The Variable Screening Approach for Growth-at-Risk Prediction

Hongqi Chen* Ji Hyung Lee †

November 2022

Abstract

This paper investigates a variable screening approach to study growth-at-risk (GaR) forecasting with high-dimensional predictors. Unlike the existing studies focusing on a few predictors, we use a high-dimensional Fred-QD dataset that can retain useful information on GaR forecasting. To do this, we refine and extend the quantile partial correlation (QPC) based variable screening method by Ma, Li, and Tsai (2017) so that the method can employ time series data. A set of Monte Carlo simulations confirms the validity of QPC under weak dependence, and the empirical application on variable selection for GaR forecasting illustrates the benefit of the method. Some labor market factors are shown to be particularly useful in predicting GaR.

Keywords: variable screening; high dimensional time series; quantile method; growth-at-risk.

1 Introduction

With the increasing interest in the stability of economic growth, defining, measuring, and forecasting the tail risks of economic activity attract more attention in macroeconomic and policy research. Recently a new terminology, growth-at-Risk (GaR), is proposed by Adrian, Boyarchenko, and Giannone (2019). The GaR denotes the conditional quantile of GDP growth rate at the lower 5% level, indicating possible economic recession levels with a 5% probability.

The prediction of conditional quantiles and distributions of the economic time series is challenging. The nonlinear structure in modeling the distributional relationship between macro variables makes it nontrivial to formulate a tractable model. The potential nonlinear causal relations are difficult to identify, thereby complicating the choice of proper covariates. In the context of GaR forecasting, there is a debate about whether GaR is predictable or not. Adrian, Boyarchenko, and Giannone (2019) consider the National Financial Conditions Index (NFCI) as the main predictor of GaR in the United States by using quantile regressions, and conclude that this financial market indicator is useful to predict future GaR. Following this paper, the quantile regression approach using NFCI becomes popular in the GaR literature. Adrian, Grinberg, Liang, and Malik (2020) extend this analysis into 11 advanced economies and construct the term

*Department of Economics, University of Illinois Urbana-Champaign. Email: hongqic2@illinois.edu
†Department of Economics, University of Illinois Urbana-Champaign. Email: jihyung@illinois.edu

structure of GaR using panel quantile regressions, and Figueres and Jarociński (2020) study the vulnerable growth and GaR in the European area. Brownlees and Souza (2021) find that the GARCH model outperforms the quantile regression in the GaR forecasting, and Plagborg-Møller, Reichlin, Ricco, and Hasenzagl (2020) show that the GaR is poorly estimated if the extracted factors are employed, such as NFCI, because the NFCI is endogenous to the economic conditions; see also Hasenzagl, Reichlin, and Ricco (2020). This discussion highlights the need for careful investigation on which variables to choose when predicting GaR and, generally, when predicting the distributions of economic time series.

Variable screening and variable selection play important roles in modern research of economics and statistics. In the era of big-data, forecasters often encounter a huge number of potential predictors such as hundreds or thousands of macroeconomic and financial time series data. It is therefore an important step to identify certain key predictors by variable screening and selection. We discuss the popular l_1-penalization methods and motivate our variable screening approach in high dimensional data environment. First, there is computational advantage of the variable screening approach to the l_1-penalization methods when the dimension of data is huge. Second, when the sample size is smaller than the number of parameters, the variable selection consistency of the l_1-penalization needs a set of stringent conditions, while the variable screening technique does not require such restrictions. In the presence of the ultra-high dimensional predictors, forecasters typically consider a two-step procedure, screening the variables first and select the most important predictors using information criteria or penalization methods after the first screening. The selection consistency of the variable screening method could be beneficial in this two-step procedure.

There is rich literature on the variable screening. Fan and Lv (2008) introduce sure independence screening (SIS henceforth) , and Wang (2009) shows forward selection algorithm can be utilized for variable screening. While most studies focus on the least square regression, Ma, Li, and Tsai (2017) employ quantile partial correlation (QPC henceforth) from Li, Li, and Tsai (2015) and propose three screening algorithms in quantile regressions. In particular, the algorithm 3 in Ma, Li, and Tsai (2017) extends Wang (2009)'s approach into QPC framework, and show the convergence of sample QPC to the population QPC, confirming variable screening consistency. Kong, Li, and Zerom (2019); Zhang and Zhou (2018) propose two other types of quantile correlations to handle specific model structures.

Many existing papers using the variable screening for mean or quantile regressions are designed for independent data. Yousuf (2018) is one exception, extending SIS to the high dimensional time series model under Wu (2005)'s functional dependence framework. Another popular algorithm for variable screening is forward selection, and there are a few papers on forward selection algorithm with dependent data. Ing (2020) and Chiou, Guo, and Ing (2020) consider orthogonal greedy algorithm for linear regression with dependent observations. Sancetta et al. (2016) discuss different greedy algorithms to select relevant covariates for prediction under a mixing condition.

This paper adopts and extends the QPC approach of Ma, Li, and Tsai (2017), and proves its asymptotic properties with the $\beta-$mixing processes. There is no research for the variable screening in time series quantile regression with dependent data, to the best of our knowledge. We first provide a new convergence bounds of QPC for the i.i.d data by Ma, Li, and Tsai (2017). Second, we derive a new probabilistic bounds of

QPC under the $\beta-$mixing condition, and confirm that variable screening property still holds. We thereby extend the empirical ambit of QPC method designed for the independent processes, enabling many time series applications. For example, we can apply QPC to GaR prediction using Fred-QR data of McCracken and Ng (2020), which is our empirical application.

The rest of the paper is organized as follows. Section 2 introduces the model and variable screening in quantile regression model. Section 3 proves the theoretical results under stationary $\beta-$mixing framework. Section 4 shows simulation studies confirming that QPC has good performance for variable screening in our time series setup. In section 5, we provide the empirical variable selection results for GaR forecasting with high-dimensional dataset. The last section concludes.

We use the following notations through the paper. We denote $\rho_\tau(u) = u(\tau - 1(u < 0))$ as the quantile loss function, and let $\psi_\tau(u) = \tau - 1(u < 0)$. We use $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ as the largest and smallest eigenvalues of matrix A. $\|a\|$ denotes the L_2-norm of vector a, $\|A\|$ denotes the Frobenius norm of matrix A, and C and C_i for $i = 1, 2, \ldots$ denote generic constants across the paper.

2 Model and QPC screening approach

We consider weakly dependent random variables $\{Y_{i+1}, X_i\}_{i=1}^n$. $\{Y_{i+1}\}_{i=1}^n$ is a scalar sequence of the response variable while $\{X_i\}_{i=1}^n$ is a stationary $\beta-$mixing sequence of $p-$dimensional covariates. We can represent $X_i = (X_{i,1}, \ldots, X_{i,p})^T$. We allow a potential high-dimensional number of covariates $(p > n)$. We assume the conditional quantile of Y_{i+1} given X_i follows the linear form, in the sense that

$$Q_{Y_{i+1}}(\tau | X_i) = X_i^T \beta_\tau$$

for any $i = 1, \ldots, n$. For simplicity, we omit the τ subscript in $\beta\tau$ below. We use $S \subseteq \{1, \ldots, p\}$ denoting a generic index set of the covariate X_i. To avoid the confounding in notation, we only use j rather than i on X_j to denote a specific covariate $(X_{1,j}, \ldots, X_{n,j})^T$ for $j = 1, \ldots, p$. For a index set S, $X_{i,S}$ indicates the covariates with the index set S on $i-$th observation. Our objective is to select the relevant covariates X_js for the response variable.

To find those covatiates, a typical approach is variable screening. In Ma, Li, and Tsai (2017), they propose three different screening algorithms with QPC. This paper focuses on their algorithm 3 to illustrate the usage of QPC in variable screening with dependent data. The algorithm relies on the forward selection technique to choose the relevant covariates in every step. The criterion in the forward selection is based on the QPC of each covariate. In a specific step j, we select the covariate X_j with the largest QPC conditioning on the previously selected index set S_j. The definition of quantile partial correlation at τ-quantile is defined as follows for random variables $\{Y_{i+1}, X_{i,j}, X_{i,S_j}\}$:

$$qpcor_\tau\left(Y_{i+1}, X_{i,j} | X_{i,S_j}\right) = \frac{cov\left(\psi_\tau\left(Y_{i+1} - X_{i,S_j}^T \alpha_j^0\right), X_{i,j} - X_{i,S_j}^T \theta_j^0\right)}{\sqrt{var\left(\psi_\tau\left(Y_{i+1} - X_{i,S_j}^T \alpha_j^0\right) var\left(X_{i,j} - X_{i,S_j}^T \theta_j^0\right)\right)}}$$

$$= \frac{E\left[\psi_\tau\left(Y_{i+1} - X_{i,S_j}^T \alpha_j^0\right)\left(X_{i,j} - X_{i,S_j}^T \theta_j^0\right)\right]}{\sqrt{\tau\left(1-\tau\right)\sigma_{ij}^2}}$$

where $\alpha_j^0 = \arg\min_{\alpha_j} E\left(\rho_\tau\left(Y_{i+1} - X_{i,S_j}^T \alpha_j\right)\right)$, $\theta_j^0 = \arg\min_{\theta_j} E\left(\left(X_{i,j} - X_{i,S_j}^T \theta_j\right)^2\right)$, and $\sigma_{ij}^2 = var\left(X_{i,j} - X_{i,S_j}^T \theta_j\right)$.

QPC is closely related to the population quantile regression coefficients. Since we assume the linear conditional quantile form of Y_{i+1} on X_i, we can define the minimizer for the quantile loss

$$\left(\beta_1^0, \ldots, \beta_j^0, \ldots, \beta_p^0\right) = \underset{(\beta_1,\ldots,\beta_j,\ldots,\beta_p)}{\arg\min} E\left(\rho_\tau\left(Y_{i+1} - \beta_1 X_{i,1} - \cdots - \beta_j X_{i,j} - \cdots - \beta_p X_{i,p}\right)\right)$$

Given α_j^0 and S_j, β_j^* is defined in the following equation,

$$\beta_j^* = \underset{\beta_j}{\arg\min} E\left(\rho_\tau\left(Y_{i+1} - X_{i,S_j}^T \alpha_j^0 - X_{i,j}\beta_j\right)\right).$$

From Li, Li, and Tsai (2015), we can express $qpcor_\tau\left(Y_{i+1}, X_{i,j}|X_{i,S_j}\right) = \rho\left(\beta_j^*\right)$ where $\rho\left(\cdot\right)$ is a continuous increasing function such that $\rho\left(\beta_j^*\right) = 0$ if and only if $\beta_j^* = 0$.

Since we have one-to-one correspondence between $qpcor_\tau\left(Y_{i+1}, X_{i,j}|X_{i,S_j}\right)$ and β_j^*, in addition, we know $\beta_j^* = 0$ if and only if $\beta_j^0 = 0$ by the lemma 1 below, then we can use $qpcor_\tau\left(Y_{i+1}, X_{i,j}|X_{i,S_j}\right)$ as a substitution of β_j^0 to order the importance among covariates.

Moreover, this QPC approach has an advantage when the dimension of p is relatively large: A direct estimation of β_j^0 is impossible without a penalization in the high dimensional case ($p > n$), but we can always obtain $qpcor_\tau\left(Y_{i+1}, X_{i,j}|X_{S_{i,j}}\right)$ through iterated algorithms. Therefore, the QPC is a useful tool for variable screening in quantile regression framework.

With the sample data, we can define the sample quantile partial correlation $\widehat{qpcor}_\tau$ as the following

$$\widehat{qpcor}_\tau\left(Y_{i+1}, X_{i,j}|X_{i,S_j}\right) = \frac{\frac{1}{n}\sum_{i=1}^n\left(\psi_\tau\left(Y_{i+1} - X_{i,S_j}^T \hat{\alpha}_j\right)\left(X_{i,j} - X_{i,S_j}^T \hat{\theta}_j\right)\right)}{\sqrt{\tau\left(1-\tau\right)\hat{\sigma}_{ij}^2}}$$

where $\hat{\alpha}_j = \arg\min \frac{1}{n}\sum_{i=1}^n\left(\rho_\tau\left(Y_{i+1} - X_{i,S_j}^T \alpha_j\right)\right)$, $\hat{\theta}_j = \arg\min \frac{1}{n}\sum_{i=1}^n\left(X_{i,j} - X_{i,S_j}^T \theta_j\right)^2$, and $\hat{\sigma}_{ij}^2 = \frac{1}{n}\sum_{i=1}^n\left(X_{i,j} - X_{i,S_j}^T \hat{\theta}_j\right)^2$.

With the sample QPC, we use the following forward selection algorithm to screen variables (algorithm 3 in Ma, Li, and Tsai (2017)) and obtain the coefficients for the conditional quantile of Y_{i+1}.

Algorithm 1

1. Initialize the active set of variables $S_1 = \emptyset$.

2. For $d = 1, \ldots, D$

   (a) Compute $j^* = \arg\max_{j \notin S_d} \left|\widehat{qpcor}_\tau \left(Y_{i+1}, X_{i,j} | X_{i,S_d}\right)\right|$

   (b) Update $S_{d+1} = S_d \cup \{j^*\}$

3. Given a specific choice of D (selected by information criterion, e.g. BIC or EBIC), we obtain $\hat{\beta}_{S_D} = \arg\min_{\beta_{S_D}} \frac{1}{n} \sum_{i=1}^{n} \rho_\tau \left(Y_{i+1} - X_{i,S_D}^T \beta_{S_D}\right)$.

3 Theory of QPC

In this section, we show two theoretical properties of QPC: the uniform convergence of sample QPC to the population QPC and the variable screening property. We follow the proof procedure in Ma, Li, and Tsai (2017). Before extending the results into the β-mixing framework, we first restate some results in the original proof of Ma, Li, and Tsai (2017).

3.1 Theoretical properties under the i.i.d scenario

We first review the result of the uniform convergence of $\widehat{qpcor}_\tau \left\{Y_{i+1}, X_{i,j} | X_{i,S_j}\right\}$ to the population QPC under the i.i.d framework as in Ma, Li, and Tsai (2017).

Theorem. *(Theorem 1 in Ma, Li, and Tsai (2017)) Under conditions (C1) and (C2) in Ma, Li, and Tsai (2017), for some universal constant C, $0 < \kappa < \frac{1}{2}$, and $r_n = Cn^\omega$ for some $0 \leq \omega < \min\left((1 - 2\kappa), 2\kappa\right)$, we have*

$$
\begin{aligned}
&P\left(\sup_{1 \leq j \leq p_n} \left|\widehat{qpcor}_\tau \left(Y_{i+1}, X_{i,j} | X_{i,S_j}\right) - qpcor_\tau \left(Y_{i+1}, X_{i,j} | X_{i,S_j}\right)\right| \geq C r_n^{\frac{1}{2}} n^{-\kappa}\right) \\
\leq & p_n \left(C r_n^2 e^{-\frac{Cn^{1-2\kappa}}{r_n}} + C r_n^2 e^{-C\frac{n}{r_n^2}}\right)
\end{aligned}
$$

Remark 1. This theorem from Ma, Li, and Tsai (2017) proves the convergence of sample QPC to the population counterpart. We detected some minor mistakes in the proof of Ma, Li, and Tsai (2017). We summarize the issue in the footnote* below. Moreover, theorem 2 in their paper should be revised accordingly.

*

1. In the (S.5) inequality and the inequality before (S.5) of Ma, Li, and Tsai (2017)'s supplemental materials, the correct modifications of the inequalities are

$$
P\left(\left|\lambda_{\min}\left(n^{-1}\sum_{i=1}^{n} X_{i,S_j} X_{i,S_j}^T\right) - \lambda_{\min}\left(E X_{i,S_j} X_{i,S_j}^T\right)\right| \geq c_3^* r_n n^{-1} \delta_1^*\right) \leq 2exp\left(-c_4^* {\delta_1^*}^2 n^{-1}\right)
$$

   and

$$
P\left(\left\|n^{-1}\sum_{i=1}^{n} X_{i,S_j} X_{i,S_j}^T - E\left(X_{i,S_j} X_{i,S_j}^T\right)\right\| \geq c_3^* r_n n^{-1} \delta_1^*\right) \leq 2exp\left(-c_4^* {\delta_1^*}^2 n^{-1}\right)
$$

   Their lemma 2, 4, and 5 should change accordingly.

2. In the (S.12) inequality of Ma, Li, and Tsai (2017)'s supplemental materials. The Bernstein's inequality should lead to

3.2 Theoretical properties under mixing processes

In this subsection, we state the following conditions to show the uniform convergence of $\widehat{qpcor}_\tau \left\{Y_{i+1}, X_{i,j}|X_{i,S_j}\right\}$ to the population QPC for $\beta-$mixing sequences.

Assumption 3.1. *Let the predictor $\{X_{i,j}\}_{i=1}^n$ is stationary $\beta-$mixing sequences with mixing coefficients $\beta(n) \le e^{-c_0 n}$ for some constant c_0 with every $j = 1, \dots, p$.*[†]

Assumption 3.2. *Assume that $\sup_{i,j}|X_{i,j}| \le M_1$, $\sup_{i,j}\left|X_{i,S_j}^T\theta_j^0\right| \le M_2$, $\sup_{i,j}\left|X_{i,S_j}^T\pi_j^0\right| \le M_3$, $\sup_j \left\|\frac{1}{n}\sum_{i=1}^n X_{i,S_j}X_{i,j}\right\| \le M_4$*

Assumption 3.3. *For every $1 \le i \le n$ and $1 \le j \le p$, we assume*

$$m \le \lambda_{\min}\left(E\left(X_{i,S_j}X_{i,S_j}^T\right)\right) \le \lambda_{\max}\left(E\left(X_{i,S_j}X_{i,S_j}^T\right)\right) \le M$$

Assumption 3.4. *For every $1 \le i \le n$ the conditional density of Y_{i+1} on X_i, $f_{Y_{i+1}|X_i}(y)$, is 1-Lipschiz and is bounded above and below from 0.*

Assumption 3.5. *There exists a sequence $r_n = \max_{1\le j\le p}|S_j| = Cn^\gamma$ for some $0 \le \gamma < 2\kappa$ and $0 < \kappa < \frac{1}{4}$.*

Assumption 3.6. $\min_{j\in M_*}\left|qpcor_\tau\left(Y_{i+1}, X_{i,j}|X_{i,S_j}\right)\right| \ge C_0 r_n^{\frac{1}{2}} n^{-\kappa}$ *for some $0 < \kappa < \frac{1}{2}$ and some constant C_0, where $M_* = \left\{j : \beta_j^0 > 0,\ 1 \le j \le p\right\}$.*

Remark 2. Assumption 3.1 imposes $\beta-$mixing condition on the covariates. Assumption 3.2 collects the boundedness conditions for the predictors, comparable to the condition (C2) in Ma, Li, and Tsai (2017). These conditions are common in the high-dimensional literature. Assumption 3.3 imposes the bounds on the maximal and minimal eigenvalues of the Gram matrix $EX_{i,S_j}X_{i,S_j}^T$ for the stationary process X_{i,S_j}. Assumption 3.4 implies a regular conditional density of Y_{i+1}, which is standard in the quantile regression literature. Assumption 3.5 imposes a rate condition for r_n, which is used to control the convergence rate on the sample QPC. Assumption 3.6 is required in the variable screening consistency. It relates the magnitude of the population QPC with respect to the covaraite X_j whose coefficient β_j^0 is positive. The last two assumptions are adopted from assumption (C3) in Ma, Li, and Tsai (2017).

From the previous instruction, we know that $qpcor_\tau\left(Y_{i+1}, X_{i,j}|X_{i,S_j}\right)$ is related with β_j^*. The next lemma from Ma, Li, and Tsai (2017) shows the connection between β_j^* and β_j^0.

Lemma 1. *Suppose $\beta^0 = \left(\beta_0^0, \dots, \beta_p^0\right)^T$ is the unique minimizer of $E\left(\rho_\tau\left(Y_{i+1} - \beta_0 - \beta_1 X_{i,1} - \cdots - \beta_p X_{i,p}\right)\right)$. Suppose α_j^0 are unique minimizers of $E\left(\rho_\tau\left(Y_{i+1} - X_{i,-j}^T\alpha\right)\right)$ and $\left(\beta_0^*, \beta_j^*\right)^T$ are unique minimizers of $E\left(\rho_\tau\left(Y_{i+1}^* - \beta_0 - X_{i,j}\beta_{j\tau}\right)\right)$ where $Y_{i+1}^* = Y_{i+1} - X_{i,-j}^T\alpha_j^0$. Then $\beta_j^0 = 0$ if and only if $\beta_j^* = 0$.*

---

the following result

$$P\left(\left|\bar{\omega}_n\left(\pi_j^0\right) - \bar{\omega}\left(\pi_j^0\right) \ge \frac{1}{2}c_8^* n^{-2\kappa}\right|\right) \le 2exp\left(-c_{12}^* n^{1-4\kappa}\right)$$

So their lemma 3 and the range of κ in theorem 1 should change correspondingly.

[†]The $\beta-$mixing condition implies $\alpha-$mixing (strong mixing) condition. Therefore, the theoretical properties of probabilistic bounds under $\alpha-$mixing condition from Merlevède, Peligrad, and Rio (2009) hold in our proof.

Our goal is using QPC to screen variables, and we can estimate the sample QPC, $\widehat{qpcor_\tau}\left(Y_{i+1}, X_{i,j}|X_{i,S_j}\right)$ from the data. Therefore, we establish the uniform convergence of $\widehat{qpcor_\tau}\left(Y_{i+1}, X_{i,j}|X_{i,S_j}\right)$ to the population QPC, $qpcor_\tau\left(Y_{i+1}, X_{i,j}|X_{i,S_j}\right)$. To show this uniform convergence, we first prove the convergences results of $\hat{\theta}_j$, $\hat{\pi}_j$, $\frac{1}{n}\sum_{i=1}^n \psi_\tau\left(Y_{i+1} - X_{i,S_j}^T\hat{\pi}\right)\left(X_{i,j} - X_{i,S_j}^T\hat{\theta}_j\right)$, and $\hat{\sigma}_j^2$ in the following lemma 2 to lemma 5.

Lemma 2. *Under assumption 3.1 and 3.2, by assuming $n^{-1}\delta_n = o(1)$ and $n^{-1}\delta_1 = o(1)$, denote*

$$\hat{\theta}_j = \left(\frac{1}{n}\sum_{i=1}^n X_{i,S_j}X_{i,S_j}^T\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^n X_{i,S_j}X_{i,j}\right)$$

and

$$\theta_j^0 = E\left(X_{i,S_j}X_{i,S_j}^T\right)^{-1} E\left(X_{i,S_j}X_{i,j}\right)$$

as the sample estimated coefficient and population coefficient of X_{i,S_j} on $X_{i,j}$. Then for some positive constants $C_1, C_2, C_3, C_4, C_5, C_6, C_7$, we have

$$P\left(\left\|\hat{\theta}_j - \theta_j^0\right\| \geq \frac{C_1 r_n^{\frac{1}{2}}\delta_n}{mn} + \frac{C_2 r_n \delta_1}{n}\right) \leq r_n e^{-\frac{C_3\delta_n^2}{n}} + e^{-\frac{C_5 n}{r_n^2}} + e^{-\frac{C_4\delta_1^2}{n}}$$

Moreover, if $\delta_1 = \delta_n$, we have

$$P\left(\left\|\hat{\theta}_j - \theta_j^0\right\| \geq C_6\frac{r_n\delta_n}{n}\right) \leq C_7 r_n e^{-\frac{C_3\delta_n^2}{n}} + e^{-\frac{C_5 n}{r_n^2}}$$

Lemma 3. *Under assumption 3.1, 3.2, and 3.3, for any $1 \leq j \leq p_n$ and some constant $C_8, C_9, C_{10}, C_{11}, C_{12} > 0$, we have for some $0 < \kappa < \frac{1}{4}$,*

$$P\left(\left\|\hat{\pi}_j - \pi_j^0\right\| \geq C_8 n^{-\kappa}\right) \leq e^{-C_9 n^{1-4\kappa}} + e^{-C_{10} r_n^{-1} n^{1-2\kappa}} + C_{12} n^{\frac{1}{2}} e^{-C_{11} n^{\frac{1}{2}}}$$

where $\hat{\pi}_j = \arg\min \bar{\omega}_n(\pi_j) = \arg\min \frac{1}{n}\sum_{i=1}^n\left(\rho_\tau\left(Y_{i+1} - X_{i,S_j}^T\pi_j\right) - \rho_\tau(Y_{i+1})\right)$ and $\pi_j^0 = \arg\min\bar{\omega}(\pi_j) = \arg\min E\left(\rho_\tau\left(Y_{i+1} - X_{i,S_j}^T\pi_j\right) - \rho_\tau(Y_{i+1})\right)$.

Remark 3. Under the weak dependence, we have three terms for the probabilistic bound of $\left\|\hat{\pi}_j - \pi_j^0\right\| \geq C_8 n^{-\kappa}$. This is different from the i.i.d case, where only the first two terms appear. The additional third term $C_{12} n^{\frac{1}{2}} e^{-C_{11} n^{\frac{1}{2}}}$ does not depend on the mixing coefficient $\beta(a) = e^{-c_0 a}$. While using theorem 1 in Mohri and Rostamizadeh (2009), as long as we let $\delta > 2(\mu - 1)e^{-c_0 a}$, the probabilistic bound is controlled by δ. We can obtain a different bound by a different choice of δ.

Lemma 4. *Under assumption 3.1, 3.2, and 3.3, for any $1 \leq j \leq p_n$ and $c_{10} > 0$, we have some universal positive constants $C_{13}, C_{14}, C_{15}, C_{16}$ such that*

$$P\left(\left|\frac{1}{n}\sum_{i=1}^n \psi_\tau\left(Y_{i+1} - X_{i,S_j}^T\hat{\pi}\right)\left(X_{i,j} - X_{i,S_j}^T\hat{\theta}_j\right) - E\left[\psi_\tau\left(Y_{i+1} - X_{i,S_j}^T\pi_j^0\right)X_{i,j}\right]\right| \geq C_{13} r_n^{\frac{1}{2}} n^{-\kappa}\right)$$

$$\leq C_{14}e^{-C_9 n^{1-4\kappa}} + e^{-C_{10} r_n^{-1} n^{1-2\kappa}} + C_{12} n^{\frac{1}{2}} e^{-C_{11} n^{\frac{1}{2}}} + C_{16} r_n e^{-C_{15}\frac{n}{r_n^2}}$$

Lemma 5. *Under assumptions 3.1, 3.2, and 3.3, for any $1 \leq j \leq p_n$ and some positive constants $C_{17}, C_{18}, C_{19}, C_{20}, C_{21}, C_{22}$, there exists some universal constants C such that*

$$P\left(\left|\hat{\sigma}_j^2 - \sigma_j^2\right| \geq C_{17} r_n^{\frac{1}{2}} n^{-\kappa}\right) \leq e^{-\frac{C_{18} n}{r_n}} + C_{20} e^{-\frac{C_{19} n}{r_n^2}} + C_{22} r_n e^{-\frac{C_{21} n^{1-2\kappa}}{r_n}}$$

In addition, with assumption 3.4, we have for $a \in (0,1)$,

$$P\left(\left|\hat{\sigma}_j^2 - \sigma_j^2\right| \geq a\sigma_j^2\right) \leq e^{-\frac{C_{18} n}{r_n}} + C_{20} e^{-\frac{C_{19} n}{r_n^2}} + C_{22} r_n e^{-\frac{C_{21} n^{1-2\kappa}}{r_n}}$$

Based on the convergence of different components in $\widehat{qpcor}_\tau\left(Y_{i+1}, X_{i,j}|X_{i,S_j}\right) - qpcor_\tau\left(Y_{i+1}, X_{i,j}|X_{i,S_j}\right)$ above, we now state our theorem 1 below.

Theorem 1. *Under the assumption 3.1, 3.2, 3.3, and 3.4, for some positive constants $C_1^*, C_2^*, \ldots, C_{10}^*$ and a sufficiently large n, we have for $0 < \kappa < \frac{1}{4}$ and $r_n = \max_{1 \leq j \leq p} |S_j| = Cn^\gamma$ with $0 \leq \gamma < 2\kappa$,*

$$\begin{aligned}
&P\left(\sup_{1 \leq j \leq p_n} \left|\widehat{qpcor}_\tau\left(Y_{i+1}, X_{i,j}|X_{i,S_j}\right) - qpcor_\tau\left(Y_{i+1}, X_{i,j}|X_{i,S_j}\right)\right| \geq C_1^* r_n^{\frac{1}{2}} n^{-\kappa}\right) \\
\leq& p_n\left(e^{-\frac{C_2^* n}{r_n}} + C_4^* r_n e^{-C_3^* \frac{n}{r_n^2}} + C_6^* r_n e^{-\frac{C_5^* n^{1-2\kappa}}{r_n}} + C_8^* e^{-C_7^* n^{1-4\kappa}} + C_{10}^* n^{\frac{1}{2}} e^{-C_9^* n^{\frac{1}{2}}}\right)
\end{aligned}$$

Proof. We know

$$\begin{aligned}
\left|\widehat{qpcor}_\tau\left(Y_{i+1}, X_{i,j}|X_{i,S_j}\right) - qpcor_\tau\left(Y_{i+1}, X_{i,j}|X_{i,S_j}\right)\right| &= \left|\left(\hat{\sigma}_j^2\right)^{-1}\left(\phi_{jn} - \phi_j\right) - \left(\hat{\sigma}_j^2\right)^{-1}\left(\sigma_j^2\right)^{-1}\phi_j\left(\hat{\sigma}_j^2 - \sigma_j^2\right)\right| \\
&\leq \left(\hat{\sigma}_j^2\right)^{-1}\left|\phi_{jn} - \phi_j\right| + \left(\hat{\sigma}_j^2\right)^{-1}\left(\sigma_j^2\right)^{-1}\left|\phi_j\right|\left|\hat{\sigma}_j^2 - \sigma_j^2\right|
\end{aligned}$$

where $\phi_{jn} = \frac{1}{n}\sum_{i=1}^n \psi_\tau\left(Y_{i+1} - X_{i,S_j}^T \hat{\pi}_j\right)\left(X_{i,j} - X_{i,S_j}^T \hat{\theta}_j\right)$ and $\phi_j = E\left(\psi_\tau\left(Y_{i+1} - X_{i,S_j}^T \pi_j^0\right) X_{i,j}\right)$

Using the fact that $\left|a^{-1} - b^{-1}\right| \geq k^* b^{-1} \implies |a - b| \geq kb$ with $k^* = \frac{1}{1-k} - 1$ for any $k \in (0,1)$, we can show for some constant $c_\sigma^{-1} \geq \left(\sigma_j^2\right)^{-1}$

$$\begin{aligned}
P\left(\left(\hat{\sigma}_j^2\right)^{-1} \geq (1 + k^*) c_\sigma^{-1}\right) &\leq P\left(\left(\hat{\sigma}_j^2\right)^{-1} \geq (1 + k^*)\left(\sigma_j^2\right)^{-1}\right) \\
&\leq P\left(\left|\left(\hat{\sigma}_j^2\right)^{-1} - \left(\sigma_j^2\right)\right| \geq k^*\left(\sigma_j^2\right)^{-1}\right) \\
&\leq P\left(\left|\hat{\sigma}_j^2 - \sigma_j^2\right| \geq k\sigma_j^2\right) \\
&\leq e^{-\frac{C_{18} n}{r_n}} + C_{20} e^{-\frac{C_{19} n}{r_n^2}} + C_{22} r_n e^{-\frac{C_{21} n^{1-2\kappa}}{r_n}}
\end{aligned}$$

where the last inequality is based on lemma 5.

From lemma 4, we know

$$P\left(|\phi_{jn}-\phi_j| \geq C_{13} r_n^{\frac{1}{2}} n^{-\kappa}\right) \leq C_{14} e^{-C_9 n^{1-4\kappa}} + e^{-C_{10} r_n^{-1} n^{1-2\kappa}} + C_{12} n^{\frac{1}{2}} e^{-C_{11} n^{\frac{1}{2}}} + C_{16} r_n e^{-C_{15}\frac{n}{r_n^2}}$$

Together with the previous inequality, we show

$$\begin{aligned}
&P\left(\left(\hat{\sigma}_j^2\right)^{-1} |\phi_{jn}-\phi_j| \geq (1+k^*)\, c_\sigma^{-1} C_{13} r_n^{\frac{1}{2}} n^{-\kappa}\right) \\
\leq & e^{-\frac{C_{18} n}{r_n}} + C_{20} e^{-\frac{C_{19} n}{r_n^2}} + C_{22} r_n e^{-\frac{C_{21} n^{1-2\kappa}}{r_n}} + C_{14} e^{-C_9 n^{1-4\kappa}} + e^{-C_{10} r_n^{-1} n^{1-2\kappa}} + C_{12} n^{\frac{1}{2}} e^{-C_{11} n^{\frac{1}{2}}} + C_{16} r_n e^{-C_{15}\frac{n}{r_n^2}} \\
\leq & C_{14} e^{-C_9 n^{1-4\kappa}} + c_{42} r_n e^{-c_{41} r_n^{-1} n^{1-2\kappa}} + C_{12} n^{\frac{1}{2}} e^{-C_{11} n^{\frac{1}{2}}} + c_{40} r_n e^{-c_{39}\frac{n}{r_n^2}}
\end{aligned}$$

Moreover, using lemma 5 with $|\phi_j| \leq M_1$ and $\left(\sigma_j^2\right)^{-1} \leq c_\sigma^{-1}$,

$$\begin{aligned}
&P\left(\left(\hat{\sigma}_j^2\right)^{-1} \left(\sigma_j^2\right)^{-1} |\phi_j| \left|\hat{\sigma}_j^2 - \sigma_j^2\right| \geq (1+k^*)\, c_\sigma^{-1} c_\sigma^{-1} M_1 C_{17} r_n^{\frac{1}{2}} n^{-\kappa}\right) \\
\leq & e^{-\frac{C_{18} n}{r_n}} + C_{20} e^{-\frac{C_{19} n}{r_n^2}} + C_{22} r_n e^{-\frac{C_{21} n^{1-2\kappa}}{r_n}}
\end{aligned}$$

Hence, by combining two probabilistic bounds above and rearranging those constants into $C_2^*, C_3^*, C_4^*, C_5^*, C_6^*, C_7^*, C_8^*, C_9^*, C_{10}^*$, we obtain

$$\begin{aligned}
&P\left(\left|qp\hat{c}or_\tau\left(Y_{i+1}, X_{i,j}|X_{i,S_j}\right) - qpcor_\tau\left(Y_{i+1}, X_{i,j}|X_{i,S_j}\right)\right| \geq C_1^* r_n^{\frac{1}{2}} n^{-\kappa}\right) \\
\leq & e^{-\frac{C_2^* n}{r_n}} + C_4^* r_n e^{-C_3^*\frac{n}{r_n^2}} + C_6^* r_n e^{-\frac{C_5^* n^{1-2\kappa}}{r_n}} + C_8^* e^{-C_7^* n^{1-4\kappa}} + C_{10}^* n^{\frac{1}{2}} e^{-C_9^* n^{\frac{1}{2}}}
\end{aligned}$$

Therefore,

$$\begin{aligned}
&P\left(\sup_{1\leq j\leq p_n} \left|\widehat{qpcor}_\tau\left(Y_{i+1}, X_{i,j}|X_{i,S_j}\right) - qpcor_\tau\left(Y_{i+1}, X_{i,j}|X_{i,S_j}\right)\right| \geq C_1^* r_n^{\frac{1}{2}} n^{-\kappa}\right) \\
\leq & p_n\left(e^{-\frac{C_2^* n}{r_n}} + C_4^* r_n e^{-C_3^*\frac{n}{r_n^2}} + C_6^* r_n e^{-\frac{C_5^* n^{1-2\kappa}}{r_n}} + C_8^* e^{-C_7^* n^{1-4\kappa}} + C_{10}^* n^{\frac{1}{2}} e^{-C_9^* n^{\frac{1}{2}}}\right)
\end{aligned}$$

∎

The uniform convergence bound for sample QPC under the mixing conditions have more terms than the convergence bound under the i.i.d case. This result is due to the proof of lemma 3. The convergence bound depend on p_n, n, and r_n. Depending on the choice of r_n, the dominating term varies. Moreover, similarly to common variable screening results, as long as the number of covariates $\log p_n$ has the same order with the number of observations n, the bound shrinks to zero for sufficiently large n.

As for the variable screening property, we define the selected set $\hat{M}_{v_n}$ as

$$\hat{M}_{v_n} = \left\{j : \widehat{qpcor}_\tau\left(Y_{i+1}, X_{i,j}|X_{i,S_j}\right) \geq v_n \ for\ 1 \leq j \leq p\right\}$$

for some threshold v_n converging to 0 and the original nonzero coefficients set $M_* = \left\{j : \beta^0_{j\tau} > 0 \ for \ 1 \leq j \leq p\right\}$. The screening property now follows.

Theorem 2. *Under the assumption 3.6 and conditions of theorem 1, for some universal constant C and the κ defined above, if $v_n = C_1^* r_n^{\frac{1}{2}} n^{-\kappa}$, we have*

$$P\left(M_* \subset \hat{M}_{v_n}\right) \geq 1 - s_n \left(e^{-\frac{C_2^* n}{r_n}} + C_4^* r_n e^{-C_3^* \frac{n}{r_n^2}} + C_6^* r_n e^{-\frac{C_5^* n^{1-2\kappa}}{r_n}} + C_8^* e^{-C_7^* n^{1-4\kappa}} + C_{10}^* n^{\frac{1}{2}} e^{-C_9^* n^{\frac{1}{2}}}\right)$$

where s_n is the number of nonzero coefficients.

Proof. If we consider the set

$$A_n = \left\{\sup_{j \in M_*} \left|\widehat{qpcor_\tau}\left(Y_{i+1}, X_{i,j} | X_{i,S_j}\right) - qpcor_\tau\left(Y_{i+1}, X_{i,j} | X_{i,S_j}\right)\right| \leq \frac{1}{2} C_1^* r_n^{\frac{1}{2}} n^{-\kappa}\right\}$$

with the assumption 3.6 that $\min_{j \in M_*} \left|qpcor_\tau\left(Y_{i+1}, X_{i,j} | X_{i,S_j}\right)\right| \geq \frac{1}{2} C_1^* r_n^{\frac{1}{2}} n^{-\kappa}$. We have $\left|qp\hat{c}or_\tau\left(Y_{i+1}, X_{i,j} | X_{i,S_j}\right)\right| \geq C_1^* r_n^{\frac{1}{2}} n^{-\kappa}$. Since $v_n = C_1^* r_n^{\frac{1}{2}} n^{-\kappa}$, by the definition of $\hat{M}_{v_n}$, we know $M_* \subset \hat{M}_{v_n}$ on the set A_n. Therefore, after knowing that s_n is the number of nonzero coefficients, we have

$$\begin{aligned} P\left(M_* \subset \hat{M}_{v_n}\right) &\geq 1 - P\left(A_n^C\right) \\ &\geq 1 - s_n \left(e^{-\frac{C_2^* n}{r_n}} + C_4^* r_n e^{-C_3^* \frac{n}{r_n^2}} + C_6^* r_n e^{-\frac{C_5^* n^{1-2\kappa}}{r_n}} + C_8^* e^{-C_7^* n^{1-4\kappa}} + C_{10}^* n^{\frac{1}{2}} e^{-C_9^* n^{\frac{1}{2}}}\right) \end{aligned}$$

by the Bonferroni bound and theorem 1.

∎

Theorem 2 shows that the variable screening property of QPC also holds with weakly dependent data. The significant variables will be contained in the selected set of our screening procedure with probability approaching to one as sample size diverges to infinity.

4 Monte Carlo simulation

In this section, we present Monte Carlo simulations to evaluate the performance of QPC under two stage selection procedure with high dimensional time series data. The reason to use two stage selection procedure is that the variable screening consistency only implies the true nonzero coefficients set is eventually contained in the selected set with probability approaching one. There is no theoretical guarantee that these two sets exactly coincide in finite samples. In the two stage selection procedure, we first use QPC to screen variables and narrow down our selection objectives across covariates based on the screening consistency property. Next, we use extended BIC (EBIC) type methods and l_1 penalization, as in Belloni and Chernozhukov (2011), to obtain our final set of chosen variables. For EBIC-type approach, we choose the optimal $D \in \{1, \ldots, D_{max}\}$ which minimize the EBIC criterion: $\ln\left(\frac{1}{n}\sum_{t=1}^n \rho\left(Y_t - X_t^T \hat{\beta}_\tau\right)\right) + D\frac{\log n}{2n}\log p$. For the l_1 penalization approach, the tuning parameter λ is selected by BIC as given in the R package "rqPen". In our simulation, we fix our setup as $n = 200$, $p = 1000$ and the number of nonzero coefficients $s = 4$. During the QPC screening step, we rely on the quantile forward regression (QFR) algorithm and set our maximal steps $D_{\max} = \left\lfloor \frac{n}{\log n} \right\rfloor$ as in Ma, Li, and Tsai (2017). We repeat our simulation 200 times.

To demonstrate the performance of QPC variable screening, we use the following criteria.

1. MQE: Mean quantile prediction error ;
2. Crate: Correctly selection rate across all simulations ($0 \sim 200$);
3. TP: Average number of true positive covariates for all simulations ($0 \sim 4$);
4. FP: Average number of false positive covariates for all simulations.

The most important criterion is "TP". If "TP" is four, the simulation result confirms the variable screening consistency. For the simulation design, we generate two types of the linear models considered by Ma, Li, and Tsai (2017) and Yousuf (2018).

4.1 DGP based on Yousuf (2018)

We consider the following DGP for every t

$$y_t = \sum_{i=1}^{p} \beta_i x_{it} + \varepsilon_t$$

and the predictors $X_t = (x_{1t}, \ldots, x_{pt})^T$ follow the stationary AR(1) processes

$$X_t = \phi X_{t-1} + \eta_t$$

We let $\varepsilon_t \sim z_t - \sigma \Phi^{-1}(\tau)$, and $z_t \sim N\left(0, \sigma^2\right)$ where $\Phi^{-1}(\cdot)$ is the inverse of standard normal cumulative density function with $\sigma = 1$. For the innovation η_t in the AR(1) process of X_t, we set $\eta_t \sim N(0, \Sigma_\eta)$ where $\Sigma_\eta = \left\{\sigma_{\eta_{ij}}\right\} = \left\{\rho^{|i-j|}\right\}$ is a $p \times p$ covariance matrix of η_t. For β_i, we design $\beta_i = 1$ for $i = 1, 2, 3, 4$, otherwise $\beta_i = 0$ for $i > 4$. We then show the results for different $\tau = 0.2, 0.5, 0.8$ to investigate the different quantile levels. We show our simulation with the different choices of $\phi = 0.2, 0.5, 0.8$ representing the persistent level of the data. In addition, we try different $\rho = 0, 0.5$ to impose zero and moderate degree of correlations among the covariates.

As in Table 1, four true nonzero coefficients are included after variable screening in almost all cases. The correct selection rate for the EBIC approach is very high when the persistence level of data is low or moderate. Although l_1 penalization can also achieve the same true positive rate as the EBIC approach, it shows an issue of the high false positive rate.

Table 1: Performance of 2-step QPC screening procedure

ρ	τ	0.2				0.5				0.8			
		MQE	Crate	TP	FP	MQE	Crate	TP	FP	MQE	Crate	TP	FP
0	$\phi = 0.2$												
	QPC-EBIC	0.284	198	4	0.01	0.386	200	4	0	0.289	199	4	0.005
	QPC-l_1	0.337	0	4	25.905	0.492	0	4	32.29	0.356	0	4	26.39
	$\phi = 0.5$												
	QPC-EBIC	0.286	198	4	0.01	0.391	199	4	0.005	0.288	199	4	0.005
	QPC-l_1	0.324	0	4	24.285	0.512	0	4	29.91	0.357	0	4	25.13
	$\phi = 0.8$												
	QPC-EBIC	0.291	98	4	0.85	0.398	178	4	0.115	0.285	91	4	0.85
	QPC-l_1	0.356	2	4	16.765	0.496	3	4	17.43	0.365	4	4	17.535
0.5	$\phi = 0.2$												
	QPC-EBIC	0.284	200	4	0	0.389	200	4	0	0.287	198	4	0.01
	QPC-l_1	0.319	0	4	23.69	0.509	0	4	30.94	0.342	1	4	24.05
	$\phi = 0.5$												
	QPC-EBIC	0.291	198	4	0.01	0.394	200	4	0	0.284	199	4	0.005
	QPC-l_1	0.322	4	4	20.59	0.478	4	4	25.92	0.339	1	4	21.285
	$\phi = 0.8$												
	QPC-EBIC	0.294	183	4	0.085	0.399	200	4	0	0.288	186	4	0.07
	QPC-l_1	0.329	17	4	10.835	0.461	16	4	10.485	0.36	14	4	12.5

4.2 DGP based on Ma, Li, and Tsai (2017)

Consider the following DGP for every t

$$y_t = \beta x_{1t} + \beta x_{2t} + \beta x_{3t} - 3\sqrt{\rho}\beta x_{4t} + \varepsilon_t$$

and the predictors X_t follow the stationary AR(1) processes

$$X_t = \phi X_{t-1} + \eta_t$$

Here, we set $\varepsilon_t \sim z_t - \sigma\Phi^{-1}(\tau)$, and $z_t \sim N\left(0, \sigma^2\right)$ where $\Phi^{-1}(\cdot)$ is the inverse of the standard normal cumulative density function with $\sigma = 1$. For the innovation η_t in the AR(1) process of X_t, we set $\eta_t \sim N(0, \Sigma_\eta)$ where $\Sigma_\eta = \left\{\sigma_{\eta_{ij}}\right\}$ is a $p \times p$ covariance matrix of η_t. We impose $\sigma_{\eta_{ii}} = 1$ and $\sigma_{\eta_{ij}} = \rho\left(1 - \phi^2\right)$ for all $i \neq j$ and $i, j \neq 4$. We let $\sigma_{\eta_{4j}} = \sigma_{\eta_{i4}} = \sqrt{\rho}\left(1 - \phi^2\right)$ in order to rule out the marginal correlation between x_{4t} and y_t in population. For the coefficients, we assign $\beta = 2.5\left(1 + |\tau - 0.5|\right)$ as Ma, Li, and Tsai (2017).

Under this setup, we report the results for different $\tau = 0.2, 0.5, 0.8$ to accommodate different quantile levels. We show our simulation with different choices of $\phi = 0.2, 0.5, 0.8$ which represent the persistent level of the data. In addition, we try different $\rho = 0.5, 0.95$ to impose moderate and high correlation among covariates.

Table 2: Performance of 2-step QPC screening procedure

ρ	τ	0.2 MQE	0.2 Crate	0.2 TP	0.2 FP	0.5 MQE	0.5 Crate	0.5 TP	0.5 FP	0.8 MQE	0.8 Crate	0.8 TP	0.8 FP
0.50	$\phi = 0.2$												
	QPC-EBIC	0.287	110	4	0.595	0.292	159	4	0.24	0.292	49	4	1.305
	QPC-l_1	0.353	0	4	25.225	0.338	0	4	19.02	0.430	10	4	7.785
	$\phi = 0.5$												
	QPC-EBIC	0.389	154	4	0.26	0.393	199	4	0.005	0.397	171	4	0.17
	QPC-l_1	0.520	0	4	31.54	0.479	0	4	25.645	0.500	4	4	9.78
	$\phi = 0.8$												
	QPC-EBIC	0.282	102	4	0.59	0.289	161	4	0.215	0.290	48	4	1.37
	QPC-l_1	0.368	0	4	25.045	0.327	0	4	19.32	0.399	8	4	8.43
0.95	$\phi = 0.2$												
	QPC-EBIC	0.292	187	4	0.065	0.291	157	4	0.215	0.298	55	4	1.135
	QPC-l_1	0.357	0	4	26.005	0.347	2	4	17.975	0.478	6	4	7.385
	$\phi = 0.5$												
	QPC-EBIC	0.393	197	4	0.015	0.397	194	4	0.03	0.399	181	4	0.095
	QPC-l_1	0.523	0	4	32.49	0.492	0	4	25.18	0.550	9	4	7.995
	$\phi = 0.8$												
	QPC-EBIC	0.286	189	4	0.055	0.291	148	4	0.265	0.287	65	4	1.055
	QPC-l_1	0.352	0	4	25.79	0.355	1	4	18.975	0.464	15	4	6.79

In table 2, we summarize the performance of the 2-step procedure. For both methods, all four true nonzero coefficients are included in the screening step and remain after the selection step. This result verifies the variable screening consistency property of QPC. In addition, we figure out that using EBIC in the variable selection step can produce more promising results than the l_1 penalization method. In the EBIC approach, the numbers of false positive covariates are very low (less than 2) across all simulations and the exact selection rate is relatively high, especially at $\tau = 0.5$, while l_1 penalization tends to choose more irrelevant predictors. This indicates the 2-step QPC with EBIC is a useful tool for variable selection in the high dimensional linear quantile models with time-series data.

5 Empirical application: macroeconomic variable screening for growth-at-risk

Measuring and predicting the downside risk of economic growth are attracting more and more attention. Adrian, Boyarchenko, and Giannone (2019) propose a new measurement, growth-at-risk (GaR), to evaluate the risk. The GaR estimates the conditional quantile of the real GDP growth rate at 5% level. Many researchers have proposed improved forecasts on GaR since then. There are also debates on whether GaR is predictable and which economic indicator can predict GaR. Adrian, Boyarchenko, and Giannone (2019) first use National Financial Condition Index (NFCI) as the main predictor, arguing that tighter financial condition increases the downside risk. They also outline the evolution of GaR. Brownlees and Souza (2021) backtest GaR forecasting with quantile regressions and GARCH models. Their finding suggests that standard volatility models such as GARCH can produce more accurate predictions. Plagborg-Møller, Reichlin, Ricco, and Hasenzagl (2020) consider using extracted factors to predict GaR. They suggest that the forecasting capability of financial variables is limited.

In this section, we consider this problem from another aspect. Unlike the existing literature relying on a few indexes or factors to forecast GaR, we focus on data-dependent macroeconomic variable selection. Our question is: which macroeconomic variable is important for GaR prediction? To answer this question, we utilize our QPC technique with a large number of economic covariates. From the theoretical aspect, QPC has variable screening property; hence it is a reliable tool to decide which variables to include. This variable screening property works well under the high-dimensional setup. In addition, our variable screening approach can rank the importance of the variables. From the empirical aspect, there will be a loss of information if researchers only consider indexes or factors. In QPC approach, every variable is scrutinized, and the maximal information from the dataset is exploited.

There is another desirable feature of QPC for variable screening. The QPC for every variable is related to the corresponding coefficient of multivariate quantile regression. From the section 2 of Ma, Li, and Tsai (2017), we can denote $qpcor_\tau\left(Y_{i+1}, X_{i,j}|X_{i,S_j}\right) = \rho\left(\beta_j^*\right)$ where ρ is a continuous and increasing function and $\beta_j^* = \arg\min_{\beta_j} E\left[\rho_\tau\left(Y_{i+1} - X_{i,S_j}^T\alpha_j^0 - X_j\beta_{j\tau}\right)\right]$ given α_j^0 and X_{i,S_j}. We also have $\rho\left(\beta_j^*\right) = 0$ if and only if $\beta_j^* = 0$. This indicates that if a macroeconomic variable is selected by our QPC procedure (i.e., $qpcor_\tau\left(Y_{i+1}, X_{i,j}|X_{i,S_j}\right) \neq 0$), this variable has none-zero impact on the GaR prediction. Moreover, since QPC is a measurement of partial correlation, it eliminates the impact on the response variable from other predictors.

Throughout our empirical application, we use a truncated dataset containing the U.S. GDP growth rate and 243 predictors from Fred-QD. The Fred-QD dataset is provided by McCracken and Ng (2020). It is a quarterly dataset of the U.S. macroeconomic time series. A key feature of this dataset is that it contains a huge amount of indicators, but the total number of the quarter span is relatively short. Our selected time span is from 1987Q3 to 2021Q4, which contains 138 quarters. We consider one-step forward prediction using the quantile regression model,

$$Q_\tau\left(Y_{t+1}|X_t\right) = \beta_1 X_{1t} + \cdots + \beta_p X_{pt}$$

with $\tau = 0.05$. Following the instruction in McCracken and Ng (2016), we first transform all variables into stationary time series. Next, we use the QPC forward selection procedure to select variables with EBIC. We document the frequency patterns of the selected predictors. Therefore, we adopt fixed window recursive forecasts with different window lengths. We set the window length l as 40, 60, 80, and 100 quarters. The number of corresponding forecasting periods is 98, 78, 58, and 38 quarters (from 1997Q3, 2002Q3, 2007Q3, and 2012Q3 to 2021Q4). Table 3 summarizes the top 10 most frequently selected macroeconomic variables under different forecasting window lengths.

Our result shows that "AWHMAN", "COMPAPFF" and "CLAIMSx" appear in all cases. "AWHMAN" is the average weekly hours of production and non-supervisory employees: manufacturing (hours). "CLAIMSx" is the initial claim. They are employment-related variables. "COMPAPFF" is the 3-month commercial pa-

Table 3: Frequency table of top-10 selected macroeconomic variables (quarterly)

l	40		60	
Forecasting Periods	1997Q3-2021Q4: 98		2002Q3-2021Q4: 78	
Rank	Variable Name	Prob	Variable Name	Prob
1	AWHMAN	0.6020	AWHMAN	0.9615
2	COMPAPFF	0.3673	COMPAPFF	0.5256
3	IMPGSC1	0.1224	IMPGSC1	0.1923
4	CUSR0000SAS	0.1224	IPNMAT	0.1795
5	CMRMTSPLx	0.0714	TLBSNNCBBDIx	0.1154
6	HWIURATIOx	0.0714	CONSPIx	0.1026
7	CLAIMSx	0.0714	CLAIMSx	0.0897
8	FEDFUNDS	0.0510	DRIWCIL	0.0641
9	REVOLSLx	0.0510	UEMP15T26	0.0385
10	OUTNFB	0.0306	PERMITNE	0.0385
l	80		100	
Forecasting Periods	2007Q3-2021Q4: 58		2012Q3-2021Q4: 38	
Rank	Variable Name	Prob	Variable Name	Prob
1	AWHMAN	0.9655	AWHMAN	1.0000
2	IPNMAT	0.4828	CPF3MTB3Mx	0.4211
3	DRIWCIL	0.1897	CMRMTSPLx	0.3684
4	USSTHPI	0.1897	CLAIMSx	0.1842
5	CLAIMSx	0.1379	COMPAPFF	0.1842
6	COMPAPFF	0.1207	UEMP15T26	0.0789
7	CPF3MTB3Mx	0.0690	UEMP27OV	0.0789
8	PRFIx	0.0345	CPILFESL	0.0526
9	IMPGSC1	0.0345	HWIx	0.0263
10	UNRATESTx	0.0345	DRIWCIL	0.0263

per minus federal funds rate. It is an indicator of short-term interest rates. Since our application considers one-step ahead forecasting, it is reasonable that the short-term interest rate is a significant predictor. In addition, “IPNMAT” (Industrial production: materials), “CONSPIx” (Nonrevolving consumer credit to personal income), “TLBSNNCBBDlx” (Nonfinancial corporate business sector liabilities to disposable business income), and “DRIWCIL” (Federal reserve bank senior loans officer opinion survey: net percentage of domestic respondents reporting increased willingness to make consume installment loans) also have predictive powers from time to time.

For robustness check, we apply the same procedure with monthly data. Since the GDP growth rate is reported at a quarterly frequency, we use Chicago Feb National Activity Index (CFNAI) as a monthly substitution for it. CFNAI documents the monthly U.S. economic activity and is released by the Federal Reserve Bank of Chicago. It is a weighted average of a bunch of economic activity indicates, so it mimics the behavior of the GDP growth rate closely at the monthly level. For predictors, we use the Fred-MD dataset from McCracken and Ng (2016), which contains 127 monthly macroeconomic time series. We follow the same procedure as above and consider the fixed window length l as 120, 180, and 240 months. The corresponding forecasting periods are 240, 180, and 120 (from 2002M4, 2007M4, and 2012M4 to 2022M3). Similarly, we

Table 4: Frequency table of top-10 selected macroeconomic variables (monthly)

l	120		180		240	
Forecasting Periods	2002M4-2022M3: 240		2007M4-2022M3: 180		2012M4-2022M3: 120	
Rank	Variable Name	Prob	Variable Name	Prob	Variable Name	Prob
1	CES0600000007	0.5708	CP3Mx	0.7167	COMPAPFFx	0.8250
2	AWHMAN	0.3958	COMPAPFFx	0.7056	AWHMAN	0.6750
3	COMPAPFFx	0.3417	CES0600000007	0.5056	CP3Mx	0.5667
4	VIXCLSx	0.2875	AWHMAN	0.4944	USGOOD	0.4917
5	NDMANEMP	0.2542	NDMANEMP	0.3889	CES0600000007	0.3333
6	CP3Mx	0.2542	HWIURATIO	0.2333	CLAIMSx	0.2000
7	TB3MS	0.2167	CLAIMSx	0.2056	PAYEMS	0.1500
8	CLAIMSx	0.1458	VIXCLSx	0.1389	EXCAUSx	0.1500
9	FEDFUNDS	0.0917	FEDFUNDS	0.0667	USGOVT	0.1417
10	BUSINVx	0.0667	MANEMP	0.0444	MANEMP	0.1250

summarize the top 10 most frequently selected macroeconomic variables in the following table 4.

With monthly data, the result shows that "CES060000007", "AWHMAN", "CP3Mx", "COMPAPFF", and "CLAIMSx" are selected in all scenarios. "AWHMAN", "COMPAPFF", and "CLAIMSx" are consistent with quarterly results. Our monthly exercise further confirms those three key predictors. As for the other two variables, "CP3Mx" is the 3-Month AA financial commercial paper rate, which is closely related to "COMPAPFF". "CES060000007" is the average weekly hours for good producing. It is also an employment variable.

Surprisingly, both quarterly and monthly results complement the GaR literature from another perspective. While most research considers financial conditions as indicators for GaR prediction, e.g., Adrian, Boyarchenko, and Giannone (2019) and Brownlees and Souza (2021), our results suggest both indicators from the labor market and short-term interest rate from the money market contribute most. For financial condition variables, our result is more specific than the NFCI from Adrian, Boyarchenko, and Giannone (2019). We suggest the short-term commercial paper premium is useful, and this indicator is closely related to recession from the existing literature, see Gertler and Gilchrist (2018) and reference therein. Meanwhile, our evidence also shows labor market indicators like initial claims for unemployment and average working hours cannot be neglected. This finding is also consistent with existing research in Schmidt (2022). Schmidt (2022) points out that the initial claims, as a proxy for the labor market risk index, are a robust predictor of broad market returns, especially for stock return predictability. Therefore, the labor market indicators will also play an important role in GaR prediction. More specifically, the predictability of the financial condition index on GaR will be partly absorbed by labor market indicators. This intuition is confirmed by our QPC approach since the labor market indicators are selected more frequently, and fewer financial indicators are chosen conditioning on labor market indicators.

Another complementary result of the QPC approach is to check whether NFCI indeed has predictive power or not. In the existing literature, Adrian, Boyarchenko, and Giannone (2019) and Plagborg-Møller, Reichlin, Ricco, and Hasenzagl (2020) hold diametrically opposed views. The former insists NFCI is useful

Table 5: Frequency table of top-10 selected macroeconomic variables (quarterly)

l	40		60	
Forecasting Periods	1997Q3-2021Q4: 98		2002Q3-2021Q4: 78	
Rank	Variable Name	Prob	Variable Name	Prob
1	AWHMAN	0.6020	AWHMAN	0.9615
2	COMPAPFF	0.3673	COMPAPFF	0.5128
3	IMPGSC1	0.1224	IMPGSC1	0.1923
4	CUSR0000SAS	0.1224	IPNMAT	0.1795
5	CMRMTSPLx	0.0714	TLBSNNCBBDIx	0.1154
6	HWIURATIOx	0.0714	CONSPIx	0.1026
7	CLAIMSx	0.0714	CLAIMSx	0.0897
8	FEDFUNDS	0.0510	DRIWCIL	0.0641
9	REVOLSLx	0.0510	UEMP15T26	0.0385
10	OUTNFB	0.0306	PERMITNE	0.0385
l	80		100	
Forecasting Periods	2007Q3-2021Q4: 58		2012Q3-2021Q4: 38	
Rank	Variable Name	Prob	Variable Name	Prob
1	AWHMAN	0.9655	AWHMAN	1.0000
2	IPNMAT	0.4655	CPF3MTB3Mx	0.4211
3	USSTHPI	0.1897	CMRMTSPLx	0.3684
4	DRIWCIL	0.1724	CLAIMSx	0.1842
5	CLAIMSx	0.1379	COMPAPFF	0.1842
6	COMPAPFF	0.1207	UEMP15T26	0.0789
7	CPF3MTB3Mx	0.0690	UEMP27OV	0.0789
8	NFCI	0.0517	CPILFESL	0.0526
9	IMPGSC1	0.0345	HWIx	0.0263
10	UNRATESTx	0.0345	DRIWCIL	0.0263

in GaR prediction, while the latter argues it is not. Therefore, we include NFCI in our dataset and see if it is selected by our QPC approach. If NFCI is selected, then it still has predictive power, even conditioning on other selected predictors. If it is not selected, then the result indicates that NFCI is actually endogenous and relates to other macroeconomic variables. Hence other macroeconomic variables matter. The following tables 5 and 6 are the quarterly and monthly results.

Our findings show that NFCI is not significantly selected with quarterly data but is ranked in the top 1 or 2 with monthly data. This result complements both sides of the GaR debates. With quarterly data, the NFCI is endogenous and contemporaneous to real economic activities, as Plagborg-Møller, Reichlin, Ricco, and Hasenzagl (2020) suggests. On the other hand, for monthly data, NFCI is a significant predictor, confirming the results in Adrian, Boyarchenko, and Giannone (2019) at the monthly level.

6 Conclusion

This paper extends the quantile partial correlation approach for variable screening with weakly dependent data. The theoretical results are shown to hold under $\beta-$mixing condition. Monte Carlo simulations confirm

Table 6: Frequency table of top-10 selected macroeconomic variables (monthly)

l	120		180		240	
Forecasting Periods	2002M4-2022M3: 240		2007M4-2022M3: 180		2012M4-2022M3: 120	
Rank	Variable Name	Prob	Variable Name	Prob	Variable Name	Prob
1	CES0600000007	0.5708	NFCI	0.8167	NFCI	1.0000
2	NFCI	0.5125	CES0600000007	0.5056	NDMANEMP	0.6667
3	AWHMAN	0.3958	AWHMAN	0.4944	AWHMAN	0.6667
4	TB3MS	0.2167	HWIURATIO	0.2833	CES0600000007	0.3333
5	WPSID62	0.1042	NDMANEMP	0.2333	CLAIMSx	0.2000
6	NDMANEMP	0.1000	CLAIMSx	0.1333	USGOVT	0.1333
7	CLAIMSx	0.0958	VIXCLSx	0.0944	MANEMP	0.1250
8	FEDFUNDS	0.0917	FEDFUNDS	0.0667	CONSPI	0.0417
9	HWIURATIO	0.0875	USGOVT	0.0444	GS10	0.0333
10	BUSINVx	0.0667	PERMITS	0.0278	T10YFFM	0.0250

that the QPC approach performs well in the high-dimensional times-series model. We employ our QPC variable screening approach to study growth-at-risk(GaR) forecasting with many predictors. Our new empirical evidence suggests that the labor market indicators are also valuable for GaR prediction besides the financial market indicator.

A Proofs

A.1 Proofs of lemma 2

beginproof We can represent $\hat{\theta}_j - \theta_j^0$ as

$$\begin{aligned}\hat{\theta}_j - \theta_j^0 =& E\left(X_{i,S_j} X_{i,S_j}^T\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^{n} X_{i,S_j} X_{i,j}^T - E\left(X_{i,S_j} X_{i,j}\right)\right) \\ &+ \left(\left(\frac{1}{n}\sum_{i=1}^{n} X_{i,S_j} X_{i,S_j}^T\right)^{-1} - E\left(X_{i,S_j} X_{i,S_j}^T\right)^{-1}\right)\left(\frac{1}{n}\sum_{i=1}^{n} X_{i,S_j} X_{i,j}\right) \\ :=& \Gamma_{1j} + \Gamma_{2j}\end{aligned}$$

Now, consider $T_{ikj} = X_{i,k}X_{i,j} - E\left(X_{i,k}X_{i,j}\right)$ for $k \in \{0\} \cup S_j$. Under assumption 3.1, we have $\{T_{ijk}\}_{i=1}^n$ is $\alpha-$mixing with $\alpha(n) \leq e^{-c_0 n}$. Using assumption 3.2, we have $|T_{ijk}| \leq 2M_1^2$. Use Theorem 1 in Merlevède, Peligrad, and Rio (2009), we have, for some universal positive constant C, C_1, C_3,

$$\begin{aligned}P\left(\frac{1}{n}\left|\sum_{i=1}^{n} T_{ikj}\right| \geq \frac{C_1 \delta_n}{n}\right) &\leq e^{-\frac{CC_1^2\delta_n^2}{n(2M_1^2)^2 + (2M_1^2)C_1\delta_n(\log n)(\log\log n)}} \\ &\leq e^{-\frac{C_3\delta_n^2}{n}}\end{aligned}$$

for sufficiently large n. The last inequality is based on $n^{-1}\delta_n = o(1)$.‡

Since $r_n = \max_{1\leq j\leq p}|S_j|$, using the Bonferroni bound, we have

$$P\left(\left\|\frac{1}{n}\sum_{i=1}^{n} X_{i,S_j}X_{i,j} - E\left(X_{i,S_j}X_{i,j}\right)\right\| \geq \frac{C_1 r_n^{\frac{1}{2}}\delta_n}{n}\right) \leq r_n e^{-\frac{C_3\delta_n^2}{n}}$$

Using the assumption 3.3, we obtain

$$P\left(\|\Gamma_{1j}\| \geq \frac{C_1 r_n^{\frac{1}{2}}\delta_n}{mn}\right) \leq r_n e^{-\frac{C_3\delta_n^2}{n}}$$

Now we consider $D_j := \frac{1}{n}\sum_{i=1}^{n} X_{i,S_j}X_{i,S_j}^T - E\left(X_{i,S_j}X_{i,S_j}^T\right)$ and $D_{j,kk'} = \frac{1}{n}\sum_{i=1}^{n} X_{i,k}X_{i,k'} - E\left(X_{i,k}X_{i,k'}\right)$ for $k, k' \in \{0\} \cup S_j$. Hence, under assumption 3.1, we have $\left\{D_{j,kk'}\right\}_j$ is $\alpha-$mixing with $\alpha(n) \leq e^{-c_0 n}$. Moreover, we have $\left|X_{i,k}X_{i,k'}\right| \leq 2M_1^2$ and $var\left(X_{i,k}X_{i,k'}\right) \leq M_1^4$. Use Theorem 1 in Merlevède, Peligrad, and Rio (2009), we have, for some positive constants C, c_1, C_4,

‡

Remark. If $\frac{\delta_n}{n} = O(1)$, we can only obtain $P\left(\frac{1}{n}\left|\sum_{i=1}^{n} T_{ikj}\right| \geq \frac{c_1\delta_n}{n}\right) \leq e^{-\frac{C\delta_n^2}{n\log n\log\log n}}$. Here, we assume $r_n = Cn^\gamma$ for some constant γ, then $\frac{\delta_n}{n} = o(1)$ is satisfied in lemma 4.

$$P\left(\left|D_{j,kk'}\right|\geq\frac{c_1\delta_n}{n}\right)\leq e^{-\frac{Cc_1^2\delta_n^2}{n(2M_1^2)^2+(2M_1^2)c_1\delta_n(\log n)(\log\log n)}}$$
$$\leq e^{-\frac{C_4\delta_n^2}{n}}$$

for suffiently large n and all $\delta_n,\ k,k^{'}$. The last inequality is from $n^{-1}\delta_n=o\left(1\right)$.

In addition, we have

$$\begin{aligned}\left\|\frac{1}{n}\sum_{i=1}^{n}X_{i,S_j}X_{i,S_j}^T-E\left(X_{i,S_j}X_{i,S_j}^T\right)\right\|&=\lambda_{\max}\left(\frac{1}{n}\sum_{i=1}^{n}X_{i,S_j}X_{i,S_j}^T-E\left(X_{i,S_j}X_{i,S_j}^T\right)\right)\\&\leq\sum_{k\in\{0\}\cup S_j}\left(\frac{1}{n}\sum_{i=1}^{n}X_{i,kk}X_{i,kk}-E\left(X_{i,k}X_{i,k}\right)\right)\\&\leq r_n\max_k\left|D_{j,kk}\right|\end{aligned}$$

and, using Weyl's inequality, we can show that

$$\begin{aligned}\left|\lambda_{\min}\left(\frac{1}{n}\sum_{i=1}^{n}X_{i,S_j}X_{i,S_j}^T\right)+\lambda_{\min}\left(-E\left(X_{i,S_j}X_{i,S_j}^T\right)\right)\right|&\leq\left|\lambda_{\min}\left(\frac{1}{n}\sum_{i=1}^{n}X_{i,S_j}X_{i,S_j}^T-E\left(X_{i,S_j}X_{i,S_j}^T\right)\right)\right|\\&\leq\left|\left\|\frac{1}{n}\sum_{i=1}^{n}X_{i,S_j}X_{i,S_j}^T-E\left(X_{i,S_j}X_{i,S_j}^T\right)\right\|\right|\\&\leq r_n\max_k\left|D_{j,kk}\right|\end{aligned}$$

Therefore, we obtain

$$P\left(\left|\lambda_{\min}\left(\frac{1}{n}\sum_{i=1}^{n}X_{i,S_j}X_{i,S_j}^T\right)+\lambda_{\min}\left(-E\left(X_{i,S_j}X_{i,S_j}^T\right)\right)\right|\geq r_n\frac{c_1\delta_1^*}{n}\right)\leq e^{-\frac{C_4{\delta_1^*}^2}{n}}$$

and

$$P\left(\left\|\frac{1}{n}\sum_{i=1}^{n}X_{i,S_j}X_{i,S_j}^T-E\left(X_{i,S_j}X_{i,S_j}^T\right)\right\|\geq r_n\frac{c_1\delta_1}{n}\right)\leq e^{-\frac{C_4\delta_1^2}{n}}$$

for some δ_1^*,δ_1. Now we consider $\delta_1^*=\frac{c_2nm}{c_1r_n}$ for some positive constant c_2 and together with assumption 3.3, we show that

$$P\left(\left|\lambda_{\min}\left(\frac{1}{n}\sum_{i=1}^{n}X_{i,S_j}X_{i,S_j}^T\right)-\lambda_{\min}\left(E\left(X_{i,S_j}X_{i,S_j}^T\right)\right)\right|\geq c_2\lambda_{\min}\left(E\left(X_{i,S_j}X_{i,S_j}^T\right)\right)\right)\leq e^{-\frac{C_4c_2^2m^2n}{c_1^2r_n^2}}$$

Using the fact that $\left|a^{-1}-b^{-1}\right| \geq k^{*} b^{-1} \implies |a-b| \geq kb$ with $k^{*}=\frac{1}{1-k}-1$ for any $k \in (0,1)$, we have

$$P\left(\left|\lambda_{\min}^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} X_{i,S_j}X_{i,S_j}^{T}\right)-\lambda_{\min}^{-1}\left(E\left(X_{i,S_j}X_{i,S_j}^{T}\right)\right)\right| \geq \left(\frac{1}{1-c_2}-1\right)\lambda_{\min}^{-1}\left(E\left(X_{i,S_j}X_{i,S_j}^{T}\right)\right)\right) \leq e^{-\frac{C_4 c_2^2 m^2 n}{c_1^2 r_n^2}}$$

which indicates

$$\begin{aligned}
P\left(\lambda_{\min}^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} X_{i,S_j}X_{i,S_j}^{T}\right) \geq \left(\frac{1}{1-c_2}\right)\lambda_{\min}^{-1}\left(E\left(X_{i,S_j}X_{i,S_j}^{T}\right)\right)\right) &\leq e^{-\frac{C_4 c_2^2 m^2 n}{c_1^2 r_n^2}} \\
P\left(\lambda_{\min}^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} X_{i,S_j}X_{i,S_j}^{T}\right) \geq \left(\frac{1}{1-c_2}\right)m^{-1}\right) &\leq e^{-\frac{C_4 c_2^2 m^2 n}{c_1^2 r_n^2}} \\
P\left(\left\|\left(\frac{1}{n}\sum_{i=1}^{n} X_{i,S_j}X_{i,S_j}^{T}\right)^{-1}\right\| \geq \left(\frac{1}{1-c_2}\right)m^{-1}\right) &\leq e^{-\frac{C_4 c_2^2 m^2 n}{c_1^2 r_n^2}} \\
P\left(\left\|\left(\frac{1}{n}\sum_{i=1}^{n} X_{i,S_j}X_{i,S_j}^{T}\right)^{-1}\right\| \left\|E\left(X_{i,S_j}X_{i,S_j}^{T}\right)^{-1}\right\| \geq \left(\frac{1}{1-c_2}\right)m^{-2}\right) &\leq e^{-\frac{C_4 c_2^2 m^2 n}{c_1^2 r_n^2}}
\end{aligned}$$

Therefore, since by definition

$$\|\Gamma_{2j}\| = \left\|\left(\frac{1}{n}\sum_{i=1}^{n} X_{i,S_j}X_{i,S_j}^{T}\right)^{-1} E\left(X_{i,S_j}X_{i,S_j}^{T}\right)^{-1}\left(E\left(X_{i,S_j}X_{i,S_j}^{T}\right)-\frac{1}{n}\sum_{i=1}^{n} X_{i,S_j}X_{i,S_j}^{T}\right)\left(\frac{1}{n}\sum_{i=1}^{n} X_{i,S_j}X_{i,j}\right)\right\|$$

we can show

$$\begin{aligned}
\|\Gamma_{2j}\| \leq &\left\|\left(\frac{1}{n}\sum_{i=1}^{n} X_{i,S_j}X_{i,S_j}^{T}\right)^{-1}\right\| \left\|E\left(X_{i,S_j}X_{i,S_j}^{T}\right)^{-1}\right\| \\
&\times \left\|\frac{1}{n}\sum_{i=1}^{n} X_{i,S_j}X_{i,S_j}^{T}-E\left(X_{i,S_j}X_{i,S_j}^{T}\right)\right\| \left\|\frac{1}{n}\sum_{i=1}^{n} X_{i,S_j}X_{i,j}\right\|
\end{aligned}$$

Hence, using assumption 3.2 and 3.3, we have

$$\begin{aligned}
P\left(\|\Gamma_{2j}\| \geq \left(\frac{1}{1-c_2}\right)m^{-2} r_n \frac{c_1\delta_1}{n} M_4\right) &\leq e^{-\frac{C_4 c_2^2 m^2 n}{c_1^2 r_n^2}} + e^{-\frac{C_4 \delta_1^2}{n}} \\
P\left(\|\Gamma_{2j}\| \geq \frac{C_6 r_n \delta_1}{n}\right) &\leq e^{-\frac{C_7 n}{r_n^2}} + e^{-\frac{C_4 \delta_1^2}{n}}
\end{aligned}$$

where $C_6=\left(\frac{1}{1-c_2}\right)m^{-2}c_1 M_4$ and we denote some postive constant $C_5=\frac{C_4 c_2^2 m^2}{c_1^2}$.

Hence, by the triangular inequality and Bonferroni bound, we have

$$P\left(\left\|\hat{\theta}_j-\theta_j^0\right\| \geq \frac{C_1 r_n^{\frac{1}{2}} \delta_n}{mn}+\frac{C_2 r_n \delta_1}{n}\right) \leq r_n e^{-\frac{C_3 \delta_n^2}{n}}+e^{-\frac{C_5 n}{r_n^2}}+e^{-\frac{C_4 \delta_1^2}{n}}$$

Thus, we have shown the probabilistic bound for $\left\|\hat{\theta}_j-\theta_j^0\right\|$.

In addition, if $\delta_1=\delta_n$, we can directly obtain

$$P\left(\left\|\hat{\theta}_j-\theta_j^0\right\| \geq C_6 \frac{r_n \delta_n}{n}\right) \leq C_7 r_n e^{-\frac{C_3 \delta_n^2}{n}}+e^{-\frac{C_5 n}{r_n^2}}$$

for some positive constant C_7. endproof

A.2 Proofs of lemma 3

beginproof Based on the inequality (S.9) in Ma, Li, and Tsai (2017), we have

$$P\left(\left\|\hat{\pi}_j-\pi_j^0\right\| \geq C_8 n^{-\kappa}\right) \leq P\left(\sup_{\left\|\pi_j-\pi_j^0\right\| \leq C_8 n^{-\kappa}}\left|\bar{\omega}_n\left(\pi_j\right)-\bar{\omega}\left(\pi_j\right)\right| \geq \frac{1}{2} \inf_{\left\|\pi_j-\pi_j^0\right\|=C_8 n^{-\kappa}} \bar{\omega}\left(\pi_j\right)-\bar{\omega}\left(\pi_j^0\right)\right)$$

Consider the case that $\pi_j=\pi_j^0+C_8 n^{-\kappa} u$ with some u satisfying $\|u\|=1$. By the Knight's identity from Knight (1998), we show

$$\begin{aligned}
& \bar{\omega}\left(\pi_j\right)-\bar{\omega}\left(\pi_j^0\right) \\
= & E\left(\rho_\tau\left(Y_{i+1}-X_{i, S_j}^T \pi_j^0-X_{i, S_j}^T C_8 n^{-\kappa} u\right)-\rho_\tau\left(Y_{i+1}-X_{i, S_j}^T \pi_j^0\right)\right) \\
= & E\left(-C_8 n^{-\kappa} X_{i, S_j}^T u\left(\tau-1\left(Y_{i+1}-X_{i, S_j}^T \pi_j^0 \leq 0\right)\right)\right) \\
& +E\left(\int_0^{X_{i, S_j}^T C_8 n^{-\kappa} u} 1\left(Y_{i+1}-X_{i, S_j}^T \pi_j^0 \leq s\right)-1\left(Y_{i+1}-X_{i, S_j}^T \pi_j^0 \leq 0\right) d s\right) \\
= & E_X\left(\int_0^{X_{i, S_j}^T C_8 n^{-\kappa} u} f_{Y_{i+1} \mid X_i}(\zeta) s d s\right)
\end{aligned}$$

where $\zeta \in\left(X_{i, S_j}^T \pi_j^0, X_{i, S_j}^T \pi_j^0+s\right)$.

Using assumption 3.1, 3.2, and 3.3 with $\|u\|=1$, we have

$$\bar{\omega}\left(\pi_j\right)-\bar{\omega}\left(\pi_j^0\right)=C_{10} E\left(\left(X_{i, S_j}^T C_8 n^{-\kappa} u\right)^2\right) \geq C_{10} C_8 m n^{-2 \kappa}$$

for some positive constant C_{10}.

Hence, we show that for some positive constant c_3,

$$\inf_{\left\|\pi_j-\pi_j^0\right\|=C_8 n^{-\kappa}} \bar{\omega}\left(\pi_j\right)-\bar{\omega}\left(\pi_j^0\right) \geq c_3 n^{-2 \kappa}$$

Therefore, by the triangular inequality, we know

$$
\begin{aligned}
P\left(\left\|\hat{\pi}_j-\pi_j^0\right\| \geq C_8 n^{-\kappa}\right) &\leq P\left(\sup_{\left\|\pi_j-\pi_j^0\right\| \leq C_8 n^{-\kappa}}\left|\bar{\omega}_n\left(\pi_j\right)-\bar{\omega}\left(\pi_j\right)\right| \geq \frac{1}{2} \inf_{\left\|\hat{\pi}_j-\pi_j^0\right\|=C_8 n^{-\kappa}} \bar{\omega}\left(\pi_j\right)-\bar{\omega}\left(\pi_j^0\right)\right) \\
&\leq P\left(\sup_{\left\|\pi_j-\pi_j^0\right\| \leq C_8 n^{-\kappa}}\left|\bar{\omega}_n\left(\pi_j\right)-\bar{\omega}\left(\pi_j\right)\right| \geq \frac{1}{2} c_3 n^{-2\kappa}\right) \\
&\leq P\left(\left|\bar{\omega}_n\left(\pi_j^0\right)-\bar{\omega}\left(\pi_j^0\right)\right| \geq \frac{1}{2} c_3 n^{-2\kappa}\right) \\
&+\left(\sup_{\left\|\pi_j-\pi_j^0\right\| \leq C_8 n^{-\kappa}}\left|\bar{\omega}_n\left(\pi_j\right)-\bar{\omega}_n\left(\pi_j^0\right)-\bar{\omega}\left(\pi_j\right)+\bar{\omega}\left(\pi_j^0\right)\right| \geq \frac{1}{2} c_3 n^{-2\kappa}\right) \\
&:=\Gamma_3+\Gamma_4
\end{aligned}
$$

For Γ_3, by Knight (1998)'s identity and assumption 3.2, for some positive constant c_4, we have

$$
\begin{aligned}
\rho_\tau\left(Y_{i+1}-X_{i,S_j}^T \pi_j^0\right)-\rho_\tau\left(Y_{i+1}\right) &= -X_{i,S_j}^T \pi_j^0\left(\tau-1\left(Y_{i+1} \leq 0\right)\right)+\int_0^{X_{i,S_j}^T \pi_j^0} 1\left(Y_{i+1} \leq s\right)-1\left(Y_{i+1} \leq 0\right) ds \\
&\leq c_4 \sup_{i,j}\left|X_{i,S_j}^T \pi_j^0\right| \\
&\leq c_4 M_3
\end{aligned}
$$

Since $\rho_\tau\left(Y_{i+1}-X_{i,S_j}^T \pi_j^0\right)-\rho_\tau\left(Y_{i+1}\right)$ is an $\alpha-$mixing sequence, using theorem 1 in Merlevède, Peligrad, and Rio (2009), for some universal positive constants C, c_5, we obtain

$$
\begin{aligned}
\Gamma_3 &= P\left(\left|\bar{\omega}_n\left(\pi_j^0\right)-\bar{\omega}\left(\pi_j^0\right)\right| \geq \frac{1}{2} c_3 n^{-2\kappa}\right) \\
&= P\left(\left|\sum_{i=1}^n \rho_\tau\left(Y_{i+1}-X_{i,S_j}^T \pi_j^0\right)-\rho_\tau\left(Y_{i+1}\right)-E\left(\rho_\tau\left(Y_{i+1}-X_{i,S_j}^T \pi_j^0\right)-\rho_\tau\left(Y_{i+1}\right)\right)\right| \geq \frac{1}{2} c_3 n^{1-2\kappa}\right) \\
&\leq e^{-\frac{C \frac{1}{4} c_3^2 n^{2-4\kappa}}{n c_4^2 M^2+c_4 M_3 \frac{1}{2} c_3 n^{1-2\kappa} \log n \log\log n}} \\
&\leq e^{-c_5 n^{1-4\kappa}}
\end{aligned}
$$

For Γ_4, we consider $V_{ij}\left(\pi_j\right)=\rho_\tau\left(Y_{i+1}-X_{i,S_j}^T \pi_j\right)-\rho_\tau\left(Y_{i+1}-X_{i,S_j}^T \pi_j^0\right)$. Using Knight (1998)'s identity again, we show

$$
\begin{aligned}
V_{ij}\left(\pi_j\right)= & -\left(X_{i,S_j}^T \pi_j-X_{i,S_j}^T \pi_j^0\right)\left(\tau-1\left(Y_{i+1}-X_{i,S_j}^T \pi_j^0 \leq 0\right)\right) \\
& +\int_0^{X_{i,S_j}^T \pi_j-X_{i,S_j}^T \pi_j^0} 1\left(Y_{i+1}-X_{i,S_j}^T \pi_j^0 \leq s\right)-1\left(Y_{i+1}-X_{i,S_j}^T \pi_j^0 \leq 0\right) ds
\end{aligned}
$$

Therefore, by assumption 3.2, we have

$$\sup_{\left\|\pi_j-\pi_j^0\right\|\leq C_8 n^{-\kappa}} \left|V_{ij}\left(\pi_j\right)\right| \leq \sup_{\left\|\pi_j-\pi_j^0\right\|\leq C_8 n^{-\kappa}} 2\left|X_{i,S_j}^T\pi_j - X_{i,S_j}^T\pi_j^0\right| \leq 2\sup_{i,j}\left\|X_{i,S_j}\right\|\left\|\pi_j-\pi_j^0\right\| \leq 2r_n^{\frac{1}{2}}M_1C_8n^{-\kappa}$$

Then we can represent

$$\Gamma_4 = P\left(\sup_{\left\|\pi_j-\pi_j^0\right\|\leq C_8 n^{-\kappa}}\left|\frac{1}{n}\sum_{i=1}^n V_{ij}\left(\pi_j\right) - EV_{ij}\left(\pi_j\right)\right| \geq \frac{1}{2}c_3n^{-2\kappa}\right)$$

Now, since we have $V_{ij}\left(\pi_j\right) - EV_{ij}\left(\pi_j\right)$ is β-mixing with mixing coefficient $\beta\left(n\right) \leq e^{-c_0n}$ and $\sup_{\left\|\pi_j-\pi_j^0\right\|\leq C_8 n^{-\kappa}}\left|V_{ij}\left(\pi_j\right)\right| \leq 2r_n^{\frac{1}{2}}M_1C_8n^{-\kappa}$, using theorem 1 in Mohri and Rostamizadeh (2009) §, we have

$$\begin{aligned}
&P\{E\left(\sup_{\left\|\pi_j-\pi_j^0\right\|\leq C_8 n^{-\kappa}}\left|\frac{1}{n}\sum_{i=1}^n V_{ij}\left(\pi_j\right) - EV_{ij}\left(\pi_j\right)\right|\right) \\
&\leq E\sup_{\left\|\pi_j-\pi_j^0\right\|\leq C_8 n^{-\kappa}}\left|\frac{1}{n}\sum_{i=1}^n \epsilon_iV_{ij}\left(\pi_j\right)\right| + 2r_n^{\frac{1}{2}}M_1C_8n^{-\kappa}\frac{\sqrt{\log\frac{2}{\delta-2(\mu-1)e^{-c_0a}}}}{\sqrt{2\mu}}\} \\
\geq& 1-\delta
\end{aligned}$$

for any $\mu, a > 0$ with $2\mu a = n$ and $\delta > 2\left(\mu-1\right)e^{-c_0a}$.

For some universal constant C, c_6, c_7, if we let $\mu = Cn^{\frac{1}{2}}$, $a = \frac{1}{2C}n^{\frac{1}{2}}$, and $\delta = 3\left(\mu-1\right)e^{-c_0a} := c_6n^{\frac{1}{2}}e^{-c_7n^{\frac{1}{2}}}$ we have

§

Remark. For i.i.d case, using the symmetrization technique, we have

$$E\left(\sup_{\left\|\pi_j-\pi_j^0\right\|\leq C_8 n^{-\kappa}}\left|\frac{1}{n}\sum_{i=1}^n V_{ij}\left(\pi_j\right) - EV_{ij}\left(\pi_j\right)\right|\right) \leq 2E\sup_{\left\|\pi_j-\pi_j^0\right\|\leq C_8 n^{-\kappa}}\left|\frac{1}{n}\sum_{i=1}^n \epsilon_iV_{ij}\left(\pi_j\right)\right|$$

where ϵ_i follows the Rademacher random variable.

By the Talagrand contraction principle (page 95 in Ledoux and Talagrand (2013)), we have

$$\begin{aligned}
E\left(\sup_{\left\|\pi_j-\pi_j^0\right\|\leq C_8 n^{-\kappa}}\left|\frac{1}{n}\sum_{i=1}^n V_{ij}\left(\pi_j\right) - EV_{ij}\left(\pi_j\right)\right|\right) &\leq 2E\sup_{\left\|\pi_j-\pi_j^0\right\|\leq C_8 n^{-\kappa}}\left|\frac{1}{n}\sum_{i=1}^n \epsilon_i2r_n^{\frac{1}{2}}M_1C_8n^{-\kappa}\right| \\
&\leq 4r_n^{\frac{1}{2}}M_1c_7n^{-\kappa}\frac{C}{\sqrt{n}} \\
&= Cr_n^{\frac{1}{2}}n^{-\kappa-\frac{1}{2}}
\end{aligned}$$

where the last step uses the fact that $E\left|\frac{1}{n}\sum_{i=1}^n\epsilon_i\right| \leq \frac{C}{\sqrt{n}}$. So we have the expectation bound. The last step to show the probabilistic bound in lemma 3 is based on the Hoeffding's inequality. So the proof for lemma 3 under i.i.d case is complete.

$$
\begin{aligned}
&P\{E\left(\sup_{\left\|\pi_j-\pi_j^0\right\|\le c_7 n^{-\kappa}}\left|\frac{1}{n}\sum_{i=1}^{n}V_{ij}\left(\pi_j\right)-EV_{ij}\left(\pi_j\right)\right|\right)\\
&\le E\sup_{\left\|\pi_j-\pi_j^0\right\|\le C_8 n^{-\kappa}}\left|\frac{1}{n}\sum_{i=1}^{n}\epsilon_i V_{ij}\left(\pi_j\right)\right|+2r_n^{\frac{1}{2}}M_1C_8Cn^{-\kappa}\}\\
\ge&1-c_6n^{\frac{1}{2}}e^{-c_7n^{\frac{1}{2}}}
\end{aligned}
$$

Using the Talagrand contraction principle from page 95 in Ledoux and Talagrand (2013), we know

$$
E\sup_{\left\|\pi_j-\pi_j^0\right\|\le C_8 n^{-\kappa}}\left|\frac{1}{n}\sum_{i=1}^{n}\epsilon_i V_{ij}\left(\pi_j\right)\right|\le c_8 r_n^{\frac{1}{2}}n^{-\kappa-\frac{1}{2}}
$$

for some positive constant c_8

Hence, denoting $c_9=2M_1C_8C$, we have

$$
P\left\{E\left(\sup_{\left\|\pi_j-\pi_j^0\right\|\le C_8 n^{-\kappa}}\left|\frac{1}{n}\sum_{i=1}^{n}V_{ij}\left(\pi_j\right)-EV_{ij}\left(\pi_j\right)\right|\right)\le c_8 r_n^{\frac{1}{2}}n^{-\kappa-\frac{1}{2}}+c_9r_n^{\frac{1}{2}}n^{-\kappa}\right\}\ge 1-c_6n^{\frac{1}{2}}e^{-c_7n^{\frac{1}{2}}}
$$

$$
P\left\{E\left(\sup_{\left\|\pi_j-\pi_j^0\right\|\le C_8 n^{-\kappa}}\left|\frac{1}{n}\sum_{i=1}^{n}V_{ij}\left(\pi_j\right)-EV_{ij}\left(\pi_j\right)\right|\right)\le c_{10}r_n^{\frac{1}{2}}n^{-\kappa}\right\}\ge 1-c_6n^{\frac{1}{2}}e^{-c_7n^{\frac{1}{2}}}
$$

where $c_{10}=\max\left(c_8,c_9\right)$.

Now, we denote $V=\sup_{\left\|\pi_j-\pi_j^0\right\|\le C_8 n^{-\kappa}}\left|\frac{1}{n}\sum_{i=1}^{n}V_{ij}\left(\pi_j\right)-EV_{ij}\left(\pi_j\right)\right|$, then we obtain

$$
\begin{aligned}
\Gamma_4=&P\left(V-EV\ge\frac{1}{2}c_3n^{-2\kappa}-EV\right)\\
&\le P\left(V-EV\ge\frac{1}{2}c_3n^{-2\kappa}-c_{10}r_n^{\frac{1}{2}}n^{-\kappa},EV\le c_{10}r_n^{\frac{1}{2}}n^{-\kappa}\right)\\
&+P\left(V-EV\ge\frac{1}{2}c_3n^{-2\kappa}-EV,EV>c_{10}r_n^{\frac{1}{2}}n^{-\kappa}\right)\\
&\le P\left(V-EV\ge\frac{1}{2}c_3n^{-2\kappa}-c_{10}r_n^{\frac{1}{2}}n^{-\kappa}\right)+P\left(EV>c_{10}r_n^{\frac{1}{2}}n^{-\kappa}\right)\\
&\le P\left(V-EV\ge\frac{1}{2}c_3n^{-2\kappa}-c_{10}r_n^{\frac{1}{2}}n^{-\kappa}\right)+c_6n^{\frac{1}{2}}e^{-c_7n^{\frac{1}{2}}}
\end{aligned}
$$

Using the Hoeffding's inequality under dependent data from van de Geer (2002) with $\sup_{\left\|\pi_j-\pi_j^0\right\|\le C_8 n^{-\kappa}}\left|V_{ij}\left(\pi_j\right)-EV_{ij}\left(\pi_j\right)\right|\le Cr_n^{\frac{1}{2}}n^{-\kappa}$ for some positive constant C on the first term, we show that

$$
\begin{aligned}
\Gamma_4 &\leq e^{-\frac{2\left(n\left(\frac{1}{2}c_3 n^{-2\kappa}-c_{10} r_n^{\frac{1}{2}} n^{-\kappa}\right)\right)^2}{n\left(C r_n^{\frac{1}{2}} n^{-\kappa}\right)^2}} + c_6 n^{\frac{1}{2}} e^{-c_7 n^{\frac{1}{2}}} \\
&\leq e^{-c_{11} r_n^{-1} n^{1-2\kappa}} + c_6 n^{\frac{1}{2}} e^{-c_7 n^{\frac{1}{2}}}
\end{aligned}
$$

Hence,

$$
\begin{aligned}
P\left(\left\|\hat{\pi}_j - \pi_j^0\right\| \geq C_8 n^{-\kappa}\right) &\leq e^{-c_5 n^{1-4\kappa}} + e^{-c_{11} r_n^{-1} n^{1-2\kappa}} + c_6 n^{\frac{1}{2}} e^{-c_7 n^{\frac{1}{2}}} \\
&\leq e^{-C_9 n^{1-4\kappa}} + e^{-C_{10} r_n^{-1} n^{1-2\kappa}} + C_{12} n^{\frac{1}{2}} e^{-C_{11} n^{\frac{1}{2}}}
\end{aligned}
$$

in which we denote $C_9 = c_5$, $C_{10} = c_{11}$, $C_{11} = c_7$ and $C_{12} = c_6$.

endproof

A.3 Proofs of lemma 4

beginproof

We first rewrite

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} \psi_\tau\left(Y_{i+1} - X_{i,S_j}^T \hat{\pi}_j\right)\left(X_{i,j} - X_{i,S_j}^T \hat{\theta}_j\right) - E\left[\psi_\tau\left(Y_{i+1} - X_{i,S_j}^T \pi_j^0\right) X_{i,j}\right] \\
=&\Delta_{1j} + \Delta_{2j} + \Delta_{3j}
\end{aligned}
$$

where

$$
\begin{aligned}
\Delta_{1j} &= \frac{1}{n}\sum_{i=1}^{n} \psi_\tau\left(Y_{i+1} - X_{i,S_j}^T \pi_j^0\right) X_{i,j} - E\left[\psi_\tau\left(Y - X_{i,S_j}^T \pi_j^0\right) X_{i,j}\right] \\
\Delta_{2j} &= \frac{1}{n}\sum_{i=1}^{n} \left(\psi_\tau\left(Y_{i+1} - X_{i,S_j}^T \hat{\pi}_j\right) - \psi_\tau\left(Y_{i+1} - X_{i,S_j}^T \pi_j^0\right)\right) X_{i,j} \\
\Delta_{3j} &= -\frac{1}{n}\sum_{i=1}^{n} \psi_\tau\left(Y_{i+1} - X_{i,S_j}^T \hat{\pi}_j\right) X_{i,S_j}^T \hat{\theta}_j
\end{aligned}
$$

By assumption 3.2, we have $\left|\psi_\tau\left(Y_{i+1} - X_{i,S_j}^T \pi_j^0\right) X_{i,j}\right| \leq M_1$. Since $\psi_\tau\left(Y_{i+1} - X_{i,S_j}^T \pi_j^0\right) X_{i,j}$ is $\alpha-$mixing, using theorem 1 in Merlevède, Peligrad, and Rio (2009), for some positive constants C, c_{12}, c_{13}, we have

$$
\begin{aligned}
P\left(\left|\Delta_{1j}\right| \geq c_{12} n^{-\kappa}\right) &\leq e^{-\frac{C c_{12}^2 n^{2-2\kappa}}{n M_1^2 + M_1 c_{12} \log n \log\log n}} \\
&\leq e^{-c_{13} n^{1-2\kappa}}
\end{aligned}
$$

For Δ_{2j}, we first consider there exists u^* with $\|u^*\| \leq 1$ and some positive constant c_{13}, then we use Cauchy-Schwarz inequality with assumption 3.1 and 3.4

$$
\begin{aligned}
&\sup_{\|u\|\leq 1}\left|\left(\psi_\tau\left(Y_{i+1}-X_{i,S_j}^T\left(\pi_j^0+c_{13}n^{-\kappa}u\right)\right)-\psi_\tau\left(Y_{i+1}-X_{i,S_j}^T\pi_j^0\right)\right)X_{i,j}\right|\\
=&\sup_{\|u\|\leq 1}\left|\left(\psi_\tau\left(Y_{i+1}-X_{i,S_j}^T\left(\pi_j^0+c_{13}n^{-\kappa}u^*\right)\right)-\psi_\tau\left(Y_{i+1}-X_{i,S_j}^T\pi_j^0\right)\right)X_{i,j}\right|\\
\leq&\sup_{\|u\|\leq 1}\left|\int_{X_{i,S_j}^T\pi_j^0}^{X_{i,S_j}^T\left(\pi_j^0+c_{13}n^{-\kappa}u^*\right)} f_{Y_{i+1}|X_i}\left(y\right)dy\right|\left|X_{i,j}\right|\\
\leq& c_{13}n^{-\kappa}\left|X_{i,S_j}^T u^*\right|\\
\leq& c_{13}M_1^2 r_n^{\frac{1}{2}}n^{-\kappa}
\end{aligned}
$$

We define

$$
\Pi_{ij}=\sup_{\|u\|\leq 1}\left|\left(\psi_\tau\left(Y_{i+1}-X_{i,S_j}^T\left(\pi_j^0+c_{13}n^{-\kappa}u\right)\right)-\psi_\tau\left(Y_{i+1}-X_{i,S_j}^T\pi_j^0\right)\right)X_{i,j}\right|
$$

Using theorem 1 in Merlevède, Peligrad, and Rio (2009), for some positive constants C, c_{14}, c_{15}we obtain

$$
\begin{aligned}
&P\left(\left|\frac{1}{n}\sum_{i=1}^n \Pi_{ij}-E\Pi_{ij}\right|\geq c_{14}r_n^{\frac{1}{2}}n^{-\kappa}\right)\\
\leq& e^{-\frac{C\left(c_{14}r_n^{\frac{1}{2}}n^{1-\kappa}\right)^2}{n\left(c_{13}M_1^2r_n^{\frac{1}{2}}n^{-\kappa}\right)^2+\left(c_{13}M_1^2r_n^{\frac{1}{2}}n^{-\kappa}\right)c_{14}r_n^{\frac{1}{2}}n^{-\kappa}\log n\log\log n}}\\
\leq& e^{-c_{15}n}
\end{aligned}
$$

Therefore, since we have $E\Pi_{ij}\leq c_{13}M_1^2r_n^{\frac{1}{2}}n^{-\kappa}$, we show that

$$
\begin{aligned}
&P\left(\left|\frac{1}{n}\sum_{i=1}^n \Pi_{ij}\right|\geq c_{16}r_n^{\frac{1}{2}}n^{-\kappa}\right)\\
\leq& P\left(\left|\frac{1}{n}\sum_{i=1}^n \Pi_{ij}-E\Pi_{ij}\right|+E\Pi_{ij}\geq c_{16}r_n^{\frac{1}{2}}n^{-\kappa}\right)\\
\leq& P\left(\left|\frac{1}{n}\sum_{i=1}^n \Pi_{ij}-E\Pi_{ij}\right|\geq c_{17}r_n^{\frac{1}{2}}n^{-\kappa}\right)\\
\leq& e^{-c_{18}n}
\end{aligned}
$$

for some positive constants c_{16}, c_{17}, c_{18} and the last inequality follows the same procedure as above.

Hence, we have for some positive constant c_{19}, c_{20}

$$
P\left(\left|\Delta_{2j}\right|\geq c_{19}r_n^{\frac{1}{2}}n^{-\kappa}\right)
$$

$$
\begin{aligned}
&\leq P\left(|\Delta_{2j}| \geq c_{19} r_n^{\frac{1}{2}} n^{-\kappa}, \left\|\hat{\pi}_j - \pi_j^0\right\| < c_{13} n^{-\kappa}\right) \\
&\quad + P\left(|\Delta_{2j}| \geq c_{19} r_n^{\frac{1}{2}} n^{-\kappa}, \left\|\hat{\pi}_j - \pi_j^0\right\| \geq c_{13} n^{-\kappa}\right) \\
&\leq P\left(\left|\frac{1}{n}\sum_{i=1}^{n} \Pi_{ij}\right| \geq c_{19} r_n^{\frac{1}{2}} n^{-\kappa}\right) + P\left(\left\|\hat{\pi}_j - \pi_j^0\right\| \geq c_{13} n^{-\kappa}\right) \\
&\leq e^{-c_{20} n} + e^{-C_9 n^{1-4\kappa}} + e^{-C_{10} r_n^{-1} n^{1-2\kappa}} + C_{12} n^{\frac{1}{2}} e^{-C_{11} n^{\frac{1}{2}}}
\end{aligned}
$$

where the last two inequalities rely on the result of lemma 3 with $C_8 = c_{13}$.

For Δ_{3j}, we first define

$$
g(\pi_j) = \frac{1}{n}\sum_{i=1}^{n} \rho_\tau\left(Y_{i+1} - X_{i,S_j}^T \pi_j\right)
$$

and its subdifferential

$$
\partial g(\pi_j) = \left\{\partial g_k(\pi_j) : k \in \{0\} \cup S_j\right\}^T
$$

with

$$
\partial g_k(\pi_j) = -\frac{1}{n}\sum_{i=1}^{n} \psi_\tau\left(Y_{i+1} - X_{i,S_j}^T \pi_j\right) X_{i,k} - \frac{1}{n}\sum_{i=1}^{n} 1\left(Y_{i+1} = X_{i,S_j}^T \pi_j\right) v_i X_{i,j}
$$

where $v_i \in [\tau - 1, \tau]$.

Since $\hat{\pi}_j = \arg\min \frac{1}{n}\sum_{i=1}^{n} \rho_\tau\left(Y_{i+1} - X_{i,S_j}^T \pi_j\right)$, we know there exists $v_i^* \in [\tau - 1, \tau]$ such that $\partial g_k(\hat{\pi}_j) = 0$. Therefore, we have

$$
\Delta_{3j} = \frac{1}{n}\sum_{i=1}^{n} 1\left(Y_{i+1} = X_{i,S_j}^T \hat{\pi}_j\right) v_i^* X_{i,S_j}^T \hat{\theta}_j
$$

and. by the triangular inequality and assumption 3.2, we know

$$
\begin{aligned}
&\left|\frac{1}{n}\sum_{i=1}^{n} 1\left(Y_{i+1} = X_{i,S_j}^T \hat{\pi}_j\right) v_i^* X_{i,S_j}^T \hat{\theta}_j\right| \\
\leq& \frac{1}{n}\sum_{i=1}^{n} 1\left(Y_{i+1} = X_{i,S_j}^T \hat{\pi}_j\right) \left|X_{i,S_j}^T \hat{\theta}_j\right| \\
\leq& \frac{1}{n}\sum_{i=1}^{n} 1\left(Y_{i+1} = X_{i,S_j}^T \hat{\pi}_j\right) \left(\left|X_{i,S_j}^T \theta_j^0\right| + \left|X_{i,S_j}^T \hat{\theta}_j - X_{i,S_j}^T \theta_j^0\right|\right) \\
\leq& \frac{1}{n}\sum_{i=1}^{n} 1\left(Y_{i+1} = X_{i,S_j}^T \hat{\pi}_j\right) \left(M_2 + r_n^{\frac{1}{2}} M_1 \left\|\hat{\theta}_j - \theta_j^0\right\|\right)
\end{aligned}
$$

Using lemma 2 and let $\delta_n = \left(\frac{r_n}{n}\right)^{-1}$, we know for some positive constant c_{21}, c_{22}, c_{23}

$$
P\left(\left\|\hat{\theta}_j - \theta_j^0\right\| \geq c_{21}\right) \leq C_7 r_n e^{-\frac{C_3 \delta_n^2}{n}} + e^{-\frac{C_5 n}{r_n^2}} = c_{23} r_n e^{-c_{22}\frac{n}{r_n^2}}
$$

Hence, we have

$$P\left(M_2 + r_n^{\frac{1}{2}} M_1 \left\|\hat{\theta}_j - \theta_j^0\right\| \geq M_2 + c_{21} M_1 r_n^{\frac{1}{2}}\right) \leq c_{23} r_n e^{-c_{22}\frac{n}{r_n^2}}$$

Since $P\left(Y_{i+1} = X_{i,S_j}^T \hat{\pi}_j\right) = 0$ and $P\left(\frac{1}{n}\sum_{i=1}^n 1\left(Y_{i+1} = X_{i,S_j}^T \hat{\pi}_j\right) > \varepsilon\right) = 0$ for any $\varepsilon > 0$. We let $\varepsilon = \frac{1}{r_n^{\frac{1}{2}} n}$, then we have

$$\begin{aligned}
&P\left(|\Delta_{3j}| \geq \frac{1}{r_n^{\frac{1}{2}} n}\left(M_2 + c_{21} M_1 r_n^{\frac{1}{2}}\right)\right) \\
\leq &P\left(\left|\frac{1}{n}\sum_{i=1}^n 1\left(Y_{i+1} = X_{i,S_j}^T \hat{\pi}_j\right)\right| \left|\left(M_2 + r_n^{\frac{1}{2}} M_1 \left\|\hat{\theta}_j - \theta_j^0\right\|\right)\right| \geq \frac{1}{r_n^{\frac{1}{2}} n}\left(M_2 + c_{21} M_1 r_n^{\frac{1}{2}}\right)\right) \\
\leq &c_{23} r_n e^{-c_{22}\frac{n}{r_n^2}}
\end{aligned}$$

Finally, we combine the probabilistic bounds for Δ_{1j}, Δ_{2j} and Δ_{3j}, so we obtain

$$\begin{aligned}
&P\left\{\left|\frac{1}{n}\sum_{i=1}^n \psi_\tau\left(Y_{i+1} - X_{i,S_j}^T \hat{\pi}_j\right)\left(X_{i,j} - X_{i,S_j}^T \hat{\theta}_j\right) - E\left[\psi_\tau\left(Y_{i+1} - X_{i,S_j}^T \pi_j^0\right) X_{i,j}\right]\right| \geq C_{13} r_n^{\frac{1}{2}} n^{-\kappa}\right\} \\
\leq &P\{\left|\frac{1}{n}\sum_{i=1}^n \psi_\tau\left(Y_{i+1} - X_{i,S_j}^T \hat{\pi}_j\right)\left(X_{i,j} - X_{i,S_j}^T \hat{\theta}_j\right) - E\left[\psi_\tau\left(Y_{i+1} - X_{i,S_j}^T \pi_j^0\right) X_{i,j}\right]\right| \\
&\geq c_{12} n^{-\kappa} + c_{22} r_n^{\frac{1}{2}} n^{-\kappa} + \frac{1}{r_n^{\frac{1}{2}} n}\left(M_2 + c_{21} M_1 r_n^{\frac{1}{2}}\right)\} \\
\leq &e^{-c_{13} n^{1-2\kappa}} + e^{-c_{20} n} + e^{-C_9 n^{1-4\kappa}} + e^{-C_{10} r_n^{-1} n^{1-2\kappa}} + C_{12} n^{\frac{1}{2}} e^{-C_{11} n^{\frac{1}{2}}} + c_{23} r_n e^{-c_{22}\frac{n}{r_n^2}} \\
\leq &C_{14} e^{-C_9 n^{1-4\kappa}} + e^{-C_{10} r_n^{-1} n^{1-2\kappa}} + C_{12} n^{\frac{1}{2}} e^{-C_{11} n^{\frac{1}{2}}} + C_{16} r_n e^{-C_{15}\frac{n}{r_n^2}}
\end{aligned}$$

where $C_{15} = c_{22}$ and $C_{16} = c_{23}$.

endproof

A.4 Proofs of lemma 5

beginproof

Since $\sigma_j^2 = var\left(X_{i,j} - X_{i,S_j}^T \theta_j^0\right)$ and $\hat{\sigma}_j^2 = \frac{1}{n}\sum_{i=1}^n \left(X_{i,j} - X_{i,S_j}^T \hat{\theta}_j\right)^2$, we can rewrite

$$\left|\hat{\sigma}_j^2 - \sigma_j^2\right| \leq \Gamma_{5j} + \Gamma_{6j}\left(\hat{\theta}_j\right)$$

where

$$\begin{aligned}
\Gamma_{5j} &= \left|\frac{1}{n}\sum_{i=1}^n \left(X_{i,j} - X_{i,S_j}^T \theta_j^0\right)^2 - E\left(X_{i,j} - X_{i,S_j}^T \theta_j^0\right)^2\right| \\
\Gamma_{6j}\left(\hat{\theta}_j\right) &= \left|\frac{1}{n}\sum_{i=1}^n \left(X_{i,j} - X_{i,S_j}^T \hat{\theta}_j\right)^2 - \frac{1}{n}\sum_{i=1}^n \left(X_{i,j} - X_{i,S_j}^T \theta_j^0\right)^2\right|
\end{aligned}$$

For Γ_{5j}, we use theorem 1 in Merlevède, Peligrad, and Rio (2009) and we know

$$\left|\left(X_{i,j} - X_{i,S_j}^T \theta_j^0\right)^2 - E\left(X_{i,j} - X_{i,S_j}^T \theta_j^0\right)^2\right| \leq M$$

for some universal constant M by assumption 3.2. So we have

$$\begin{aligned} P\left(\Gamma_{5j} \geq c_{24} r_n^{\frac{1}{2}} n^{-\kappa}\right) &\leq e^{-\frac{C\left(nc_{24} r_n^{\frac{1}{2}} n^{-\kappa}\right)^2}{-nM + Mnc_{24} r_n^{\frac{1}{2}} n^{-\kappa} \log n \log\log n}} \\ &\leq e^{-c_{25} r_n n^{1-2\kappa}} \end{aligned}$$

for some constants C, c_{24}, c_{25}.

For $\Gamma_{6j}\left(\hat{\theta}_j\right)$, we can rewrite it as

$$\begin{aligned} \Gamma_{6j}\left(\hat{\theta}_j\right) &= \left|\frac{1}{n}\sum_{i=1}^n \left(\left(X_{i,j} - X_{i,S_j}^T \hat{\theta}_j\right) + \left(X_{i,j} - X_{i,S_j}^T \theta_j^0\right)\right)\left(X_{i,S_j}^T\left(\hat{\theta}_j - \theta_j^0\right)\right)\right| \\ &= \left|\frac{1}{n}\sum_{i=1}^n \left(2\left(X_{i,j} - X_{i,S_j}^T \theta_j^0\right) + X_{i,S_j}^T\left(\hat{\theta}_j - \theta_j^0\right)\right)\left(X_{i,S_j}^T\left(\hat{\theta}_j - \theta_j^0\right)\right)\right| \\ &\leq \left|\frac{1}{n}\sum_{i=1}^n \left(2\left(X_{i,j} - X_{i,S_j}^T \theta_j^0\right) X_{i,S_j}^T\left(\hat{\theta}_j - \theta_j^0\right)\right)\right| + \left(\hat{\theta}_j - \theta_j^0\right)^T \left(\frac{1}{n}\sum_{i=1}^n X_{i,S_j} X_{i,S_j}^T\right)\left(\hat{\theta}_j - \theta_j^0\right) \\ &:= \Gamma_{7j} + \Gamma_{8j}\left(\hat{\theta}_j\right) \end{aligned}$$

For $\Gamma_{8j}\left(\hat{\theta}_j\right)$, from the proof of lemma 2, we know

$$P\left(\left\|\frac{1}{n}\sum_{i=1}^n X_{i,S_j} X_{i,S_j}^T - E\left(X_{i,S_j} X_{i,S_j}^T\right)\right\| \geq r_n \frac{c_1 \delta_1}{n}\right) \leq e^{-\frac{C_4 \delta_1^2}{n}}$$

Together with the inequality $\left|\lambda_{\max}(A) - \lambda_{\max}(B)\right| \leq \|A - B\|$ for symmetric matrices A and B, we obtain

$$P\left(\left\|\lambda_{\max}\left(\frac{1}{n}\sum_{i=1}^n X_{i,S_j} X_{i,S_j}^T\right) - \lambda_{\max}\left(E\left(X_{i,S_j} X_{i,S_j}^T\right)\right)\right\| \geq r_n \frac{c_1 \delta_1}{n}\right) \leq e^{-\frac{c_4 \delta_1^2}{n}}$$

Similarly, we let $\delta_1 = \frac{c_{24} n}{c_1 r_n} m \leq \frac{c_{24} n}{c_1 r_n} \lambda_{\max}\left(E\left(X_{i,S_j} X_{i,S_j}^T\right)\right)$ for some constant c_{24} and denote $c_{25} = \frac{c_{24}^2}{c_1^2 c_4^{\frac{1}{2}}}$, then we obtain

$$P\left(\left|\lambda_{\max}\left(\frac{1}{n}\sum_{i=1}^n X_{i,S_j} X_{i,S_j}^T\right) - \lambda_{\max}\left(E\left(X_{i,S_j} X_{i,S_j}^T\right)\right)\right| \geq c_{24} \lambda_{\max}\left(E\left(X_{i,S_j} X_{i,S_j}^T\right)\right)\right) \leq e^{-\frac{c_{25} n}{r_n^2}}$$

which indicates

$$P\left(\lambda_{\max}\left(\frac{1}{n}\sum_{i=1}^{n}X_{i,S_j}X_{i,S_j}^T\right)\geq(1+c_{24})\,M\right)\leq e^{-\frac{c_{25}n}{r_n^2}}$$

by assumption 3.3.

From lemma 2, we know

$$P\left(\left\|\hat{\theta}_j-\theta_j^0\right\|\geq C_6\frac{r_n\delta_n}{n}\right)\leq C_7r_ne^{-\frac{C_3\delta_n^2}{n}}+e^{-\frac{C_5n}{r_n^2}}$$

We let $\delta_n=c_{26}r_n^{-\frac{1}{2}}n^{1-\kappa}$ for some positive constant c_{26} and we obtain

$$P\left(\left\|\hat{\theta}_j-\theta_j^0\right\|\geq C_6c_{26}r_n^{\frac{1}{2}}n^{-\kappa}\right)\leq C_7r_ne^{-\frac{C_3c_{26}^2n^{1-2\kappa}}{r_n}}+e^{-\frac{C_5n}{r_n^2}}$$

Therefore, we can show

$$\begin{aligned}&P\left(\Gamma_{8j}\left(\hat{\theta}_j\right)\geq(1+c_{24})\,M\left(C_6c_{26}r_n^{\frac{1}{2}}n^{-\kappa}\right)^2\right)\\ \leq&e^{-\frac{c_{25}n}{r_n^2}}+C_7r_ne^{-\frac{C_3c_{26}^2n^{1-2\kappa}}{r_n}}+e^{-\frac{C_5n}{r_n^2}}\end{aligned}$$

As for Γ_{7j}, let $\theta_j=\theta_j^0+c_{27}r_n^{\frac{1}{2}}n^{-\kappa}u$ where $c_{27}=C_6c_{26}$, $u\in\mathbb{R}^{|S_j|}$ and $\|u\|\leq 1$. We then define

$$\Phi_j\left(u\right)=\frac{1}{n}\sum_{i=1}^{n}\left(X_{i,j}-X_{i,S_j}^T\theta_j^0\right)X_{i,S_j}^T\left(\hat{\theta}_j-\theta_j^0\right)$$

From assumption 3.2, we know $\left|\left(X_{i,j}-X_{i,S_j}^T\theta_j^0\right)X_{i,S_j}^T\left(\hat{\theta}_j-\theta_j^0\right)\right|\leq Mr_nn^{-\kappa}$ for some universal constant M. Using theorem 1 in Merlevède, Peligrad, and Rio (2009), for some positive constants C,c_{28},c_{29}, we can obtain

$$\begin{aligned}P\left(\left|\Phi_j\left(u\right)\right|\geq c_{28}r_n^{\frac{1}{2}}n^{-\kappa}\right)&\leq e^{-\frac{C\left(c_{28}^2nr_n^{\frac{1}{2}}n^{-\kappa}\right)^2}{n(Mr_nn^{-\kappa})^2+Mr_nn^{-\kappa}c_{28}nr_n^{\frac{1}{2}}n^{-\kappa}\log n\log\log n}}\\&\leq e^{-\frac{c_{29}n}{r_n}}\end{aligned}$$

Then we partition $\Lambda=\left\{u:u\in\mathbb{R}^{|S_j|},\|u\|\leq 1\right\}$ as a union of l_n disjoint subsets $\Lambda_1,\ldots,\Lambda_{l_n}$. Each subset has equal spaces in each direction of u. Therefore, we know $\sup_{u,u'\in\Lambda_k}\left\|u-u'\right\|\leq\frac{\sqrt{r_n}}{l_n^{\frac{1}{|S_j|}}}$ for all $k\in\{1,\ldots,l_n\}$. Hence, for $u_k\in\Lambda_k$, we have

$$\sup_{u\in\Lambda}\left|\Phi_j\left(u\right)\right|\leq\sup_k\left|\Phi_j\left(u_k\right)\right|+\sup_k\sup_{u\in\Lambda_k}\left|\Phi_j\left(u\right)-\Phi_j\left(u_k\right)\right|$$

By the previous inequality and the Bonferroni bound, we show that

$$P\left(\sup_{k}\left|\Phi_j\left(u_k\right)\right| \geq c_{28} r_n^{\frac{1}{2}} n^{-\kappa}\right) \leq l_n e^{-\frac{c_{29} n}{r_n}}$$

Moreover, we know

$$\begin{aligned}
&\sup_{k} \sup_{u \in \Lambda_k}\left|\Phi_j(u)-\Phi_j\left(u_k\right)\right| \\
=&\sup_{k} \sup_{u \in \Lambda_k}\left|\frac{1}{n} \sum_{i=1}^{n}\left(X_{i,j}-X_{i,S_j}^T \theta_j^0\right) X_{i,S_j}^T\left(c_{27} r_n^{\frac{1}{2}} n^{-\kappa}\left(u-u_k\right)\right)\right| \\
\leq&\left(M_1+M_2\right) M_1 c_{27} r_n n^{-\kappa} \sup_{k} \sup_{u \in \Lambda_k}\left\|u-u_k\right\| \\
\leq& c_{30} r_n^{\frac{3}{2}} n^{-\kappa} \frac{1}{l_n^{\frac{1}{\left|S_j\right|}}}
\end{aligned}$$

where $c_{30}=\left(M_1+M_2\right) M_1 c_{27}$.

Letting $l_n^{\frac{1}{\left|S_j\right|}}=r_n$, we have

$$\sup_{k} \sup_{u \in \Lambda_k}\left|\Phi_j(u)-\Phi_j\left(u_k\right)\right| \leq c_{30} r_n^{\frac{1}{2}} n^{-\kappa}$$

Hence, denoting $c_{31}=c_{28}+c_{30}$, we obtain

$$P\left(\sup_{u \in \Lambda}\left|\Phi_j(u)\right| \geq c_{31} r_n^{\frac{1}{2}} n^{-\kappa}\right) \leq r_n^{r_n} e^{-\frac{c_{29} n}{r_n}} \leq e^{-\frac{c_{29} n}{r_n}+r_n \log r_n} \leq e^{-\frac{c_{32} n}{r_n}}$$

which indicates

$$P\left(\left|\Gamma_{7j}\right| \geq c_{33} r_n^{\frac{1}{2}} n^{-\kappa}\right) \leq e^{-\frac{c_{32} n}{r_n}}$$

for some positive constant c_{33} under the event $\left\|\hat{\theta}_j-\theta_j^0\right\| \leq c_{27} r_n^{\frac{1}{2}} n^{-\kappa}$.

Hence, by the probabilistic bound of $\left\|\hat{\theta}_j-\theta_j^0\right\|$, we have

$$P\left(\left|\Gamma_{7j}\right| \geq c_{33} r_n^{\frac{1}{2}} n^{-\kappa}\right) \leq e^{-\frac{c_{32} n}{r_n}}+C_7 r_n e^{-\frac{C_3 c_{26}^2 n^{1-2\kappa}}{r_n}}+e^{-\frac{C_5 n}{r_n^2}}$$

By the assumption that $r_n^{\frac{1}{2}} n^{-\kappa}=o(1)$, combining the results of Γ_{7j} and $\Gamma_{8j}\left(\hat{\theta}_j\right)$, for some positive constant $c_{34}, c_{35}, c_{36}, c_{37}, c_{38}$, we obtain

$$\begin{aligned}
&P\left(\left|\Gamma_{6j}\left(\hat{\theta}_j\right)\right| \geq c_{34} r_n^{\frac{1}{2}} n^{-\kappa}\right) \\
\leq& e^{-\frac{c_{32} n}{r_n}}+C_7 r_n e^{-\frac{C_3 c_{26}^2 n^{1-2\kappa}}{r_n}}+e^{-\frac{C_5 n}{r_n^2}}+e^{-\frac{c_{25} n}{r_n^2}}+C_7 r_n e^{-\frac{C_3 c_{26}^2 n^{1-2\kappa}}{r_n}}+e^{-\frac{C_5 n}{r_n^2}}
\end{aligned}$$

$$\leq e^{-\frac{c_{32}n}{r_n}} + c_{35}e^{-\frac{c_{36}n}{r_n^2}} + c_{37}r_n e^{-\frac{c_{38}n^{1-2\kappa}}{r_n}}$$

Therefore, we have

$$P\left(\left|\hat{\sigma}_j^2 - \sigma_j^2\right| \geq C_{17}r_n^{\frac{1}{2}}n^{-\kappa}\right) \leq e^{-\frac{C_{18}n}{r_n}} + C_{20}e^{-\frac{C_{19}n}{r_n^2}} + C_{22}r_n e^{-\frac{C_{21}n^{1-2\kappa}}{r_n}}$$

In the end, by 3.4, we know $r_n^{\frac{1}{2}}n^{-\kappa} = o(1)$ so $C_{17}r_n^{\frac{1}{2}}n^{-\kappa} \leq a\sigma_j^2$ for some positive number $a \in (0,1)$, hence we have

$$P\left(\left|\hat{\sigma}_j^2 - \sigma_j^2\right| \geq a\sigma_j^2\right) \leq e^{-\frac{C_{18}n}{r_n}} + C_{20}e^{-\frac{C_{19}n}{r_n^2}} + C_{22}r_n e^{-\frac{C_{21}n^{1-2\kappa}}{r_n}}$$

endproof

References

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

Adrian, T., F. Grinberg, N. Liang, and S. Malik (2020): "The term structure of growth-at-risk," .

Belloni, A., and V. Chernozhukov (2011): "L1-penalized quantile regression in high-dimensional sparse models," *The Annals of Statistics*, 39(1), 82–130.

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

Chiou, H.-T., M. Guo, and C.-K. Ing (2020): "Variable selection for high-dimensional regression models with time series and heteroscedastic errors," *Journal of Econometrics*, 216(1), 118–136.

Fan, J., and J. Lv (2008): "Sure independence screening for ultrahigh dimensional feature space," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(5), 849–911.

Figueres, J. M., and M. Jarociński (2020): "Vulnerable growth in the euro area: Measuring the financial conditions," *Economics Letters*, p. 109126.

Gertler, M., and S. Gilchrist (2018): "What happened: Financial factors in the great recession," *Journal of Economic Perspectives*, 32(3), 3–30.

Hasenzagl, T., L. Reichlin, and G. Ricco (2020): "Financial variables as predictors of real growth vulnerability," .

Ing, C.-K. (2020): "Model selection for high-dimensional linear regression with dependent observations," *Annals of Statistics*, 48(4), 1959–1980.

Knight, K. (1998): "Limiting distributions for L1 regression estimators under general conditions," *Annals of statistics*, pp. 755–770.

Kong, Y., Y. Li, and D. Zerom (2019): "Screening and selection for quantile regression using an alternative measure of variable importance," *Journal of Multivariate Analysis*, 173, 435–455.

Ledoux, M., and M. Talagrand (2013): *Probability in Banach Spaces: isoperimetry and processes.* Springer Science & Business Media.

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

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

McCracken, M., and S. Ng (2020): "Fred-qd: A quarterly database for macroeconomic research," Discussion paper, National Bureau of Economic Research.

McCracken, M. W., and S. Ng (2016): "FRED-MD: A monthly database for macroeconomic research," *Journal of Business & Economic Statistics*, 34(4), 574–589.

Merlevède, F., M. Peligrad, and E. Rio (2009): "Bernstein inequality and moderate deviations under strong mixing conditions," in *High dimensional probability V: the Luminy volume*, pp. 273–292. Institute of Mathematical Statistics.

Mohri, M., and A. Rostamizadeh (2009): "Rademacher complexity bounds for non-iid processes," .

Plagborg-Møller, M., L. Reichlin, G. Ricco, and T. Hasenzagl (2020): "When is growth at risk?," *Brookings Papers on Economic Activity*, 2020(1), 167–229.

Sancetta, A., et al. (2016): "Greedy algorithms for prediction," *Bernoulli*, 22(2), 1227–1277.

Schmidt, L. (2022): "Climbing and falling off the ladder: Asset pricing implications of labor market event risk," *Available at SSRN 2471342*.

van de Geer, S. A. (2002): "On Hoeffdings inequality for dependent random variables," in *Empirical process techniques for dependent data*, pp. 161–169. Springer.

Wang, H. (2009): "Forward regression for ultra-high dimensional variable screening," *Journal of the American Statistical Association*, 104(488), 1512–1524.

Wu, W. B. (2005): "Nonlinear system theory: Another look at dependence," *Proceedings of the National Academy of Sciences*, 102(40), 14150–14154.

Yousuf, K. (2018): "Variable screening for high dimensional time series," *Electronic Journal of Statistics*, 12(1), 667–702.

Zhang, S., and Y. Zhou (2018): "Variable screening for ultrahigh dimensional heterogeneous data via conditional quantile correlations," *Journal of Multivariate Analysis*, 165, 1–13.