boldsymbol{F}\boldsymbol{\Phi}^{-1}\boldsymbol{\Omega}^{(1)}_\tau\left((\boldsymbol{H}_\tau^{-1}\boldsymbol{\beta}_\tau)^{(1)} + d_\tau^{(1)} + o_p(1)\right) \\
&\quad +\boldsymbol{\varepsilon}\boldsymbol{\Omega}^{(1)}_\tau\left((\boldsymbol{H}_\tau^{-1}\boldsymbol{\beta}_\tau)^{(1)} + d_\tau^{(1)} + o_p(1)\right)\Bigg) \\
&= I\Bigg(\beta_{0,\tau}\boldsymbol{\iota} - \tilde{\beta}^{(1)}_{0\tau}\boldsymbol{\iota} + \boldsymbol{F}\left(\boldsymbol{\beta}_\tau - \boldsymbol{\Phi}^{-1}\boldsymbol{\Omega}^{(1)}_\tau\left((\boldsymbol{H}_\tau^{-1}\boldsymbol{\beta}_\tau)^{(1)} + d_\tau^{(1)} + o_p(1)\right)\right) \\
&\quad +\boldsymbol{u}^{\tau} - \boldsymbol{\varepsilon}\boldsymbol{\Omega}^{(1)}_\tau\left((\boldsymbol{H}_\tau^{-1}\boldsymbol{\beta}_\tau)^{(1)} + d_\tau^{(1)} + o_p(1)\right) > 0\Bigg) \\
&= I(\beta^{(2)}_{0,\tau}\boldsymbol{\iota} + \boldsymbol{F}\boldsymbol{\beta}^{(2)}_\tau + \boldsymbol{u}^{(2)}_\tau > 0) \qquad \text{(A.30)} \\
&= I(\beta^{(2)}_{0,\tau}\boldsymbol{\iota} + \check{\beta}^{(2)}_\tau\check{\boldsymbol{f}}^{(2)} + \boldsymbol{u}^{(2)}_\tau > 0) \\
&= \alpha^{(2)}_{0,\tau}\boldsymbol{\iota} + \bar{\nu}_2(\check{\beta}^{(2)}_\tau)\tilde{\boldsymbol{f}}^{(2)} + \bar{\nu}_{2,u}(\check{\beta}^{(2)}_\tau)\boldsymbol{u}^{(2)}_\tau + \bar{\boldsymbol{u}}^{(2)}, \\
&= \alpha^{(2)}_{0,\tau}\boldsymbol{\iota} + (\bar{\beta}^{(2)}_{1,\tau}\boldsymbol{I} + \bar{\gamma}^{(2)}_{1,\tau}\mathrm{diag}(\boldsymbol{\xi}^{\tau}))\boldsymbol{f}_1 + \cdots + (\bar{\beta}^{(2)}_{K_f^*,\tau}\boldsymbol{I} + \bar{\gamma}^{(2)}_{K_f^*,\tau}\mathrm{diag}(\boldsymbol{\xi}^{\tau}))\boldsymbol{f}_{K_f^*} + \bar{\boldsymbol{\omega}}^{(2)}, \quad \text{(A.31)}
\end{aligned}$$

where $\bar{\beta}^{(2)}_{j,\tau} = \bar{\nu}_2(\check{\beta}^{(2)}_\tau)\frac{\beta_{j,\tau}}{\check{\beta}^{(2)}_\tau}$, $\bar{\gamma}^{(2)}_{j,\tau} = \bar{\nu}_{2,u}(\check{\beta}^{(2)}_\tau)\gamma_j$ for $j = 1, \ldots, K_f^*$, $\check{\boldsymbol{f}}^{(2)} = \frac{\beta_{1,\tau}}{\check{\beta}^{(2)}_\tau}\boldsymbol{f}_1 + \cdots + \frac{\beta_{K_f^*,\tau}}{\check{\beta}^{(2)}_\tau}\boldsymbol{f}_{K_f^*}$, and $\bar{\boldsymbol{u}}^{(2)}$ is uncorrelated with $\tilde{\boldsymbol{f}}^{(2)}$ and $\boldsymbol{u}^{(2)}_\tau$. In the third equality, $\boldsymbol{\Omega}^{(1)}_\tau = \boldsymbol{J}_N\boldsymbol{W}^{(1)}_{xr}(\boldsymbol{W}^{(1)\prime}_{xr}\boldsymbol{J}_N\boldsymbol{W}^{(1)}_{xr})^{-1}\boldsymbol{W}^{(1)}_{rr}$, $\boldsymbol{W}^{(1)}_{xr} = \boldsymbol{X}'\boldsymbol{J}_T\hat{\boldsymbol{z}}_{1,\tau}$, $\boldsymbol{W}^{(1)}_{rr} = \hat{\boldsymbol{z}}'_{1,\tau}\boldsymbol{J}_T\hat{\boldsymbol{z}}_{1,\tau}$. The fourth inequality holds as $\tilde{\beta}^{(1)}_{1,\tau} \xrightarrow[N,T\to\infty]{p} \tilde{\boldsymbol{H}}_\tau^{(1)-1}\beta^{(1)}_{1,\tau}$, where the asymptotic limit is the one obtained by running a $\tau$-quantile regresssion of $\boldsymbol{y}$ on $\boldsymbol{f}^{(1)}$, where $\hat{\boldsymbol{f}}_1^{(1)} \xrightarrow[N,T\to\infty]{p} \boldsymbol{f}_1^{(1)}\tilde{\boldsymbol{H}}_\tau^{(1)}$. The fifth inequality holds by Eq.(A.26), a result obtained in Angrist et al., 2006. Also, the superscript (1) in $(\boldsymbol{H}_\tau^{-1}\boldsymbol{\beta}_\tau)^{(1)}$ indicates one element of the corresponding vector. In Eq.(A.30), $\boldsymbol{\beta}_\tau^{(2)} = \boldsymbol{\beta}_\tau - \boldsymbol{\Phi}^{-1}\boldsymbol{\Omega}_\tau^{(1)}\Big((\boldsymbol{\beta}_\tau\boldsymbol{H}_\tau^{-1})^{(1)} + d_\tau^{(1)} + o_p(1)\Big)$, $\beta_{0,\tau}^{(2)} = \beta_{0,\tau} - \tilde{\beta}_{0\tau}^{(1)}$ and $\boldsymbol{u}_\tau^{(2)} = \boldsymbol{u}^\tau - \boldsymbol{\varepsilon}\boldsymbol{\Omega}_\tau^{(1)}\Big((\boldsymbol{\beta}_\tau\boldsymbol{H}_\tau^{-1})^{(1)} + d_\tau^{(1)} + o_p(1)\Big)$. Also, from Eq.(A.30) we can verify that $\boldsymbol{\beta}_\tau^{(2)}$ and $\boldsymbol{u}_\tau^{(2)}$ only concern the relevant factors as $\boldsymbol{\Omega}_\tau^{(1)}$ is constructed using $\hat{\boldsymbol{z}}_{1,\tau}$ automatic-proxy and $d_\tau^{(1)}$ only depends on the relevant factors by Lemma 5. In the ninth equality we have $\check{\boldsymbol{f}}^{(2)} = \frac{\beta_{1,\tau}}{\check{\beta}_\tau^{(2)}}\boldsymbol{f}_1 + \cdots + \frac{\beta_{K_f^*,\tau}}{\check{\beta}_\tau^{(2)}}\boldsymbol{f}_{K_f^*}$. In the tenth equality we use $\text{qcov}_\tau(\boldsymbol{y}, \check{\boldsymbol{f}}^{(2)}) = \nu_2(\check{\beta}_\tau^{(2)}) \neq 0$ by Lemma 1 in Li et al., 2015, and $\text{qcov}_\tau(\boldsymbol{y}, \boldsymbol{u}_\tau^{(2)}) = \nu_{2,u}(\check{\beta}_\tau^{(2)}) \neq 0$ by Lemma 4. Threfore, based on Eq.(A.31), it is clear that the loadings of $\hat{\boldsymbol{z}}_{2,\tau}$ corresponding to the irrelevant factors are equal to zero.

The induction step proceeds as follows. By hypothesis, suppose we have $k < K_f^*$ automatic-proxies with factor loadings $[\boldsymbol{\Lambda}_{f,\tau}^k, \mathbf{0}]$, where $\boldsymbol{\Lambda}_{f,\tau}^k$ is a $k \times K_f^*$ matrix and full row rank. The $k$-automatic-proxy obtained with Qcov3PRF is $\hat{\boldsymbol{z}}_{k,\tau} = I(\boldsymbol{y} - \tilde{\beta}_{0,\tau}^{(k-1)}\boldsymbol{\iota} - \tilde{\beta}_{1,\tau}^{(k-1)}\hat{\boldsymbol{f}}_1^{(k-1)} - \cdots - -\tilde{\beta}_{k-1,\tau}^{(k-1)}\hat{\boldsymbol{f}}_{k-1}^{(k-1)} > 0)$, which has zero population covariance with irrelevant factors by the same argument given in the $k = 2$ case.

The second thing to show is that the proxy $\hat{\boldsymbol{z}}_{k,\tau}$ vector of loadings on relevant factors is linearly independent of the rows of $\boldsymbol{\Lambda}_{f,\tau}^{(k)}$. First, note that following similar steps to those used to obtain Eq.(A.31) we get that $\hat{z}_{k,t}^\tau = \alpha_{0,\tau}^{(k)} + \boldsymbol{\Psi}_{f,t}^{(k)}\bar{\boldsymbol{\beta}}_\tau^{(k)}\boldsymbol{f}_t + \bar{\omega}_t^{(k)}$, where $\boldsymbol{\Psi}_{f,t}^{(k)} = (1 + \frac{\bar{\gamma}_{1,\tau}^{(k)}}{\bar{\beta}_{1,\tau}^{(k)}}\xi_{t+1}^\tau, \ldots, 1 + \frac{\bar{\gamma}_{K_f^*,\tau}^{(k)}}{\bar{\beta}_{K_f^*,\tau}^{(k)}}\xi_{t+1}^\tau)$, the relevant-factor loadings are given by $\boldsymbol{\Lambda}_{f,\tau}^k = \bar{\boldsymbol{\beta}}_{f,\tau}^{(k)} = \boldsymbol{\beta}_{f,\tau} - \boldsymbol{\Phi}^{-1}\boldsymbol{\Omega}_\tau^{(k)}\Big((\boldsymbol{H}_{f,\tau}^{-1}\boldsymbol{\beta}_{f,\tau})^{(k)} + \boldsymbol{d}_\tau^{(k)} + o_p(1)\Big)$, and the superscript $(k)$ in $(\boldsymbol{H}_{f,\tau}^{-1}\boldsymbol{\beta}_{f,\tau})^{(k)}$ indicates some $k$ elements of the corresponding vector. Note that $(\boldsymbol{F}\boldsymbol{H}_\tau)^{(k)} = [(\boldsymbol{f}\tilde{\boldsymbol{H}}_{f,\tau})^{(k)}, \mathbf{0}]$ and $(\boldsymbol{H}_\tau^{-1}\boldsymbol{\beta}_\tau)^{(k)} = [(\boldsymbol{H}_{f,\tau}^{-1}\boldsymbol{\beta}_{f,\tau})^{(k)}, \mathbf{0}]$. Also note that as part of the induction hypothesis, $\boldsymbol{\Omega}_\tau^{(k)}$ is constructed based on $(\hat{\boldsymbol{z}}_{1,\tau}, \ldots, \hat{\boldsymbol{z}}_{k,\tau})$.

Now, using a similar argument done in the proof for Theorem 7 in the Appendix in Kelly and Pruitt, 2015, project $\hat{z}_{k,t}^\tau$'s relevant-factor loadings onto the column space of $\boldsymbol{\Lambda}_{f,\tau}^{(k-1)'}$. The residual's loading vectors are linearly independent of $\boldsymbol{\Lambda}_{f,\tau}^{(k-1)'}$ if the difference between it and its projection on $\boldsymbol{\Lambda}_{f,\tau}^{(k-1)'}$ is non-zero. This difference is then equal to $\left(\boldsymbol{I} - \boldsymbol{\Lambda}_{f,\tau}^{(k-1)'}(\boldsymbol{\Lambda}_{f,\tau}^{(k-1)}\boldsymbol{\Lambda}_{f,\tau}^{(k-1)'})^{-1}\boldsymbol{\Lambda}_{f,\tau}^{(k-1)}\right)\bar{\boldsymbol{\beta}}_{f,\tau}^{(k)}$. Because $\bar{\boldsymbol{\beta}}_{f,\tau}^{(k)} \neq \boldsymbol{0}$ with probability one (as $\breve{\boldsymbol{\beta}}_{\tau}^{(k)}$ is different from zero), this difference is zero only when $\boldsymbol{\Lambda}_{f,\tau}^{(k-1)'}(\boldsymbol{\Lambda}_{f,\tau}^{(k-1)}\boldsymbol{\Lambda}_{f,\tau}^{(k-1)'})^{-1}\boldsymbol{\Lambda}_{f,\tau}^{(k-1)} = \boldsymbol{I}$. But the induction hypothesis ensures that this is not the case as long as $k < K_f^*$. Therefore, the difference between the $k$th automatic proxy's loading vector and its projection onto the column space of $\boldsymbol{\Lambda}_{f,\tau}^{(k-1)'}$ is nonzero, which implies that its loading vector is linearly independent of the rows of $\boldsymbol{\Lambda}_{f,\tau}^{(k-1)}$. Therefore, we have shown that the automatic-proxies obtained by Qcov3PRF satisfy Assumption 1.2.

Finally, it is left to check that the automatic proxies satisfy Assumptions 1.3, 3.4, 4.3, 4.4 and 4.5 when the remaining parts of Assumptions 1–5 hold. Recall that $\hat{z}_{k,t}^{\tau} = \alpha_{0,\tau}^{(k)} + \boldsymbol{\Psi}_{f,t}^{(k)}\bar{\boldsymbol{\beta}}_{\tau}^{(k)}\boldsymbol{f}_t + \bar{\omega}_t^{(k)}$, where $\boldsymbol{\Psi}_{f,t}^{(k)} = (1 + \frac{\bar{\gamma}_{1,\tau}^{(K)}}{\bar{\beta}_{1,\tau}^{(k)}}\xi_{t+1}^{\tau}, \ldots, 1 + \frac{\bar{\gamma}_{K_f^*,\tau}^{(k)}}{\bar{\beta}_{K_f^*,\tau}^{(k)}}\xi_{t+1}^{\tau})$, $\bar{\omega}_t^{(k)} = \bar{u}_t^{(k)} + \bar{\nu}_{k,u}(\breve{\beta}_{\tau}^{(k)})\left(-\boldsymbol{\varepsilon}_t\boldsymbol{\Omega}_{\tau}^{(k-1)}\left((\boldsymbol{H}_{\tau}^{-1}\boldsymbol{\beta}_{\tau})^{(k)} + d_{\tau}^{(k)} + o_p(1)\right) + \sigma_u\xi_{t+1}\right)$, $\boldsymbol{\Omega}_{\tau}^{(k)} = \boldsymbol{J}_N\boldsymbol{W}_{xr}^{(k)}(\boldsymbol{W}_{xr}^{(k)'}\boldsymbol{J}_N\boldsymbol{W}_{xr}^{(k)})^{-1}\boldsymbol{W}_{rr}^{(k)}$, $\boldsymbol{W}_{xr}^{(k)} = \boldsymbol{X}'\boldsymbol{J}_T\hat{\boldsymbol{z}}_{1:k,\tau}$, $\boldsymbol{W}_{rr}^{(k)} = \hat{\boldsymbol{z}}_{1:k,\tau}'\boldsymbol{J}_T\hat{\boldsymbol{z}}_{1:k,\tau}$. Assumption 1.3 is satisfied as $\mathbb{E}[\psi_{f,jt}^{(k)}] = 1 + \frac{\bar{\gamma}_{j,\tau}^{(k)}}{\bar{\beta}_{j,\tau}^{(k)}}\mathbb{E}[\xi_{t+1}^{\tau}] \neq 0$ and $\mathbb{E}[|\psi_{f,jt}^{(k)}|^2] = 1 + 2|\frac{\bar{\gamma}_{j,\tau}^{(k)}}{\bar{\beta}_{j,\tau}^{(k)}}|\mathbb{E}[\xi_{t+1}^{\tau}] + \frac{\bar{\gamma}_{j,\tau}^{(k)2}}{\bar{\beta}_{j,\tau}^{(k)2}}\mathbb{E}[\xi_{t+1}^{\tau 2}] < \infty$, $j = 1, \ldots, K_f^*$. Fourth and fifth expressions in Assumption 3.4 are satisfied by Assumption 1.3 and Assumption 3.6 concerning the stationarity and ergodicity of $y_{t+1}$ (and then of $\xi_{t+1}$).

By similar limiting arguments as those derived in Theorem 1, the error term $\bar{\omega}_t^{(k)}$ can be expressed as $\bar{u}_t^{(k)} + N^{-1}\boldsymbol{a}'\boldsymbol{\varepsilon}_t + b\xi_{t+1}$, where $\boldsymbol{a}$ is an $N \times 1$ vector of order $\boldsymbol{O}_p(1)$ and b is of order $O(1)$. We consider that the term $\bar{u}_t^{(k)}$ is of lower magnitude compared to $\boldsymbol{\varepsilon}_t$ and $\xi_{t+1}$ so we disregard it and avoid to check whether it satisfies the remaining assumptions.[2] Also, by Assumption 3.5, the $\xi_{t+1}$ and $\boldsymbol{\varepsilon}_t$ components of $\bar{\omega}_t^{(k)}$ are independent and can be handled separately. By Assumption 3.6 and Lemma 1.4 in Kelly and Pruitt, 2015, the $\xi_{t+1}$

---

[2] Indeed, it must be the case that $\bar{u}_t^{(k)}$ satisfies the first three expressions in Assumption 3.4 by Assumption 3.6. Assumptions 4.4 and 4.5 are satisfied as $\bar{u}_t^{(k)}$ is uncorrelated with $\boldsymbol{\varepsilon}_t$ and $\xi_{t+1}$. However, we only have that $\boldsymbol{\varepsilon}_t$ is uncorrelated with $\boldsymbol{F}_t$, and then, Assumption 4.3 is not generally satisfied for this term. However, we do not see unrealistic to consider that this term is independent of $\boldsymbol{F}_t$. Alternatively, we can relax Assumption 4.3 by replacing it with $\mathbb{E}|F_t^2(m)\bar{u}_t^{2(k)}| = \mathbb{E}|F_t^2(m)|\mathbb{E}|\bar{u}_t^{2(k)}|$, in this case.

component satisfies the first three expressions in Assumption 3.4. By Assumption 3.5, the $\xi_{t+1}$ term satisfies Assumptions 4.3, 4.4. By Assumption 3.6, the $\xi_{t+1}$ term satisfies the fist expression in Assumption 4.5 given that $\psi_{f,jt}^{(k)} = 1 + c_j \xi_{t+1}$. The second expression in Assumption 4.5 is satisfied by the same arguments.

By Assumption 3.3, Lemma 1.3 in Kelly and Pruitt, 2015, and the stationarity and ergodicity of $\boldsymbol{\varepsilon}_t$ in Assumption 3.6, the $\boldsymbol{\varepsilon}_t$ component satisfies the first three expressions in Assumption 3.4, respectively. By Assumption 3.3, the $\boldsymbol{\varepsilon}_t$ term satisfies Assumption 4.3. Assumption 4.4 holds as the term $\mathbb{E}|N^{-1}T^{-1/2}\sum_{j,t} a_j \varepsilon_{it}\varepsilon_{jt}|^2$ is satisfied by Assumption 4.2. The first expression in Assumption 4.5 is satisfied by the independence between $\xi_{t+1}$ and $\varepsilon_{it}$ (Assumption 3.5).

Together these results imply that the $K_f^*$ automatic-proxies generated by Qcov3PRF satisfy the conditions of Theorems 1 and 2. □

## A.3 How PLS for conditional quantile prediction and Qcov3PRF are related

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 for $j = 1, \ldots, K_f$, denoted as $\hat{\boldsymbol{y}}^{(j)}$, aims to solve the following optimization problem:

$$\boldsymbol{\phi}^{(j)} = \underset{\boldsymbol{\phi}, \|\boldsymbol{\phi}\|=1}{\arg\max} \{\mathrm{cov}(\boldsymbol{X}^{(j-1)}\boldsymbol{\phi}, \boldsymbol{y}^{(j-1)})' \mathrm{cov}(\boldsymbol{X}^{(j-1)}\boldsymbol{\phi}, \boldsymbol{y}^{(j-1)})\}, \tag{A.32}$$

where $\boldsymbol{y} = (y_2, \ldots, y_{T+1})'$, $\boldsymbol{X} = (\boldsymbol{x}_1', \ldots, \boldsymbol{x}_T')$, $\boldsymbol{\phi}$ an $N \times 1$ vector, and $\widehat{\mathrm{cov}}(\boldsymbol{X}^{(j-1)}\boldsymbol{\phi}, \boldsymbol{X}^{(l-1)}\boldsymbol{\phi}^{(l)}) = 0$ for $l = 1, \ldots, j-1$. This restriction is equivalent to normalize the latent factors to be orthogonal following the standard PCA method. $\boldsymbol{X}^{(j-1)}$ denotes the deflated underlying set of predictors, and $\boldsymbol{y}^{(j-1)}$ is the deflated target variable. Specifically, $\boldsymbol{X}^{(j-1)}$ is 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 $j$th iteration in PLS the predictors $\boldsymbol{X}^{(j)}$ are deflated by the $\boldsymbol{f}_j$ factor. The PLS procedure is presented in Algorithm A.1.

**Algorithm A.1** Partial Least Squares (PLS)

```
Let $\boldsymbol{X} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_N)$. Standardize each $\boldsymbol{x}_i$, $i = 1, \dots, N$ to have mean zero and variance one.
Set $\boldsymbol{X}^{(0)} = \boldsymbol{X}$ and $\boldsymbol{y}^{(0)} = \boldsymbol{y}$.
for $j = 1, \dots, K_f$ do
    1. Compute $\phi_i^{(j-1)} = \mathrm{cov}(\boldsymbol{x}_i^{(j-1)}, \boldsymbol{y}^{(j-1)})$. Then, $\boldsymbol{\phi}^{(j-1)} = (\phi_1^{(j-1)}, \dots, \phi_N^{(j-1)'})$.
    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 $\boldsymbol{p}_j = \frac{\boldsymbol{X}^{(j-1)'}\boldsymbol{f}_j}{\boldsymbol{f}_j'\boldsymbol{f}_j}$, and the loading of $\boldsymbol{y}$ as $q_j = \frac{\boldsymbol{y}^{(j-1)'}\boldsymbol{f}_j}{\boldsymbol{f}_j'\boldsymbol{f}_j}$.
    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 $\boldsymbol{y}^{(j)} = \boldsymbol{y}^{(j-1)} - q_j\boldsymbol{f}_j$.
```

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 get consistent forecasts. This is because it is required to have $K_f$ deflated target residuals $\boldsymbol{y}^{(j)}$ that are linearly independent and spanned only by the $K_f$ relevant factors.

The particular way of *deflating* and the use of mean covariance in Step 1 in Algorithm A.1 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[3], 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

---

[3] Regarding Qcov3PRF, the non-orthogonality of the estimated factors is clear by looking at Lemma 2.

objective function (A.32) using *qcov* (Li et al., 2015). The new objective function is given by:

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

where $\text{cov}(\boldsymbol{X}^{(j-1)}\boldsymbol{\phi}_{\tau}, \boldsymbol{X}^{(l-1)}\boldsymbol{\phi}_{\tau}^{(l)}) = 0$ for $l = 1, \ldots, j-1$. Méndez-Civieta et al., 2022 adapt Algorithm A.1 by changing the mean covariance with quantile-covariance in Step 1. 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:

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

where $\boldsymbol{x}_i$ is a $T \times 1$ vector. By looking at fPQR, the objective in each iteration $j$ is deflating $\boldsymbol{X}^{(j)}$ and $\boldsymbol{y}_{\tau}^{(j)}$ by focusing at the measures $\text{cov}(\boldsymbol{X}^{(j)}, \boldsymbol{f}_j)$ and $\text{qcov}_{\tau}(\boldsymbol{y}_{\tau}^{(j)}, \boldsymbol{f}_j)$, respectively.[4] The former covariance still needs to be taken with respect to the mean as restriction $\text{cov}(\boldsymbol{X}^{(j-1)}\boldsymbol{\phi}_{\tau}, \boldsymbol{X}^{(l-1)}\boldsymbol{\phi}_{\tau}^{(l)}) = 0$ continues to hold. However, the latter covariance needs to consider the quantile component of $\boldsymbol{y}_{\tau}^{(j)}$ in its relationship with $\boldsymbol{f}_j$. Then, it is required to consider some partial quantile-covariance measure. In this case, from Li et al., 2015 we consider the partial quantile-covariance (*qpcov*) defined as follows:

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

---

[4] In fact, in each iteration in PLS we have $\text{cov}(\boldsymbol{X}^{(j)}, \boldsymbol{f}_j) = 0$ and $\text{cov}(\boldsymbol{y}^{(j)}, \boldsymbol{f}_j) = 0$. However, in general $\text{qcov}_{\tau}(\boldsymbol{y}_{\tau}^{(j)}, \boldsymbol{f}_j) \neq 0$ even though the factors $\boldsymbol{f}_j$ are orthogonal.

Then, we deflat $\boldsymbol{X}^{(j-1)}$ in the same way as in Step 3 in Algorithm A.1. This is by setting $\boldsymbol{X}^{(j)}$ equal to the residuals obtained by regressing each $\boldsymbol{x}_i^{(j-1)}$ on $\boldsymbol{f}_j$ using least squares. Whereas, regarding $\text{qcov}_\tau(\boldsymbol{y}_\tau^{(j)}, \boldsymbol{f}_j)$, we set $\boldsymbol{y}_\tau^{(j)}$ equal to the residuals obtained by regressing $\boldsymbol{y}_\tau^{(j-1)}$ on $\boldsymbol{f}_j$ using quantile regression. Therefore, a PLS extension for quantile regression incorporates $\text{cov}(\boldsymbol{X}^{(j)}, \boldsymbol{f}_j)$ and $\text{qcov}_\tau(\boldsymbol{y}_\tau^{(j)}, \boldsymbol{f}_j)$. Based on *qcov*, the deflated targets are given by $\boldsymbol{y}_\tau^{(j)} = \boldsymbol{y}_\tau^{(j-1)} - \hat{q}_{j,\tau}\boldsymbol{f}_j$ where $\hat{q}_{j,\tau}$ is the estimate of running a $\tau$-quantile regression of $\boldsymbol{y}_\tau^{(j-1)}$ on $\boldsymbol{f}_j$. Deflating $\boldsymbol{y}_\tau^{(j-1)}$ with quantile regression motivates the automatic proxies generated in Qcov3PRF. Moreover, $K_f^* \geq K$ relevant factors need to be considered by Assumption 1, as it is shown in the main text. In Theorem 3 we show that the generated $K_f^*$ transformed deflated target residuals $I(\boldsymbol{y}_\tau^{(j)} > 0)$ are linearly independent and spanned only by the $K_f^*$ relevant factors.

To illustrate the relationship between Qcov3PRF and PLS (for quantile prediction) with $\boldsymbol{y}_\tau^{(j)}$ generated by virtue of using *qcov* we consider the case where a single predictive index is constructed. Applying PLS (for quantile prediction) we have:

1. Set $\phi_{i,\tau} = \boldsymbol{x}_i'\boldsymbol{y}_\tau$, and $\boldsymbol{\phi}_\tau = (\phi_{1,\tau}, \ldots, \phi_{N,\tau})'$, where $\boldsymbol{y}_\tau = I(\boldsymbol{y} - Q_y^\tau > 0)$.
2. Set $f_t = \boldsymbol{x}_t'\boldsymbol{\phi}_\tau$, and $\boldsymbol{f} = (f_1, \ldots, f_T)'$.
3. The forecast for the conditional $\tau$ quantile of $y_{t+1}$ is obtained from running a quantile regression of $y_{t+1}$ on $f_t$.

This forecast is the same we obtain with Qcov3PRF when $K_f^* = 1$.

## A.4 Additional simulation results

Additional simulation experiments focus on checking 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:

$$
\begin{aligned}
y_{t+1} &= -f_{1,t} - 0.5 f_{2,t} + (0.5 f_{2,t} + f_{3,t})\xi_{t+1}, && \text{(A.33)} \\
\boldsymbol{x}_t &= \boldsymbol{\phi}_f' \boldsymbol{f}_t + \boldsymbol{\phi}_g' \boldsymbol{g}_t + \boldsymbol{\varepsilon}_t, && \text{(A.34)}
\end{aligned}
$$

where $\varepsilon_{it} \sim \mathcal{N}(0,1)$, $\phi_{f,i} \sim \mathcal{N}(0,1)$, $\phi_{g,i} \sim \mathcal{N}(0,1)$, $\xi_t \sim \mathcal{N}(0,1)$ i.i.d. We consider two data generating processes where there are six factors and three are relevant, i.e., $K = 6$ and $K_f^* = 3$. The first alternative (A.34.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 (A.34.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, $\sigma_i$ 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 A.1, for $\tau \in \{0.05, 0.10, 0.25\}$, we see that the MAE and the MSE approach to zero, and the correlation approaches to one as $T$ and $N$ increase for both alternatives in DGP (A.33)-(A.34), 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 A.1, we see that both MAE and 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, MAE, MSE and correlation provide better estimates using Qcov3PRF3 and PCQR6, which indicates the inconsistency of PQR and Qcov3PRF1.[5]

---

[5] Indeed, we know from Kelly and Pruitt, 2015 that the forecast estimates with one relevant factor

Table A.1: In-sample MAE, MSE and correlations with DGP (A.33)-(A.34).

| | | | DGP (A.34.a) | | | | DGP (A.34.b) | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $T = N$ | $\tau$ | 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 |

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#.

## A.5 Additional empirical results

Complementing the empirical results presented in the main text where the empirical results are presented using expanding windows, Table A.2 contains the out-of-sample $R^2_\tau$ with IP growth as the dependent variable for $h = \{3, 12\}$ using rolling windows. We present the results for Qcov3PRF for one and two factors[6] and PCQR for up to three factors. We also report the $R^2_\tau$ from using QR when the predictors are the NFCI,

---

when $K_f \geq 2$ are inconsistent albeit they result in in-sample and out-of-sample $R^2_\tau$ 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.

[6] We do not consider three factors in Qcov3PRF as it faces problems of multicollinearity in some estimations as a consequence of using a small number of observations in rolling windows.

the Risk, Credit, Leverage and Nonfinancial Leverage, respectively. We consider rolling windows of 50 observations evaluating the out-of-sample samples 2011M08-2023M12 and 2015M10-2023M12.

Table A.2: Out-of-sample $R^2_\tau(\%)$ IP growth with rolling window.

| | $\tau$=0.05 | $\tau$=0.1 | $\tau$=0.25 | $\tau$=0.5 | $\tau$=0.75 | $\tau$=0.9 | $\tau$=0.95 |
|---|---|---|---|---|---|---|---|
| $h = 3$ months and evaluating over the OOS sample 2011M08-2023M12. | | | | | | | |
| QR with NFCI | -0.5 | 1.5 | -0.3 | -8.1 | -15.1 | -12.8 | -8.7 |
| QR with Risk | -39.1 | -1.6 | -3.0 | -6.1 | -15.2 | -19.6 | -7.1 |
| QR with Credit | 7.1 | 2.1 | 1.9 | -5.1 | -16.8 | -8.2 | 0.2 |
| QR with Leverage | 3.0 | 9.4 | 5.7 | 0.9 | -7.2 | -19.8 | 14.1 |
| QR with 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 |
| 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 |
| $h = 3$ months and evaluating over the OOS sample 2015M10-2023M12. | | | | | | | |
| QR with NFCI | -7.6 | -8.6 | -4.7 | -7.9 | -17.0 | -13.8 | -8.0 |
| QR with Risk | -54.4 | -15.2 | -6.6 | -3.2 | -16.6 | -22.2 | -5.8 |
| QR with Credit | 0.1 | -11.4 | -4.8 | -6.0 | -20.1 | -9.3 | 1.0 |
| QR with Leverage | -3.8 | 0.5 | 4.0 | 3.3 | -5.5 | -22.7 | 16.3 |
| QR with 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 |
| 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 |
| $h = 12$ months and evaluating over the OOS sample 2011M08-2023M12. | | | | | | | |
| QR with NFCI | 1.0 | 11.2 | 5.4 | -5.8 | -10.6 | -6.5 | 1.1 |
| QR with Risk | -5.2 | 11.3 | 5.2 | -2.1 | -11.9 | -8.4 | -4.4 |
| QR with Credit | 6.8 | 10.2 | 4.4 | -4.5 | -10.4 | 1.4 | 7.3 |
| QR with Leverage | 5.6 | 11.5 | 7.2 | -5.1 | -2.3 | -10.0 | 12.8 |
| QR with 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 |
| $h = 12$ months and evaluating over the OOS sample 2015M08-2023M12. | | | | | | | |
| QR with NFCI | -13.3 | 0.9 | 3.7 | -4.5 | -14.8 | -13.5 | -5.9 |
| QR with Risk | -22.5 | -1.3 | 1.2 | 0.5 | -16.5 | -15.1 | -12.0 |
| QR with Credit | -1.4 | -0.9 | 3.1 | -3.7 | -15.6 | -4.3 | 1.3 |
| QR with Leverage | -8.9 | -3.3 | 1.4 | -2.4 | -1.4 | -13.3 | 13.4 |
| QR with 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 |

Note: The out-of-sample $R^2_\tau$ is reported. It evaluates the performance using the samples 2011M08-2023M12 and 2015M08-2023M12. 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 $h$ months before. Qcov3PRF# and PCQR# denote Qcov3PRF and PCQR implementation with # number of factors, respectively. The predictors are standardized.

# References

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.

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

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

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.

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

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

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

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