Partial Time-Varying Regression Modelling under General Heterogeneity

Liudas Giraitis, George Kapetanios, Yufei Li, Tien Chuong Nguyen

Working Paper No. 985 December 2024 ISSN 1473-0278

School of Economics and Finance

Queen Mary
University of London

Partial Time-Varying Regression Modelling under General Heterogeneity

Liudas Giraitis[1], George Kapetanios[2], Yufei Li[2], Tien Chuong Nguyen[3]

[1]Queen Mary University of London

[2]King's College London

[3]Vietnam National University

December 18, 2024

Abstract

This paper explores a semiparametric version of a time-varying regression, where a subset of the regressors have a fixed coefficient and the rest a time-varying one. We provide an estimation method and establish associated theoretical properties of the estimates and standard errors in extended for heterogeneity regression space. In particular, we show that the estimator of the fixed regression coefficient preserves the parametric rate of convergence, and that, despite of general heterogenous environment, the asymptotic normality property for components of regression parameters can be established and the estimators of standard errors have the same form as those given by White (1980). The theoretical properties of the estimator and good finite sample performance are confirmed by Monte Carlo experiments and illustrated by an empirical example on forecasting.

JEL classification: C13, C14, C50

Keywords: structural change, time-varying parameters, non-parametric estimation

1 Introduction

Many empirical studies in applied economics and finance rely on regressions with stationary or mixing covariates. The literature on structural change in regression parameters is vast. A more recent strand of research turns the attention to regression space. The investigation of regression space that permits estimation and inference on regression models with fixed or stochastic parameters, has received increasing interest, see e.g. Phillips, Li, Gao (2017), Hu, Kasparis, Wang (2024), and Giraitis, Kapetanios, Li (2024) and references therein. Phillips, Li, Gao (2017) study extensions of regression modelling to non-stationary region for $I(1)$ covariates, and theoretical framework in Hu, Kasparis, Wang (2024) allows for a wide range of stationary regressors, which can be strongly dependent, non-mixing and may exhibit long memory. Giraitis, Kapetanios, Li (2024) contribute by shifting the focus from structural change in parameters to regression covariates which is a new addition to regression literature. They show that regression space permits both a stationary covariate η_t and its linear transform $\mu_t + g_t\eta_t$ when the shift μ_t and the scale g_t variables are independent of η_t, and other assumptions on μ_t, g_t are minimal. Such covariates still allow building confidence intervals and estimation of standard errors for components of regression parameter (fixed or time-varying).

The time series and regression modelling with deterministic smoothly varying parameters has a long pedigree in statistics, starting with the work of Priestley (1965) and has been followed up by Robinson (1989), Robinson (1991), Dahlhaus (1997), Chen and Hong (2012) and others. This approach, while popular in statistics, has been less prominent in applied macroeconometrics where random coefficient models dominate, see e.g. Muller and Watson (2008), Kapetanios and Yates (2008) and Muller and Petalas (2010). The estimation of locally stationary time series models with deterministic parameters is well investigated, see e.g. Dahlhaus and Giraitis (1998). Building on previous work, Giraitis, Kapetanios, and Yates (2014) have developed a framework for the estimation of time series models with smoothly-varying stochastic parameters.

A number of tests for the presence of parameter breaks exist in the literature, see Chow (1960), Brown, Durbin, Evans (1974), Ploberger and Kramer (1992), and for on going smooth change, see Kristensen (2012), Chen and Hong (2012) and Chen (2015). Testing for change of time varying deterministic parameter was investigated in the recent work by Hu, Kasparis, Wang (2024).

This research builds on our previous work, Giraitis, Kapetanios, Li (2024), on Robinson (1988) who introduced semiparametric regression modelling and Kristensen (2012) who provided parameter estimates and a test for stability of time-varying parameter in regression model. This paper contributes by developing estimation theory for *partially time-varying regression* (PTVR) model in extended for heterogeneity regression space. PTVR methods allows simultaneous estimation of the fixed regression parameter (with parameteric rate) and time-varying parameter (with non-parametric rate) which can be deterministic or stochastic. Simultaneous estimation of regression parameters requires smooth evolution of the time varying parameter and scale factors. This leads to significant challenges and differences from Giraitis, Kapetanios, Li (2024), where regression models with fixed and time-varying parameters were considered separately.

We show rigorously that the PTVR estimates of a single component of parameters have desirable theoretical properties of consistency and asymptotic normality. Under our general framework, this requires significant technical effort. Further theoretical and methodological contribution is showing that the estimators of the robust standard errors have the same form as those given by White (1980), and in earlier work by Eicker (1963). The conditions we use are rather weak. Under general heterogeneity covered by our regression setting, stochastic parameters, regressors and scale factors require 8-th moments, no mixing assumption is used and regressors can take a very general non-stationary form. It is worth noting that a model specification in Kristensen (2012) uses non-stationary mixing covariates, and mixing assumption is common in modelling smooth structural change, see e.g. Chen and Hong (2012), Giraitis, Kapetanios, and Marcellino (2021) and Dendramis, Giraitis, and Kapetanios (2021).

The paper is structured as follows. In Section 2, we present the main results. In Section 3 we use Monte Carlo simulations to show that the theoretical properties extend to finite samples. Section 4 provides an empirical illustration and Section 5 concludes. Proofs and further simulation findings are provided in the Online Supplement.

Below $\to_d$, $\to_p$ stand for convergence in distribution and probability. We use notation $||A||$ to denote Frobenius norm of a matrix A. We will write $a_n \asymp_p b_n$ if $a_n = O_p(b_n)$ and $b_n = O_p(a_n)$.

2 Partially time-varying regression

In this paper we discuss the estimation of a partially time-varying regression model (PTVR) for a univariate variable

$$y_t = \alpha' x_t + \beta_t' z_t + u_t, \quad t = 1, ..., n, \tag{1}$$

which combines a regression model with a fixed parameter and a regression model with a time-varying parameter. Regressors in (1) can be divided into two groups,

$x_t = (x_{1t}, .., x_{qt})'$ and $z_t = (z_{1t}, .., z_{pt})'$, where regression parameter $\alpha = (\alpha_1, .., \alpha_q)'$ at x_t is fixed and regression parameter $\beta_t = (\beta_{1t}, .., \beta_{pt})'$ at z_t is time varying. We suppose that the regression noise u_t is serially uncorrelated. Other assumptions on x_t, z_t and u_t will be specified later.

The setting of our PTVR model is indebted to innovatory work by Robinson (1988) on $\sqrt{n}$-consistent semiparametric regression and also hinges upon semiparametric analysis and regression settings by Kristensen (2012) and Fan and Huang (2005). It allows for general heterogeneity in regressors and noise and structural change of time-varying regression coefficient over time.

Our research builds on the recent work by Giraitis, Kapetanios, Li (2024) which extends the existing literature on regression estimation in two directions. Firstly, it shows that regression estimation of fixed parameter remains valid in very general heterogenous environment, and under weak conditions it still permits developing asymptotic theory, computation of robust standard errors and building confidence intervals for a single component α_k of the fixed regression parameter α. Secondly, it shows that the same general environment allows point-wise kernel estimation of the components β_{kt} of time-varying parameter β_t. β_t is assumed to be smoothly-varying and it can be stochastic or deterministic.

In this paper, we focus on simultaneous estimation of the fixed parameter α and time-varying parameter β_t. Our objective is to outline the setting and develop a practical estimation procedure, where at once the fixed parameter can be estimated with parametric rate, the time-varying parameter with non-parametric rate, the asymptotic normality for components of parameters can be established and the standard errors computed. Simultaneous estimation requires slightly stronger assumptions in comparison to Giraitis, Kapetanios, Li (2024). Nevertheless, it offers practical estimation of partially time-varying regression model (1) for regressors under very general types of heteroskedasticity. Our primarily interest is also developing a rigourous estimation theory equipped with complete proofs. Although this requires considerable technical effort, it notably validates and extends the use of partially time-varying regression modelling in applied work.

First we derive closed-form OLS estimators for simultaneous estimation of α and β_t in partially time-varying regression model (1). Introduce notation of a time-varying estimator

$$\widehat{\beta}_{t,\alpha} = \Big(\sum_{j=1}^{n} b_{n,tj} z_j z_j'\Big)^{-1}\Big(\sum_{j=1}^{n} b_{n,tj} z_j (y_j - \alpha' x_j)\Big) \tag{2}$$

of β_t based on $y_j - \alpha' x_j$. Assumptions on the weights $b_{n,tj}$ will be specified below.

We define the OLS estimator $\widehat{\alpha}$ of α as the minimizer of the following objective function:

$$\widehat{\alpha} = \operatorname*{argmin}_{\alpha} L(\alpha), \quad L(\alpha) = \sum_{t=1}^{n} \big(y_t - \alpha' x_t - \widehat{\beta}'_{t,\alpha} z_t \big)^2.$$

We obtain the estimator $\widehat{\beta}_t$ of β_t by setting $\widehat{\beta}_t = \widehat{\beta}_{t,\widehat{\alpha}}$ in equation (2).

Introduce notation

$$\widehat{\beta}_{zx,t} = S^{-1}_{zz,t} S_{zx,t}, \quad \widehat{\beta}_{zy,t} = S^{-1}_{zz,t} S_{zy,t}, \tag{3}$$

where

$$S_{zz,t} = \sum_{j=1}^{n} b_{n,tj} z_j z'_j, \quad S_{zx,t} = \sum_{j=1}^{n} b_{n,tj} z_j x'_j, \quad S_{zy,t} = \sum_{j=1}^{n} b_{n,tj} z_j y_j. \tag{4}$$

Lemma 2.1. *The estimators $\widehat{\alpha}$ and $\widehat{\beta}_t = \widehat{\beta}_{t,\widehat{\alpha}}$ of α and β_t in (1) take the form:*

$$\widehat{\alpha} = \Big(\sum_{t=1}^{n} (x_t - \widehat{\beta}'_{zx,t} z_t)(x_t - \widehat{\beta}'_{zx,t} z_t)' \Big)^{-1} \Big(\sum_{t=1}^{n} (x_t - \widehat{\beta}'_{zx,t} z_t)(y_t - z'_t \widehat{\beta}_{zy,t}) \Big), \tag{5}$$

$$\widehat{\beta}_t = \Big(\sum_{j=1}^{n} b_{n,tj} z_j z'_j \Big)^{-1} \Big(\sum_{j=1}^{n} b_{n,tj} z_j (y_j - x'_j \widehat{\alpha}) \Big). \tag{6}$$

These are closed-form estimators for the fixed parameter α and time-varying parameter β_t and they are easy to compute. The proof of Lemma 2.1 is given in the Supplemental Material.

Assumptions. Regression noise variables

$$u_t = h_t \varepsilon_t, \tag{7}$$

are uncorrelated and can be written as a product of a stationary martingale difference noise $\{\varepsilon_t\}$ and a stochastic or deterministic scale factor $\{h_t\}$ which is independent of $\{\varepsilon_t\}$. More specifically, they have the following properties.

Assumption 2.1. *$\{\varepsilon_t\}$ is a stationary martingale difference (m.d.) sequence with respect to some σ-field filtration $\mathcal{F}_t$:*

$$\mathbb{E}[\varepsilon_t | \mathcal{F}_{t-1}] = 0, \quad \mathbb{E}\varepsilon_t^2 = 1.$$

$\{\varepsilon_t\}$ is independent of $\{h_t\}$. Moreover, variable ε_1 has probability distribution density

$f(x)$ and $f(x) \le c < \infty$ when $|x| \le x_0$ for some $x_0 > 0$.

The information set $\mathcal{F}_t$ will be generated by the past history $\mathcal{F}_t = \sigma(\varepsilon_s, z_s,\ s \le t)$ and possibly other variables.

We postulate that regressors $x_t = (x_{1t}, ..., x_{qt})'$, $z_t = (z_{1t}, ..., z_{pt})'$ can be written as a product of a scale factor and a stationary process: for $k = 1, ..., q$ and $j = 1, ..., p$,

$$x_{kt} \;=\; g_{xk,t}\eta_{xk,t}, \quad z_{jt} = g_{zj,t}\eta_{zj,t}, \tag{8}$$

where $\eta_{xt} = (\eta_{x1,t}, ..., \eta_{xq,t})'$, $\eta_{zt} = (\eta_{z1,t}, ..., \eta_{zp,t})'$ are stationary sequences, and $g_{xt} = (g_{x1,t}, ..., g_{xq,t})'$, $g_{zt} = (g_{z1,t}, ..., g_{zp,t})'$ are deterministic or stochastic scale factors.

We assume that $\{g_{xt}, g_{zt}, h_t\}$ are independent of $\{\eta_{xt}, \eta_{zt}, \varepsilon_t\}$ and sequences η_{xt}, η_{zt} in (8) may have non-zero mean.

Definition 2.1. *We say that a (univariate) covariance stationary sequence $\{\xi_t\}$ has short memory (SM) if $\sum_{h=-\infty}^{\infty} |\mathrm{cov}(\xi_h, \xi_0)| < \infty$.*

This setting becomes workable by imposing the following assumptions on stationary sequences $\{\eta_{xt}\}$, $\{\eta_{zt}\}$, scale factors and time-varying parameter β_t.

Assumption 2.2. *$\eta_{xt} = (\eta_{x1,t}, ..., \eta_{xq,t})'$, $\eta_{zt} = (\eta_{z1,t}, ..., \eta_{zp,t})'$ are $\mathcal{F}_{t-1}$ measurable sequences, $E[\eta^2_{xk,t}] = 1$, $E[\eta^2_{zj,t}] = 1$ and $E[\eta^8_{xk,t}] < \infty$, $E[\eta^8_{zj,t}] < \infty$, $E[\varepsilon^8_t] < \infty$.*

(i) $\{\eta_{zk,t}\}$, $\{\eta_{xk,t}\}$, $\{\eta_{zk,t}\eta_{z\ell,t}\}$, $\{\eta_{zk,t}\eta_{x\ell,t}\}$ and $\{\eta_{xk,t}\eta_{x\ell,t}\}$ are covariance stationary SM sequences.

(ii) $E[\eta_{z1}\eta'_{z1}] = \Sigma_{zz}$ is a positive definite matrix.

The scale factors in (8) are smoothly varying in the following terms.

Assumption 2.3. (i) *The scale factors $h_t \ge c_0$, $g_{xk,t} \ge c_0$, $g_{zj,t} \ge c_0$ are deterministic or stochastic random variables bounded away from 0 by $c_0 > 0$, and $Eh_t^8 \le C$, $Eg^8_{xk,t} \le C$, $Eg^8_{zj,t} \le C$ where c_0, C do not depend on t, n, k and j.*

(ii) *For some $\gamma_1 \in (3/4, 1]$, for $t, s = 1, .., n$,*

$$(E||g_{xt} - g_{xs}||^8)^{1/8} \le C\Big(\frac{|t-s|}{n}\Big)^{\gamma_1}, \quad (E||g_{zt} - g_{zs}||^8)^{1/8} \le C\Big(\frac{|t-s|}{n}\Big)^{\gamma_1}, \tag{9}$$

where $C < \infty$ does not depend on t, s and n.

The triangular arrays of scale factors $h_t = h_{n,t}$, $g_{xt} = g_{nx,t}$, $g_{zt} = g_{nz,t}$ may vary with n. We skip the subindex n for the brevity of notation. Assumptions 2.2 and 2.3 imply that

$$Eu_t^8 \le C, \quad E||x_t||^8 \le C, \quad E||z_t||^8 \le C, \tag{10}$$

where $C < \infty$ does not depend on t, n. Triangular arrays of parameters $\beta_t = \beta_{nt}$, $t = 1, ..., n$ in (1) are deterministic or stochastic processes. They satisfy the following property.

Assumption 2.4. *For some* $\gamma_2 \in (3/4, 1]$,

$$(E||\beta_t - \beta_s||^4)^{1/4} \;\; \le \;\; C\big(\frac{|t-s|}{n}\big)^{\gamma_2}, \quad E||\beta_t||^4 \le C, \quad t, s, = 1, ..., n, \tag{11}$$

where $C < \infty$ *does not depend on* t, s *and* n.

It is worth noting that no restrictions on dependence between scale factors g_{xt}, g_{zt}, h_t and the time-varying parameter β_t are imposed and no smoothness restrictions on the scale factor h_t in regression noise $u_t = h_t\varepsilon_t$ are required. In Kristensen (2012), regression space is limited to deterministic and smooth h_t and h_t needs to be estimated from the data. In addition, he imposes the assumption that $E[\varepsilon_t^2|x_t, z_t] = 1$ which restricts mutual dependence between ε_t and η_{xt}, η_{zt}. As a consequence, estimation procedures suggested in Kristensen (2012) are not robust to heterogeneity permitted by our regression space.

In the estimation of the regression parameters we use the following weights:

$$b_{n,tj} = K\big(\frac{|t-j|}{H}\big). \tag{12}$$

We suppose that the kernel function K on its support satisfies the following property: for some $d > 4$ and $C < \infty$,

$$K(x) \;\; \le \;\; C(1 + x^d)^{-1}, \;\; |(d/dx)K(x)| \le C(1 + x^d)^{-1}, \;\; x \ge 0. \tag{13}$$

Examples of kernel weights satisfying this assumption include the flat kernel, $K(x) = (1/2)I(|x| \le 1)$ and Gaussian kernel, $K(x) = (1/\sqrt{2\pi})e^{-x^2/2}$.

In the estimation of the fixed parameter α we assume that $\widehat{\beta}_{zx,t}$, $\widehat{\beta}_{zy,t}$ in $\widehat{\alpha}$ are computed with the bandwidth H that has the property that

$$n^a \le H = O(n^{2/3}), \quad a > 1/2. \tag{14}$$

Subsequently, the estimator $\widehat{\beta}_t$ in (6) can be computed using the weights $b_{n,tj}$ either with the same bandwidth H as in $\widehat{\alpha}$ or with a different bandwidth $H_z \to \infty$, $H_z = o(n)$.

2.1 Estimation of the fixed parameter

This section contains results on estimation of the fixed parameter $\alpha = (\alpha_1, ..., \alpha_q')$ in partially time-varying regression model (1) by the estimator $\widehat{\alpha} = (\widehat{\alpha}_1, ..., \widehat{\alpha}_q')$ given in (5). This estimators can be written as

$$\widehat{\alpha} = S_{\widehat{v}\widehat{v}}^{-1} S_{\widehat{v}\widetilde{y}_x} \tag{15}$$

using notation

$$\begin{aligned} \widehat{v}_t &= x_t - \widehat{\beta}'_{zx,t} z_t, \quad \widetilde{y}_{xt} = y_t - \widehat{\beta}'_{zy,t} z_t, \\ S_{\widehat{v}\widehat{v}} &= \sum_{t=1}^n \widehat{v}_t \widehat{v}'_t, \quad S_{\widehat{v}\widetilde{y}_x} = \sum_{t=1}^n \widehat{v}_t \widetilde{y}'_{xt}. \end{aligned} \tag{16}$$

The fact that we allow for a very general regression setting rules out standard asymptotic normality theory results that are common in regression literature. However, we show that this setting admits estimation of a single component α_k of the fixed parameter $\alpha = (\alpha_1, ..., \alpha_q)'$. We show that the asymptotic normality property for the estimator $\widehat{\alpha}_k$ of the component α_k can be established which allows to build confidence intervals for α_k.

We need an additional assumption. Denote

$$\nu_t = \eta_{xt} - E[\eta_{xt}\eta'_{zt}](E[\eta_{zt}\eta'_{zt}])^{-1}\eta_{zt}. \tag{17}$$

Assumption 2.5. *(i) For any k, ℓ, the products $w_t = \eta_{zk,t}\eta_{z\ell,t}$ and $w_t = \eta_{zk,t}\eta_{x\ell,t}$ have the following properties:*

(i) $\{\varepsilon_t^2\}$, $\{w_t \varepsilon_t^2\}$ are covariance stationary SM sequences.

(ii) $E[\nu_1 \nu_1']$ is a positive definite matrix.

(iii) There exists a sequence $m = m_n = O(\log n)$ such that for $j, i > t + m$,

$$\begin{aligned} E[w_j|\mathcal{F}_t] &= E[w_j] + r_{mt,j}, \quad (Er^4_{mt,j})^{1/4} \le Cn^{-2}, \\ E[w_j w_i|\mathcal{F}_t] &= E[w_j w_i] + r_{mt,ji}, \quad (Er^2_{mt,ji})^{1/2} \le Cn^{-2}, \end{aligned} \tag{18}$$

where $C < \infty$ does not depend on j, i, t and m.

Remark 2.1. Observe that under Assumption 2.2, the components of $\nu_t = (\nu_{1t}, ..., \nu_{qt})'$ and of $\nu_t \nu_t' = (\nu_{kt}\nu_{\ell t})$ are a covariance stationary SM sequences. Moreover, under Assumption 2.5, the product $w_t = \nu_{zk,t}\nu_{z\ell,t}$ of components of ν_t also satisfies Assumption 2.5, and $w_t \varepsilon_t^2$ is a covariance stationary SM sequence.

Remark 2.2. Stationary processes η_{zt} and η_{xt} satisfy Assumption 2.5(iii) in the following cases: (i) $\{\eta_{zt}\}$, $\{\eta_{xt}\}$ are mutually independent of $\{\varepsilon_t\}$.

(ii) η_{zj}, η_{xi} are independent of ε_t for $j, i \geq t + L$ for some $L > 0$.

(iii) $\{\eta_{zt}\}$, $\{\eta_{xt}\}$ are stationary linear processes as in Lemma 2.2 below.

To describe the standard error in the estimation of component α_k of the parameter $\alpha = (\alpha_1, ..., \alpha_q)'$ we introduce additional notation:

$$\begin{aligned} v_t &= x_t - E[x_t z_t' | \mathcal{F}_n^*](E[z_t z_t' | \mathcal{F}_n^*])^{-1} z_t, \\ S_{vv} &= \textstyle\sum_{t=1}^n v_t v_t', \quad S_{vvuu} = \sum_{t=1}^n v_t v_t' u_t^2, \\ \Omega_{\alpha,n} &= (E[S_{vv}|\mathcal{F}_n^*])^{-1} E[S_{vvuu}|\mathcal{F}_n^*](E[S_{vv}|\mathcal{F}_n^*])^{-1} = (\omega_{jk}) \end{aligned} \tag{19}$$

where $\mathcal{F}_n^* = \sigma\big(h_t, g_{xt}, g_{zt}, t = 1, ..., n\big)$ is the information set generated by scales.

The next theorem focuses on estimation of components of the parameter $\alpha = \big(\alpha_1, ..., \alpha_q\big)'$ and derives the asymptotic normality property for the estimator $\widehat{\alpha}_k$ of α_k.

Theorem 2.1. *Suppose Assumptions 2.1-2.5 are satisfied and (14) holds.*

Then, the t-statistic for the parameter α_k, $k = 1, ..., q$ has property:

$$\frac{\widehat{\alpha}_k - \alpha_k}{\sqrt{\omega_{kk}}} \to_d \mathcal{N}(0,1), \qquad \sqrt{\omega_{kk}} \asymp_p n^{-1/2}. \tag{20}$$

In practical applications, the standard error $\sqrt{\omega_{kk}}$ can be estimated by the diagonal element $\sqrt{\widehat{\omega}_{kk}}$ of the matrix

$$\widehat{\Omega}_{\alpha,n} = S_{\widehat{v}\widehat{v}}^{-1} S_{\widehat{v}\widehat{v}\widehat{u}\widehat{u}} S_{\widehat{v}\widehat{v}}^{-1} = (\widehat{\omega}_{jk}), \quad \widehat{u}_t = y_t - \widehat{\alpha}' x_t - \widehat{\beta}_t' z_t. \tag{21}$$

The estimator $\widehat{\beta}_t$ using residuals $\widehat{u}_t$ is computed with the same bandwidth H as in estimator $\widehat{\alpha}$.

Corollary 2.1. *Under the assumptions of Theorem 2.1, for $k = 1, ..., q$,*

$$\frac{\widehat{\alpha}_k - \alpha_k}{\sqrt{\widehat{\omega}_{kk}}} \to_d \mathcal{N}(0,1), \quad \frac{\widehat{\omega}_{kk}}{\omega_{kk}} = 1 + o_p(1). \tag{22}$$

This result allows to compute robust standard errors and build robust confidence intervals for the components α_k of α.

The estimator of robust standard errors $\widehat{\Omega}_{\alpha,n}$ has the same form as the estimator for heteroskedasticity-consistent standard errors by White (1980). When regressors x_t and z_t are stationary processes and $\{u_t\}$ is an i.i.d. noise independent of $\{x_t,\, z_t\}$, $\widehat{\Omega}_{\alpha,n}$ can be replaced by

$$\widehat{\Omega}^{(st)}_{\alpha,t} = S^{-1}_{\widehat{v}\widehat{v},t}\,\widehat{\sigma}^2_u, \quad \widehat{\sigma}^2_u = n^{-1}\sum_{j=1}^{n}\widehat{u}^2_j. \tag{23}$$

Unlike $\widehat{\Omega}_{\alpha,t}$, this estimator (23) is not robust to presence of heterogeneity in data, see examples in the Monte Carlo study in the Online Supplement.

The following lemma provides conditions when a stationary linear process satisfies Assumption 2.5(iii).

Lemma 2.2. *Suppose that components of* $\eta_{zt} = (\eta_{z1,t}, ..., \eta_{zq,t})'$ *and* $\eta_{xt} = (\eta_{x1,t}, ..., \eta_{xp,t})'$ *are stationary linear processes*

$$\eta_{zk,t} \;=\; \sum_{i=0}^{\infty} b_{zk,i}\xi_{zk,t-i}, \qquad \eta_{x\ell,t} = \sum_{i=0}^{\infty} b_{x\ell,i}\xi_{x\ell,t-i} \tag{24}$$

with exponentially decaying weights $b_{zk,i}, b_{x\ell,i}$*:*

$$|b_{zk,i}| \le C\rho^i, \quad |b_{x\ell,i}| \le C\rho^i, \qquad (0<\rho<1). \tag{25}$$

$\{\xi_{zk,i}\}$, $\{\xi_{x\ell,i}\}$ *are uncorrelated stationary noises with the* 8*-th finite moment, and* C, ρ *do not depend on* k, ℓ, i*. Suppose that* $\mathcal{F}_t = \sigma(e_s,\, s \le t)$ *and variables* $\xi_{zk,i}$*,* $\xi_{x\ell,i}$ *are independent of* $e_t, e_{t-1}, ...$ *when* $i \ge t + L$ *for some* $L \ge 0$*. Then Assumption 2.5(iii) holds with* $m = b\log n$ *for large enough* b*.*

The estimation of PTVR parameters in Theorem 2.1 requires smooth change of scale factors $g_{zk,t}, g_{xk,t}$ and time-varying parameter β_t, see Assumptions 2.3 and 2.4, and no smoothness restrictions on the scale factor h_t in regression noise $u_t = h_t\varepsilon_t$ are imposed. Typical examples of smooth deterministic or stochastic change are as follows.

Example 2.1. A deterministic sequence $g_t = f(t/n)$, $t = 1, ..., n$, where $f(\cdot) \ge 0$ is a Lipschitz smooth function, $|f(x) - f(y)| \le C|x-y|^\gamma$ with parameter $\gamma \in (1/2, 1]$, has property $|g_t - g_s| \le C(|t-s|/n)^\gamma$, and is a standard example of is a scale factor satisfying smoothness assumption (9) with parameter $\gamma \in (1/2, 1]$.

An example of stochastic smoothly varying scale factor is a stochastic process

$$g_t = n^{-\nu}\Big|\sum_{j=1}^{t}\xi_j\Big|, \quad t=1,...,n,$$

where $\{\xi_j\}$ is a stationary ARFIMA$(0,d,0)$ process with parameter $d\in(0,1/2)$ and zero mean, see e.g. Chapter 7 in Giraitis, Koul and Surgailis (2012). It satisfies smoothness property (9) with $\gamma=1/2+d$. Indeed, then for $t>s$,

$$\begin{aligned}|g_t-h_s| &= n^{-\gamma}\Big|\big|\textstyle\sum_{j=1}^{t}\xi_j\big|-\big|\sum_{j=1}^{s}\xi_j\big|\Big| \le n^{-\gamma}\big|\sum_{j=s+1}^{t}\xi_j\big| \\ &\le (|t-s|/n)^{\gamma}|S_{t,s}|, \quad S_{t,s}=(t-s)^{-\gamma}\textstyle\sum_{j=s+1}^{t}\xi_j.\end{aligned}$$

If ARFIMA process is generated by i.i.d. innovations with $p\ge 2$ finite moments, then by stationarity and Propositions 4.4.3 and 3.3.1 in Giraitis, Koul and Surgailis (2012),

$$E|S_{t,s}|^p = E|S_{t-s,0}|^p \le (ES^2_{t-s,0})^{p/2}\le C.$$

Then, $(E|g_t-g_s|^p)^{1/2}\le C(|t-s|/n)^{\gamma}$ where $C<\infty$ does not depend on t,s and n.

2.2 Estimation of the time-varying parameter

This subsection outlines results on estimation of the time-varying parameter $\beta_t = (\beta_{1t},...,\beta_{pt})'$ in partially time-varying regression model (1). The estimator $\widehat{\beta}_t=(\widehat{\beta}_{1t},...,\widehat{\beta}_{pt})'$ given in (6) can be written as

$$\widehat{\beta}_t = S^{-1}_{zz,t}S_{z\widetilde{y}_z,t} \tag{26}$$

using notation:

$$S_{zz,t}=\textstyle\sum_{j=1}^{n} b_{n,tj}z_jz_j', \quad S_{z\widetilde{y},t}=\sum_{j=1}^{n} b_{n,tj}z_j\widetilde{y}_{zj}, \quad \widetilde{y}_{zj}=y_j-x_j'\widehat{\alpha}. \tag{27}$$

We compute the weights $b_{n,tj}$ in the estimator $\widehat{\beta}_t$ with bandwidth parameter $H_z\to\infty$ which can be different from the bandwidth H used in the estimation of the fixed parameter α. The bandwidth H_z satisfies $H_z=o(n)$, $H_z\to\infty$. It is required to satisfy assumption (14) only in the estimation of standard errors in Corollary 2.2.

We consider the point-wise estimation of components β_{kt} of the parameter vector β_t for $t=1,...,n$. In particular, we derive the asymptotic normality property and estimation procedure for standard errors for the estimator $\widehat{\beta}_{kt}$ of β_{kt} at time t.

Standard errors for the estimator $\widehat{\beta}_{kt}$ will be described using the diagonal elements $\omega_{\beta kk,t}$ of the matrix

$$\begin{aligned} \Omega_{\beta,t} &= (E[S_{zz,t}|\mathcal{F}_n^*])^{-1} E[S_{zzuu,t}|\mathcal{F}_n^*](E[S_{zz,t}|\mathcal{F}_n^*])^{-1} = (\omega_{jk,t}), \\ S_{zzuu,t} &= \textstyle\sum_{j=1}^n b_{n,tj}^2 z_j z_j' u_j^2. \end{aligned} \tag{28}$$

In the next theorem we establish the consistency rate and the asymptotic normality property for estimation of the component β_{kt} of the time-varying parameter β_t by the estimator $\widehat{\beta}_{kt}$ in partially time-varying regression model (1).

Theorem 2.2. *Suppose assumptions of Theorem 2.1 are satisfied. Then, for* $1 \le t = t_n \le n$ *and* $k = 1, ..., p$, *the following holds:*

$$\widehat{\beta}_{kt} - \beta_{kt} = O_p\big(H_z^{-1/2} + (H_z/n)^{\gamma_2}\big). \tag{29}$$

Moreover, if $H_z = o(n^{2\gamma_2/(2\gamma_2+1)})$, *then:*

$$\frac{\widehat{\beta}_{kt} - \beta_{kt}}{\sqrt{\omega_{kk,t}}} \to_d \mathcal{N}(0,1), \quad \sqrt{\omega_{kk,t}} \asymp_p H_z^{-1/2}. \tag{30}$$

Furthermore, the unknown standard error $\sqrt{\omega_{kk,t}}$ in the normal approximation above can be estimated using the diagonal element $\widehat{\omega}_{kk,t}$ of the matrix:

$$\widehat{\Omega}_{\beta,t} = S_{zz,t}^{-1} S_{zz\widehat{u}\widehat{u},t} S_{zz,t}^{-1} = (\widehat{\omega}_{jk,t}), \quad \widehat{u}_j = y_j - \widehat{\alpha}' x_j - \widehat{\beta}_j' z_j. \tag{31}$$

Corollary 2.2. *Assume that bandwidth* H_z *satisfies property (14). Then, under assumptions of Theorem 2.2, for k=1, ..., p, the following holds:*

$$\frac{\widehat{\beta}_{kt} - \beta_{kt}}{\sqrt{\widehat{\omega}_{kk,t}}} \to_d \mathcal{N}(0,1), \qquad \frac{\widehat{\omega}_{kk,t}}{\omega_{kk,t}} = 1 + o_p(1). \tag{32}$$

Computation of standard errors $\sqrt{\widehat{\omega}_{kk,t}}$ in estimation of partially time-varying model is straightforward. The estimator $\widehat{\Omega}_{\beta,t}$ of robust standard errors in (31) is a time-varying version of heteroskedasticity-consistent standard errors by White (1980).

The robust estimation of standard errors by $\widehat{\Omega}_{\beta,t}$ differs from the standard estimation by

$$\widehat{\Omega}^{(st)}_{\beta,t} = (K_{2,t}/K_t)S^{-1}_{zz,t}\widehat{\sigma}^2_u, \quad \widehat{\sigma}^2_u = n^{-1}\sum_{j=1}^{n}\widehat{u}^2_j, \qquad \widehat{u}_j = y_j - \widehat{\alpha}'x_j - \widehat{\beta}'z_j, \tag{33}$$
$$K_{2,t} = \textstyle\sum_{j=1}^n b^2_{n,tj}, \quad K_t = \sum_{j=1}^n b_{n,tj}.$$

Estimator (33) is applicable when regressors are stationary processes and independent of regression noise u_t which is a stationary martingale difference noise. It is not robust to heterogeneity, and fails to estimate standard errors under heteroskedasticity settings, see examples considered in the Monte Carlo study in the Online Supplement.

3 Monte Carlo study

In this section, we use Monte Carlo simulations to examine the theoretical properties of the estimators of parameters of the partially time-varying regression model, established in Section 2. The theory shows that the fixed parameter α can be estimated with the parametric rate $\sqrt{n}$ and the time-varying parameter β_t with a non-parametric rate which is slower than $\sqrt{n}$. In particular, we explore the validity of the asymptotic normality property of t-statistics established for the components of parameters α and β_t in finite samples and its robustness to heterogeneity. We consider a variety of scale factors, noises and time-varying parameters β_t allowed by our model setting, see also Section 8 in the Online Supplement.

The simulations confirm the validity of our theoretical results. The estimators show good finite sample performances, reveal robustness to heterogeneity under different regression settings and confirm the ease of practical application.

We set the sample size to $n = 1500$, conduct 1000 replications, and use the bandwidth $H = n^h$ and $H_z = n^h$, $h = 0.4, 0.5, 0.6, 0.7$. (Estimation results for $n = 200, 800$ are available upon request).

3.1 Estimation of PTVR model

We generate arrays of samples y_t, $t = 1, ..., n$ of a partially time-varying regression model

$$\begin{aligned} y_t &= \alpha x_t + \beta_t' z_t + u_t, \quad u_t = h_t\varepsilon_t \\ &= \beta_{1t} + \alpha x_t + \beta_{2t}z_{2t} + u_t, \end{aligned} \tag{34}$$

with a fixed parameter $\alpha = 0.5$, time-varying parameter $\beta_t = (\beta_{1t}, \beta_{2t})'$ and regressors $z_t = (1, z_{2t})'$, where $\beta_{1t} = 0.5 \sin(\pi t/n) + 1,\ \ t = 1, ..., n$ is a time-varying intercept.

The regression noise $u_t = h_t \varepsilon_t$ is a product of an i.i.d. $\mathcal{N}(0,1)$ noise ε_t and a deterministic scale factor

$$h_t = 0.5 \sin(0.8\pi t/n) + 1, \quad t = 1, ..., n. \tag{35}$$

The regressors $\{x_t\}$ and $\{z_{2t}\}$ are univariate, and are products of scale factors g_{xt}, g_{zt} and stationary MA(1) processes η_{xt}, η_{zt},

$$\begin{aligned} x_t &= g_{xt}\eta_{xt}, \qquad \eta_{xt} = 0.2 + \epsilon_{xt} + 0.5\epsilon_{x,t-1}, \\ z_t &= g_{zt}\eta_{zt}, \qquad \eta_{zt} = 0.2 + \epsilon_{zt} + 0.5\epsilon_{z,t-1}, \end{aligned} \tag{36}$$

where $\{\epsilon_{xt}\}$, $\{\epsilon_{zt}\}$ are mutually independent i.i.d. $\mathcal{N}(0,1)$ noises and mutually independent of $\{g_{xt}, g_{zt}\}$.

We consider two cases of scale factors g_{xt}, g_{zt}, $t = 1, ..., n$:

$$\text{Deterministic}: \quad g_{xt} = 0.5 \sin(0.3\pi t/n) + 1, \quad g_{zt} = 0.5 \sin(0.4\pi t/n) + 1, \tag{37}$$

$$\text{Stochastic}: \quad g_{xt} = |n^{-\gamma} \sum_{i=1}^{t} \upsilon_{xi}| + 0.2, \quad g_{zt} = |n^{-\gamma} \sum_{i=1}^{t} \upsilon_{zi}| + 0.2, \tag{38}$$

where $\{\upsilon_{xi}\}, \{\upsilon_{zi}\}$ are stationary ARFIMA$(0, d, 0)$ processes with memory parameter $d = 0.4$.

We centre on two cases of time-varying parameter β_t:

$$\text{Deterministic}: \quad \beta_{2t} = 0.5 \sin(0.5\pi t/n) + 1, \ \ t = 1, ..., n, \tag{39}$$

$$\text{Stochastic}: \quad \beta_{2t} = |n^{-\gamma} \sum_{i=1}^{t} e_i| + 0.2, \tag{40}$$

where e_i is an ARFIMA$(0, d, 0)$ process with parameter $d = 0.4$.

The stochastic processes g_{xt}, g_{zt} in (38) satisfy the smoothness assumption (9) with parameter $\gamma_1 = 0.5+d = 0.9$, and the stochastic parameter β_{2t} in (38) satisfies smoothness assumption (11) with parameter $\gamma_2 = 0.5 + d = 0.9$, see Example 2.1.

We consider two partially time-varying regression models. Model 3.1 combines deterministic scale factors and time-varying parameter β_{2t}, while Model 3.2 is based on stochastic scale factors and β_{2t}.

Model 3.1. *y_t, $t = 1, ..., n$ follows model (34) with deterministic scale factors $\{g_{xt}, g_{zt}\}$*

as in (37) and parameter β_{2t} *as (39).*

Model 3.2. y_t, $t = 1, ..., n$ *follows model (34) with stochastic scale factors* $\{g_{xt}, g_{zt}\}$ *as in (38) and parameter* β_{2t} *as in (40).*

More complex simulation examples, that verify the robustness of our estimation and inference approach, can be found in the Online Supplement.

We start with the analysis of the deterministic setting of Model 3.1. Table 1 reports the bias, RMSE and coverage rate (in %) for 95% confidence intervals for the fixed parameter α. The estimation results confirm the good performance of the PTVR estimation method for the fixed parameter α, achieving small bias, RMSE and good coverage rates for bandwidth $H = n^{0.6}$ and $H_z = n^{0.4}, n^{0.5}, ..., n^{0.7}$. We find that PTVR estimation shows little difference over combinations of H, H_z. So, in PTVR estimation of parameters α, one can consider pre-selecting H, e.g. in this simulation study we set $H = n^{0.6}$. The impact of H_z on the quality of estimation of α is also minimal. Hence in estimation of α once could set $H = H_z$ as recommended by the theory.

Table 1: Estimation of α in Model 3.1.

h	Bias	RMSE	CP	SD
0.4	0.00014	0.02735	94.6	0.02735
0.5	0.00017	0.02679	94.1	0.02679
0.6	0.00025	0.02645	94.5	0.02645
0.7	0.00025	0.02639	94.2	0.02639

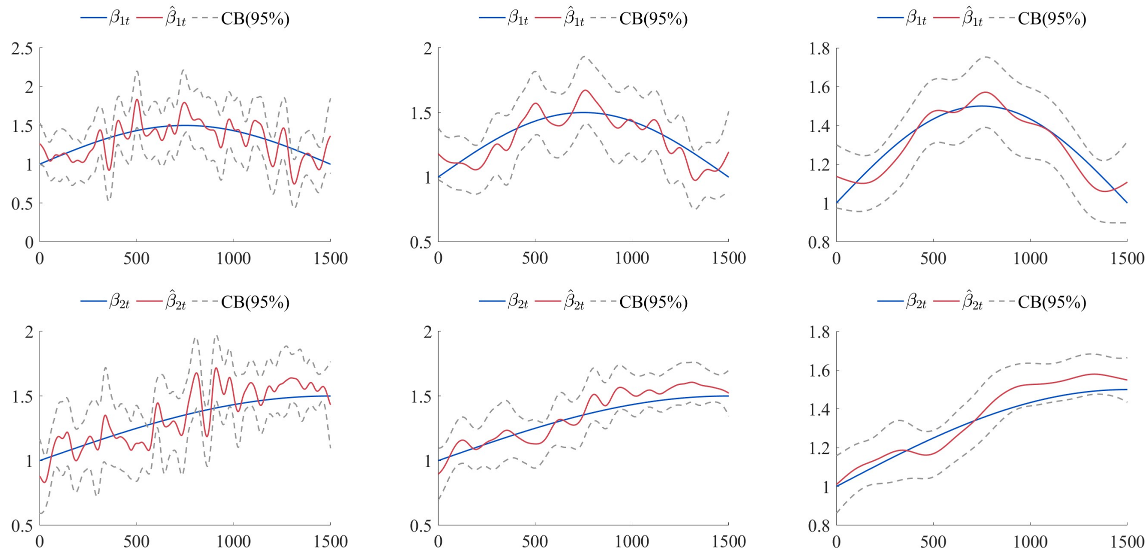
Table 2: Estimation of α in Model 3.2.

h	Bias	RMSE	CP	SD
0.4	0.00017	0.06834	93.5	0.06834
0.5	0.00021	0.06736	94.1	0.06736
0.6	0.00020	0.06690	94.3	0.06690
0.7	0.00035	0.06660	94.3	0.06660

Figure 1 displays point-wise estimation results for parameter β_t for a single sample of Model 3.1 for the bandwidths $H = n^{0.6}, H_z = n^{0.5}$ and sample size $n = 1500$. It reports the true parameters β_t (blue solid line), the estimates $\widehat{\beta}_t$ (red solid line) and their point-wise 95% confidence bands (gray dash lines). The first row of panels reports estimation results for the deterministic time-varying intercept β_{1t}. The 95% confidence band covers β_{1t} and the estimator $\widehat{\beta}_{1t}$ follows closely the path of the true intercept β_{1t}. The time-varying parameter β_{2t}, is also covered by the 95% confidence interval almost everywhere and the estimator $\widehat{\beta}_{2t}$ tracks the changes in β_{2t}.

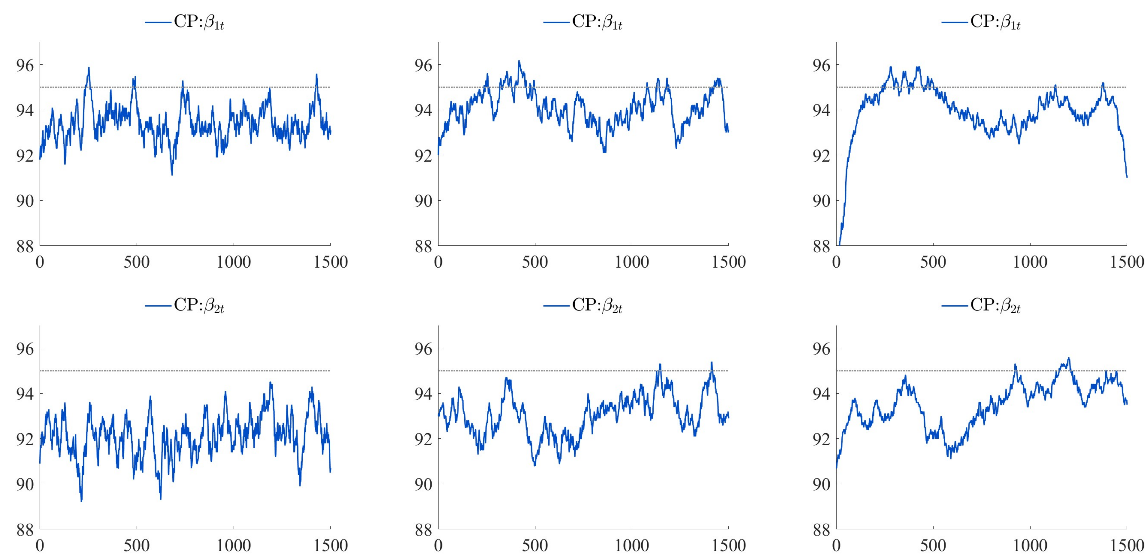
The three panels of Figure 1 (a) (b) (c) spell out the importance of selection of H_z in the estimation of time-varying parameter β_t. For $H_z = n^{0.4}, n^{0.5}, n^{0.6}$, the true parameter β_t is well-covered by the 95% confidence bands including the end points. However, the confidence intervals become unstable over the time and wide, for small bandwidth

$H_z = n^{0.4}$. As the bandwidth increases, the paths of the estimates $\widehat{\beta}_t$ become more and more smooth, and the confidence bands for β_t become narrower. In addition, we also find that confidence bands for the time-varying parameter β_t are wider than for a fixed parameter α.



(a) $H_z = n^{0.4}$ (b) $H_z = n^{0.5}$ (c) $H_z = n^{0.6}$

Figure 1: Robust 95% confidence bands for time-varying parameters β_{1t}, β_{2t} in Model 3.1: $n = 1500$, bandwidth $H = n^{0.6}$, $H_z = n^h$, $h = 0.4, 0.5, 0.6$. Single replication.



(a) $H_z = n^{0.4}$ (b) $H_z = n^{0.5}$ (c) $H_z = n^{0.6}$

Figure 2: Coverage rates (in %) of robust 95% confidence intervals for time-varying parameters β_{1t}, β_{2t} in Model 3.1: $n = 1500$, bandwidth $H = n^{0.6}$, $H_z = n^h$, $h = 0.4, 0.5, 0.6$.

Figure 2 displays the empirical coverage rate (in %) of 95% confidence intervals for time-varying parameter β_t in Model 3.1. It allows the evaluation of the validity of the asymptotic normal approximation for components of the parameter β_t in the point-wise estimation of β_t by $\widehat{\beta}_t$ for sample size $n = 1500$. The bandwidth $H = n^{0.6}$ is fixed. For example, it shows that in panel (b), the coverage rates are close to the nominal 95%. Estimation with $H_z = n^{0.5}$ achieves slightly better coverage rate than with $H_z = n^{0.4}$, and we see more coverage distortions when $H_z = n^{0.6}$ which is consistent with our previous finding that the confidence intervals become narrower in estimation with larger bandwidth.

The bias of the PTVR estimator $\widehat{\beta}_t$ is close to zero, and the RMSE becomes smaller as H_z increases. (These results are available upon request).

The second Model 3.2 focuses on the stochastic setting. Again, we pre-select the bandwidth parameter $H = n^{0.6}$, in estimation of the fixed parameter α. Table 2 reports estimation outcomes for the fixed parameter α which confirm good performance of the PTVR estimator $\widehat{\alpha}$ and excellent coverage rate for 95% confidence intervals.

Figure 3 displays PTVR estimation results for the time-varying parameter β_t for a single sample from Model 3.2 for bandwidth parameters $H_z = n^{0.4}, n^{0.5}, n^{0.6}$. The 95% confidence intervals cover the path of the true parameter β_t for most of the times, and the estimator $\widehat{\beta}_t$ tracks the path of β_t.

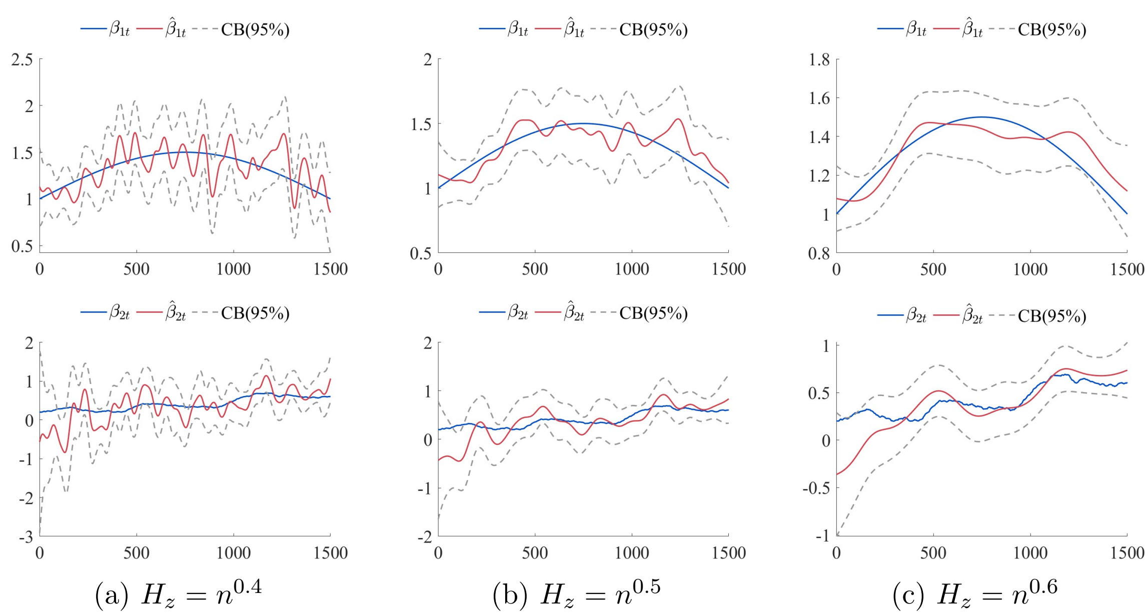
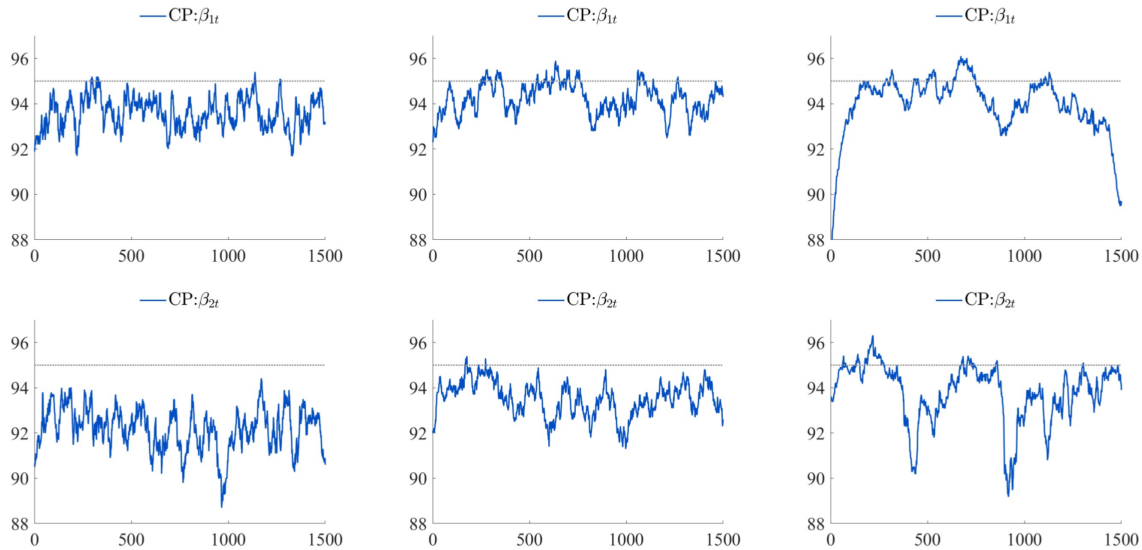


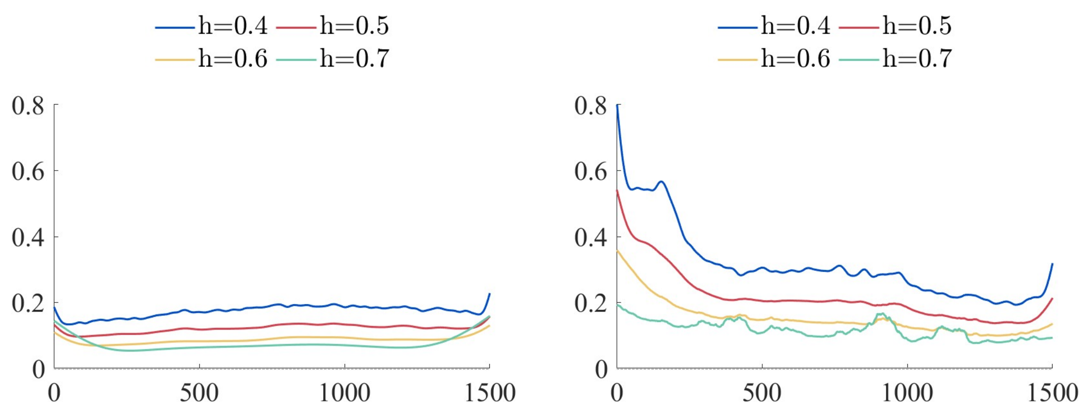
Figure 3: Robust 95% confidence bands for time-varying parameters β_{1t}, β_{2t} in Model 3.2: $n = 1500$, bandwidth $H = n^{0.6}$, $H_z = n^h$, $h = 0.4, 0.5, 0.6$. Single replication.



(a) $H_z = n^{0.4}$ (b) $H_z = n^{0.5}$ (c) $H_z = n^{0.6}$

Figure 4: Coverage rates (in %) of robust 95% confidence intervals for time-varying parameters β_{1t}, β_{2t} in Model 3.2: $n = 1500$, bandwidth $H = n^{0.6}$, $H_z = n^h$, $h = 0.4, 0.5, 0.6$.

Figure 4 displays empirical coverage rate of 95% confidence intervals in PTVR estimation of the time-varying parameter β_t in Model 3.2. With bandwidth $H_z = n^{0.5}$, the coverage rate is very close to the nominal 95%. Figure 5 shows that the RMSE is small and becomes smaller when bandwidth increases. However, the RMSE can rise when there is a lot of variability in β_t.



(a) β_{1t} (b) β_{2t}

Figure 5: RMSE for time-varying parameters β_{1t}, β_{2t} in Model 3.2: $n = 1500$, bandwidth $H = n^{0.6}$, $H_z = n^h$, $h = 0.4, 0.5, 0.6$.

3.2 Forecasting using PTVR model

In this section we consider a forecasting exercise using four competing methods, including a PTVR model. We generate a sample $y_1, ..., y_n$ of $n = 1500$ observations from the PTVR model given by

$$y_t = \alpha x_t + \beta_t z_t + u_t, \quad t = 1, \cdots, n, \tag{41}$$

with fixed parameter $\alpha = 0.5$, time-varying parameter $\beta_t = (1/2)\sin(2\pi t/n) + 1$, and i.i.d. $\mathcal{N}(0,1)$ noise u_t. The regressors $x_t = 0.5x_{t-1} + \eta_{xt}$ and $z_t = 0.7z_{t-1} + \eta_{zt}$ are stationary AR(1) processes generated by uncorrelated noises $\eta_{xt} = v_{1t} + v_{3t}$, $\eta_{zt} = v_{2,t} + v_{3,t}$ which are cross-correlated, where $\{v_{1t}\}$, $\{v_{2t}\}$, $\{v_{3t}\}$ and $\{u_t\}$ are mutually independent i.i.d. $\mathcal{N}(0,1)$ noises.

To compute the 1-step ahead forecast $\widehat{y}_{t|t-1}$ of y_t, we use the fitted PTVR model

$$\widehat{y}_{t|t-1} = \widehat{\alpha}\, x_{t-1} + \widehat{\beta}_{t-1} z_{t-1},$$

where parameters α, β_{t-1} are estimated using $y_1, ..., y_{t-1}$. We compare this forecast with forecasts $\widehat{y}_{t|t-1}$ obtained using a method where all parameters are estimated as time varying (FTVR), the OLS method where all parameters are estimated as fixed, and an $AR(1)$ forecast.

To evaluate the quality of the forecast, for each method we conduct in-sample forecasting of y_t by $\widehat{y}_{t|t-1}$ for $t = t_0, ..., n$ and evaluate mean square forecast error (MSFE):

$$\text{MSFE} = \frac{1}{n - t_0} \sum_{t=t_0+1}^{n} (y_t - \widehat{y}_{t|t-1})^2.$$

In our simulation, $t_0 = 750$. We will use MSFE to determine the best forecast method and bandwidth H that minimises forecast error. For each forecast method, we compute also the MSFE ratio =MSFE/MSFE$_{PTVR}$.

Table 3 reports MSFE and MSFE ratio results for all four forecasting methods and bandwidths $H = H_z = n^{0.4}, n^{0.5}, n^{0.6}, n^{0.7}$. Regardless of bandwidth, the PTVR and FTVR methods produce a smaller forecast error than OLS and AR(1), and achieve the smallest MSFE at $H \sim n^{0.5}, n^{0.6}$ which is close to 1, or the variance of the regression noise u_t, which suggests high quality for the forecast. The MSFE ratio for the FTVR method is close to 1, which implies that PVTR and FTVR methods are comparable. They produce significantly better forecasts than the remaining two methods, OLS (ratio > 1.5) and AR(1) (ratio > 3).

Table 3: MSFE and MSFE ratio

Bandwidth		PTVR	FTVR	OLS	AR(1)
$H = n^{0.4}$	MSFE	1.101	1.154	1.788	3.454
	MSFE (ratio)	1	1.048	1.642	3.137
$H = n^{0.5}$	MSFE	1.078	1.108	1.788	3.454
	MSFE (ratio)	1	1.028	1.659	3.204
$H = n^{0.6}$	MSFE	1.081	1.097	1.788	3.454
	MSFE (ratio)	1	1.015	1.654	3.195
$H = n^{0.7}$	MSFE	1.133	1.140	1.788	3.454
	MSFE (ratio)	1	1.006	1.578	3.049

To further evaluate the optimal bandwidth H for the PTVR method, we plot the MSFE for a grid of bandwidths $H = n^{0.05}, n^{0.1}, ..., n^{0.95}, n$. Figure 6 shows that the PTVR forecast achieves the smallest MSFE at $H = n^{0.55}$.

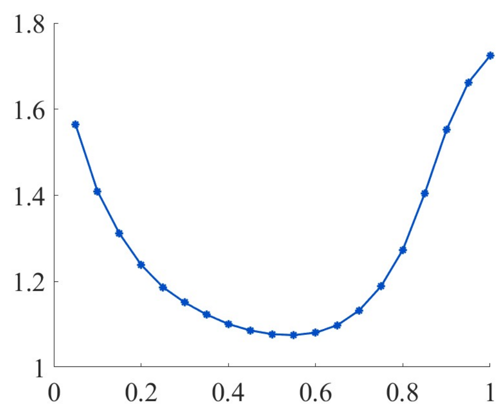


Figure 6: MSFE for PTVR method, $H = n^h$, h on the horizontal axis.

4 Empirical illustration

In this section, we assess the performance of the PTVR forecast when applied to a set of U.S. macro-economic variables. Our purpose is not to construct the best forecast method for this particular data set, but to examine the usefulness of the partially time-varying regression approach in an empirical forecasting context. The dataset is composed of 8 quarterly time series spanning from 1949Q1 to 2018Q4. Data are obtained from the Federal Reserve Economic Database of St. Louis Federal Reserve Bank. All variables are transformed following standard practice (see Stock and Watson

(2012)) and are described in Table 4. In this experiment, we use the PTVR approach for one-quarter-ahead forecasts of Real Gross Domestic Product (GDPC1).

To evaluate the quality of the PTVR forecasts, we start forecasting at time $t_0 = 100$ and continue until the entire sample is used.

Table 4: Data description

Variable	Description	Transformation form
GDPC1	Real Gross Domestic Product	$\Delta \log x_t$
GPDIC1	Real Gross Private Domestic Investment	$\Delta \log x_t$
PCEC96	Real Personal Consumption Expenditures	$\Delta \log x_t$
PAYEMS	All Employees, Total Nonfarm	$\Delta \log x_t$
AWHMAN	Average Weekly Hours of Production and Nonsupervisory Employees, Manufacturing	Δx_t
UNRATE	Unemployment Rate	$\Delta^2 x_t$
CPIAUCSL	Consumer Price Index for All Urban Consumers	$\Delta^2 \log x_t$
INDPRO	Industrial Production	$\Delta \log x_t$

Figure 7 shows that the GDPC1 series we want to forecast exhibits different patterns of fluctuation in different time periods. We first test for the presence of serial correlation in y_t (GDPC1) using standard and robust tests, see Giraitis, Li, Phillips (2024). Figure 8 confirms the presence of significant autocorrelation in GDPC1.

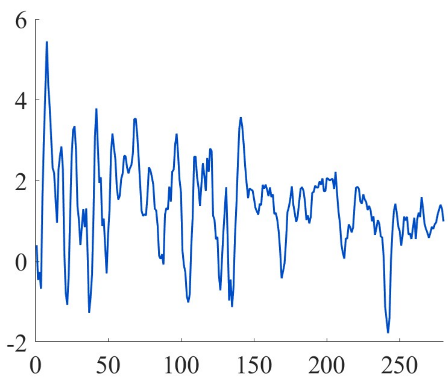


Figure 7: Plot of series GDPC1 (y_t) (1949Q1-2018Q4)

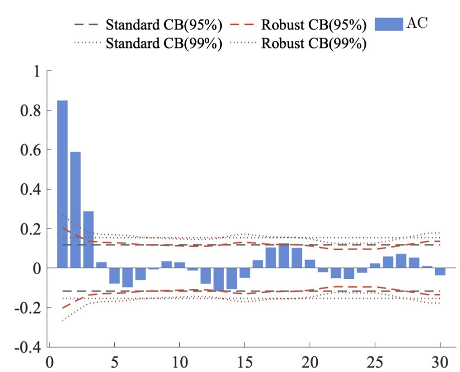


Figure 8: Correlogram for GDPC1 (y_t).

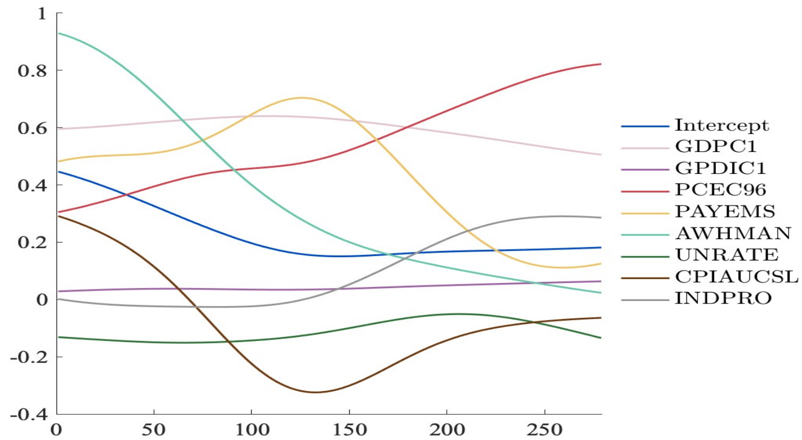


Figure 9: FTVR estimates $\widehat{\beta}_t$ of regression parameters, $H_z = n^{0.7}$.

Next, we use our partially time-varying regression model, PTVR, to regress the variable of interest GDPC1, y_t, on 8 regressors, which include a time-varying intercept, lagged dependent variable y_{t-1} and the remaining 7 variables from the Table 4 all lagged by one period.

In order to divide regression parameters into fixed and time-varying sets, we first fit to y_t the FTVR regression model with regressors (all one period lagged) which estimates all parameters as time-varying. Since the sample size $n = 279$ is small, we use a bandwidth value of $H_z = n^{0.7}$. In Figure 9, lines depict the paths of the estimates of parameters of all regressors under consideration. We notice that the intercept, and the coefficients at regressors GDPC1, GPDIC1, and UNRATE are almost constant. So, in our PTVR regression we will treat these parameters as constant.

Subsequently, we fit to y_t a PTVR regression model where the intercept and the coefficients of GDPC1, GPDIC1, UNRATE (one period lagged) are constant, and the coefficients of the remaining regressors (one period lagged) are time-varying. Here we use the same bandwidth for estimation of fixed and time-varying parameter, i.e. $H = H_z$.

Figure 10 presents the plot of residuals $\widehat{u}_t$ of PTVR regression. It shows that in the first half of the sample, residuals have larger volatility than in the second one, i.e. the variance of residuals is not constant. Figure 11 reports testing results for correlation in residuals for this period. We find significant autocorrelation at some lags at 5% significance level, and no correlation at 1% significance level. Comparing to the correlogram of y_t in Figure 8, correllation in residuals is significantly reduced and the noise u_t in PTVR regression seems uncorrelated.

To evaluate the quality of the forecasts, similarly to Section 3.2, for each forecasting method, PTVR, FTVR, OLS and AR(1), we conduct in-sample forecasting over period $t = 100, ..., 279$ and compute MSFE and MSFE ratio. (Here $t_0 = 100$, $n = 279$). The results are shown in Table 5.

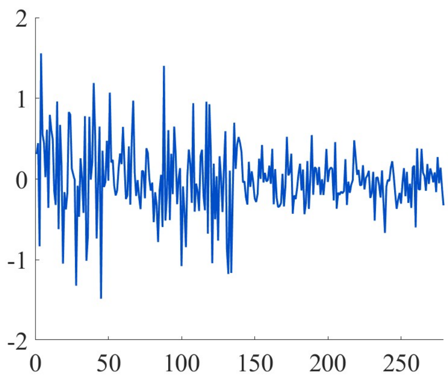


Figure 10: Plot of PTVR residuals $\widehat{u}_t$, $H = n^{0.7}$

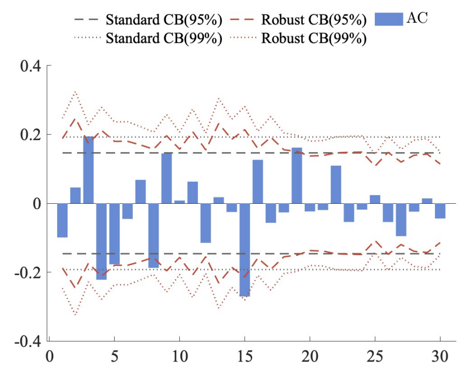


Figure 11: Correlogram of $\widehat{u}_t$, $t = 101, \cdots, 279$, $H = n^{0.7}$.

Figure 12 shows the plot of the MSFEs of the PTVR forecasts using a denser grid of bandwidths. It is clear that the optimal bandwidth is around $H = n^{0.8}$.

Overall, PTVR and FTVR methods produce smaller forecast errors compared with OLS and AR(1), and PTVR performs somewhat better than FTVR. As the bandwidth increases, the forecast MSE of PTVR and FTVR methods achieve their minimum around $H = n^{0.8}$. Hence, $H = n^{0.8}$ can be a good bandwidth choice in this empirical exercise and for these sample sizes.

Table 5: MSFE and MSFE ratio

Bandwidth		PTVR	FTVR	OLS	AR(1)
$H = n^{0.5}$	MSFE	0.178	0.198	0.143	0.198
	MSFE (ratio)	1	1.112	0.803	1.112
$H = n^{0.6}$	MSFE	0.151	0.165	0.143	0.198
	MSFE (ratio)	1	1.093	0.947	1.311
$H = n^{0.7}$	MSFE	0.140	0.142	0.143	0.198
	MSFE (ratio)	1	1.014	1.021	1.414
$H = n^{0.8}$	MSFE	0.137	0.134	0.143	0.198
	MSFE (ratio)	1	0.978	1.044	1.445
$H = n^{0.9}$	MSFE	0.139	0.135	0.143	0.198
	MSFE (ratio)	1	0.971	1.029	1.425

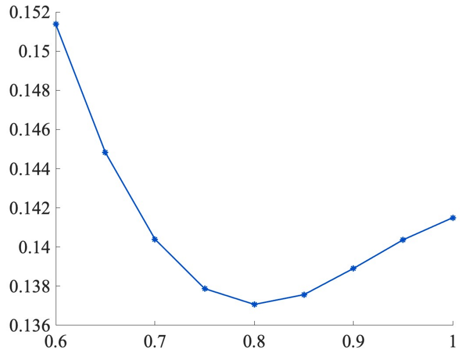


Figure 12: MSFE of PTVR forecast for $H = n^h$.

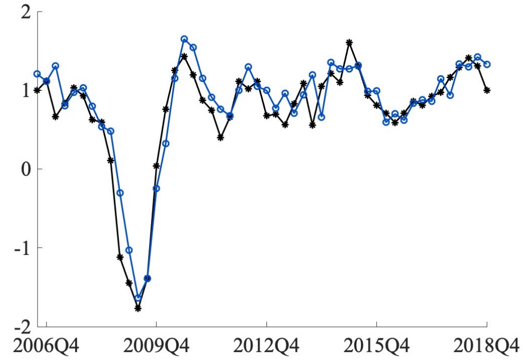


Figure 13: PTVR forecasts, $H = n^{0.8}$

Figure 13 displays the plots of the last 50 true values of y_t (GDPC1) and predicted values $y_{t|t-1}$ forecasted using the PTVR method. The predicted values by PTVR (blue solid line) and the true values (black solid line) are close and almost coincide.

5 Conclusion

In this paper we develop estimation and inference theory for a new general partially time-varying regression model. The setting of the model permits for general heterogeneity in regressors and noise and structural change of time-varying regression coefficients over time. The asymptotic estimation theory for this model has a number of novelties. In particular, the fixed parameter can be estimated with parametric rate and standard errors can be easily computed. Unlike the rest of the literature, we allow stochastic scale and parameter processes. Our assumptions on scales, parameters, regressors and noise are considerably milder that in previous work. The Monte Carlo study confirms the excellent performance of parameter estimation and inference in finite samples and in forecasting on simulated data. We present an empirical illustration, where we apply PTVR modelling to forecasting, leads also to promising results. The theoretical and empirical findings of this paper demonstrate the clear potential of regression models that combine fixed and time-varying parameters. Future work can be focused on methods that determine what subsets of regressors should have fixed or time-varying parameters.

References

BROWN, R. L., J. DURBIN, AND J. M. EVANS (1974): "Techniques for testing the constancy of regression relationships over time," *Journal of the Royal Statistical Association, Series A*, 138, 149–163.

CHEN, B. (2015): "Modeling and testing smooth structural changes with endogenous regressors," *Journal of Econometrics*, 185, 196–215.

CHEN, B., AND Y. HONG (2012): "Testing for Smooth Structural Changes in Time Series Models via Nonparametric Regression," *Econometrica*, 80(3), 1157–1183.

CHOW, A. (1960): "Tests of equality between sets of coefficients in two linear regressions," *Econometrica*, 28, 591–605.

DAHLHAUS, R. (1997): "Fitting time series models to nonstationary processes," *Annals of Statistics*, 25, 1–37.

DAHLHAUS, R., AND L. GIRAITIS (1998): "On the optimal segment length for parameter estimates for locally stationary time series," *Journal of Time Series Analysis*, 19, 629–655.

DENDRAMIS, Y., L. GIRAITIS, AND G. KAPETANIOS (2021): "Estimation of time-varying covariance matrices for large data sets," *Econometric Theory*, 37, 1100–1134.

EICKER, F. (1963): "Asymptotic normality and consistency of the last squares estimators for families of linear regressions," *Annals of Mathematical Statistics*, 34, 447 – 456.

FAN, J., AND T. HUANG (2005): "Profile likelihood inferences on semiparametric varying-coefficient partially linear models," *Bernoulli*, 11, 1031—57.

GIRAITIS, L., G. KAPETANIOS, AND T. YATES (2014): "Inference on stochastic time-varying coefficient models," *Journal of Econometrics*, 179, 46–65.

GIRAITIS, L., G. KAPETANIOS, AND M. MARCELLINO (2021): "Time-varying instrumental variable estimation," *Journal of Econometrics*, 224, 394–415.

GIRAITIS, L., KOUL, H.L., AND D. SURGAILIS (2012): *Large Sample Inference for Long memory Processes*. Imperial College Press, London.

GIRAITIS, L., Y. LI, AND P.C.B. PHILLIPS (2024): "Robust inference on correlation under general heterogeneity," *Journal of Econometrics*, 240(1), 105691.

GIRAITIS, L., G. KAPETANIOS, AND Y. LI (2024): "Regression modelling under general heterogeneity," *Preprint*, No. 983, Queen Mary University of London, https://www.qmul.ac.uk/sef/research/workingpapers/2024/items/983.html

Hu, Z., I. Kasparis, and Q. Wang (2024): "Time-varying parameter regressions with stationary presistent data," *Econometric Theory*, 1-17, published online.

Kapetanios, G., and T. Yates (2008): "An analysis of inflation persistence and the Great Moderation using a model of deterministic structural change," *Mimeo, Bank of England.*

Kristensen, D. (2012): "Non-parametric detection and estimation of structural change," *Econometrics Journal*, 15, 420–461.

Muller, U. K., and P. M. Petalas (2010): "Efficient estimation of the parameter path in unstable time series models," *Review of Economic Studies*, 77, 1508–1539.

Muller, U. K., and M. W. Watson (2008): "Testing models of low-frequency variability," *Econometrica*, 76, 979–1016.

Ploberger, W., and W. Kramer (1992): "The CUSUM test with OLS residuals," *Econometrica*, 60, 271–285.

Phillips, P.C.B., D. Li, and J. Gao (2017): "Estimating smooth structural change in cointegration models," *Journal of Econometrics*, 196(1), 180–195.

Priestley, M. (1965): "Evolutionary spectra and nonstationary processes," *Journal of Royal Statistical Society, Series B*, 27, 204–237.

Robinson, P. M. (1988): "Root-N-consistent semiparametric regression," in *Econometrica*, 56, 931–54.

Robinson, P. M. (1989): "Nonparametric estimation of time varying parameters," in *Economic Structural Change: Analysis and Forecasting*, ed. by P. Hackl, pp. 253–264. Springer Berlin.

Robinson, P.M. (1991): "Time varying nonlinear regression," in *Economic Structural Change: Analysis and Forecasting*, ed. by P. Hackl, pp. 179–190. Springer Berlin.

White, H. (1980): "A heteroskedasticity-honsistent covariance matrix estimator and a direct test for heteroskedasticity," in *Econometrica*, *48*(4), 817–838.

Supplement to “Partial Time-Varying Regression Modelling under General Heterogeneity”

Liudas Giraitis[1], George Kapetanios[2], Yufei Li[2], Tien Chuong Nguyen[3]

[1]Queen Mary University of London

[2]King's College London

[3]Vietnam National University

December 18, 2024

This Supplement provides proofs of the results given in the text of the main paper. It is organised as follows: Section 6 provide proofs of the main theorems. Section 7 contains auxiliary technical lemmas used in the proofs. Section 8 provides some additional Monte Carlo simulations which are not covered by the main paper.

Formula numbering in this supplement includes the section number, e.g. (6.1), and references to lemmas are signified as “Lemma 6.#”, e.g. Lemma 6.1. Theorem references to the main paper include section number and are signified, e.g. as Theorem 2.1, while equation references do not include section number, e.g. (1), (2).

In the proofs, C stands for a generic positive constant which may assume different values in different contexts.

6 Appendix. Proofs

6.1 Proofs of Lemma 2.1 and Theorem 2.1 and 2.2

In this section we provide proofs of Lemma 2.1, Theorem 2.1 and Corollary 2.1, and Lemma 2.2 for the estimator $\widehat{\alpha}$ of the fixed parameter α, and of Theorem 2.2 and Corollary 2.2 for the estimator $\widehat{\beta}_t$ of the time-varying parameter β_t.

Proof of Lemma 2.1. We will find the minimizer $\widehat{\alpha}$ by solving the equation for the gradient,

$$\nabla L(\alpha)|_{\alpha=\widehat{\alpha}} = 0.$$

Notice that $\widehat{\beta}_{t,\alpha} = \widehat{\beta}_{zy,t} - \widehat{\beta}_{zx,t}\alpha.$,

$$y_t - x_t'\alpha - z_t'\widehat{\beta}_{t,\alpha} \;=\; (y_t - z_t'\widehat{\beta}_{zy,t}) - (x_t' - z_t'\widehat{\beta}_{zx,t})\alpha, \tag{6.1}$$

$$L(\alpha) \;=\; \sum_{t=1}^{n}\Big((y_t - z_t'\widehat\beta_{zy,t}) - (x_t' - z_t'\widehat\beta_{zx,t})\alpha\Big)^2.$$

Hence,

$$\begin{aligned}\nabla L(\alpha) &= -2\sum_{t=1}^{n}\Big((y_t - z_t'\widehat\beta_{zy,t}) - (x_t' - z_t'\widehat\beta_{zx,t})\alpha\Big)(x_t' - z_t'\widehat\beta_{zx,t})' \\ &= -2\sum_{t=1}^{n}(x_t - \widehat\beta_{zx,t}'z_t)(y_t - z_t'\widehat\beta_{zy,t}) + 2\sum_{t=1}^{n}(x_t - \widehat\beta_{zx,t}'z_t)(x_t - \widehat\beta_{zx,t}'z_t)'\alpha,\end{aligned}$$

which implies (5). The claim (6) follows setting $\alpha = \widehat\alpha$ in $\beta_{t,\alpha}$, $\widehat\beta_t = \beta_{t,\widehat\alpha}$. □

Proof of Theorem 2.1. Recall notation,

$$\begin{aligned}\beta_{zx,t} &= q_{zz,t}^{-1}q_{zx,t}, \quad q_{zz,t} = E[z_tz_t'|\mathcal{F}_n^*], \quad q_{zx,t} = E[z_tx_t'|\mathcal{F}_n^*], && (6.2)\\ \widehat\beta_{zx,t} &= S_{zz,t}^{-1}S_{zx,t}, \\ \widehat\beta_{zy,t} &= S_{zz,t}^{-1}S_{zy,t}, \\ S_{zz,t} &= \sum_{j=1}^{n} b_{n,tj}z_jz_j', \quad S_{zx,t} = \sum_{j=1}^{n} b_{n,tj}z_jx_j', && (6.3)\\ S_{zy,t} &= \sum_{j=1}^{n} b_{n,tj}z_jy_j, \quad S_{z\zeta,t} = \sum_{j=1}^{n} b_{n,tj}z_j\zeta_j,\end{aligned}$$

where $\mathcal{F}_n^* = \sigma\big(h_t, g_{xt}, g_{zt}, t = 1, ..., n\big)$ is the information set generated by the scales.

Introduce the regression model

$$\zeta_j \;=\; \beta_j'z_j + u_j \tag{6.4}$$

with a dependent variable ζ_j and time-varying parameter β_j. Denote

$$\widetilde\beta_t = S_{zz,t}^{-1}S_{z\zeta,t} \tag{6.5}$$

the time-varying OLS estimator of parameter β_t in model (6.4).

Set

$$\begin{aligned}\widehat\xi_j &= y_j - z_j'\widehat\beta_{zy,j}, && (6.6)\\ v_j &= x_j - \beta_{zx,j}'z_j, \\ \widehat v_j &= x_j - \widehat\beta_{zx,j}'z_j.\end{aligned}$$

Notice that we can write $y_j = \alpha'x_j + \beta_j'z_j + u_j$ in (1) as

$$y_j \;=\; x_j'\alpha + \zeta_j. \tag{6.7}$$

Therefore,

$$\begin{aligned}\widehat{\beta}_{zy,t} &= S_{zz,t}^{-1}S_{zy,t} = S_{zz,t}^{-1}S_{zx,t}\alpha + \widehat{S}_{zz,t}^{-1}S_{z\zeta,t} \\ &= \widehat{\beta}_{zx,t}\alpha + \widetilde{\beta}_t.\end{aligned}$$

Then

$$\begin{aligned}\widehat{\xi}_j = y_j - z_j'\widehat{\beta}_{zy,j} &= \{x_j'\alpha + z_j'\beta_t + u_j\} - \{z_j'\widehat{\beta}_{zx,j}\alpha + \widetilde{\beta}_j\} \\ &= (x_j - \widehat{\beta}_{zx,j}' z_j)'\alpha + u_j + z_j'(\beta_j - \widetilde{\beta}_j) \\ &= \widehat{v}_j'\alpha + u_j + z_j'(\beta_j - \widetilde{\beta}_j).\end{aligned}$$

So, we can write the estimator $\widehat{\alpha}$, given in (5), as

$$\begin{aligned}\widehat{\alpha} = S_{\widehat{v}\widehat{v}}^{-1}S_{\widehat{v}\widehat{\xi}} &= \alpha + S_{\widehat{v}\widehat{v}}^{-1}S_{\widehat{v}u} + S_{\widehat{v}\widehat{v}}^{-1}R_n, \quad R_n = \sum_{j=1}^{n}\widehat{v}_j z_j'(\beta_j - \widetilde{\beta}_j) \\ &= \alpha + S_{vv}^{-1}S_{vu} + \{S_{\widehat{v}\widehat{v}}^{-1}S_{\widehat{v}u} - S_{vv}^{-1}S_{vu}\} + S_{\widehat{v}\widehat{v}}^{-1}R_n, \qquad (6.8)\end{aligned}$$

where

$$\begin{aligned}&S_{\widehat{v}\widehat{v}} = \sum_{t=1}^{n}\widehat{v}_t\widehat{v}_t', \quad S_{\widehat{v}u} = \sum_{t=1}^{n}\widehat{v}_t u_t, \quad S_{vv} = \sum_{t=1}^{n}v_t v_t', \qquad (6.9) \\ &S_{vu} = \sum_{t=1}^{n}v_t u_t, \quad S_{\widehat{v}\widehat{\xi}} = \sum_{t=1}^{n}\widehat{v}_t\widehat{\xi}_t.\end{aligned}$$

In (6.89) of Lemma 6.5 we show that

$$S_{\widehat{v}\widehat{v}}^{-1}S_{\widehat{v}u} - S_{vv}^{-1}S_{vu} = o_p(n^{-1/2}),$$

and in Lemma 6.6 it is shown that $S_{\widehat{v}\widehat{v}}^{-1}R_n = o_p(n^{-1/2})$. Then,

$$\widehat{\alpha} - \alpha = S_{vv}^{-1}S_{vu} + o_p(n^{-1/2}). \qquad (6.10)$$

Observe that the OLS estimator $\widetilde{\alpha}$ of parameter α in OLS estimation of fixed parameter in the regression model

$$y_j^* = \alpha' v_j + u_j \qquad (6.11)$$

has property

$$\widetilde{\alpha} - \alpha = S_{vv}^{-1}S_{vy^*} - \alpha = S_{vv}^{-1}S_{vu} \qquad (6.12)$$

where $S_{vy^*} = \sum_{t=1}^{n} v_t y_t^*$.

The regressors v_j in (6.11) have the following properties, see (6.15). Denote

$$\nu_t = \eta_{xt} - E[\eta_{xt}\eta_{zt}'](E[\eta_{zt}\eta_{zt}'])^{-1}\eta_{zt}. \tag{6.13}$$

By (8), we can write

$$\begin{aligned} x_t &= I_{xt}\eta_{xt}, \quad I_{xt} = \mathrm{diag}(g_{x1,t}, ..., g_{xq,t}), \\ z_t &= I_{zt}\eta_{zt}, \quad I_{zt} = \mathrm{diag}(g_{z1,t}, ..., g_{zp,t}), \end{aligned} \tag{6.14}$$

where $\{\eta_{xk,t}\}$, $\{\eta_{zj,t}\}$ are stationary sequences which may have non-zero mean. Hence,

$$\begin{aligned} v_t &= x_t - \beta_{zx,t}' z_t = x_t - E[x_t z_t' | \mathcal{F}_n^*](E[z_t z_t' | \mathcal{F}_n^*])^{-1} z_t \\ &= I_{xt}\eta_{xt} - I_{xt}E[\eta_{xt}\eta_{zt}']I_{zt}I_{zt}^{-1}(E[\eta_{zt}\eta_{zt}'])^{-1}I_{zt}^{-1}I_{zt}\eta_{zt} \\ &= I_{xt}\{\eta_{xt} - E[\eta_{xt}\eta_{zt}'](E[\eta_{zt}\eta_{zt}'])^{-1}\eta_{zt}\} = I_{xt}\nu_t, \\ v_t v_t' &= I_{xt}\nu_t\nu_t' I_{xt}. \end{aligned} \tag{6.15}$$

Under assumption of theorem, ν_t is a covariance stationary sequence, see Remark 2.1, and the scale factor I_{xt} is independent of $\{\nu_t\}$. It is easy to see that regressors v_t has similar structure as x_j and satisfy assumptions of the corresponding regression model (1) considered in Giraitis, Kapetanios, Li (2024). Theorem 2.1 together with Lemma 2.1 and Corollary 2.1 of that paper imply (20) and $\sqrt{\omega_{kk}} \asymp_p n^{-1/2}$.

This completes the proof of the Theorem 2.1. □

Proof of Corollary 2.1. We will show that

$$n(\widehat{\Omega}_{\alpha,n} - \Omega_{\alpha,n}) = o_p(1) \tag{6.16}$$

which implies $n(\widehat{\omega}_{kk} - \omega_{kk}) = o_p(1)$. By (20), $\omega_{kk} \asymp_p n^{-1}$. Therefore,

$$\frac{\widehat{\omega}_{kk}}{\omega_{kk}} = 1 + \frac{\widehat{\omega}_{kk} - \omega_{kk}}{\omega_{kk}} = 1 + \frac{n(\widehat{\omega}_{kk} - \omega_{kk})}{n\omega_{kk}} = 1 + o_p(1).$$

This, together with (20) of Theorem 2.1 implies

$$\frac{\widehat{\alpha}_k - \alpha_k}{\sqrt{\widehat{\omega}_{kk}}} = \frac{\sqrt{\omega_{kk}}}{\sqrt{\widehat{\omega}_{kk}}}\frac{\widehat{\alpha}_k - \alpha_k}{\sqrt{\omega_{kk}}} \to_d \mathcal{N}(0,1)$$

which proves the claim (22) of the corollary.

It remains to prove (6.16). Recall that

$$\begin{aligned} \Omega_{\alpha,n} &= (E[S_{vv}|\mathcal{F}_n^*])^{-1}E[S_{vvuu}|\mathcal{F}_n^*](E[S_{vv}|\mathcal{F}_n^*])^{-1} = (\omega_{jk}), \\ \widehat{\Omega}_{\alpha,n} &= S_{\widehat{v}\widehat{v}}^{-1}S_{\widehat{v}\widehat{v}\widehat{u}\widehat{u}}S_{\widehat{v}\widehat{v}}^{-1} = (\widehat{\omega}_{jk}), \quad \widehat{u}_t = y_t - \widehat{\alpha}' x_t - \widehat{\beta}_t' z_t. \end{aligned}$$

By (6.56) and (6.55) of Lemma 6.3,

$$nS_{\widehat{v}\widehat{v}}^{-1} = n(E[S_{vv}|\mathcal{F}_n^*])^{-1} + o_p(1), \quad n(E[S_{vv}|\mathcal{F}_n^*])^{-1} = O_p(1).$$

By (6.61) and (6.60) of Lemma 6.3,

$$n^{-1}S_{\widehat{v}\widehat{v}\widehat{u}\widehat{u}} = n^{-1}E[S_{vvuu}|\mathcal{F}_n^*] + o_p(1), \quad n^{-1}E[S_{vvuu}|\mathcal{F}_n^*] = O_p(1).$$

Then,

$$\begin{aligned} n\widehat{\Omega}_{\alpha,n} &= \{nS_{\widehat{v}\widehat{v}}^{-1}\}\{n^{-1}S_{\widehat{v}\widehat{v}\widehat{u}\widehat{u}}\}\{nS_{\widehat{v}\widehat{v}}^{-1}\} \\ &= \{n(E[S_{vv}|\mathcal{F}_n^*])^{-1} + o_p(1)\}\{n^{-1}E[S_{vvuu}|\mathcal{F}_n^*] + o_p(1)\}\{(E[S_{vv}|\mathcal{F}_n^*])^{-1} + o_p(1)\} \\ &= n\Omega_{\alpha,n} + o_p(1) \end{aligned}$$

which proves (6.16). This completes the proof of the corollary. □

Proof of Theorem 2.2. By (1), $y_j = x_j'\alpha + z_j'\beta_j + u_j$. Denote by

$$\zeta_j = z_j'\beta_j + u_j \tag{6.17}$$

the regression model with a time-varying parameter. We can write

$$\begin{aligned} y_j - \widehat{\alpha}'x_j &= \{z_j'\beta_j + u_j\} + x_j'(\alpha - \widehat{\alpha}) \\ &= \zeta_j + x_j'(\alpha - \widehat{\alpha}). \end{aligned}$$

Then the estimator $\widehat{\beta}_t$ in (6), with weights $b_{n,tj}$ computed using bandwidth H_z, can be written as

$$\begin{aligned} \widehat{\beta}_t &= \big(\sum_{j=1}^n b_{n,tj}z_jz_j'\big)^{-1}\big(\sum_{j=1}^n b_{n,tj}z_j\big(y_j - \widehat{\alpha}'x_j\big)\big) \\ &= \big(\sum_{j=1}^n b_{n,tj}z_jz_j'\big)^{-1}\big(\sum_{j=1}^n b_{n,tj}z_j\big(\zeta_j + x_j'(\alpha - \widehat{\alpha})\big)\big) \\ &= S_{zz,t}^{-1}S_{z\zeta,t} + S_{zz,t}^{-1}S_{zx,t}(\alpha - \widehat{\alpha}) \\ &= \widetilde{\beta}_t + S_{zz,t}^{-1}S_{zx,t}(\alpha - \widehat{\alpha}), \quad \widetilde{\beta}_t = S_{zz,t}^{-1}S_{z\zeta,t}. \end{aligned} \tag{6.18}$$

Hence,

$$\widehat{\beta}_t - \beta_t = \widetilde{\beta}_t - \beta_t + S_{zz,t}^{-1}S_{zx,t}(\alpha - \widehat{\alpha}). \tag{6.19}$$

Bound

$$||S_{zz,t}^{-1}S_{zx,t}(\alpha - \widehat{\alpha})|| \le ||K_tS_{zz,t}^{-1}||\,||K_t^{-1}S_{zx,t}||\,||\alpha - \widehat{\alpha}||, \quad K_t = \sum_{j=1}^n b_{n,tj}. \tag{6.20}$$

By Theorem 2.1, $||\alpha - \widehat{\alpha}|| = O_p(n^{-1/2})$. In (6.28) of Lemma 6.1 below it is shown that $||K_t S_{zz,t}^{-1}|| = O_p(1)$ and in (6.27) we show that $||K_t^{-1} S_{zx,t}|| = O_p(1)$. Hence, $S_{zz,t}^{-1} S_{zx,t}(\alpha - \widehat{\alpha}) = O_p(n^{-1/2})$ which together with (6.19), yields

$$\widehat{\beta}_t - \beta_t = \widetilde{\beta}_t - \beta_t + O_p(n^{-1/2}). \tag{6.21}$$

Notice that $H_z = o(n)$ implies $n^{-1/2} = o(H_z^{-1/2})$.

Observe that the regression model $\zeta_j = z_j' \beta_j + u_j$, (6.17), with a time-varying parameter β_j is a special case of the corresponding regression model (18) considered in Giraitis, Kapetanios, Li (2024) and $\widetilde{\beta}_t$ in (6.18) is a kernel estimator of time-varying parameter β_t which asymptotic properties were established in that paper. Our model (6.4) satisfies assumptions of Theorem 3.1 in Giraitis, Kapetanios, Li (2024) which implies the claims (29) and (30) of our Theorem 2.2.

This completes the proof of the theorem. □

Proof of Corollary 2.2. The proof follows the same pattern as in the proof of Corollary 2.1. For completeness, we include a detailed proof. It suffices to show that

$$H_z(\widehat{\Omega}_{\beta,t} - \Omega_{\beta,t}) = o_p(1). \tag{6.22}$$

Then, $H_z(\widehat{\omega}_{kk,t} - \omega_{kk,t}) = o_p(1)$ which together with (30), $\omega_{kk} \asymp_p H_z^{-1}$, implies

$$\frac{\widehat{\omega}_{kk,t}}{\omega_{kk,t}} = 1 + \frac{\widehat{\omega}_{kk,t} - \omega_{kk,t}}{\omega_{kk,t}} = 1 + \frac{H_z(\widehat{\omega}_{kk,t} - \omega_{kk,t})}{H_z \omega_{kk,t}} = 1 + o_p(1).$$

Together with (30) of Theorem 2.2, this verifies the claim (32) of corollary:

$$\frac{\widehat{\beta}_{kt} - \beta_{kt}}{\sqrt{\widehat{\omega}_{kk,t}}} = \frac{\sqrt{\omega_{kk,t}}}{\sqrt{\widehat{\omega}_{kk,t}}} \frac{\widehat{\beta}_{kt} - \beta_{kt}}{\sqrt{\omega_{kk,t}}} \to_d \mathcal{N}(0,1).$$

Next we verify (6.22). Recall that

$$\begin{aligned} \Omega_{\beta,t} &= (E[S_{zz,t}|\mathcal{F}_n^*])^{-1} E[S_{zzuu,t}|\mathcal{F}_n^*] (E[S_{zz,t}|\mathcal{F}_n^*])^{-1} = (\omega_{jk,t}), \\ \widehat{\Omega}_{\beta,t} &= S_{zz,t}^{-1} S_{zz\widehat{u}\widehat{u},t} S_{zz,t}^{-1} = (\widehat{\omega}_{jk,t}), \quad \widehat{u}_t = y_t - \widehat{\alpha}' x_t - \widehat{\beta}_t' z_t. \end{aligned}$$

Let $K_t = \sum_{j=1}^n b_{n,tj}$ where $b_{n,tj}$ are defined using bandwidth H_z. Then, by (6.40), $K_t \asymp H_z$. By (6.29) of Lemma 6.1,

$$K_t S_{zz,t}^{-1} = K_t (E[S_{zz,t}|\mathcal{F}_n^*])^{-1} + o_p(1).$$

By (6.63) and (6.62) of Lemma 6.3,

$$K_t^{-1} S_{zz\widehat{u}\widehat{u},t} = K_t^{-1} E[S_{zzuu,t}|\mathcal{F}_n^*] + o_p(1), \quad K_t^{-1} E[S_{zzuu,t}|\mathcal{F}_n^*] = O_p(1).$$

Then,

$$\begin{aligned}K_t\,\widehat{\Omega}_{\beta,t} &= \{K_tS_{zz,t}^{-1}\}\{K_t^{-1}S_{zz\widehat{u}\widehat{u},t}\}\{K_tS_{zz,t}^{-1}\}\\ &= \{K_t(E[S_{zz,t}|\mathcal{F}_n^*])^{-1}+o_p(1)\}\{K_t^{-1}E[S_{zzuu,t}|\mathcal{F}_n^*]+o_p(1)\}\{(E[S_{zz,t}|\mathcal{F}_n^*])^{-1}+o_p(1)\}\\ &= K_t\,\Omega_{\beta,t}+o_p(1).\end{aligned}$$

This proves (6.22) and completes the proof of the corollary. □

Proof of Lemma 2.2. We prove the first claim in (19). (The proof of the secon claim is similar.)

Clearly, it suffices to prove (19) for

$$w_j=\eta_{z1,j}\eta_{z1,j}=\sum_{i_1,i_2=0}^{\infty}a_{z1,i_1}a_{z1,i_2}\xi_{z1,j-i_1}\xi_{z1,j-i_2}.$$

Denote

$$c_{j,i_1i_2}=a_{z1,i_1}a_{z1,i_2},\quad \zeta_{j,i_1i_2}=\xi_{z1,j-i_1}\xi_{z1,j-i_2}-E[\xi_{z1,j-i_1}\xi_{z1,j-i_2}].$$

Then, we can write

$$w_j-E[w_j]\;=\;\sum_{i_1,i_2=0}^{\infty}c_{j,i_1i_2}\zeta_{j,i_1i_2}.$$

Set $m=b\log n$ where b is such that $b(\log\rho)/2\le -8$. Then $\rho^{m/2}=\exp(b(\log\rho)\log n)\le\exp(-8\log n)\le n^{-8}$. Write

$$\begin{aligned}w_j-E[w_j] &= \sum_{i_1,i_2=0}^{m/2}c_{j,i_1i_2}\zeta_{j,i_1i_2}+\sum_{i_1,i_2=0:\max(i_1,i_2)>m/2}c_{j,i_1i_2}\zeta_{j,i_1i_2}\\ &= r^+_{mt,j}+r_{mt,j}.\end{aligned}$$

Under assumptions of lemma, for $j\ge m$ and $i_1,i-2\le m/2$, $E[\xi_{z1,j-i_1}\xi_{z1,j-i_2}|\mathcal{F}_t]=E[\xi_{z1,j-i_1}\xi_{z1,j-i_2}]$ and hence, $E[\zeta_{j,i_1i_2}|\mathcal{F}_t]=0$ and $E[r^+_{mt,j}|\mathcal{F}_t=0$.

On the other hand, $E\zeta^4_{j,i_1i_2}\le C$ where $C<\infty$ does not depend on j,i_1,i_2. Hence, it is easy to see that,

$$\begin{aligned}Er^4_{mt,j} &\le C\sum_{i_1,\dots,i_8=0:i_1>m/2}^{\infty}|c_{j,i_1i_2}c_{j,i_3i_4}c_{j,i_5i_6}c_{j,i_7i_8}|\\ &\le C\sum_{i=m/2+1}^{\infty}|a_{z1,i}|\Big(\sum_{i=0}^{\infty}|a_{z1,i}|\Big)^7\le C\sum_{i=m/2+1}^{\infty}\rho^i\le C\rho^{m/2}\le Cn^{-8}.\end{aligned}$$

This proves the first claim in (19): $(Er^4_{mt,j})^{1/4}\le Cn^{-2}$. □

6.2 Auxiliary Lemmas

This section contains auxiliary lemmas used in the proofs of Theorem 2.1 and 2.2.

Recall notation K_t given in (6.20). Denote

$$\begin{aligned} Q_{zz,t} &= K_t^{-1} S_{zz,t}, \quad Q_{zx,t} = K_t^{-1} S_{zx,t}, \\ q_{zz,t} &= E[z_t z_t' | \mathcal{F}_n^*], \quad q_{zx,t} = E[z_t x_t' | \mathcal{F}_n^*]. \end{aligned} \tag{6.23}$$

Lemma 6.1. *Suppose that the assumptions of Theorem 2.1 are satisfied. Then for bandwidth H such that $H = o(n)$, $H \to \infty$, the following holds:*

$$||q_{zz,t}^{-1}|| \le C, \quad E||\beta_{zx,t}||^8 \le C, \quad E||q_{zx,t}||^4 \le C, \tag{6.24}$$

$$E||Q_{zz,t} - q_{zz,t}||^4 \le C\big(H^{-2}m^2 + (H/n)^{4\gamma_1}\big), \tag{6.25}$$

$$E||Q_{zx,t} - q_{zx,t}||^4 \le C\big(H^{-2}m^2 + (H/n)^{4\gamma_1}\big), \tag{6.26}$$

$$E||Q_{zx,t}||^4 \le C, \tag{6.27}$$

$$||Q_{zz,t}^{-1}|| = O_p(1), \quad \textit{for } t = t_n \in \{1, ..., n\}, \tag{6.28}$$

$$Q_{zz,t}^{-1} = E[Q_{zz,t}|\mathcal{F}_n^*]^{-1} + o_p(1). \tag{6.29}$$

where $C < \infty$ does not depend on t, n, and m is the same as in Assumption 2.5 (iii).

In addition, if H satisfies assumption (14), then

$$\max_{t=1,\dots,n} ||Q_{zz,t} - q_{zz,t}|| = o_p(1), \quad \max_{t=1,\dots,n} ||Q_{zx,t} - q_{zx,t}|| = o_p(1), \tag{6.30}$$

$$\max_{t=1,\dots,n} ||Q_{zz,t}^{-1}|| = O_p(1). \tag{6.31}$$

Proof of Lemma 6.1.

Proof of (6.24). Recall (6.14). By Assumption 2.2(ii), $E[\eta_{zt}\eta_{zt}'] = E[\eta_{z1}\eta_{z1}'] = \Sigma_{zz}$ is a positive definite matrix, so that $||\Sigma_{zz}^{-1}|| < \infty$. By definition of $\mathcal{F}_n^*$,

$$\begin{aligned} q_{zz,t} &= E[z_t z_t' | \mathcal{F}_n^*] = I_{zt} E[\eta_{zt}\eta_{zt}'] I_{zt} = I_{zt}\Sigma_{zz} I_{zt}, \\ q_{zx,t} &= E[z_t x_t' | \mathcal{F}_n^*] = I_{zt} E[\eta_{zt}\eta_{xt}'] I_{xt} = I_{zt} E[\eta_{z1}\eta_{x1}'] I_{xt}. \end{aligned} \tag{6.32}$$

Observe that $I_{zt}^{-1} = \mathrm{diag}(g_{z1,t}^{-1}, ..., g_{zp,t}^{-1})$ and by Assumption 2.3, $g_{zj,t} \ge c_0 > 0$ for $j = 1, ..., p$ and all t. Therefore,

$$||I_{zt}^{-1}||^2 = \sum_{j=1}^{p} g_{zj,t}^{-2} \le c_0^{-2} p.$$

Hence,

$$q_{zz,t}^{-1} = I_{xt}^{-1}\Sigma_{zz}^{-1} I_{xt}^{-1},$$

$$||q_{zz,t}^{-1}|| \le ||I_{zt}^{-1}||\,||\Sigma_{zz}^{-1}||\,||I_{zt}^{-1}|| \le c_0^{-2}p||\Sigma_{zz}^{-1}|| \le c < \infty$$

where c does not depend on t, n which proves the first claim in (6.24).

On the other hand, $\beta_{zx,t} = q_{zz,t}^{-1}q_{zx,t}$. Therefore, by (6.32),

$$\begin{aligned}
\beta_{zx,t} &= \{I_{zt}E[\eta_{zt}\eta_{zt}']I_{zt}\}^{-1}\{I_{zt}E[\eta_{zt}\eta_{xt}']I_{xt}\} = I_{zt}^{-1}E[\eta_{zt}\eta_{zt}']^{-1}E[\eta_{zt}\eta_{xt}']I_{xt},\\
||\beta_{zx,t}|| &\le ||I_{zt}^{-1}||\,||\Sigma_{zz,t}^{-1}||\,||E[\eta_{zt}\eta_{xt}']||\,||I_{xt}|| \le C||I_{xt}||,
\end{aligned}$$

where $C < \infty$ does not depend on t. Thus,

$$\begin{aligned}
E||\beta_{zx,t}||^8 &\le CE||I_{xt}||^8 \le C,\\
E||q_{zx,t}||^4 &= E||I_{zt}E[\eta_{z1}\eta_{x1}']I_{xt}||^4\\
&\le E[||I_{zt}||^4||I_{xt}||^4]\,||E[\eta_{z1}\eta_{x1}']^4|| \le C(E||I_{zt}||^8 + E||I_{xt}||^8)(E||\eta_{z1}||^8 + E||\eta_{x1}||^8) \le C
\end{aligned}$$

by Assumption 2.3(i) where $C < \infty$ does not depend on t, n. This proves the second and third claim in (6.24).

Proof of (6.25). (Proof of (6.26) is similar). Write

$$\begin{aligned}
z_jz_j' - E[z_tz_t'|\mathcal{F}_n^*] &= \{z_jz_j' - E[z_jz_j'|\mathcal{F}_n^*]\} + \{E[z_jz_j'|\mathcal{F}_n^*] - E[z_tz_t'|\mathcal{F}_n^*]\},\\
&= \{I_{zj}(\eta_{zj}\eta_{zj}' - E[\eta_{zj}\eta_{zj}'])I_{zj}\} + \{I_{zj}E[\eta_{z1}\eta_{z1}']I_{zj} - I_{zt}E[\eta_{z1}\eta_{z1}']I_{zt}\}.
\end{aligned}$$

By Assumption 2.2(i), $\{\eta_{zt}\eta_{zt}'\}$ is a stationary time series with finite 4-th moment. Therefore, $E[\eta_{zj}\eta_{zj}'] = E[\eta_{z1}\eta_{z1}']$ for any $j \ge 1$. Hence,

$$\begin{aligned}
Q_{zz,t} - q_{zz,t} &= Q_{zz,t}^* + R_{zz,t}^*, \qquad (6.33)\\
Q_{zz,t}^* &= K_t^{-1}\sum_{j=1}^n b_{n,tj}I_{zj}(\eta_{zj}\eta_{zj}' - E[\eta_{zj}\eta_{zj}'])I_{zj} = \{q_{\ell m,t}\},\\
R_{zz,t}^* &= K_t^{-1}\sum_{j=1}^n b_{n,tj}\Big(I_{zj}E[\eta_{z1}\eta_{z1}']I_{zj} - I_{zt}E[\eta_{z1}\eta_{z1}']I_{zt}\Big) = \{r_{\ell m,t}\}.
\end{aligned}$$

Then,

$$\begin{aligned}
||Q_{zz,t} - q_{zz,t}||^2 &\le ||Q_{zz,t}^*||^2 + ||R_{zz,t}^*||^2, \qquad (6.34)\\
||Q_{zz,t} - q_{zz,t}||^4 &\le 2||Q_{zz,t}^*||^4 + 2||R_{zz,t}^*||^4,\\
||Q_{zz,t}^*||^4 &= \big(\sum_{\ell,m=1}^p q_{\ell m,t}^2\big)^2 \le p^2\sum_{\ell,m=1}^p q_{\ell m,t}^4,\\
||R_{zz,t}^*||^4 &= \big(\sum_{\ell,m=1}^p r_{\ell m,t}^2\big)^2 \le p^2\sum_{\ell,m=1}^p r_{\ell m,t}^4.
\end{aligned}$$

We will show that

$$Eq^4_{\ell k,t} \le CH^{-2}m^2, \tag{6.35}$$

$$Er^4_{\ell k,t} \le C(H/n)^{4\gamma_1}, \tag{6.36}$$

where C does not depend on t, n. Clearly, together with (6.34) this implies (6.25):

$$\begin{aligned} E||Q^*_{zz,t}||^4 &\le CH^{-2}m^2, \quad E||R^*_{zz,t}||^4 \le C(H/n)^{4\gamma_1}, \\ E||Q_{zz,t} - q_{zz,t}||^4 &\le C(H^{-2}m^2 + (H/n)^{4\gamma_1}). \end{aligned} \tag{6.37}$$

Notice that

$$\begin{aligned} q_{\ell m,t} &= K_t^{-1} \sum_{j=1}^n b_{n,tj} g_{z\ell,j} g_{zm,j} (\eta_{z\ell,j}\eta_{zm,j} - E[\eta_{z\ell,j}\eta_{zm,j}]), \\ r_{\ell m,t} &= K_t^{-1} \sum_{j=1}^n b_{n,tj} \big(g_{z\ell,j} g_{zm,j} - g_{z\ell,t} g_{zm,t}\big) E[\eta_{z\ell,1}\eta_{zm,1}]. \end{aligned}$$

Proof of (6.35). Denote

$$f_j = g_{z\ell,j} g_{zm,j}, \qquad \omega_j = \eta_{z\ell,j}\eta_{zm,j} - E[\eta_{z\ell,j}\eta_{zm,j}]. \tag{6.38}$$

By Assumptions 2.2(i), $\{\omega_j\}$ is a zero mean stationary sequence which has 4 finite moments and satisfies Assumption 2.5(iii). Moreover, by Assumption 2.3(ii), $\max_j Ef_j^4 \le C$ and the sequences $\{f_j\}$ and $\{\omega_j\}$ are mutually independent. Thus,

$$Ef_j^4\omega_j^4 = Ef_j^4 E\omega_j^4 \le C.$$

Hence by (7.4) of Lemma 7.2,

$$\begin{aligned} Eq^4_{\ell m,t} &\le E\big(K_t^{-1}\sum_{j=1}^n b_{n,tj} f_j \omega_j\big)^4 \le \big(K_t^{-1}\sum_{j=1}^n b_{n,tj}\big(E[f_j^4\omega_j^4]\big)^{1/4}\big)^4 \\ &\le C\big(\sum_{j=1}^n K_t^{-1} b_{n,tj}\big)^2 \big(\max_{t,j=1,\dots,n} K_t^{-1} b_{n,tj}\big)^2 m^2 \le C(\max_{t=1,\dots,n} K_t^{-1})^2 m^2, \end{aligned} \tag{6.39}$$

because $K^{-1}\sum_{j=1}^n b_{n,tj} = 1$ and $b_{n,tj} \le C$.

Observe that under assumption (13), it holds

$$\max_{t=1,\dots,n} K_t^{-1} \le CH^{-1}, \qquad \max_{t=1,\dots,n} K_t \le CH, \tag{6.40}$$

where $C < \infty$ does not depend on n which together with (6.39) implies (6.35).

Proof of (6.36). Bound,

$$\begin{aligned}|f_j - f_t| &= |g_{z\ell,j} g_{zm,j} - g_{z\ell,t} g_{zm,t}| \\ &\le |(g_{z\ell,j} - g_{z\ell,t}) g_{zm,j}| + |g_{z\ell,t}(g_{zm,j} - g_{zm,t})|.\end{aligned}$$

Under Assumption 2.3 on the scale factors $g_{z\ell,t}$, $Eg^8_{z\ell,t} \le C$, and

$$\begin{aligned}E|g_{z\ell,j} - g_{z\ell,t}|^8 &\le C\big(|j-t|/n\big)^{8\gamma_1}, \\ E(|g_{z\ell,j} - g_{z\ell,t}|g_{zm,j})^4 &\le (E|g_{z\ell,j} - g_{z\ell,t}|^8)^{1/2}(Eg^8_{zm,j})^{1/2} \le C\big(|j-t|/n\big)^{4\gamma_1},\end{aligned}$$

where $C < \infty$ does not depend on t, j, n. This implies

$$E|f_j - f_t|^4 \le C\big(|j-t|/n\big)^{4\gamma_1}.$$

Since $|E[\eta_{z\ell,1}\eta_{zm,1}]| < \infty$, we can bound

$$\begin{aligned}E|r_{\ell m,t}|^4 &\le CE\Big(H^{-1}\sum_{j=1}^n b_{n,tj}|f_j - f_t|\Big)^4 \\ &\le C\big(H^{-1}\sum_{j=1}^n b_{n,tj}\{E|f_j - f_t|^4\}^{1/4}\big)^4 \le C\{H^{-1}\sum_{j=1}^n b_{n,tj}\big(|t-j|/n\big)^{\gamma_1}\}^4 \\ &\le C(H/n)^{4\gamma_1}\{H^{-1}\sum_{j=1}^n b_{n,tj}\big(|t-j|/H\big)^{\gamma_1}\}^4 \le C(H/n)^{4\gamma_1}\end{aligned}$$

since under assumption (13), kernel weights $b_{n,tj}$ have property

$$H^{-1}\sum_{j=1}^n b_{n,tj}\Big(\frac{|j-t|}{H}\Big)^{\gamma_1} \le C, \tag{6.41}$$

where $C < \infty$ does not depend on t, H, n. This competes the proof of (6.36).

Proof of (6.27). By (6.26) and (6.24),

$$E||Q_{zx,t}||^4 \le 4E||Q_{zx,t} - q_{zx,t}||^4 + 4E||q_{zx,t}||^4 \le C$$

which proves (6.27).

Proof of (6.28). We have

$$\begin{aligned}Q^{-1}_{zz,t} - q^{-1}_{zz,t} &= -Q^{-1}_{zz,t}(Q_{zz,t} - q_{zz,t})q^{-1}_{zz,t}, \\ ||Q^{-1}_{zz,t} - q^{-1}_{zz,t}|| &\le ||Q^{-1}_{zz,t}||\,||Q_{zz,t} - q_{zz,t}||\,||q^{-1}_{zz,t}||.\end{aligned} \tag{6.42}$$

Therefore, we can bound

$$\begin{aligned}||Q_{zz,t}^{-1}|| &\leq ||q_{zz,t}^{-1}|| + ||Q_{zz,t}^{-1} - q_{zz,t}^{-1}|| \\ &\leq ||q_{zz,t}^{-1}|| + ||Q_{zz,t}^{-1}||\, ||Q_{zz,t} - q_{zz,t}||\, ||q_{zz,t}^{-1}||.\end{aligned} \tag{6.43}$$

By (6.24) and (6.25), $||q_{zz,t}^{-1}|| \leq C$ and $||Q_{zz,t} - q_{zz,t}|| = o_p(1)$. Hence,

$$||Q_{zz,t}^{-1}|| \leq ||q_{zz,t}^{-1}||(1 - ||Q_{zz,t} - q_{zz,t}||)^{-1} = O(1)(1 - o_p(1))^{-1} = O_p(1) \tag{6.44}$$

which proves (6.28).

Proof of (6.29). Write

$$Q_{zz,t}^{-1} - E[Q_{zz,t}|\mathcal{F}_n^*]^{-1} = \{Q_{zz,t}^{-1} - q_{zz,t}^{-1}\} + \{q_{zz,t}^{-1} - E[Q_{zz,t}|\mathcal{F}_n^*]^{-1}\}.$$

We will show that

$$Q_{zz,t}^{-1} - q_{zz,t}^{-1} = o_p(1), \tag{6.45}$$

$$q_{zz,t}^{-1} - E[Q_{zz,t}|\mathcal{F}_n^*]^{-1} = o_p(1), \tag{6.46}$$

which implies (6.29): $Q_{zz,t}^{-1} - E[Q_{zz,t}|\mathcal{F}_n^*]^{-1} = o_p(1)$. To prove (6.45), notice that by (6.42),

$$||Q_{zz,t}^{-1} - q_{zz,t}^{-1}|| \leq ||Q_{zz,t}^{-1}||\, ||Q_{zz,t} - q_{zz,t}||\, ||q_{zz,t}^{-1}|| = o_p(1)$$

since by (6.24), (6.25) and (6.28), $||q_{zz,t}^{-1}|| \leq C$, $||Q_{zz,t} - q_{zz,t}|| = o_p(1)$ and $||Q_{zz,t}^{-1}|| = O_p(1)$, which proves (6.45).

To prove (6.46), write

$$\begin{aligned}E[Q_{zz,t}|\mathcal{F}_n^*] &= q_{zz,t} + \{E[Q_{zz,t}|\mathcal{F}_n^*] - q_{zz,t}\} \\ &= q_{zz,t}\big(1 + q_{zz,t}^{-1}\{E[Q_{zz,t}|\mathcal{F}_n^*] - q_{zz,t}\}\big) = q_{zz,t}\big(1 + o_p(1)\big)\end{aligned}$$

since

$$||q_{zz,t}^{-1}\{E[Q_{zz,t}|\mathcal{F}_n^*] - q_{zz,t}\}|| \leq ||q_{zz,t}^{-1}||\, ||E[Q_{zz,t}|\mathcal{F}_n^*] - q_{zz,t}|| = o_p(1)$$

because $||q_{zz,t}^{-1}|| \leq C$ by (6.24), and by (6.33) and (6.37),

$$\begin{aligned}E[Q_{zz,t}|\mathcal{F}_n^*] - q_{zz,t} &= K_t^{-1}\sum_{j=1}^{n} b_{n,tj}\big(I_{zj}E[\eta_{z1}\eta_{z1}']I_{zj} - I_{zt}E[\eta_{z1}\eta_{z1}']I_{zt}\big) \\ &= R_{zz,t}^* = O_p(H/n)^{2\gamma_1} = o_p(1).\end{aligned}$$

This proves (6.46) and completes the proof of (6.29).

Proof of (6.30) first claim. (Proof of second claim is similar.) Denote $i_n = \max_{t=1,\ldots,n} ||Q_{zz,t} -$

$q_{zz,t}||$. It suffices to show that for any $\epsilon > 0$,

$$P(i_n \geq \epsilon) \to 0, \quad n \to \infty. \tag{6.47}$$

Observe that

$$P(i_n \geq \epsilon) \;\; \leq \;\; E\big[\sum_{j=1}^{n} I\big(||Q_{zz,j} - q_{zz,j}|| \geq \epsilon\big)\big] \leq \epsilon^{-4} \sum_{j=1}^{n} E||Q_{zz,j} - q_{zz,j}||^4. \tag{6.48}$$

By (6.26) of Lemma 6.1,

$$E||Q_{zz,t} - q_{zz,t}||^4 \leq C\big(H^{-2}m^2 + (H/n)^{4\gamma_1}\big) = o(n^{-1}),$$

where C does not depend on t, n. The last equality holds because by assumption $m = O(\log n)$, $\gamma_1 > 3/4$ and $n^a \leq H = O(n^{2/3})$ for some $a > 1/2$ by assumption (14). This proves (6.30):

$$\sum_{j=1}^{n} E||Q_{zz,j} - q_{zz,j}||^4 \leq C\big(H^{-2}m^2 + (H/n)^{4\gamma_1}\big)n = o(1).$$

Recall that by (6.33), $Q_{zz,j} - q_{zz,j} = Q^*_{zz,t} + R^*_{zz,t}$ and $Q^*_{zz,t}$ and $R^*_{zz,t}$ have property (6.37). Thus, by the same argument as in the proof (6.47) it follows that

$$\max_{t=1,\dots,n} ||Q^*_{zz,t}|| = o_p(1), \quad \max_{t=1,\dots,n} ||R^*_{zz,t}|| = o_p(1). \tag{6.49}$$

We will use this property in the proofs below.

Proof of (6.31). By (6.43),

$$\begin{aligned} ||Q^{-1}_{zz,t}|| \;\; &\leq \;\; q_{n,1} + ||Q^{-1}_{zz,t}||\, q_{n2}, \\ & \qquad q_{n,1} = \max_{t=1,\dots,n} ||q^{-1}_{zz,t}||, \quad q_{n,2} = \max_{t=1,\dots,n} \{||Q_{zz,t} - q_{zz,t}||\, ||q^{-1}_{zz,t}||\}. \end{aligned}$$

By (6.24) and (6.30), $q_{n,1} = O(1)$ and $q_{n,2} = o_p(1)$. Hence we obtain that

$$\max_{t=1,\dots,n} ||Q^{-1}_{zz,t}|| \;\; \leq \;\; q_{n,1}(1 - q_{n,2})^{-1} = O(1)(1 - o_p(1))^{-1} = O_p(1).$$

This proves (6.31) and completes the proof of the lemma. □

Let $\beta_{zx,t}$ and $\widehat{\beta}_{zx,t}$ be as in (6.2).

Lemma 6.2. *Suppose that assumptions of Theorem 2.1 are satisfied. Then,*

$$||\widehat{\beta}_{zx,t} - \beta_{zx,t}|| \;\; \leq \;\; C_n\{||Q_{zz,t} - q_{zz,t}||\, ||\beta_{zx,t}|| + ||Q_{zx,t} - q_{zx,t}||\}, \tag{6.50}$$

$$||\widehat{\beta}_{zx,t} - \beta_{zx,t}|| \;\; \leq \;\; c_n(||\beta_{zx,t}|| + 1), \quad t = 1, \dots, n \tag{6.51}$$

where $C_n = O_p(1)$, $c_n = o_p(1)$ and C_n, c_n do not depend on t.

Moreover, the following equality holds:

$$\widehat{\beta}_{zx,t} - \beta_{zx,t} = Q_{zz,t}^{-1}(Q_{zx,t} - q_{zx,t}) + Q_{zz,t}^{-1}(q_{zz,t} - Q_{zz,t})\beta_{zx,t}. \tag{6.52}$$

Proof of Lemma 6.2. Using (6.23), we can write

$$\widehat{\beta}_{zx,t} = S_{zz,t}^{-1} S_{zx,t} = Q_{zz,t}^{-1} Q_{zx,t}, \quad \beta_{zx,t} = (E[z_t z_t | \mathcal{F}_n^*])^{-1} E[z_t x_t' | \mathcal{F}_n^*] = q_{zz,t}^{-1} q_{zx,t}.$$

So, we obtain

$$\begin{aligned}
\widehat{\beta}_{zx,t} - \beta_{zx,t} &= Q_{zz,t}^{-1} Q_{zx,t} - q_{zz,t}^{-1} q_{zx,t} \qquad (6.53)\\
&= Q_{zz,t}^{-1}(Q_{zx,t} - q_{zx,t}) + (Q_{zz,t}^{-1} - q_{zz,t}^{-1}) q_{zx,t}\\
&= Q_{zz,t}^{-1}(Q_{zx,t} - q_{zx,t}) + Q_{zz,t}^{-1}(q_{zz,t} - Q_{zz,t}) q_{zz,t}^{-1} q_{zx,t}\\
&= Q_{zz,t}^{-1}(Q_{zx,t} - q_{zx,t}) + Q_{zz,t}^{-1}(q_{zz,t} - Q_{zz,t}) \beta_{zx,t}
\end{aligned}$$

which implies (6.52). Recall that we denote by $C_n = O_p(1)$ a generic random variable which may change from line to line and does not depend on t. Then, by (6.24) and (6.31) of Lemma 6.1,

$$\max_{t=1,\dots,n} ||q_{zz,t}^{-1}|| \le C_n, \qquad \max_{t=1,\dots,n} ||Q_{zz,t}^{-1}|| \le C_n.$$

Thus, by (6.53),

$$\begin{aligned}
||\widehat{\beta}_{zx,t} - \beta_{zx,t}|| &\le ||Q_{zz,t}^{-1}|| \, ||Q_{zx,t} - q_{zx,t}|| + ||Q_{zz,t}^{-1}|| \, ||q_{zz,t} - Q_{zz,t}|| \, ||\beta_{zx,t}||\\
&\le C_n \{ ||Q_{zx,t} - q_{zx,t}|| + ||Q_{zz,t} - q_{zz,t}|| \, ||\beta_{zx,t}|| \}.
\end{aligned}$$

This implies (6.50) with $C_n = O_p(1)$.

Next, by (6.26) and (6.25) of Lemma 6.1.

$$\max_{t=1,\dots,n} ||Q_{zz,t} - q_{zz,t}|| = o_p(1), \qquad \max_{t=1,\dots,n} ||Q_{zx,t} - q_{zx,t}|| = o_p(1)$$

which implies (6.51):

$$||\widehat{\beta}_{zx,t} - \beta_{zx,t}|| \le c_n(||\beta_{zx,t}|| + 1).$$

This completes the proof of the lemma. □

Recall notation $v_t, \widehat{v}_t$ given in (6.6). Denote

$$\begin{aligned}
S_{vv} &= \sum_{t=1}^n v_t v_t', \quad S_{\widehat{v}\widehat{v}} = \sum_{t=1}^n \widehat{v}_t \widehat{v}_t', \quad S_{vvuu} = \sum_{t=1}^n v_t v_t' u_t^2, \quad S_{\widehat{v}\widehat{v}\widehat{u}\widehat{u}} = \sum_{t=1}^n \widehat{v}_t \widehat{v}_t' \widehat{u}_t^2, \qquad (6.54)\\
V_{vv} &= n^{-1} S_{vv}, \quad V_{\widehat{v}\widehat{v}} = n^{-1} S_{\widehat{v}\widehat{v}}, \quad V_{\widehat{v}\widehat{v}} = n^{-1} S_{\widehat{v}\widehat{v}} \quad V_{vvuu} = n^{-1} S_{vvuu}, \quad V_{\widehat{v}\widehat{v}\widehat{u}\widehat{u}} = n^{-1} S_{\widehat{v}\widehat{v}\widehat{u}\widehat{u}},
\end{aligned}$$

$$S_{zzuu,t} = \sum_{j=1}^n b_{n,tj}^2 z_j z_j' u_j^2, \quad S_{zz\widehat{u}\widehat{u},t} = \sum_{j=1}^n b_{n,tj}^2 z_j z_j' \widehat{u}_j^2.$$

Lemma 6.3. *Under the assumptions of Theorem 2.1,*

$$||V_{vv}^{-1}|| = O_p(1), \tag{6.55}$$

$$||V_{\widehat{v}\widehat{v}} - V_{vv}|| = o_p(1), \tag{6.56}$$

$$||V_{\widehat{v}\widehat{v}}^{-1}|| = O_p(1), \tag{6.57}$$

$$||V_{\widehat{v}\widehat{v}}^{-1} - V_{vv}^{-1}|| = o_p(1), \tag{6.58}$$

$$||V_{vu}|| = O_p(n^{-1/2}). \tag{6.59}$$

Under the assumptions of Corollary 2.1,

$$||E[V_{vvuu}|\mathcal{F}_n^*]|| = O_p(1), \tag{6.60}$$

$$||V_{\widehat{v}\widehat{v}\widehat{u}\widehat{u}} - E[V_{vvuu}|\mathcal{F}_n^*]|| = o_p(1). \tag{6.61}$$

Under the assumptions of Corollary 2.2,

$$||K_t^{-1} E[S_{zzuu,t}|\mathcal{F}_n^*]|| = O_p(1), \tag{6.62}$$

$$||K_t^{-1}\{S_{zz\widehat{u}\widehat{u},t} - E[S_{zzuu,t}|\mathcal{F}_n^*]\}|| = o_p(1). \tag{6.63}$$

Proof of Lemma 6.3.

Proof of (6.55). Recall (6.15). Then,

$$\begin{aligned} V_{vv} &= n^{-1}\sum_{t=1}^n v_t v_t' = n^{-1}\sum_{t=1}^n I_{xt}\nu_t\nu_t' I_{xt} \\ &= n^{-1}\sum_{t=1}^n I_{xt}E[\nu_t\nu_t']I_{xt} + n^{-1}\sum_{t=1}^n I_{xt}\{\nu_t\nu_t' - E[\nu_t\nu_t']\}I_{xt} \\ &= V_{vv}^* + R_{vv}^*. \end{aligned}$$

We will show below that

$$V_{vv}^{*\,-1} = O_p(1), \tag{6.64}$$

$$||R_{vv}^*|| = o_p(1). \tag{6.65}$$

Then, using (6.64) and (6.65), similarly as in (6.43) and (6.44) we obtain:

$$\begin{aligned} ||V_{vv}^{-1}|| &\le ||V_{vv}^{*\,-1}|| + ||V_{vv}^{-1} - V_{vv}^{*\,-1}|| \\ &\le ||V_{vv}^{*\,-1}|| + ||V_{vv}^{*\,-1}(V_{vv}^* - V_{vv})V_{vv}^{-1}|| \\ &\le ||V_{vv}^{*\,-1}|| + ||V_{vv}^{*\,-1}||\,||R_{vv}^*||\,||V_{vv}^{-1}|| \end{aligned}$$

which implies (6.55):

$$||V_{vv}^{-1}|| \quad \leq \quad \frac{||V_{vv}^{*-1}||}{1 - ||R_{vv}^{*}||\, ||V_{vv}^{*-1}||} = \frac{O_p(1)}{1 - o_p(1)} = O_p(1).$$

Proof of (6.64). By Assumption 2.5, ν_t is a short memory stationary sequence, and $\Sigma_{\nu\nu} = E[\nu_t \nu_t'] = E[\nu_1 \nu_1']$ is a positive definite matrix. Then, there exists $b_0 > 0$ such that for any $a = (a_1, ..., a_q)'$, $||a||^2 = 1$,

$$a'\Sigma_{\nu\nu} a \geq b_0.$$

Moreover, by Assumption 2.3(i), $g_{xk,t} \geq c_0 > 0$ for some $c_0 > 0$. Hence,

$$||I_{xt}a||^2 = \sum_{k=1}^{p} g_{xk,t}^2 a_k^2 \geq c_0^2 \sum_{t=1}^{p} a_k^2 = c_0^2 ||a||^2 = c_0^2.$$

So,

$$\begin{aligned} a'V_{vv}^{*}a &= a'n^{-1}\sum_{t=1}^{n} I_{xt}\Sigma_{\nu\nu}I_{xt}a \\ &\leq n^{-1}\sum_{t=1}^{n}(I_{xt}a)'\Sigma_{\nu\nu}(I_{xt}a) \geq n^{-1}\sum_{t=1}^{n}||I_{xt}a||^2 b_0 \\ &\geq n^{-1}\sum_{t=1}^{n} c_0 b_0 = c_0 b_0 = b > 0. \end{aligned}$$

This implies that the largest eigenvalue of $V_{vv}^{*-1} \leq b^{-1}$ does not exceed $1/b$ and proves (6.64).

Proof of (6.65). Observe that the elements of

$$R_{vv}^{*} = n^{-1}\sum_{t=1}^{n} I_{xt}\{\nu_t\nu_t' - E[\nu_t\nu_t']\}I_{xt} = \{s_{n,\ell k}\}$$

are of the form

$$s_{n,\ell k} \quad = \quad n^{-1}\sum_{t=1}^{n} g_{x\ell,t}g_{xk,t}\{\nu_{x\ell,t}\nu_{xk,t} - E[\nu_{x\ell,t}\nu_{xk,t}]\}.$$

Under Assumption 2.2, $\omega_t = \nu_{x\ell,t}\nu_{xk,t} - E[\nu_{x\ell,t}\nu_{xk,t}]$ is a covariance stationary short memory sequence with zero mean, sequence $\{\beta_t = g_{x\ell,t}g_{xk,t}\}$ is independent of $\{\omega_t\}$ and $E\beta_t^2 \leq Eg_{x\ell,t}^4 + Eg_{xk,t}^4 \leq C$ where $C < \infty$ does not depend on t, k, ℓ. Hence, by Lemma 7.1(i),

$$Es_{n,\ell k}^2 \quad \leq \quad Cn^{-2}\sum_{t=1}^{n} E\beta_t^2 \leq Cn^{-1} = o(1)$$

which implies $s_{n,\ell k} = o_p(1)$ and proves (6.65).

Proof of (6.56). We have

$$\widehat{v}_t\widehat{v}_t' - v_tv_t' = (\widehat{v}_t - v_t)(\widehat{v}_t - v_t)' + (\widehat{v}_t - v_t)v_t' - v_t(\widehat{v}_t - v_t)',$$

where $\widehat{v}_t - v_t = (\beta_{zx,t} - \widehat{\beta}_{zx,t})'z_t$. Then,

$$\begin{aligned} ||\widehat{v}_t\widehat{v}_t' - v_tv_t'|| &\leq ||\widehat{v}_t - v_t||^2 + 2||\widehat{v}_t - v_t||\,||v_t|| \\ &\leq ||\beta_{zx,t} - \widehat{\beta}_{zx,t}||^2||z_t||^2 + 2||\beta_{zx,t} - \widehat{\beta}_{zx,t}||\,||z_t||\,||v_t||. \end{aligned}$$

Using the bound (6.51) for $||\beta_{zx,t} - \widehat{\beta}_{zx,t}||$, we obtain

$$\begin{aligned} ||\widehat{v}_t\widehat{v}_t' - v_tv_t'|| &\leq c_n^2(||\beta_{zx,t}|| + 1)^2||z_t||^2 + c_n(||\beta_{zx,t}|| + 1)||z_t||\,||v_t|| \\ &\leq 2(c_n^2 + c_n)(||\beta_{zx,t}|| + 1)^2\big(||z_t||^2 + ||v_t||^2\big) \\ &\leq 2(c_n^2 + c_n)\big((||\beta_{zx,t}|| + 1)^4 + ||z_t||^4 + ||v_t||^2\big), \end{aligned} \tag{6.66}$$

where $c_n = o_p(1)$. Therefore,

$$\begin{aligned} ||V_{\widehat{v}\widehat{v}} - V_{vv}|| &\leq n^{-1}\sum_{t=1}^{n}||\widehat{v}_t\widehat{v}_t' - v_tv_t'|| \leq (c_n^2 + c_n)r_n, \\ r_n &= n^{-1}\sum_{t=1}^{n}\{(||\beta_{zx,t}|| + 1)^4 + ||z_t||^4 + ||v_t||^2\}. \end{aligned} \tag{6.67}$$

By (6.24), $E||\beta_{zx,t}||^4 \leq C$, by (10), $E||z_t||^4 \leq C$, and by (6.15), $E||v_t||^2 \leq E||I_{xt}\nu_t\nu_t'I_{xt}||^2 \leq E||I_{xt}||^2E||\nu_t||^2 \leq C$, where C does not depend on t, n. Therefore,

$$Er_n = n^{-1}\sum_{t=1}^{n}E\big((||\beta_{zx,t}|| + 1)^4 + ||z_t||^4 + ||v_t||^2\big) \leq C$$

which implies $r_n = O(1)$ and proves the required claim: $||V_{\widehat{v}\widehat{v}} - V_{vv}|| \leq o_p(1)$.

Proof of (6.57). Similarly as in the proof of (6.55), we obtain

$$\begin{aligned} ||V_{\widehat{v}\widehat{v}}^{-1}|| &\leq ||V_{vv}^{-1}|| + ||V_{\widehat{v}\widehat{v}}^{-1} - V_{vv}^{-1}|| \\ &\leq ||V_{vv}^{-1}|| + ||V_{\widehat{v}\widehat{v}}^{-1}||\,||V_{\widehat{v}\widehat{v}} - V_{vv}||\,||V_{vv}^{-1}||. \end{aligned}$$

Using (6.55) and (6.56), this implies (6.57):

$$||V_{\widehat{v}\widehat{v}}^{-1}|| \leq \frac{||V_{vv}^{-1}||}{1 - ||V_{\widehat{v}\widehat{v}} - V_{vv}||\,||V_{vv}^{-1}||} = \frac{O_p(1)}{1 - o_p(1)} = O_p(1).$$

Proof of (6.58). Bound,

$$||V_{\widehat{v}\widehat{v}}^{-1} - V_{vv}^{-1}|| = ||V_{vv}^{-1}(V_{\widehat{v}\widehat{v}} - V_{vv})V_{\widehat{v}\widehat{v}}^{-1}||$$

$$\leq \quad ||V_{vv}^{-1}||\, ||V_{\widehat{v}\widehat{v}} - V_{vv}||\, ||V_{\widehat{v}\widehat{v}}^{-1}||.$$

Together with (6.57), (6.56) and (6.55), this implies (6.58):

$$||V_{\widehat{v}\widehat{v}}^{-1} - V_{vv}^{-1}|| \quad = \quad O_p(1)o_p(1)O_p(1) = o_p(1)$$

Proof of (6.59). By (6.15), $v_t = I_{xt}\nu_t$. Hence

$$V_{vu} = n^{-1}\sum_{t=1}^{n} v_t u_t = n^{-1}\sum_{t=1}^{n}\{h_t I_{xt}\}\{\nu_t\varepsilon_t\} = (v_{n1}, ..., v_{nq})'$$

where

$$v_{n\ell} = n^{-1}\sum_{t=1}^{n}\{h_t g_{x\ell,t}\}\{\nu_{\ell t}\varepsilon_t\}.$$

Under assumptions of lemma, $h_t g_{x\ell,t}$ and $\nu_{\ell t}\varepsilon_t$ are mutually independent variables, $\{\nu_{\ell,t}\varepsilon_t\}$ is a martingale difference sequence and $E[h_t^2 g_{x\ell,t}^2] \leq E[h_t^4] + E[g_{x\ell,t}^4] \leq C$, $E[\nu_{\ell t}^2\varepsilon_t^2] \leq C$ where $C < \infty$ does not depend on t. Hence,

$$Ev_{n\ell}^2 = n^{-2}\sum_{t=1}^{n} E(h_t^2 g_{x\ell,t}^2)E(\nu_{\ell t}^2\varepsilon_t^2) \leq Cn^{-1}$$

which implies that $||V_{vu}|| = O_p(n^{-1/2})$.

Proof of (6.60). Under the assumptions of the lemma,

$$\begin{aligned} ||v_t v_t' u_t^2|| \quad &\leq \quad ||v_t||^2 u_t^2, \\ E||v_t v_t' u_t^2|| \quad &\leq \quad E||v_t||^4 + Eu_t^4 \leq C \end{aligned}$$

where $C < \infty$ does not depend on t. Hence,

$$\begin{aligned} E||E[v_t v_t' u_t^2|\mathcal{F}_n^*]|| \quad &\leq \quad E\big[E[||v_t v_t' u_t^2||\,\big|\mathcal{F}_n^*]\big] \\ &= \quad E[||v_t v_t' u_t^2|| \leq C. \end{aligned}$$

Then,

$$E||E[V_{vvuu}|\mathcal{F}_n^*]|| \leq n^{-1}\sum_{t=1}^{n} E||E[v_t v_t' u_t^2|\mathcal{F}_n^*]|| \leq Cn^{-1}\sum_{t=1}^{n} 1 \leq C$$

which implies (6.60).

Proof of (6.61). We have

$$||V_{\widehat{v}\widehat{v}\widehat{u}\widehat{u}} - E[V_{vvuu}|\mathcal{F}_n^*]|| \leq ||V_{vvuu} - E[V_{vvuu}|\mathcal{F}_n^*]|| + ||V_{\widehat{v}\widehat{v}\widehat{u}\widehat{u}} - V_{vvuu}||.$$

We will show that

$$||V_{vvuu} - E[V_{vvuu}|\mathcal{F}_n^*]|| = o_p(1), \tag{6.68}$$

$$||V_{\widehat{v}\widehat{v}\widehat{u}\widehat{u}} - V_{vvuu}|| = o_p(1), \tag{6.69}$$

which implies (6.61): $||V_{\widehat{v}\widehat{v}\widehat{u}\widehat{u}} - E[V_{vvuu}|\mathcal{F}_n^*]|| = o_p(1)$.

Proof of (6.68). By (6.15), $v_t v_t' = I_{xt}\nu_t\nu_t' I_{xt}$. Hence,

$$v_t v_t' u_t^2 = I_{xt}\nu_t\nu_t'\varepsilon_t^2 I_{xt}h_t^2, \quad E[v_t v_t' u_t^2|\mathcal{F}_n^*] = I_{xt}E[\nu_t\nu_t'\varepsilon_t^2]I_{xt}h_t^2,$$
$$v_t v_t' u_t^2 - E[v_t v_t' u_t^2|\mathcal{F}_n^*] = I_{xt}(\nu_t\nu_t'\varepsilon_t^2 - E[\nu_t\nu_t'\varepsilon_t^2])I_{xt}h_t^2.$$

Therefore, we can write

$$V_{vvuu} - E[V_{vvuu}|\mathcal{F}_n^*] = n^{-1}\sum_{t=1}^n A_t W_t B_t$$
$$A_t = I_{xt}, \quad W_t = \nu_t\nu_t'\varepsilon_t^2 - E[\nu_t\nu_t'\varepsilon_t^2], \quad B_t = I_{xt}h_t^2.$$

Under Assumption 2.2(iii), see also Remark 2.1, of Theorem 2.1, the elements of W_t are covariance stationary SM sequences, and the elements of A_t and B_t are independent of the elements of W_t. Hence, by (7.2) of Lemma 7.1,

$$E||V_{vvuu} - E[V_{vvuu}|\mathcal{F}_n^*]||^2 = E||n^{-1}\sum_{t=1}^n A_t W_t B_t||^2 \le Cn^{-2}\sum_{t=1}^n E\big[||A_t||^2||B_t||^2\big].$$

Observe that under assumptions of lemma,

$$E\big[||A_t||^2||B_t||^2\big] \le E[||I_{xt}||^4 h_t^4] \le E[||I_{xt}||^8] + E[h_t^8] \le C \tag{6.70}$$

where $C < \infty$ does not depend on t, n. Therefore,

$$E||V_{vvuu} - E[V_{vvuu}|\mathcal{F}_n^*]||^2 \le n^{-2}\sum_{t=1}^n 1 \le Cn^{-1} = o(1)$$

which implies (6.68).

Proof of (6.69). We have

$$||V_{vvuu} - V_{\widehat{v}\widehat{v}\widehat{u}\widehat{u}}|| \le n^{-1}\sum_{t=1}^n ||\widehat{v}_t\widehat{v}_t'\widehat{u}_t^2 - v_t v_t' u_t^2||. \tag{6.71}$$

Bound

$$||\widehat{v}_t\widehat{v}_t'\widehat{u}_t^2 - v_t v_t' u_t^2|| \le ||(\widehat{v}_t\widehat{v}_t' - v_t v_t')u_t^2|| + ||v_t v_t'(\widehat{u}_t^2 - u_t^2)||$$

$$\leq \quad ||\widehat{v}_t\widehat{v}_t' - v_tv_t'||u_t^2 + ||v_t||^2|\widehat{u}_t^2 - u_t^2|.$$

Thus,

$$\begin{aligned} ||V_{vvuu} - V_{\widehat{v}\widehat{v}\widehat{u}\widehat{u}}|| &\leq n^{-1}\sum_{t=1}^n ||\widehat{v}_t\widehat{v}_t' - v_tv_t'||u_t^2 + n^{-1}\sum_{t=1}^n ||v_t||^2|\widehat{u}_t^2 - u_t^2| \\ &= q_{n1} + q_{n2}. \end{aligned}$$

It suffices to show that

$$q_{n1} = o_p(1), \qquad q_{n2} = o_p(1). \tag{6.72}$$

To evaluate q_{n1}, recall that by (6.66),

$$||\widehat{v}_t\widehat{v}_t' - v_tv_t'|| \leq c_n(||\beta_{zx,t}|| + 1)^2\big(||z_t||^2 + ||v_t||^2\big), \quad c_n = o_p(1).$$

Hence

$$q_{n1} \leq c_nq_{n1}^*, \quad q_{n1}^* = n^{-1}\sum_{t=1}^n(||\beta_{zx,t}|| + 1)^2\big(||z_t||^2 + ||v_t||^2\big)u_t^2.$$

Notice that under the assumptions of the lemma, by (6.24) and (10),

$$\begin{aligned} &E[\{(||\beta_{zx,t}|| + 1)^2\big(||z_t||^2 + ||v_t||^2\big)\}u_t^2] \\ &\quad \leq E(||\beta_{zx,t}|| + 1)^4 + E\big[\big(||z_t||^2 + ||v_t||^2\big)^2u_t^4\big] \\ &\quad \leq E(||\beta_{zx,t}|| + 1)^4 + E\big(||z_t||^2 + ||v_t||^2\big)^4 + E[u_t^8] \leq C \end{aligned}$$

where $C < \infty$ does not depend on t, n. Hence $Eq_{n1}^* \leq C$ which implies $q_{n1}^* = O_p(1)$. Thus,

$$q_{n1} \leq c_nq_{n1}^* = o_p(1)O_p(1) = o_p(1).$$

Next we evaluate q_{n2}. Notice that

$$\widehat{u}_t^2 - u_t^2 \quad = \quad (\widehat{u}_t - u_t)^2 + 2(\widehat{u}_t - u_t)u_t.$$

We have,

$$\begin{aligned} \widehat{u}_t - u_t &= (\alpha - \widehat{\alpha})'x_t + (\beta_t - \widehat{\beta}_t)'z_t, \\ |\widehat{u}_t - u_t| &\leq ||x_t||\,||\alpha - \widehat{\alpha}|| + ||z_t||\,||\beta_t - \widehat{\beta}_t||, \\ (\widehat{u}_t - u_t)^2 &\leq 2||x_t||^2\,||\alpha - \widehat{\alpha}||^2 + 2||(\beta_t - \widehat{\beta}_t)'z_t||^2. \end{aligned} \tag{6.73}$$

We will denote by C_n and c_n the random variables $C_n = O_p(1)$ and $c_n = o_p(1)$ which do not depend on t. By Theorem 2.1, $||\alpha - \widehat{\alpha}|| = O_p(n^{-1/2}) = c_n$. Recall notation

$$Q_{zz,t} = K_t^{-1} S_{zz,t}, \quad Q_{zu,t} = K_t^{-1} S_{zu,t}, \quad Q_{zz\beta,t} = K_t^{-1} \sum_{j=1}^n b_{n,tj} z_j z_j' (\beta_j - \beta_t).$$

Recall that by (6.4), $\zeta_j = z_j' \beta_j + u_j$ and thus,

$$\begin{aligned} S_{z\zeta,t} &= \sum_{j=1}^n b_{n,tj} z_j \zeta_j = \sum_{j=1}^n b_{n,tj} z_j z_j' \beta_j + \sum_{j=1}^n b_{n,tj} z_j u_j \\ &= \sum_{j=1}^n b_{n,tj} z_j z_j' (\beta_j - \beta_t) + S_{zz,t} \beta_t + \sum_{j=1}^n b_{n,tj} z_j u_j, \\ S_{zz,t}^{-1} S_{z\zeta,t} - \beta_t &= Q_{zz,t}^{-1} \{ Q_{zz\beta,t} + Q_{zu,t} \}. \end{aligned}$$

Therefore, by (6.18),

$$\begin{aligned} \widehat{\beta}_t - \beta_t &= \{ S_{zz,t}^{-1} S_{z\zeta,t} - \beta_t \} + S_{zz,t}^{-1} S_{zx,t} (\alpha - \widehat{\alpha}) \\ &= Q_{zz,t}^{-1} \{ Q_{zz\beta,t} + Q_{zu,t} \} + Q_{zz,t}^{-1} Q_{zx,t} (\alpha - \widehat{\alpha}), \\ ||\widehat{\beta}_t - \beta_t|| &\leq ||Q_{zz,t}^{-1}|| \{ ||Q_{zz\beta,t}|| \, || + ||Q_{zu,t}|| + ||Q_{zx,t}|| \, ||\alpha - \widehat{\alpha}|| \} \\ &\leq C_n \{ ||Q_{zz\beta,t}|| + ||Q_{zu,t}|| \} + c_n \, ||Q_{zx,t}||, \end{aligned} \tag{6.74}$$

since by (6.28), $\max_{t=1,\dots,n} ||Q_{zz,t}^{-1}|| \leq C_n = O_p(1)$. Hence,

$$||\widehat{\beta}_t - \beta_t||^2 \leq C_n \{ ||Q_{zz\beta,t}||^2 + ||Q_{zu,t}||^2 \} + c_n \, ||Q_{zx,t}||^2. \tag{6.75}$$

Hence, using (6.74), (6.75) and (6.73), we obtain the following bounds:

$$\begin{aligned} |\widehat{u}_t - u_t| &\leq c_n ||x_t|| + ||z_t|| \, ||\beta_t - \widehat{\beta}_t|| \\ &\leq c_n (||x_t|| + ||z_t|| \, ||Q_{zx,t}||) + C_n ||z_t|| \, \{ ||Q_{zz\beta,t}|| + ||Q_{zu,t}|| \}, \\ |\widehat{u}_t - u_t|^2 &= c_n (||x_t||^2 + ||z_t||^2 ||Q_{zx,t}||^2) + C_n ||z_t||^2 \, \{ ||Q_{zz\beta,t}||^2 + ||Q_{zu,t}||^2 \}. \end{aligned}$$

Hence, we can bound

$$\begin{aligned} q_{n2} &= n^{-1} \sum_{t=1}^n ||v_t||^2 |\widehat{u}_t^2 - u_t^2| \\ &\leq n^{-1} \sum_{t=1}^n ||v_t||^2 \big((\widehat{u}_t - u_t)^2 + 2 |\widehat{u}_t - u_t| \, |u_t| \big) \\ &\leq c_n q_{n2,1} + C_n q_{n2,2} + C_n q_{n2,3}, \end{aligned} \tag{6.76}$$

where

$$q_{n2,1} = n^{-1} \sum_{t=1}^{n} ||v_t||^2 \{||x_t||^2 + ||z_t||^2 ||Q_{zx,t}||^2 + u_t^2\},$$
$$q_{n2,2} = n^{-1} \sum_{t=1}^{n} ||v_t||^2 ||z_t||^2 \{||Q_{zz\beta,t}||^2 + ||Q_{zu,t}||^2\},$$
$$q_{n2,3} = n^{-1} \sum_{t=1}^{n} ||v_t||^2 \, ||z_t|| \, |u_t| \{||Q_{zz\beta,t}|| + ||Q_{zu,t}||\},$$

and $c_n = o_p(1)$, $C_n = O_p(1)$. We will show that

$$q_{n2,1} = O_p(1), \quad q_{n2,2} = o_p(1), \quad q_{n2,3} = o_p(1), \tag{6.77}$$

which together with (6.76) implies $q_{n2} = o_p(1)$ and proves (6.72).

(1) First, we bound $q_{n2,1}$. Observe that

$$||v_t||^2 ||z_t||^2 ||Q_{zx,t}||^2 \le ||v_t||^4 ||z_t||^4 + ||Q_{zx,t}||^4 \le ||v_t||^8 + ||z_t||^8 + ||Q_{zx,t}||^4.$$

Then, under assumptions of lemma and by (6.27),

$$\begin{aligned} & E[||v_t||^2 \{||x_t||^2 + ||z_t||^2 \, ||Q_{zx,t}||^2 + u_t^2\}] \\ & \quad \le 3E[||v_t||^4 + ||x_t||^4 + ||v_t||^8 + ||z_t||^8 + ||Q_{zx,t}||^4 + u_t^4] \le C \end{aligned} \tag{6.78}$$

where $C < \infty$ does not depend on t, n. Hence, $Eq_{n2,1} \le C$, and $q_{n2,1} = O_p(1)$.

(2) Next we bound $q_{n2,2}$. Denote $i_n = \max_{t=1,\dots,n} ||z_t||^2$. Then,

$$\begin{aligned} ||Q_{zz\beta,t}|| &\le K_t^{-1} \sum_{j=1}^{n} b_{n,tj} ||z_j z_j' (\beta_j - \beta_t)|| \le K_t^{-1} \sum_{j=1}^{n} b_{n,tj} ||z_j||^2 ||\beta_j - \beta_t|| \\ &\le i_n Q_{\beta,t}, \qquad Q_{\beta,t} = K_t^{-1} \sum_{j=1}^{n} b_{n,tj} ||\beta_j - \beta_t||. \end{aligned} \tag{6.79}$$

Then,

$$\begin{aligned} q_{n2,2} &\le n^{-1} \sum_{t=1}^{n} ||v_t||^2 ||z_t||^2 \big\{||Q_{zz\beta,t}||^2 + ||Q_{zu,t}||^2\big\} \\ &\le i_n^2 \{n^{-1} \sum_{t=1}^{n} ||v_t||^2 ||z_t||^2 ||Q_{\beta,t}||^2\} + \{n^{-1} \sum_{t=1}^{n} ||v_t||^2 ||z_t||^2 ||Q_{zu,t}||^2\} \qquad (6.80) \\ &= i_n^2 j_{n1} + j_{n2}. \qquad (6.81) \end{aligned}$$

Observe that

$$\begin{aligned} E\big[||v_t||^2||z_t||^2||Q_{\beta,t}||^2\big] &\leq (E\big[||v_t||^4||z_t||^4\big])^{1/2}(E||Q_{\beta,t}||^4)^{1/2} \leq C(E||Q_{\beta,t}||^4)^{1/2}, \\ E\big[||v_t||^2||z_t||^2||Q_{zu,t}||^2\big] &\leq (E\big[||v_t||^4||z_t||^4\big])^{1/2}(E||Q_{zu,t}|||^4)^{1/2} \leq C(E||Q_{zu,t}||^4)^{1/2} \end{aligned}$$

since under the assumptions of the Corollary 2.1, $E\big[||v_t||^4||z_t||^4\big] \leq E||v_t||^8 + E||z_t||^8 \leq C$ where $C < \infty$ does not depend on t. Moreover, by (6.116) and (6.120),

$$E||Q_{\beta,t}||^4 \leq C(H/n)^{4\gamma_2}, \quad E||Q_{zu,t}||^4 \leq CH^{-1}.$$

By assumption of corollary, $H = O(n^{2/3})$, and by Assumption 2.4, $\gamma_2 \geq 3/4 + \delta$ for some $\delta > 0$. Hence $(H/n)^{4\gamma_2} \leq C(n^{-1/3})^{3+4\delta} \leq Cn^{-1-\delta}$. Therefore,

$$\begin{aligned} Ej_{n1} &\leq n^{-1}\sum_{t=1}^{n} E[||v_t||^2||z_t||^2||Q_{\beta,t}||^2] \leq C(H/n)^{2\gamma_2} \leq Cn^{-1/2-\delta/2}, \quad (6.82) \\ Ej_{n2} &\leq n^{-1}\sum_{t=1}^{n} E[||v_t||^2||z_t||^2||Q_{zu,t}||^2] \leq Cn^{-1}\sum_{t=1}^{n} H^{-1} = CH^{-1} = o(1). \end{aligned}$$

Under assumptions of lemma, the variable $\xi_t = ||z_t||^2$ has 4-th finite moments: $E\xi_t^4 = E||z_t||^8 \leq C$ where $C < \infty$ does not depend on t. Hence, by Lemma 6.4, $i_n = O_p(n^{1/4+a})$ for any $a > 0$. Suppose that $a < \delta/4$. Then,

$$q_{n2,2} \leq i_n^2 j_{n1} + j_{n2} = O_p(n^{1/2+2a})O_p(n^{-1/2-\delta/2}) + o_p(1) = o_p(1). \quad (6.83)$$

(3) Finally, to bound $q_{n2,3}$, notice that by (6.115) and (6.119),

$$E[||Q_{zz\beta,t}||^2] \leq C(H/n)^{2\gamma_2} = o(1), \quad E[||Q_{zu,t}||^2] \leq CH^{-1} = o(1).$$

Therefore,

$$\begin{aligned} &E[||v_t||^2\,||z_t||\,|u_t|\{||Q_{zz\beta,t}|| + ||Q_{zu,t}||\}] \\ &\quad \leq (E[||v_t||^4\,||z_t||^2\,|u_t|^2])^{1/2}\{(E||Q_{zz\beta,t}||^2)^{1/2} + (E||Q_{zu,t}||^2)^{1/2}\} \\ &\quad \leq C((H/n)^{\gamma_2} + H^{-1/2}) = o(1) \quad (6.84) \end{aligned}$$

which implies that $Eq_{n2,3} = o(1)$ and $q_{n2,3} = o_p(1)$. This completes the proof of (6.61).

Proof of (6.62). We have

$$\begin{aligned} E[z_j z_j' u_j^2|\mathcal{F}_n^*] &= E[I_{zj}\eta_{zj}\eta_{zj}' I_{zj} h_j^2 \varepsilon_j^2|\mathcal{F}_n^*] = E[I_{zj}E[\eta_{zj}\eta_{zj}'\varepsilon_j^2]I_{zj}h_j^2], \\ ||E[z_j z_j' u_j^2|\mathcal{F}_n^*]|| &\leq E[\{||I_{zj}||^2 h_j^2\}\{||\eta_{zj}||^2\varepsilon_j^2\}] \leq E[||I_{zj}||^2 h_j^2]E[||\eta_{zj}||^2\varepsilon_j^2] \\ &\leq \{E||I_{zj}||^4 + Eh_j^4\}\{E||\eta_{zj}||^4 + E\varepsilon_j^4\} \leq C \end{aligned}$$

where $C < \infty$ does not depend on j, see Assumptions 2.2 and 2.3. Thus,

$$
\begin{aligned}
K_t^{-1} E[S_{zzuu,t} | \mathcal{F}_n^*] &\leq K_t^{-1} \sum_{j=1}^{n} b_{n,tj} E\big[E[z_j z_j' u_j^2 | \mathcal{F}_n^*]\big] \\
&\leq C K_t^{-1} \textstyle\sum_{j=1}^{n} b_{n,tj}^2 \leq C K_t^{-1} \sum_{j=1}^{n} b_{n,tj} = C,
\end{aligned}
$$

which implies the required claim: $K_t^{-1} E[S_{zzuu} | \mathcal{F}_n^*] = O_p(1)$.

Proof of (6.63). Bound

$$
||S_{zz\widehat{u}\widehat{u},t} - E[S_{zzuu,t} | \mathcal{F}_n^*]|| \leq ||S_{zzuu,t} - E[S_{zzuu,t} | \mathcal{F}_n^*]|| + ||S_{zz\widehat{u}\widehat{u},t} - S_{zzuu}||.
$$

Similarly, as for deriving the bound (6.68) in the proof of (6.61), it suffices to show that

$$
||K_t^{-1}(S_{zzuu,t} - E[S_{zzuu,t} | \mathcal{F}_n^*])|| = o_p(1), \tag{6.85}
$$

$$
||K_t^{-1}(S_{zz\widehat{u}\widehat{u},t} - S_{zzuu,t})|| = o_p(1), \tag{6.86}
$$

which implies (6.63): $||K_t^{-1}(S_{zz\widehat{u}\widehat{u},t} - E[S_{zzuu,t} | \mathcal{F}_n^*])|| = o_p(1)$.

Proof of (6.85). We have

$$
z_j z_j' u_j^2 = I_{zj} \eta_{zj} \eta_{zj}' \varepsilon_j^2 I_{zj} h_j^2, \quad z_j z_j' u_j^2 - E[z_j z_j' u_j^2 | \mathcal{F}_n^*] = I_{zj} \big(\eta_{zj} \eta_{zj}' \varepsilon_j^2 - E[\eta_{zj} \eta_{zj}' \varepsilon_j^2]\big) I_{zj} h_j^2.
$$

Hence, we can write

$$
\begin{aligned}
K_t^{-1}(S_{zzuu,t} - E[S_{zzuu,t} | \mathcal{F}_n^*]) &= K_t^{-1} \sum_{j=1}^{n} b_{n,tj}^2 A_j W_j B_j, \\
A_j = I_{zj}, \quad W_j &= \eta_{zj} \eta_{zj}' \varepsilon_j^2 - E[\eta_{zj} \eta_{zj}' \varepsilon_j^2], \quad B_j = I_{zj} h_j^2.
\end{aligned}
$$

By Assumption 2.5(i), the elements of W_j are covariance stationary SM sequences, and the elements of A_j and B_j are independent of the elements of W_j. Hence, by (7.2) of Lemma 7.1,

$$
\begin{aligned}
E||K_t^{-1}(S_{zzuu,t} - E[S_{zzuu,t} | \mathcal{F}_n^*]||^2 &= E||K_t^{-1} \sum_{j=1}^{n} b_{n,tj}^2 A_j W_j B_j||^2 \\
&\leq C K_t^{-2} \sum_{j=1}^{n} b_{n,tj}^4 E\big[||A_j||^2 ||B_j||^2\big].
\end{aligned}
$$

Under the assumptions of the lemma, similarly as in (6.70) it holds, $E\big[||A_j||^2 ||B_j||^2\big] \leq C$, where $C < \infty$ does not depend on j, n. This implies

$$
E||K_t^{-1}(S_{zzuu,t} - E[S_{zzuu,t} | \mathcal{F}_n^*])||^2 \leq C K_t^{-2} \sum_{j=1}^{n} b_{n,tj}^4 \leq C K_t^{-1} \leq C H^{-1} = o(1),
$$

since $b_{n,tj} \le C$, $K_t^{-1}\sum_{t=1}^n b_{n,tj} = 1$ and $K_t \le CH^{-1}$ by (6.40). This implies (6.85).

Proof of (6.86). We have

$$\begin{aligned} S_{zz\widehat{u}\widehat{u},t} - S_{zzuu,t} &= \sum_{t=1}^n b_{n,tj}^2 z_j z_j' (\widehat{u}_j^2 - u_j^2), \\ ||S_{zz\widehat{u}\widehat{u},t} - S_{zzuu,t}|| &\le \sum_{t=1}^n b_{n,tj}^2 ||z_j||^2 |\widehat{u}_j^2 - u_j^2|. \end{aligned}$$

Notice that $b_{n,tj} \le C$. Then, similarly to (6.76), we obtain

$$\begin{aligned} j_n &= K_t^{-1} ||S_{zz\widehat{u}\widehat{u},t} - S_{zzuu,t}|| \\ &\le K_t^{-1} \sum_{j=1}^n b_{n,tj} ||z_j||^2 |\widehat{u}_j^2 - u_j^2| \le c_n q_{n2,1,t} + C_n q_{n2,2,t} + C_n q_{n2,3,t} \end{aligned}$$

where $c_n = o_p(1)$, $C_n = O_p(1)$ and

$$\begin{aligned} q_{n2,1,t} &= K_t^{-1} \sum_{j=1}^n b_{n,tj} ||z_j||^2 \{||x_j||^2 + ||z_j||^2 ||Q_{zx,j}||^2 + u_j^2\}, \\ q_{n2,2,t} &= K_t^{-1} \sum_{j=1}^n b_{n,tj} ||z_j||^4 \{||Q_{zz\beta,j}||^2 + ||Q_{zu,j}||^2\}, \\ q_{n2,3,t} &= K_t^{-1} \sum_{j=1}^n b_{n,tj} ||z_j||^3 \, |u_j| \{||Q_{zz\beta,j}|| + ||Q_{zu,j}||\}. \end{aligned}$$

It remains to show that

$$q_{n2,1,t} = O_p(1), \quad q_{n2,2,t} = o_p(1), \quad q_{n2,3,t} = o_p(1), \tag{6.87}$$

which implies the required claim, $j_n = o_p(1)$.

To evaluate $q_{n2,1,t}$ notice that similarly as in (6.78), $E[||z_j||^2\{||x_j||^2 + ||z_j||^2 ||Q_{zx,j}||^2 + u_j^2\}] \le C$ where $C < \infty$ does not depend on t, n. Hence,

$$E q_{n2,1,t} \le C K_t^{-1} \sum_{j=1}^n b_{n,tj} = C$$

which implies $q_{n2,1,t} = O_p(1)$.

To bound $q_{n2,2,t}$, notice that similarly to (6.81), we can write

$$q_{n2,2,t} \le K_t^{-1} \sum_{j=1}^n b_{n,tj} ||z_j||^4 \{||Q_{zz\beta,j}||^2 + ||Q_{zu,j}||^2\}$$

$$\begin{aligned}
&\le i_n^2\{K_t^{-1}\sum_{j=1}^n b_{n,tj}||z_j||^4||Q_{\beta,j}||^2\} + \{K_t^{-1}\sum_{j=1}^n b_{n,tj}||z_j||^4||Q_{zu,j}||^2\} \\
&= i_n^2 j_{n1,t} + j_{n2,t}
\end{aligned}$$

where i_n is the same as in (6.79). Since $K_t^{-1}\sum_{j=1}^n b_{n,tj} = 1$, the same argument as in the proof of (6.82) implies that $Ej_{n1,t}$, $Ej_{n2,t}$ and satisfy the same bounds (6.82) as Ej_{n1}, Ej_{n2}. Therefore $q_{n2,2,t}$ satisfies the same bound (6.87) as $q_{n2,2}$ which implies $q_{n2,2,t} = o_p(1)$. Observe that under assumptions of Corollary 2.2, the bandwidth H_z has property $H_z = O(n^{2/3})$ used in the proof of $q_{n2,2} = o_p(1)$ above. The latter follows from the assumption $H_z = o(n^{2\gamma_2/(2\gamma_2+1)})$ imposed on bandwidth H_z in Corollary 2.2 and assumption $\gamma_2 \in (3/4, 1]$.

To bound $q_{n2,3,t}$, notice that $E[||z_j||^3\,|u_j\{||Q_{zz\beta,j}|| + ||Q_{zu,j}||\}] = o(1)$ satisfies the same bound as $E[||v_t||^2\,||z_t||\,|u_t|\{||Q_{zz\beta,t}|| + ||Q_{zu,t}||\}]$ in (6.84). This implies that

$$q_{n2,3,t} \le K_t^{-1}\sum_{j=1}^n b_{n,tj}E[||z_j||^3\,|u_j\{||Q_{zz\beta,j}|| + ||Q_{zu,j}||\}] \le o(1)K_t^{-1}\sum_{j=1}^n b_{n,tj} = o(1).$$

Hence, $q_{n2,3,t} = o_p(1)$ which completes the proof of (6.87), (6.86) and the lemma. □

Lemma 6.4. *Suppose that for some integer $k \ge 2$, the random variables $\xi_t \ge 0$ have the property $E\xi_t^k \le C$ where $C < \infty$ does not depend on t. Then, for any $a > 0$,*

$$i_n = \max_{t=1,\dots,n} \xi_t = o(n^{1/k+a}). \tag{6.88}$$

Proof of Lemma 6.4. For any $\epsilon > 0$,

$$P(i_n \ge \epsilon n^{1/k+a}) \le E\big[\sum_{t=1}^n I\big(\xi_t \ge \epsilon n^{1/k+a}\big)\big] \le \epsilon^{-k}n^{-1-ka}\sum_{t=1}^n E\xi_t^k \le C\epsilon^{-k}n^{-ka} = o(1)$$

which proves (6.88). □

6.3 Lemmas 6.5 and 6.6

Lemma 6.5. *Under assumption of Theorem 2.1,*

$$S_{\widehat{v}\widehat{v}}^{-1}S_{\widehat{v}u} - S_{vv}^{-1}S_{vu} = o_p\big(n^{-1/2}\big). \tag{6.89}$$

Proof of Lemma 6.5. Denote $V_{\widehat{v}u} = n^{-1}S_{\widehat{v}u}$. We have

$$\begin{aligned}
S_{\widehat{v}\widehat{v}}^{-1}S_{\widehat{v}u} - S_{vv}^{-1}S_{vu} &= V_{\widehat{v}\widehat{v}}^{-1}V_{\widehat{v}u} - V_{vv}^{-1}V_{vu} \\
&= V_{\widehat{v}\widehat{v}}^{-1}(V_{\widehat{v}u} - V_{vu}) + (V_{\widehat{v}\widehat{v}}^{-1} - V_{vv}^{-1})V_{vu}.
\end{aligned}$$

By (6.57), $V_{\widehat{v}\widehat{v}}^{-1} = O_p(1)$, and by (6.58), $V_{\widehat{v}\widehat{v}}^{-1} - V_{vv}^{-1} = o_p(1)$, while in (6.59) it is shown that $V_{vu} = O_p(n^{-1/2})$. We will prove that

$$V_{\widehat{v}u} - V_{vu} \quad = \quad o_p(n^{-1/2}), \tag{6.90}$$

which implies (6.89):

$$V_{\widehat{v}\widehat{v}}^{-1} V_{\widehat{v}u} - V_{vv}^{-1} V_{vu} = O_p(1) o_p(n^{-1/2}) + o_p(1) O_p(n^{-1/2}) = o_p(n^{-1/2}).$$

Proof (6.90). By (6.6), we have

$$\begin{aligned} \widehat{v}_t - v_t &= (\beta_{zx,t} - \widehat{\beta}_{zx,t})' z_t = z_t'(\beta_{zx,t} - \widehat{\beta}_{zx,t}), \\ V_{\widehat{v}u} - V_{vu} &= n^{-1} \sum_{t=1}^{n} (\widehat{v}_t - v_t) u_t = -n^{-1} \sum_{t=1}^{n} z_t'(\widehat{\beta}_{zx,t} - \beta_{zz,t}) u_t. \end{aligned} \tag{6.91}$$

By (6.52),

$$\begin{aligned} &\widehat{\beta}_{zx,t} - \beta_{zx,t} = Q_{zz,t}^{-1}(Q_{zx,t} - q_{zx,t}) + Q_{zz,t}^{-1}(q_{zz,t} - Q_{zz,t})\beta_{zx,t} \\ &\quad = \{Q_{zz,t}^{-1} - q_{zz,t}^{-1}\}\Big((Q_{zx,t} - q_{zx,t}) + (q_{zz,t} - Q_{zz,t})\beta_{zx,t}\Big) \\ &\qquad + q_{zz,t}^{-1}(q_{zz,t} - Q_{zz,t})\beta_{zx,t} + q_{zz,t}^{-1}(Q_{zx,t} - q_{zx,t}) \\ &\quad = p_{1t} + p_{2t} + p_{3t}, \end{aligned}$$

where

$$\begin{aligned} p_{1t} &= \{Q_{zz,t}^{-1} - q_{zz,t}^{-1}\}\Big((Q_{zx,t} - q_{zx,t}) + (q_{zz,t} - Q_{zz,t})\beta_{zx,t}\Big), \\ p_{2t} &= q_{zz,t}^{-1}(q_{zz,t} - Q_{zz,t})\beta_{zx,t}, \\ p_{3t} &= q_{zz,t}^{-1}(Q_{zx,t} - q_{zx,t}). \end{aligned}$$

Therefore, by (6.91),

$$V_{\widehat{v}u} - V_{vu} = n^{-1} \sum_{t=1}^{n} (\widehat{v}_t - v_t) u_t = R_{n1} + R_{n2} + R_{n3}, \tag{6.92}$$

where

$$R_{nk} = -n^{-1} \sum_{t=1}^{n} z_t' p_{kt} u_t, \quad k = 1, 2, 3. \tag{6.93}$$

To prove (6.90), it remains to show that for $k = 1, 2, 3$,

$$R_{nk} = o_p(n^{-1/2}). \tag{6.94}$$

Let $k = 1$. Bound,

$$\begin{aligned} ||p_{1t}|| &\le ||Q^{-1}_{zz,t} - q^{-1}_{zz,t}||\{||Q_{zx,t} - q_{zx,t}|| + ||Q_{zz,t} - q_{zz,t}||\,||\beta_{zx,t}||\} \\ &\le ||Q^{-1}_{zz,t}||\,||Q_{zz,t} - q_{zz,t}||\,||q^{-1}_{zz,t}||\{||Q_{zx,t} - q_{zx,t}|| + ||Q_{zz,t} - q_{zz,t}||\,||\beta_{zx,t}||\}. \end{aligned} \tag{6.95}$$

By (6.24) and (6.31), $\max_{t=1,\dots,n} ||q^{-1}_{zz,t}||\,||Q^{-1}_{zz,t}|| \le C_n = O_p(1)$. Hence,

$$\begin{aligned} ||p_{1t}|| &\le C_n||Q_{zz,t} - q_{zz,t}||\{||Q_{zx,t} - q_{zx,t}|| + ||Q_{zz,t} - q_{zz,t}||\,||\beta_{zx,t}||\} \\ &\le C_n\{||Q_{zz,t} - q_{zz,t}||^2 + ||Q_{zx,t} - q_{zx,t}||^2\}(1 + ||\beta_{zx,t}||). \end{aligned} \tag{6.96}$$

Therefore,

$$||R_{n1}|| \le O_p(1) r_n,$$
$$r_n = n^{-1}\sum_{t=1}^{n} \big\{||Q_{zz,t} - q_{zz,t}||^2 + ||Q_{zx,t} - q_{zx,t}||^2\big\}(||\beta_{zx,t}|| + 1)||z_t||\,|u_t|.$$

We will show that

$$Er_n = o(n^{-1/2}), \tag{6.97}$$

which implies the required claim $R_{n1} = o_p(n^{-1/2})$. By Hölder inequality,

$$\begin{aligned} &E\big[\big(||Q_{zz,t} - q_{zz,t}||^2 + ||Q_{zx,t} - q_{zx,t}||^2\big)(||\beta_{zx,t}|| + 1)||z_t||\,|u_t|\big] \\ &\le \big\{(E\big[||Q_{zz,t} - q_{zz,t}||^4\big])^{1/2} + (E\big[||Q_{zx,t} - q_{zx,t}||^4\big])^{1/2}\big\}(E\big[(||\beta_{zx,t}|| + 1)^2||z_t||^2u_t^2\big])^{1/2}. \end{aligned}$$

By the assumptions of the lemma and (6.24),

$$E\big[(||\beta_{zx,t}|| + 1)^2||z_t||^2u_t^2\big] \le E(||\beta_{zx,t}|| + 1)^6 + E||z_t||^6 + Eu_t^6 \le C$$

where $C < \infty$ does not depend on t, whereas by (6.25) and (6.26),

$$E\big[||Q_{zz,t} - q_{zz,t}||^4\big] \le C\big(H^{-2}m^2 + (H/n)^{4\gamma_1}\big),$$
$$E\big[||Q_{zx,t} - q_{zx,t}||^4\big] \le C\big(H^{-2}m^2 + (H/n)^{4\gamma_1}\big).$$

Therefore,

$$Er_n \le Cn^{-1}\sum_{t=1}^{n}\big(H^{-2}m^2 + (H/n)^{4\gamma_1}\big)^{1/2} = C\big(H^{-2}m^2 + (H/n)^{4\gamma_1}\big)^{1/2} = o(n^{-1/2}),$$

because by assumption, $m = O(\log n)$ and $\gamma_1 > 3/4$, and by (14), $n^a \le H = O(n^{2/3})$ for some $a > 1/2$ which implies $H^{-2}m^2 = o(n^{-1})$, $(H/n)^{4\gamma_1} = o(n^{-1})$.

Let $k = 2$. (Proof for $k = 3$ is similar as for $k = 2$ and therefore it is omitted).

By (6.33), $Q_{zz,t} - q_{zz,t} = Q^*_{zz,t} + R^*_{zz,t}$. Recall that $u_t = h_t \varepsilon_t$ where $\{\varepsilon_t\}$ is a stationary martingale difference sequence. Hence, $\{\varepsilon_t\}$ is a stationary short memory sequence with zero mean. Write

$$\begin{aligned} R_{n2} &= -n^{-1}\sum_{t=1}^{n} z_t' q_{zz,t}^{-1} R^*_{zz,t}\beta_{zx,t}u_t + n^{-1}\sum_{t=1}^{n} z_t' q_{zz,t}^{-1} Q^*_{zz,t}\beta_{zx,t}u_t \\ &= -n^{-1}\sum_{t=1}^{n}\{\eta_{zt}'\varepsilon_t\}\{I_{zt}q_{zz,t}^{-1}\}R^*_{zz,t}\{\beta_{zx,t}h_t\} - n^{-1}\sum_{t=1}^{n}\{\eta_{zt}'\varepsilon_t\}\{I_{zt}q_{zz,t}^{-1}\}\{Q^*_{zz,t}\}\{\beta_{zx,t}h_t\} \\ &= -(R_{n2,1} + R_{n2,2}). \end{aligned} \tag{6.98}$$

We will show that

$$R_{n2,i} = o_p(n^{-1/2}), \quad i = 1, 2. \tag{6.99}$$

To evaluate $R_{n2,1}$, recall that under assumptions of Theorem 2.1, the variables $\{q_{zz,t}^{-1}, R^*_{zz,t}, \beta_{zx,t}, I_{zt}, h_t\}$ are independent of $\{\eta_{zt}'\varepsilon_t\}$. Hence, by Lemma 7.1(ii),

$$E||R_{n2,1}||^2 \le Cn^{-2}\sum_{t=1}^{n} E\big[||I_{zt}q_{zz,t}^{-1}||^2\,||R^*_{zz,t}||^2||h_t\beta_{zx,t}||^2\big].$$

Under the assumptions of the lemma,

$$\begin{aligned} I_{zt}q_{zz,t}^{-1} &= I_{zt}\{I_{zt}^{-1}\Sigma_{zz}^{-1}I_{zt}^{-1}\} = \Sigma_{zz}^{-1}I_{zt}^{-1}, \\ ||I_{zt}q_{zz,t}^{-1}|| &\le ||\Sigma_{zz}^{-1}||\,||I_{zt}^{-1}|| \le C \end{aligned} \tag{6.100}$$

since by Assumption 2.2(ii), Σ_{zz} is invertable, and by Assumption 2.3, $g_{zk,t} \ge c_0 > 0$. By (6.100), $||I_{zt}q_{zz,t}^{-1}|| \le C$ where $C < \infty$ does not depend on t. By (6.24), $E||\beta_{zx,t}||^8 \le C$, and by assumptions of lemma $Eh_t^8 \le C$. Hence,

$$\begin{aligned} E\big[||I_{zt}q_{zz,t}^{-1}||^2\,||R^*_{zz,t}||^2h_t^2||\beta_{zx,t}||^2\big] &\le C(E||R^*_{zz,t}||^4)^{1/2}(E[h_t^4||\beta_{zx,t}||^4])^{1/2} \\ &\le C(E||R^*_{zz,t}||^4)^{1/2} \le C(H/n)^{2\gamma_1} \end{aligned}$$

because by (6.37), $E||R^*_{zz,t}||^4 \le C(H/n)^{4\gamma_1}$ and $E[h_t^4||\beta_{zx,t}||^4] \le Eh_t^8 + E||\beta_{zx,t}||^8 \le C$. Hence,

$$E||R_{n2,1}||^2 \le C(H/n)^{2\gamma_1}n^{-2}\sum_{t=1}^{n} 1 = C(H/n)^{2\gamma_1}n^{-1} = o(n^{-1})$$

which implies $R_{n2,1} = o_p(n^{-1/2})$.

Next we evaluate

$$\begin{aligned} R_{n2} &= n^{-1}\sum_{t=1}^{n} z_t' u_t p_{2t} = n^{-1}\sum_{t=1}^{n} z_t' u_t q_{zz,t}^{-1}(Q_{zz,t} - q_{zz,t})\beta_{zx,t} \\ &= n^{-1}\sum_{t=1}^{n}\{\eta_{zt}'\varepsilon_t\}\{I_{zt}q_{zz,t}^{-1}\}(Q_{zz,t} - q_{zz,t})\{h_t\beta_{zx,t}\}. \end{aligned}$$

Denote

$$\begin{aligned} \varepsilon_t^* &= \eta_{zt}'\varepsilon_t = (\varepsilon_{1t}^*, ..., \varepsilon_{pt}^*)', \quad \varepsilon_{vt}^* = \eta_{vt}\varepsilon_t, \\ I_{zt}q_{zz,t}^{-1} &= \{\theta_{ir,t}\}, \\ Q_{zz,t}^* &= K_t^{-1}\sum_{j=1}^{n} b_{n,tj} I_{zj}\big(\eta_{zj}\eta_{zj}' - E[\eta_{zj}\eta_{zj}']\big)I_{zj} = \{s_{mv,t}\}, \\ h_t\beta_{zx,t} &= \{\delta_{k\ell,t}\}, \end{aligned}$$

where ε_t^* is a $p\times 1$ matrix with v-th element ε_{vt}^*, $I_{zt}q_{zz,t}^{-1}$ is a $p\times p$ matrix with i,r-th element $\theta_{ir,t}$, $Q_{zz,t}^*$ is a $p\times p$ matrix with m,v-th element $s_{mv,j}$, and $h_t\beta_{zx,t}$ is a $p\times q$ matrix with k,ℓ-th element $\delta_{k\ell,t}$.

It is easy to see that to prove $||R_{n2,2}|| = o_p(n^{-1/2})$ it suffices to show that for any elements $\varepsilon_{vt}^*, \theta_{ir,t}, s_{mv,t}, \delta_{k\ell,t}$, it holds

$$i_n = n^{-1}\sum_{t=1}^{n}\theta_{ir,t}\delta_{k\ell,t}\{s_{mv,t}\varepsilon_{vt}^*\} = o_p(n^{-1/2}). \tag{6.101}$$

Setting $\xi_t = \theta_{ir,t}\delta_{k\ell,t}$, we can write

$$i_n = n^{-1}\sum_{j,t=1}^{n}\xi_t\{s_{mv,j}\varepsilon_{vt}^*\} \tag{6.102}$$

where $s_{mv,t}$ can be written as

$$\begin{aligned} s_{mv,t} &= K_t^{-1}\sum_{j=1}^{n} b_{n,tj} g_{zm,j} g_{zv,j}\big(\eta_{zm,j}\eta_{zv,j} - E[\eta_{zm,j}\eta_{zv,j}]\big) \\ &= \sum_{j=1}^{n} a_{n,tj}\beta_j w_j \end{aligned}$$

with

$$w_j = \eta_{zm,j}\eta_{zv,j} - E[\eta_{zm,j}\eta_{zv,j}], \quad \beta_j = g_{zm,j}g_{zv,j}, \quad a_{n,tj} = K_t^{-1}b_{n,tj}.$$

Under assumptions of lemma, $\{\xi_t, \beta_t\}$ are independent of $\{w_t\}$. Observe that

$$E\xi_t^4 \le C, \quad E\beta_t^4 \le C,$$

where $C < \infty$ does not depend on t. Indeed, by (6.100), $||I_{zt}q_{zz,t}^{-1}|| \le C$ and therefore, $\theta_{ir,t} \le ||q_{zz,t}^{-1}|| \le C$, while by (6.24), $E||\beta_{zx,t}||^8 \le C$. Hence,

$$\begin{aligned} E\xi_t^4 &\le CE[h_t^4||\beta_{zx,t}||^4] \le C(Eh_t^8 + E||\beta_{zx,t}||^8) \le C, \\ E\beta_t^4 &\le Eg_{zm,t}^8 + Eg_{zv,t}^8 \le C, \end{aligned}$$

by Assumption 2.3(i). Hence, by (7.5) of Lemma 7.2,

$$E|i_n| \le C(\max_{t,j=1,\dots,n} a_{n,tj})m = C(K_t^{-1}\max_{t,j=1,\dots,n} b_{n,tj})m.$$

By (6.40), $K_t^{-1} \le CH^{-1}$ and $b_{n,tj} \le C$ which implies

$$E|i_n| \le CH^{-1}m = o(n^{-1/2})$$

since by assumption, $m = O(\log n)$, and by (14), $H \ge n^a$ for some $a > 1/2$. This completes the proof of (6.101) and verification of the claim $R_{n2;2} = o_p(n^{-1/2})$. That concludes the proof of the lemma. □

Lemma 6.6. *Under the assumptions of Theorem 2.1, R_n defined in (6.8) has property*

$$S_{\widehat{v}\widehat{v}}^{-1}R_n = o_p(n^{-1/2}). \tag{6.103}$$

Proof of Lemma 6.6. In (6.57) of Lemma 6.3 it is shown that $nS_{\widehat{v}\widehat{v}}^{-1} = O_p(1)$. Therefore, to prove (6.103), it suffices to show

$$n^{-1}R_n = o(n^{-1/2}). \tag{6.104}$$

By definition (6.8),

$$R_n = \sum_{j=1}^n \widehat{v}_j z_j'(\beta_j - \widetilde{\beta}_j), \qquad \widetilde{\beta}_t = S_{zz,t}^{-1}S_{z\zeta,t} = Q_{zz,t}^{-1}Q_{z\zeta,t}.$$

Recall notation

$$\begin{aligned} v_j &= x_j - \beta_{zx,j}'z_j, \quad \zeta_j = z_j'\beta_j + u_j, \\ Q_{zz,t} &= K_t^{-1}S_{zz,t}, \quad q_{zz,t} = E[z_tz_t'|\mathcal{F}_n^*], \quad Q_{zu,t} = K_t^{-1}S_{zu,t}. \end{aligned}$$

Recall also notation

$$\begin{aligned} Q_{zz\beta,t} &= K_t^{-1}\sum_{j=1}^n b_{n,tj}z_jz_j'(\beta_j - \beta_t), \\ Q_{zz\beta,t}^* &= K_t^{-1}\sum_{j=1}^n b_{n,tj}I_{zj}(\eta_{zj}\eta_{zj}' - E[\eta_{zj}\eta_{zj}'])I_{zj}(\beta_j - \beta_t), \end{aligned}$$

$$R^*_{zz\beta,t} \;=\; K_t^{-1}\sum_{j=1}^n b_{n,tj} I_{zj} E[\eta_{zj}\eta'_{zj}] I_{zj}(\beta_j - \beta_t).$$

Then, $Q_{zz\beta,t} = Q^*_{zz\beta,t} + R^*_{zz\beta,t}$, and

$$Q_{z\zeta,t} \;=\; K_t^{-1}\sum_{j=1}^n b_{n,tj} z_j \zeta_j = Q_{zz\beta,t} + Q_{zu,t} = Q^*_{zz\beta,t} + R^*_{zz\beta,t} + Q_{zu,t}.$$

Hence,

$$\begin{aligned} \widetilde{\beta}_t - \beta_t &= Q^{-1}_{zz,t} Q_{z\zeta,t} - \beta_t = Q^{-1}_{zz,t}(Q_{zz\beta,t} + Q_{zu,t}) \\ &= Q^{-1}_{zz,t}(Q^*_{zz\beta,t} + R^*_{zz\beta,t} + Q_{zu,t}). \end{aligned} \tag{6.105}$$

Write,

$$\begin{aligned} \widehat{v}_t z'_t &= (\widehat{v}_t - v_t) z'_t + v_t z'_t (1 - q^{-1}_{zz,t} Q_{zz,t}) + v_t z'_t q^{-1}_{zz,t} Q_{zz,t}, \\ \widehat{v}_t z'_t(\beta_t - \widetilde{\beta}_t) &= \{(\widehat{v}_t - v_t) z'_t (\beta_t - \widetilde{\beta}_t)\} + \{v_t z'_t (1 - q^{-1}_{zz,t} Q_{zz,t})(\beta_t - \widetilde{\beta}_t)\} \\ &\quad + \{v_t z'_t q^{-1}_{zz,t} Q^*_{zz\beta,t}\} + \{v_t z'_t q^{-1}_{zz,t} R^*_{zz\beta,t}\} + \{v_t z'_t q^{-1}_{zz,t} Q_{zu,t}\} \\ &= \rho_{1t} + \rho_{2t} + \rho_{3t} + \rho_{4t} + \rho_{5t}. \end{aligned}$$

Hence,

$$R_n = \sum_{t=1}^n \rho_{1t} + \dots + \sum_{t=1}^n \rho_{5t} = R_{n1} + \dots + R_{n5}.$$

Therefore, to prove (6.104), it suffices to show that

$$n^{-1} R_{ni} = o(n^{-1/2}), \;\; i = 1, \dots, 5. \tag{6.106}$$

(1) *Proof of (6.106) for* R_{n1}. By (6.91), $\widehat{v}_t - v_t = (\beta_{zx} - \widehat{\beta}_{zx})' z_t$, and by (6.105), $\widetilde{\beta}_t - \beta_t = Q^{-1}_{zz,t}(Q_{zz\beta,t} + Q_{zu,t})$. Therefore,

$$\begin{aligned} \rho_{1t} &= (\beta_{zx} - \widehat{\beta}_{zx})' z_t z'_t (\beta_t - \widetilde{\beta}_t) \\ &= (\beta_{zx} - \widehat{\beta}_{zx})' z_t \{z'_t q^{-1}_{zz,t}\}\{q_{zz,t} Q^{-1}_{zz,t}\}(Q_{zz\beta,t} + Q_{zu,t}), \\ ||\rho_{1t}|| &\le ||\beta_{zx} - \widehat{\beta}_{zx}||\, ||z_t||\, ||z'_t q^{-1}_{zz,t}||\, ||q_{zz,t} Q^{-1}_{zz,t}||\, ||Q_{zz\beta,t} + Q_{zu,t}||. \end{aligned}$$

Observe that under assumptions of lemma,

$$\begin{aligned} ||z'_t q^{-1}_{zz,t}|| &= ||\eta'_{zt} I_{zt}\{I^{-1}_{zt}\Sigma^{-1}_{zz} I^{-1}_{zt}\}|| = ||\eta'_{zt}\Sigma^{-1}_{zz} I^{-1}_{zt}|| \\ &\le ||\eta_{zt}||\, ||\Sigma^{-1}_{zz}||\, ||I^{-1}_{zt}|| \le C||\eta_{zt}||. \end{aligned}$$

Moreover

$$
\begin{aligned}
||q_{zz,t}Q_{zz,t}^{-1}|| &\le ||q_{zz,t}Q_{zz,t}^{-1} - 1|| + 1 = ||(q_{zz,t} - Q_{zz,t})Q_{zz,t}^{-1}|| + 1 \\
&\le (\max_{t=1,\dots,n} ||q_{zz,t} - Q_{zz,t}||)(\max_{t=1,\dots,n} ||Q_{zz,t}^{-1}||) + 1 = C_n = O_p(1),
\end{aligned}
$$

in view of the bounds (6.30) and (6.31). In addition, by (6.50),

$$
||\widehat{\beta}_{zx,t} - \beta_{zx,t}|| \le C_n(||\beta_{zx,t}|| + 1)\{||Q_{zz,t} - q_{zz,t}|| + ||Q_{zx,t} - q_{zx,t}||\}.
$$

Recall also that $||z_t|| \le ||I_{zt}||\,||\eta_{zt}||$. Thus, setting $\nu_t = (||\beta_{zx,t}|| + 1)||I_{zt}||\,||\eta_{zt}||^2$, we obtain the bound:

$$
\rho_{1t} \le C_n\rho_{1t}^*, \quad \rho_{1t}^* = \nu_t\{||Q_{zz,t} - q_{zz,t}|| + ||Q_{zx,t} - q_{zx,t}||\}\{||Q_{zz\beta,t}|| + ||Q_{zu,t}||\}.
$$

Hence,

$$
n^{-1}||R_{n1}|| \le C_n r_{n1}^*, \qquad r_{n1}^* = n^{-1}\sum_{t=1}^{n} \rho_{1t}^*,
$$

where $C_n = o_p(1)$. We will show that

$$
Er_{n1}^* = o(n^{-1/2}), \tag{6.107}
$$

which implies $r_{n1}^* = o_p(n^{-1/2})$ and proves the required claim: $n^{-1}||R_{n1}|| = o_p(n^{-1/2})$.

To verify (6.107), notice that under the assumptions of lemma and by (6.24),

$$
\begin{aligned}
E\nu_t^4 &= E\big[(||\beta_{zx,t}|| + 1)^4||I_{zt}||^4\,||\eta_{zt}||^8\big] \\
&= E\big[(||\beta_{zx,t}|| + 1)^4||I_{zt}||^4\big]\,E||\eta_{zt}||^8 \le \{E(||\beta_{zx,t}|| + 1)^8 + E||I_{zt}||^8\}E||\eta_{zt}||^8 \le C,
\end{aligned}
$$

where $C < \infty$ does not depend on t. Moreover, by (6.25), (6.26), (6.115) and (6.119),

$$
\begin{aligned}
E||Q_{zz,t} - q_{zz,t}||^4 &\le C\big(H^{-2}m^2 + (H/n)^{4\gamma_1}\big), \quad E||Q_{zx,t} - q_{zx,t}||^4 \le C\big(H^{-2}m^2 + (H/n)^{4\gamma_1}\big), \\
E||Q_{zz\beta,t}||^2 &\le C(H/n)^{2\gamma_2}, \quad E||Q_{zu,t}||^2 \le CH^{-1}
\end{aligned} \tag{6.108}
$$

where $C < \infty$ does not depend on t. Therefore, using Hölder's inequality, we can bound

$$
\begin{aligned}
&E\rho_{1t}^* \\
&\quad\le (E\nu_t^4)^{1/4}(E(||Q_{zz,t} - q_{zz,t}||^4 + ||Q_{zx,t} - q_{zx,t}||)^4)^{1/4}(E(||Q_{zz\beta,t}||^2 + ||Q_{zu,t}||)^2)^{1/2} \\
&\quad\le C\big(H^{-2}m^2 + (H/n)^{4\gamma_1}\big)^{1/4}\big(H^{-1} + (H/n)^{2\gamma_2}\big)^{1/2} \\
&\quad\le C\big(H^{-1/2}m^{1/2} + (H/n)^{\gamma_1}\big)\big(H^{-1/2} + (H/n)^{\gamma_2}\big)^{1/2} \\
&\quad= o(n^{-1/2}).
\end{aligned} \tag{6.109}
$$

The last bound in (6.109) follows noting that $\gamma_1, \gamma_2 \in (3/4, 1]$, $m = O(\log n)$, and by assumption (14), $n^a \le H = O(n^{2/3})$ for some $a > 1/2$. Then, $H^{-1}m = o(n^{-1/2})$, $H^{-1/2}m^{1/2}(H/n)^{\gamma_i} = n^{-1/2}(H/n)^{\gamma_i - 1/2}m^{1/2} = o(n^{-1/2})$ for $i = 1, 2$, and $(H/n)^{\gamma_1+\gamma_2} \le (n^{2/3}/n)^{\gamma_1+\gamma_2} = o(n^{-1/2})$.

This completes the proof of (6.106) for R_{n1}.

(2) *Proof of (6.106) for* R_{n2}. Using property (6.105) of $\widetilde{\beta}_t$, we obtain

$$\begin{aligned} \rho_{2t} &= v_t z_t'(1 - q_{zz,t}^{-1} Q_{zz,t})(\beta_t - \widetilde{\beta}_t) \\ &= v_t z_t' q_{zz,t}^{-1}(q_{zz,t} - Q_{zz,t}) Q_{zz,t}^{-1}(Q_{zz\beta,t} + Q_{zu,t}), \\ ||\rho_{2t}|| &\le ||v_t z_t'||\,||q_{zz,t}^{-1}||\,||q_{zz,t} - Q_{zz,t}||\,||Q_{zz,t}^{-1}||\,||Q_{zz\beta,t} + Q_{zu,t}|| \\ &\le C_n ||v_t z_t'||\,||q_{zz,t} - Q_{zz,t}||\,||Q_{zz\beta,t} + Q_{zu,t}||, \end{aligned}$$

since by (6.24) and (6.31), $||q_{zz,t}^{-1}||\,||Q_{zz,t}^{-1}|| \le C_n = O_p(1)$ where C_n does not depend on t. Hence,

$$n^{-1}||R_{n2}|| \le C_n r_{n2}^*, \quad r_{n2}^* = n^{-1}\sum_{t=1}^{n} ||v_t z_t'||\,||q_{zz,t} - Q_{zz,t}||\,||Q_{zz\beta,t} + Q_{zu,t}||.$$

We will show that

$$E r_{n2}^* = o(n^{-1/2}) \tag{6.110}$$

which implies $r_{n2}^* = o_p(n^{-1/2})$ and proves the required claim: $n^{-1}||R_{n2}|| = o_p(n^{-1/2})$.

It remains to show (6.110). Under assumptions of lemma, $E||v_t z_t'||^4 \le E||v_t||^8 + E||z_t'||^8 \le C$ where $C < \infty$ does not depend on t. Hence, using Hölder inequality, and the bound given in (6.108), we obtain

$$\begin{aligned} &E\big[||v_t z_t'||\,||Q_{zz,t} - q_{zz,t}||\,\{||Q_{zz\beta,t}|| + ||Q_{zu,t}||\}\big] \\ &\le (E||v_t z_t'||^4)^{1/4}(E||Q_{zz,t} - q_{zz,t}||^4)^{1/4}\{(E||Q_{zz\beta,t}||^2)^{1/2} + (E||Q_{zu,t}||^2)^{1/2}\} \\ &\le C\big(H^{-1/2}m^{1/2} + (H/n)^{\gamma_1}\big)\big((H/n)^{\gamma_2} + H^{-1/2}\big) \\ &= o(n^{-1/2}). \end{aligned} \tag{6.111}$$

where the last relation in (6.111) holds because of the same argument as in (6.109).

Thus,

$$E r_{n2}^* \le o(n^{-1/2})\{n^{-1}\sum_{t=1}^{n} 1\} = o(n^{-1/2})$$

which proves (6.110).

(3) *Proof of (6.106) for* R_{n3}. We have

$$n^{-1}||R_{n3}|| \le n^{-1}\sum_{t=1}^{n} ||q_{zz,t}^{-1}||\,||Q_{zz\beta,t}^*||\,||v_t z_t'||$$

$$\leq \quad Cn^{-1}\sum_{t=1}^{n}||Q^*_{zz\beta,t}||\,||v_t z_t'||,$$

since $||q^{-1}_{zz,t}|| \leq C$ by (6.24).

By the assumptions of the lemma, $E||v_t z_t'||^2 \leq E||v_t||^4 + E||z_t||^4 \leq C$, while by (6.118) of Lemma 6.7,

$$E||Q^*_{zz\beta,t}||^2 \quad \leq \quad CH^{-1}(H/n)^{2\gamma_2}.$$

Hence,

$$\begin{aligned}
E[||Q^*_{zz\beta,t}||\,||v_t z_t'||] &\leq (E||Q^*_{zz\beta,t}||^2)^{1/2}(E||v_t z_t'||^2)^{1/2} \leq CH^{-1/2}(H/n)^{\gamma_2}, \\
E\Big[n^{-1}\sum_{t=1}^{n}||Q^*_{zz\beta,t}||\,||v_t z_t'||\Big] &\leq n^{-1}\sum_{t=1}^{n}E[||Q^*_{zz\beta,t}||\,||v_t z_t'||] \\
&\leq CH^{-1/2}(H/n)^{\gamma_2} = Cn^{-1/2}(H/n)^{\gamma_2-1/2} = o(n^{-1/2})
\end{aligned}$$

since $\gamma_2 > 1/2$ and $H = o(n)$. This implies $n^{-1}||R_{n3}|| = o_p(n^{-1/2})$ which proves the claim (6.106) for R_{n3}.

(4) *Proof of (6.106) for R_{n4}.* By (6.15),

$$\begin{aligned}
v_t &= I_{xt}\nu_t, \;\; \nu_t = \eta_{xt} - E[\eta_{xt}\eta_{zt}'](E[\eta_{zt}\eta_{zt}'])^{-1}\eta_{zt}, \\
v_t z_t' &= I_{xt}\nu_t\eta_{zt}'I_{zt}, \quad E[\nu_t\eta_{zt}'] = 0, \\
R^*_{zz\beta,t} &= K_t^{-1}\sum_{j=1}^{n}b_{n,tj}I_{zj}E[\eta_{zj}\eta_{zj}']I_{zj}(\beta_j - \beta_t).
\end{aligned} \tag{6.112}$$

Hence, we can write $n^{-1}R_{n4}$ as

$$\begin{aligned}
n^{-1}R_{n4} &= \sum_{t=1}^{n}v_t z_t' q^{-1}_{zz,t}R^*_{zz\beta,t} = n^{-1}\sum_{t=1}^{n}I_{xt}\nu_t\eta_{zt}'\{I_{zt}q^{-1}_{zz,t}R^*_{zz\beta,t}\} \\
&= \sum_{t=1}^{n}\{n^{-1}I_{xt}\}\{\nu_t\eta_{zt}'\}\{I_{zt}q^{-1}_{zz,t}R^*_{zz\beta,t}\}.
\end{aligned}$$

Denote $i_n = \max_{t=1,\dots,n}||I_{xt}||$. Then,

$$n^{-1}R_{n4} = i_n J_n \quad J_n = \sum_{t=1}^{n}\{n^{-1}i_n^{-1}I_{xt}\}\{\nu_t\eta_{zt}'\}\{I_{zt}q^{-1}_{zz,t}R^*_{zz\beta,t}\}.$$

By Assumption 2.2(i) of theorem, the elements the matrix $\nu_t\eta_{zt}'$ are stationary short memory sequences which have zero mean, $E\nu_t\eta_{zt}' = 0$, and $\{i_n^{-1}I_{xt},\; I_{zt}q^{-1}_{zz,t}R^*_{zz\beta,t}\}$ and $\{\nu_t\eta_{zt}'\}$ are mutually independent. Hence by (7.2) of Lemma 7.1,

$$E||J_n||^2 \le Cn^{-2}\sum_{t=1}^{n} E\big[||i_n^{-1}I_{xt}||^2||I_{zt}q_{zz,t}^{-1}R^*_{zz\beta,t}||^2\big]$$
$$\le Cn^{-2}\sum_{j=1}^{n} E\big[||i_n^{-1}I_{xt}||^2||I_{zt}q_{zz,t}^{-1}||^2||R^*_{zz\beta,t}||^2\big].$$

By definition of i_n, $||i_n^{-1}I_{xt}||^2 = i_n^{-2}||I_{xt}||^2 \le 1$, and $||I_{zt}q_{zz,t}^{-1}|| \le C$ by (6.100). Hence,

$$E\big[||i_n^{-1}I_{xt}||^2||I_{zt}q_{zz,t}^{-1}||^2\,||R^*_{zz\beta,t}||^2\big] \le CE||R^*_{zz\beta,t}||^2 \le C(H/n)^{2\gamma_2}$$

by (6.117) of Lemma 6.7 where $C < \infty$ does not depend on t, n. Therefore,

$$E||J_n||^2 \le Cn^{-2}\sum_{j=1}^{n}(H/n)^{2\gamma_2} \le Cn^{-1}(H/n)^{2\gamma_2}.$$

Hence, $J_n = O_p\big(n^{-1/2}(H/n)^{\gamma_2}\big)$.

By assumption (14), $H = O(n^{2/3})$, and by Assumption 2.3(ii), $\gamma_1 \ge 3/4 + \delta$ for some $\delta > 0$. Hence $(H/n)^{\gamma_2} \le C(n^{-1/3})^{3/4+\delta} \le Cn^{-1/4-\delta/3}$. Under assumptions of lemma, the variable $||I_{zt}||^2$ has 4-finite moments: $E||I_{zt}||^8 \le C$ where $C < \infty$ does not depend on t. Hence, by Lemma 6.4, $i_n = O_p(n^{1/4+a})$ for any $a > 0$. Suppose that $a < \delta/3$. Then,

$$||n^{-1}R_{n4}||^2 \le i_nJ_n = O_p(n^{1/4+a})O_p(n^{-1/2}n^{-1/4-\delta/3}) = O_p(n^{a-\delta/3}) = o_p(1)$$

which implies (6.106) for R_{n4}.

(5) *Proof of (6.106) for* R_{n5}. Recall that $u_j = \varepsilon_j h_j$, $Q_{zu,t} = K_t^{-1}\sum_{j=1}^n b_{n,tj}z_ju_j$. Then,

$$R_{n5} = \sum_{t=1}^{n} v_tz_t'q_{zz,t}^{-1}\,Q_{zu,t} = \sum_{t=1}^{n} v_tz_t'q_{zz,t}^{-1}\{K_t^{-1}\sum_{j=1}^{n} b_{n,tj}z_ju_j\}$$
$$= \sum_{j,t=1}^{n} K_t^{-1}b_{n,tj}v_tz_t'q_{zz,t}^{-1}u_jz_j$$
$$= \sum_{t=1}^{n}\{\sum_{j=1}^{n} K_j^{-1}b_{n,tj}v_jz_j'q_{zz,j}^{-1}\}u_tz_t.$$

We need to show that

$$n^{-1}R_{n5} = o_p(n^{-1/2}). \tag{6.113}$$

By (6.112), $v_jz_j' = I_{xj}\nu_j\eta_{zj}'I_{zj}$ and $q_{zz,j} = I_{zj}\Sigma_{zz}I_{zj}$. Thus, $v_jz_j'q_{zz,j}^{-1} = I_{xj}\nu_j\eta_{zj}'\Sigma_{zz}^{-1}I_{zj}^{-1}$.

Hence,

$$R_{n5} = \sum_{t=1}^{n}\{\sum_{j=1}^{n} K_j^{-1} b_{n,tj} I_{xj}\nu_j \eta'_{zj}\Sigma_{zz}^{-1} I_{zj}^{-1}\}\{h_t I_{zt}\varepsilon_t\eta_t\}.$$

Denote $\Sigma_{zz}^{-1} = \{\sigma_{\ell m}\}$. By Assumption 2.2(ii), $||\Sigma_{zz}^{-1}|| < \infty$ which implies $|\sigma_{\ell m}| < \infty$. Observe, that $n^{-1}R_{n5}$ is a linear combination of the following type of sums:

$$i_n \quad = \quad n^{-1}\sum_{t=1}^{n}\{\sum_{j=1}^{n} K_j^{-1} b_{n,tj} g_{xk,j}\nu_{kj}\eta_{z\ell,j}\sigma_{\ell m} g_{zm,j}^{-1}\}\{h_t g_{zm,t}\varepsilon_t\eta_{zm,t}\}.$$

Clearly, to prove (6.113), it suffices to show that

$$E|i_n| = o(n^{-1/2}). \tag{6.114}$$

Setting

$$\begin{aligned} s_{nt} \quad &= \quad \sum_{j=1}^{n} K_j^{-1} b_{n,tj}\{g_{xk,j}\sigma_{\ell m} g_{zm,j}^{-1}\}\nu_{kj}\eta_{z\ell,j} \\ &= \quad \sum_{j=1}^{n} a_{n,tj}\beta_j w_j, \quad a_{n,tj} = K_j^{-1} b_{n,tj}, \;\; \beta_j = g_{xk,j}\sigma_{\ell m} g_{zm,j}^{-1}, \;\; w_j = \nu_{kj}\eta_{z\ell,j} \end{aligned}$$

we can write

$$i_n = n^{-1}\sum_{t=1}^{n}\xi_t\{s_{nt}\varepsilon_t^*\}, \quad \xi_t = h_t g_{zm,t}, \;\; \varepsilon_t^* = \varepsilon_t\eta_{zm,t}.$$

Under the assumptions of the theorem, w_j is a stationary short memory sequence with $Ew_j = 0$. Moreover, $\{w_j, \varepsilon_t^*\}$ and $\{\xi_j, \beta_j\}$ are mutually independent and $g_{zm,j} \geq c_0 > 0$ by Assumption 2.3. Hence,

$$\begin{aligned} E\xi_t^4 \quad &\leq \quad E[h_t^4 g_{zm,t}^4] \leq E[h_t^8] + E[g_{zm,t}^8] \leq C, \\ E\beta_t^4 \quad &\leq \quad E[g_{xk,j}^4\sigma_{\ell m}^4 g_{zm,j}^{-4}] \leq CE[g_{xk,j}^4] \leq C \end{aligned}$$

where $C < \infty$ does not depend on j, n.

Hence, by Lemma 7.2(iii),

$$E|i_n| \quad \leq \quad C(\max_{t,j=1,\ldots,n} a_{n,tj})m = C(\max_{t,j=1,\ldots,n} K_j^{-1} b_{n,tj})m \leq CH^{-1}m$$

because by (6.40), $K_j^{-1} \leq CH^{-1}$ and $b_{n,tj} \leq C$. Therefore,

$$Ei_n^2 \quad \leq \quad CH^{-1}m = o(n^{-1/2})$$

since $m = O(\log n)$ and by (14), $H \ge n^a$ for some $a > 1/2$. This proves (6.114), implies (6.106) for R_{n5} and completes the proof of the lemma. □

The following lemma is used in the proof of Lemma 6.6.

Lemma 6.7. *Under the assumptions of Theorem 2.1, the following holds:*

$$E||Q_{zz\beta,t}||^2 \le C(H/n)^{2\gamma_2}, \tag{6.115}$$
$$E||Q_{\beta,t}||^4 \le C(H/n)^{4\gamma_2}, \tag{6.116}$$
$$E||R^*_{zz\beta,t}||^2 \le C(H/n)^{2\gamma_2}, \tag{6.117}$$
$$E||Q^*_{zz\beta,t}||^2 \le CH^{-1}(H/n)^{2\gamma_2}, \tag{6.118}$$
$$E||Q_{zu,t}||^2 \le CH^{-1}, \tag{6.119}$$
$$E||Q_{zu,t}||^4 \le CH^{-1}, \tag{6.120}$$

where $C < \infty$ does not depend on t, n

Proof of Lemma 6.7. *Proof of (6.115).* We have

$$\begin{aligned} ||Q_{zz\beta,t}|| &\le K_t^{-1}\sum_{j=1}^{n} b_{n,tj}||z_j||^2||\beta_j - \beta_t||, \\ E||Q_{zz\beta,t}||^2 &\le K_t^{-2}\sum_{j,i=1}^{n} b_{n,tj}b_{n,ti}E\big[||z_j||^2||z_i||^2||\beta_j - \beta_t||\,||\beta_i - \beta_t||\big]. \end{aligned}$$

By (10), $E||z_j||^8 \le C$ where $C < \infty$ does not depend on j. By smoothness Assumption 2.4,

$$E||\beta_t - \beta_j||^4 \le C(|t-j|/n)^{4\gamma_2}, \quad t, j, = 1, ..., n. \tag{6.121}$$

These bounds together with Hölder inequality imply that

$$\begin{aligned} E\big[||z_j||^2||z_i||^2||\beta_j - \beta_t||\,||\beta_i - \beta_t||\big] &\le \big(E||z_j||^8E||z_i||^8E||\beta_j - \beta_t||^4E||\beta_i - \beta_t||^4\big)^{1/4} \\ &\le C(|t-j|/n)^{\gamma_2}(|t-i|/n)^{\gamma_2}. \end{aligned}$$

This, together with the bound (6.40), $K_t^{-1} \le CH^{-1}$, and the bound (6.41) which is valid also for γ_2, implies (6.115):

$$E||Q_{zz\beta,t}||^2 \le C(H/n)^{2\gamma_2}\big(H^{-1}\sum_{j=1}^{n} b_{n,tj}(|t-j|/H)^{\gamma_2}\big)^2 \le C(H/n)^{2\gamma_2}.$$

Proof of (6.116). As above,

$$\begin{aligned} E||Q_{\beta,t}||^4 &= E\Big(K_t^{-1}\sum_{j=1}^n b_{n,tj}||\beta_j-\beta_t||\Big)^4 \le \Big(K_t^{-1}\sum_{j=1}^n b_{n,tj}(E||\beta_j-\beta_t||^4)^{1/4}\Big)^4 \\ &\le C(H/n)^{4\gamma_2}\big(H^{-1}\sum_{j=1}^n b_{n,tj}((|t-j|/H)^{4\gamma_2})^{1/4}\big)^4 \le C(H/n)^{4\gamma_2}. \end{aligned}$$

This implies (6.116).

Proof of (6.117). We have

$$||R^*_{zz\beta,t}|| \le K_t^{-1}\sum_{j=1}^n b_{n,tj}||I_{zj}E[\eta_{zj}\eta'_{zj}]I_{zj}||\,||\beta_j-\beta_t||$$

where $||I_{zj}E[\eta_{zj}\eta'_{zj}]I_{zj}|| \le ||I_{zj}||^2E||\eta_{zj}||^2 \le C||I_{zj}||^2$ and $C<\infty$ does not depend on j. By Assumption 2.3(i), $E||I_{zj}||^8 \le C$. Hence,

$$||R^*_{zz\beta,t}||^2 \le CK_t^{-2}\sum_{j,i=1}^n b_{n,tj}b_{n,ti}||I_{zj}||^2||I_{zi}||^2||\beta_j-\beta_t||\,||\beta_i-\beta_t||, \tag{6.122}$$

and (6.117) follows using the same argument as in the proof of (6.115).

Proof of (6.118). Write

$$\begin{aligned} Q^*_{zz\beta,t} &= K_t^{-1}\sum_{j=1}^n b_{n,tj}I_{zj}\Big(\eta_{zj}\eta'_{zj}-E[\eta_{zj}\eta'_{zj}]\Big)I_{zj}(\beta_j-\beta_t) \\ &= \sum_{j=1}^n A_jW_jB_j, \\ & \quad A_j = K_t^{-1}b_{n,tj}I_{zj}, \quad W_j=\eta_{zj}\eta'_{zj}-E[\eta_{zj}\eta'_{zj}], \quad B_j=I_{zj}(\beta_j-\beta_t). \end{aligned}$$

Under Assumption 2.2(i) of theorem, the elements $\eta_{zk,j}\eta_{z\ell,j}-E[\eta_{zk,j}\eta_{z\ell,j}]$ of W_j are stationary short memory sequences with zero mean and $\{A_j,B_j\}$ and $\{W_j\}$ are mutually independent. Hence by (7.2) of Lemma 7.1,

$$\begin{aligned} E||Q^*_{zz\beta,t}||^2 &= E||\sum_{j=1}^n A_jW_jB_j||^2 \le C\sum_{j=1}^n E\big[||A_j||^2||B_j||^2\big] \\ &\le CK_t^{-2}\sum_{j=1}^n b^2_{n,tj}E\big[||I_{zj}||^4||\beta_j-\beta_t||^2\big]. \end{aligned}$$

Observe that

$$E\big[||I_{zj}||^4||\beta_j-\beta_t||^2\big] \le (E||I_{zj}||^8)^{1/2}(E||\beta_j-\beta_t||^4)^{1/2} \le C(|t-j|/n)^{2\gamma_2}$$

by (6.121) and because under the assumptions of the theorem, $E||I_{zt}||^8 \le C$. Moreover, $K_t^{-1} \le CH^{-1}$ and $b_{n,tj} \le C$ where $C < \infty$ does not depend on t, n. Hence,

$$E||Q^*_{zz\beta,t}||^2 \le CH^{-2}\sum_{j=1}^{n} b_{n,tj}(|t-j|/n)^{2\gamma_2}$$
$$\le CH^{-1}(H/n)^{2\gamma_2}\{H^{-1}\sum_{j=1}^{n} b_{n,tj}(|t-j|/H)^{2\gamma_2}\} \le CH^{-1}(H/n)^{2\gamma_2}$$

using the bound (6.41) which is valid also for $2\gamma_2$. This proves (6.118).

Proof of (6.119). Recall that $z_j = I_{zj}\eta_{gj}$, $u_j = h_j\varepsilon_j$. Then,

$$Q_{zu,t} = K_t^{-1}\sum_{j=1}^{n} b_{n,tj}I_{zj}\eta_{zj}\varepsilon_j h_j I_{zj} = \sum_{j=1}^{n} A_j W_j B_j,$$
$$A_j = K_t^{-1}b_{n,tj}I_{zj}, \quad W_j = \eta_{zj}\varepsilon_j, \quad B_j = h_j I_{zj}.$$

Under assumption of theorem, ε_j is a martingale difference sequence, the elements $\eta_{zk,j}\varepsilon_j$ of W_j are stationary white noise sequences with zero mean and $\{A_j, B_j\}$ and $\{W_j\}$ are mutually independent. Hence, by (7.2) of Lemma 7.1,

$$E||Q_{zu,t}||^2 \le C\sum_{j=1}^{n} E\big[||A_j||^2||B_j||^2\big] \le CK_t^{-2}\sum_{j=1}^{n} b_{n,tj}^2 E\big[||I_{zj}||^4 h_j^2\big].$$

Since $E\big[||I_{zj}||^4 h_j^2\big] \le E||I_{zj}||^8 + Eh_j^4 \le C$, this implies

$$E||Q_{zu,t}||^2 \le CK_t^{-1}\{K_t^{-1}\sum_{j=1}^{n} b_{n,tj}1\} = CK_t^{-1} \le CH^{-1}$$

which proves (6.119).

Proof of (6.120). We have

$$Q_{zu,t} = K_t^{-1}\sum_{j=1}^{n} b_{n,tj}z_j u_j = (s_{1n}, ..., s_{pn})', \quad s_{kn} = K_t^{-1}\sum_{j=1}^{n} b_{n,tj}z_{kj}u_j.$$

We will show that

$$Es_{kn}^4 = O(H^{-1}) \tag{6.123}$$

which implies the required claim $E||Q_{zu,t}||^4 = O(H^{-1})$. Write

$$s_{kn}^2 = K_t^{-1}\sum_{j,i=1}^{n} b_{n,tj}b_{n,ti}\{z_{kj}u_j\}\{z_{ki}u_i\}$$

$$= \quad K_t^{-2} \sum_{j=i=1}^{n} [...] + 2K_t^{-2} \sum_{j=1}^{n} \sum_{i=j+1}^{n} [...] = q_{kn,1} + q_{kn,2}.$$

We will verify that

$$Eq_{kn,1}^2 = O(H^{-1}), \quad Eq_{kn,2}^2 = O(H^{-1}) \tag{6.124}$$

which implies (6.123): $Es_{kn}^4 = E(q_{kn,1} + q_{kn,2})^2 \le 2Eq_{kn,1}^2 + 2Eq_{kn,2}^2 = O(H^{-1})$. We have

$$Eq_{kn,1}^2 \quad = \quad K_t^{-2} \sum_{j,i=1}^{n} b_{n,tj}^2 b_{n,ti}^2 E[z_{kj}^2 u_j^2 z_{ki}^2 u_i^2].$$

Under assumptions of lemma, $E[z_{kj}^2 u_j^2 z_{ki}^2 u_i^2] \le Ez_{kj}^8 + Eu_j^8 + Ez_{ki}^8 + Eu_i^8 \le C$ where $C < \infty$ does not depend on j, i. Hence,

$$\begin{aligned} Eq_{kn,1}^2 \quad &\le \quad CK_t^{-2} \{K_t^{-2} \sum_{j,i=1}^{n} b_{n,tj}^2 b_{n,ti}^2\} \le CK_t^{-2} \big(K_t^{-1} \sum_{j=1}^{n} b_{n,tj}\big)^2 \\ &\le \quad CK_t^{-2} \le CH^{-2} \end{aligned}$$

by (6.40), which implies (6.124) for $q_{kn,1}$.

Finally,

$$Eq_{kn,2}^2 \quad = \quad K_t^{-4} \sum_{1 \le j_1 < i_1 \le n} \sum_{1 \le j_2 < i_2 \le n} b_{n,tj_1} b_{n,ti_1} b_{n,tj_2} b_{n,ti_2} E[z_{kj_1} u_{j_1} z_{kj_2} u_{j_2} z_{ki_1} u_{i_1} z_{ki_2} u_{i_2}].$$

Observe that under assumptions of lemma,

$$|E[z_{kj_1} u_{j_1} z_{kj_2} u_{j_2} z_{ki_1} u_{i_1} z_{ki_2} u_{i_2}]| \le Ez_{kj_1}^8 + Eu_{j_1}^8 + Ez_{kj_2}^8 + Eu_{j_2}^8 + Ez_{ki_1}^8 + Eu_{i_1}^8 + Ez_{ki_2}^8 + Eu_{i_2}^8 \le C$$

where $C < \infty$ does not depend on j_1, i_1, j_2, i_2.

Suppose that $i_2 \ne i_1$. Then $j_2, j_1, i_1 < i_2$ and it is easy to see that under assumptions of the lemma, $E[z_{kj_1} u_{j_1} z_{kj_2} u_{j_2} z_{ki_1} u_{i_1} z_{ki_2} u_{i_2}] = 0$.

Therefore,

$$\begin{aligned} Eq_{kn,2}^2 \quad &\le \quad K_t^{-4} \sum_{1 \le j_1 < i_1 \le n} \sum_{1 \le j_2 < i_2 \le n} b_{n,tj_1} b_{n,ti_1} b_{n,tj_2} b_{n,ti_2} I(i_1 = i_2) \\ &\le \quad CK_t^{-1} \big(K_t^{-1} \sum_{j=1}^{n} b_{n,tj}\big)^3 \le CK_t^{-1} \le CH^{-1}. \end{aligned}$$

This proves (6.124) and completes the proof of the lemma. □

7 Auxiliary Lemmas B.

This section contains auxiliary lemmas used in the proofs of the main results in Section 2.

Lemma 7.1. (i) *Assume that sequences* $\{\beta_t\}$ *and* $\{w_t\}$ *of univariate random variables are mutually independent, and* $\{w_t\}$ *is a covariance stationary short memory sequence with zero mean. Then,*

$$E(\sum_{t=1}^{n} \beta_t w_t)^2 \le C \sum_{t=1}^{n} E\beta_t^2. \tag{7.1}$$

(ii) *Assume* $\{A_t\}$, $\{W_t\}$ *and* $\{B_t\}$ *are* $p \times m$, $m \times \ell$ *and* $\ell \times q$ *matrices which elements are random variables. Suppose that elements of* $\{W_t\}$ *are stationary short memory sequences with zero mean and* $\{A_t, B_t\}$ *and* $\{W_t\}$ *are mutually independent. Then,*

$$E||\sum_{t=1}^{n} A_t W_t B_t||^2 \le C \sum_{t=1}^{n} E\big[||A_t||^2 ||B_t||^2\big]. \tag{7.2}$$

In (7.1) and (7.2), $C < \infty$ *does not depend on* n.

Proof of Lemma 7.1. By assumption of lemma,

$$E[w_t w_s] = \mathrm{cov}(w_t, w_s) = \gamma_{w,t-s}, \qquad \sum_{k=-\infty}^{\infty} |\gamma_w(k)| < \infty.$$

Hence,

$$\begin{aligned} E(\sum_{t=1}^{n} \beta_t w_t)^2 &\le \sum_{t,s=1}^{n} E[\beta_t \beta_s] E[w_t w_s] \le \sum_{t,s=1}^{n} E[\beta_t^2 + \beta_s^2] |\gamma_{w,t-s}| \\ &\le 2 \sum_{t=1}^{n} E[\beta_t^2] \sum_{s=-\infty}^{\infty} |\gamma_{w,s}| \le C \sum_{t=1}^{n} E[\beta_t^2], \end{aligned}$$

which proves (7.1).

It is easy to see that (7.1) implies (7.2). □

In the following lemma $\eta_{xt}, \eta_{zt}, v_t$ are defined as in (8) and (6.15) and they satisfy moment conditions of Assumption 2.2. We denote by $\eta_{xk,t}, \eta_{zk,t}, v_{kt}$ a kth component of these vectors.

Lemma 7.2. (i) *Assume, that random variables* ε_t, η_{xt} *and* η_{zt} *satisfy assumptions of Theorem 2.1.*

Suppose that random variables $\{\xi_t\}$ *and* $\{\beta_t\}$ *are mutually independent of* $\{\eta_{xt}, \eta_{zt}, \varepsilon_t\}$ *and such that* $E\xi_t^4 \le C$, $E\beta_t^4 \le C$ *where* $C < \infty$ *does not depend on* t.

(i) *Define* $w_j = \eta_{zs,j}\eta_{zv,j} - E[\eta_{zs,j}\eta_{zv,j}]$ *and* $\varepsilon_j^* = \eta_{z\ell,j}$, *and set*

$$s_{nt} = \sum_{j=1}^{n} a_{n,tj}\beta_j w_j, \qquad (7.3)$$

where $a_{n,tj} \geq 0$ *are non-random weights. Then,*

$$Es_{nt}^4 \leq C\Big(\sum_{j=1}^{n} a_{n,tj}\Big)^2\big(\max_{t,j=1,\ldots,n} a_{n,tj}\big)^2 m^2, \qquad (7.4)$$

$$E\Big|n^{-1}\sum_{t=1}^{n}\xi_t s_{nt}\varepsilon_t^*\Big| \leq C\big(\max_{t,j=1,\ldots,n} a_{n,tj}\big)m, \qquad (7.5)$$

where $C < \infty$ *does not depend on* n *and* m *is the same as in Assumption 2.5(iii).*

(ii) *The bounds (7.4) and (7.5) remain valid for* $w_j = \eta_{zs,j}\eta_{xv,j} - E[\eta_{zs,j}\eta_{xv,j}]$ *and for* $w_j = v_{sj}\eta_{zv,j}$.

Proof of Lemma 7.2. We will prove (i). (The proof of (ii) is similar).

Proof of (7.4). Observe that

$$Es_{nt}^4 = \sum_{j_1,\ldots,j_4=1}^{n} a_{n,tj_1}\ldots a_{n,tj_4}E[\beta_{j_1}\ldots\beta_{j_4}]E[w_{j_1}\ldots w_{j_4}].$$

By assumptions of lemma, $|E[\beta_{j_1}\ldots\beta_{j_4}]| \leq E\beta_{j_1}^4 + \ldots + E\beta_{j_4}^4 \leq C$ and $|E[w_{j_1}\ldots w_{j_4}]| \leq Ew_{j_1}^4 + \ldots + Ew_{j_4}^4 \leq C$, where $C < \infty$ does not depend on t. Therefore,

$$\begin{aligned} Es_{nt}^4 &\leq C\sum_{1\leq j_1\leq\ldots\leq j_4\leq n} a_{n,tj_1}\ldots a_{n,tj_4}|E[w_{j_1}\ldots w_{j_4}]| \\ &\leq C\big(\max_{t,j=1,\ldots,n} a_{n,tj}\big)^2\sum_{1\leq j_1\leq\ldots\leq j_4\leq n} a_{n,tj_1}a_{n,tj_3}|E[w_{j_1}\ldots w_{j_4}]| \\ &\leq C\big(\max_{t,j=1,\ldots,n} a_{n,tj}\big)^2 \\ &\qquad\times\Big\{\sum_{1\leq j_1\leq\ldots\leq j_4\leq n\,:j_4-j_2\leq 2m}[\ldots] + \sum_{1\leq j_1\leq\ldots\leq j_4\leq n\,:j_3-j_2>m}[\ldots] + \sum_{1\leq j_1\leq\ldots\leq j_4\leq n\,:j_4-j_3>m}[\ldots]\Big\} \\ &\leq C\big(\max_{t,j=1,\ldots,n} a_{n,tj}\big)^2\{q_{nt,1} + q_{nt,2} + q_{nt,3}\}. \end{aligned} \qquad (7.6)$$

Observe that

$$\begin{aligned} q_{nt,1} &\leq C\sum_{1\leq j_1\leq\ldots\leq j_4\leq n\,:j_4-j_2\leq 2m} a_{n,tj_1}a_{n,tj_3} \\ &\leq C\Big(\sum_{j_1=1}^{n} a_{n,tj_1}\Big)\Big(\sum_{j_3=1}^{n} a_{n,tj_3}\Big\{\sum_{j_2,j_4=1:|j_2-j_3|\leq 2m,|j_4-j_3|\leq 2m}^{n} 1\Big\}\Big) \end{aligned}$$

$$\leq \quad Cm^2(\sum_{j_1=1}^{n} a_{n,tj_1})(\sum_{j_3=1}^{n} a_{n,tj_3}) = Cm^2(\sum_{j=1}^{n} a_{n,tj})^2. \tag{7.7}$$

To bound $q_{nt,2}$, recall that by Assumption 17(iii), for $j_3 - j_2 > m$,

$$E[w_{j_3}w_{j_4}|\mathcal{F}_{j_2}] = E[w_{j_3}w_{j_4}] + r_{mj_2,j_3j_4}, \quad (Er^2_{mj_2,j_3j_4})^{1/2} \leq Cn^{-2}. \tag{7.8}$$

Since for $j_1 \leq j_2$ variables w_{j_1}, w_{j_2} are $\mathcal{F}_{j_2}$ measurable, then

$$\begin{aligned} E[w_{j_1}...w_{j_4}] &= E\big[E[w_{j_1}...w_{j_4}|\mathcal{F}_{j_2}]\big] = E\big[w_{j_1}w_{j_2}E[w_{j_3}w_{j_4}|\mathcal{F}_{j_2}]\big] \\ &= E[w_{j_1}w_{j_2}]E[w_{j_3}w_{j_4}] + E\big[w_{j_1}w_{j_2}r_{mj_2,j_3j_4}\big]. \end{aligned}$$

Moreover, by the assumptions of the lemma,

$$E[w_{j_1}w_{j_2}] = \mathrm{cov}(w_{j_1}, w_{j_2}) = \gamma_w(j_2 - j_1), \quad \sum_{j=-\infty}^{\infty} |\gamma_w(j)| < \infty,$$

$$|E\big[w_{j_1}w_{j_2}r_{mj_2,j_3j_4}\big]| \leq (Ew^4_{j_1})^{1/4}(Ew^4_{j_2})^{1/4}(Er^2_{mj_2,j_3j_4})^{1/2} \leq Cn^{-2}$$

by (7.8) and since $Ew^4_j \leq C$ where $C < \infty$ does not depend on j. Therefore,

$$\begin{aligned} q_{nt,2} &\leq C \sum_{1\leq j_1\leq ...\leq j_4\leq n : j_4-j_2>2m} a_{n,tj_1}a_{n,tj_3}\{|\gamma_w(j_2-j_1)\gamma_w(j_4-j_3)| + Cn^{-2}\} \\ &\leq C\big(\sum_{j_1,j_2=1}^{n} a_{n,tj_1}|\gamma_w(j_2-j_1)|\big)^2 + C\big(\sum_{j=1}^{n} a_{n,tj}\big)^2 \\ &\leq C\big(\{\sum_{j_1=1}^{n} a_{n,tj_1}\}\{\sum_{j=-\infty}^{\infty} |\gamma_w(j)|\}\big)^2 + C\big(\sum_{j=1}^{n} a_{n,tj}\big)^2 \\ &\leq C\big(\sum_{j=1}^{n} a_{n,tj}\big)^2. \end{aligned} \tag{7.9}$$

To bound $q_{nt,3}$, notice that by the Assumption 2.5(iii), for $j_4 - j_3 > m$,

$$E[w_{j_4}|\mathcal{F}_{j_3}] = E[w_{j_4}] + r_{mj_3,j_4}, \quad (Er^4_{mj_3,j_4})^{1/4} \leq Cn^{-2}. \tag{7.10}$$

Recall that $E[w_{j_4}] = 0$. For $j_1 \leq j_2 \leq j_3$ variables $w_{j_1}, w_{j_2}, w_{j_3}$ are $\mathcal{F}_{j_3}$ measurable. Therefore,

$$\begin{aligned} E[w_{j_1}...w_{j_4}] &= E\big[E[w_{j_1}...w_{j_4}|\mathcal{F}_{j_3}]\big] = E\big[w_{j_1}w_{j_2}w_{j_3}E[w_{j_4}|\mathcal{F}_{j_3}]\big] \\ &= E\big[w_{j_1}w_{j_2}w_{j_3}r_{mj_3,j_4}\big]. \end{aligned}$$

Hence,

$$|E\big[w_{j_1}w_{j_2}w_{j_3}r_{mj_3,j_4}\big]| \leq (Ew^4_{j_1}Ew^4_{j_2}Ew^4_{j_3}Er^4_{mj_3,j_4})^{1/4} \leq Cn^{-2}$$

since under the assumption of the lemma, $Ew_j^4 \le C$ where $C < \infty$ does not depend on j. This implies

$$q_{nt,3} \le Cn^{-2} \sum_{1 \le j_1 \le \ldots \le j_4 \le n \,: j_4 - j_3 > 2m} a_{n,tj_1} a_{n,tj_3} \le C\Big(\sum_{j=1}^{n} a_{n,tj}\Big)^2. \qquad (7.11)$$

The bound (7.6) together with (7.7)-(7.11) implies (7.4).

Proof of (7.5). For $1 \le m \le k \le n$, denote

$$s_{[m,k]}(t) = \sum_{j=m}^{k} a_{n,tj}\beta_j w_j, \quad s_{(m,k]}(t) = \sum_{j=m+1}^{k} a_{n,tj}\beta_j w_j.$$

Then,

$$s_{nt} = s_{[1,t]}(t) + s_{(t,t+2m]}(t) + s_{(t+2m,n]}(t).$$

Hence,

$$\begin{aligned} i_n &= n^{-1}\sum_{t=1}^{n} \xi_t s_{nt}\varepsilon_t^* \\ &= n^{-1}\sum_{t=1}^{n} \xi_t s_{[1,t]}(t)\varepsilon_t^* + n^{-1}\sum_{t=1}^{n} \xi_t s_{(t,t+2m]}(t)\varepsilon_t^* + n^{-1}\sum_{t=1}^{n} \xi_t s_{(t+2m,n]}(t)\varepsilon_t^* \\ &= i_{n,1} + i_{n,2} + i_{n,3}. \end{aligned}$$

We will show that

$$\begin{aligned} Ei_{n,1}^2 &\le Cm(\max_{t,j=1,\ldots,n} a_{n,tj})^2, \;\; E|i_{n,2}| \le Cm(\max_{t,j=1,\ldots,n} a_{n,tj}), \\ Ei_{n,3}^2 &\le Cm(\max_{t,j=1,\ldots,n} a_{n,tj})^2, \end{aligned} \qquad (7.12)$$

which implies (7.5):

$$\begin{aligned} E|i_n| &\le E[i_{n,1} + i_{n,2} + i_{n,3}] \le (E[i_{n,1}^2])^{1/2} + E[i_{n,2}] + (E[i_{n,3}^2])^{1/2} \\ &\le Cm(\max_{t,j=1,\ldots,n} a_{n,tj}). \end{aligned}$$

We have,

$$\begin{aligned} E[i_{n,1}^2] &= n^{-2}\sum_{t,s=1}^{n} E[\xi_t\xi_s s_{[1,t]}(t)s_{[1,s]}(s)\varepsilon_t^*\varepsilon_s^*] \\ &= n^{-2}\sum_{t,s=1}^{n}\sum_{j=1}^{t}\sum_{m=1}^{s} a_{n,tj}a_{n,sm}E[\xi_t\xi_s\beta_j\beta_m]E[w_j w_m\varepsilon_t^*\varepsilon_s^*] \end{aligned}$$

$$= \quad n^{-2}\sum_{t=1}^{n} E[\xi_t^2 s_{[1,t]}^2(t)\varepsilon_t^{*\,2}], \tag{7.13}$$

because $E[w_j w_m \varepsilon_t^* \varepsilon_s^*] = 0$ for $s < t$ since $E[\varepsilon_t^*|\mathcal{F}_{t-1}] = 0$ and for $j \le t$, and w_j is $\mathcal{F}_{t-1}$ measurable.

To evaluate $E[i_{n,1}^2]$, denote $\Delta_n = m(\max_{t,j=1,\dots,n} a_{n,tj})^2$. Then,

$$\begin{aligned} &E[\xi_t^2 s_{[1,t]}^2(t)\varepsilon_t^{*\,2}] \\ &\quad = E[\xi_t^2 \varepsilon_t^{*\,2}\, s_{[1,t]}^2(t)\{I\big(\xi_t^2\varepsilon_t^{*\,2} \le (n\Delta_n)^{-1} s_{[1,t]}^2(t)\big) + I\big(\xi_t^2\varepsilon_t^{2\,*} > (n\Delta_n)^{-1} s_{[1,t]}^2(t)\big)\}] \\ &\quad = E[(n\Delta_n)^{-1}\, s_{[1,t]}^4(t)] + (n\Delta_n)E[\xi_t^4\varepsilon_t^{*\,4}]. \end{aligned}$$

By (7.4),

$$\begin{aligned} Es_{[1,t]}^4(t) \quad &\le \quad C\big(\sum_{j=1}^{n} a_{n,tj}\big)^2(\max_{t,j=1,\dots,n} a_{n,tj})^2 m^2 \\ &\le \quad Cn^2(\max_{t,j=1,\dots,n} a_{n,tj})^4 m^2 = C(n\Delta_n)^2. \end{aligned}$$

Under assumptions of lemma, $E[\xi_t^4\varepsilon_t^{*\,4}] \le E[\xi_t^4]E[\varepsilon_t^{*\,4}] \le C$, where $C < \infty$ does not depend on t, j, m, n. Hence,

$$\begin{aligned} E[i_{n,1}^2] \quad &\le \quad C\{n^{-2}\sum_{t=1}^{n}(n\Delta_n)^{-1}Es_{[1,t]}^4(t)) + n^{-2}\sum_{t=1}^{n}(n\Delta_n)E[\xi_t^4\varepsilon_t^{*\,4}]\} \qquad (7.14) \\ &\le \quad C\{n^{-2}\sum_{t=1}^{n}(n\Delta_n) + n^{-2}\sum_{t=1}^{n}(n\Delta_n)\} \le C\Delta_n \end{aligned}$$

which proves (7.12) for $i_{n,1}$.

To evaluate $Ei_{n,2}$, bound

$$|i_{n,2}| \quad \le \quad n^{-1}\sum_{t=1}^{n}\sum_{j=t+1}^{t+m} a_{n,tj}\{|\xi_t\beta_j|\}\{|w_j\varepsilon_t^*|\}.$$

Under assumptions of lemma, $E[|\xi_t\beta_j|] \le E[\xi_t^2] + E[\beta_j^2] \le C$ and $E|w_j\varepsilon_t^*| \le E[w_j^2] + E[\varepsilon_t^{*\,2}] \le C$. Then,

$$\begin{aligned} E|i_{n,2}| \quad &\le \quad n^{-1}\sum_{t=1}^{n}\sum_{j=t+1}^{t+m} a_{n,tj}E[|\xi_t\beta_j|]E[|w_t\varepsilon_j^*|] \\ &\le \quad Cn^{-1}\sum_{t=1}^{n}\sum_{j=t+1}^{t+m} a_{n,tj} \le C(\max_{t,j=1,\dots,n} a_{n,tj})m \end{aligned}$$

which proves (7.12) for $i_{n,2}$.

Finally, we estimate $E[i_{n,3}]$. We have that

$$\begin{aligned} Ei_{n,3}^2 &= E\big(n^{-1}\sum_{t=1}^n \xi_t s_{(t+2m,n]}(t)\varepsilon_t^*\big)^2 \\ &= n^{-2}\sum_{t,s=1}^n E\big[\xi_t\xi_s s_{(t+2m,n]}(t)s_{(s+2m,n]}(s)\varepsilon_t^*\varepsilon_s^*\big] \\ &= n^{-2}\sum_{t=s=1}^n [...] + 2n^{-2}\sum_{s=1}^n\sum_{t=s+1}^n [...] = q_{n1}+q_{n2}. \end{aligned}$$

The bound

$$q_{n1} = n^{-2}\sum_{t=1}^n E\big[\xi_t^2 s_{(t+2m,n]}^2(t)\varepsilon_t^{*\,2}\big] \le C\Delta_n = Cm(\max_{t,j=1,\dots,n} a_{n,tj})^2 \tag{7.15}$$

follows using the same argument as in the proof of (7.14) for $E[i_{n,1}^2]$.

Next we show that

$$q_{n2} = C(\max_{t,j=1,\dots,n} a_{n,tj})^2, \tag{7.16}$$

which together with (7.15) proves (7.12) for $i_{n,3}$.

Proof of (7.16). Let $t>s$. Then,

$$\begin{aligned} &\xi_t\xi_s s_{(t+2m,n]}(t)s_{(s+2m,n]}(s)\varepsilon_t^*\varepsilon_s^* \\ &\qquad = \xi_t\xi_s\{\sum_{j=t+2m+1}^n a_{n,tj}\beta_j w_j\}\{\sum_{i=s+2m+1}^n a_{n,si}\beta_i w_i\}\varepsilon_t^*\varepsilon_s^*, \\ &E[\xi_t\xi_s s_{(t+2m,n]}(t)s_{(s+2m,n]}(s)\varepsilon_t^*\varepsilon_s^*] \\ &\qquad = \sum_{j=t+2m+1}^n\sum_{i=s+2m+1}^n a_{n,tj}a_{n,si}E[\xi_t\xi_s\beta_j\beta_i]E[w_jw_i\varepsilon_t^*\varepsilon_s^*], \qquad (7.17) \\ &|E[\xi_t\xi_s s_{(t+2m,n]}(t)s_{(s+2m,n]}(s)\varepsilon_t^*\varepsilon_s^*]| \le C\sum_{j=t+2m+1}^n\sum_{i=s+2m+1}^n a_{n,tj}a_{n,si}|E[w_jw_i\varepsilon_t^*\varepsilon_s^*]|, \end{aligned}$$

since under assumption of lemma, $|E[\xi_t\xi_s\beta_j\beta_i]| \le E\xi_t^4+E\xi_s^4+E\beta_j^4+E\beta_i^4 \le C$ where $C<\infty$ does not depend on t,s,j and i.

We will prove that for $j\ge t+2m+1$, $i\ge s+2m+1$ and $t>s$,

$$\big|E[w_jw_i\varepsilon_t^*\varepsilon_s^*]\big| \le Cn^{-2}, \tag{7.18}$$

which together with (7.17) implies (7.16):

$$\begin{aligned} q_{n2} &\leq n^{-2}\sum_{s=1}^{n}\sum_{t=s+1}^{n}\big|E\big[\xi_t\xi_s s_{(t+2m,n]}(t)s_{(s+2m,n]}(s)\varepsilon_t^*\varepsilon_s^*\big]\big| \\ &\leq Cn^{-2}\sum_{s=1}^{n}\sum_{t=s+1}^{n}\{n^{-2}\sum_{j,i=1}^{n}a_{n,tj}a_{n,si}\} \leq (\max_{t,j=1,\dots,n} a_{n,tj})^2. \end{aligned}$$

Proof of (7.18). To bound $|E[w_jw_i\varepsilon_t^*\varepsilon_s^*]|$, we consider two cases.

a) Let $i > t+m$. Since $j > t+2m$, then by Assumption 2.5(iii),

$$E[w_jw_i\big|\mathcal{F}_{t+m}] = E[w_jw_i] + r_{m(t+m),ji}, \quad (Er^2_{m(t+m),ji})^{1/2} \leq Cn^{-2}. \tag{7.19}$$

Moreover, $t, s \leq t+m$ and therefore variables ε_t^*, ε_s^* are $\mathcal{F}_{t+m}$ measurable. Thus,

$$\begin{aligned} E[w_jw_i\varepsilon_t^*\varepsilon_s^*] &= E\big[E[w_jw_i\varepsilon_t^*\varepsilon_s^*\big|\mathcal{F}_{t+m}]\big] = E\big[\varepsilon_t^*\varepsilon_s^*E[w_jw_i\big|\mathcal{F}_{t+m}]\big] \\ &= E[\varepsilon_t^*\varepsilon_s^*]E[w_jw_i] + E[\varepsilon_t^*\varepsilon_s^*r_{m(t+m),ji}] = E[\varepsilon_t^*\varepsilon_s^*r_{m(t+m),ji}] \end{aligned}$$

since $E[\varepsilon_t^*\varepsilon_s^*] = 0$ when $t > s$. Notice that

$$|E[\varepsilon_t^*\varepsilon_s^*r_{m(t+m),ji}]| \leq (E[\varepsilon_t^{*2}\varepsilon_s^{*2}])^{1/2}(Er^2_{m(t+m),ji})^{1/2} \leq Cn^{-2}$$

since under the assumptions of the lemma, $E[\varepsilon_t^{*2}\varepsilon_s^{*2}] \leq E\varepsilon_t^{*4} + E\varepsilon_s^{*4} \leq C$ where $C < \infty$ does not depend on t, s. This implies (7.18).

b) Let $s + 2m < i \leq t+m$. Since $j > t+2m$ and $E[w_j] = 0$, then by Assumption 2.5(iii),

$$E[w_j\big|\mathcal{F}_{t+m}] = E[w_j] + r_{m(t+m),j} = r_{m(t+m),j}, \quad (Er^4_{m(t+m),j})^{1/4} \leq Cn^{-2}.$$

Moreover, for $t, s, i \leq t+m$ variables ε_t^*, ε_s^*, w_i are $\mathcal{F}_{t+m}$ measurable. Thus,

$$\begin{aligned} E[w_jw_i\varepsilon_t^*\varepsilon_s^*] &= E\big[E[w_jw_i\varepsilon_t^*\varepsilon_s^*\big|\mathcal{F}_{t+m}]\big] = E\big[\varepsilon_t^*\varepsilon_s^*w_iE[w_j\big|\mathcal{F}_{t+m}]\big] \\ &= E[\varepsilon_t^*\varepsilon_s^*w_ir_{m(t+m),j}], \\ |E[\varepsilon_t^*\varepsilon_s^*w_ir_{m(t+m),j}]| &\leq (E[\varepsilon_t^{*4}]E[\varepsilon_s^{*4}]Ew_i^4])^{1/4}(Er^4_{m(t+m),j})^{1/4} \leq Cn^{-2} \end{aligned}$$

which implies (7.18).

This completes the proof of (7.4) and of the lemma. □

8 Additional Monte Carlo Simulations

In this section, we use Monte Carlo simulations to verify the asymptotic theory established in Section 2 of the main paper for more complex data generating processes. In subsection 8.1 models permit the dependence between regressors and regression noise. In Section 8.2, we compare the performance of robust standard errors and standard errors in PTVR estimation.

8.1 Regressors generated by AR(1) processes

In Section 3 of the main paper, we consider the model

$$y_t = \alpha' x_t + \beta_t' z_t + u_t, \quad t = 1, ..., n,$$

with regressors $x_t = g_{xt}\eta_{xt},\ z_t = g_{zt}\eta_{zt}$ where η_{xt}, η_{zt} are stationary MA(1) processes.

This section explores the finite-sample performance of PTVR estimation procedures when the components η_{xt}, η_{zt} of regressors are stationary autoregressive AR(1) processes. The regressors x_t and z_t are constructed as:

$$\begin{aligned} x_t &= g_{xt}\eta_{xt}, \qquad \eta_{xt} = 0.2 + 0.5\eta_{x,t-1} + \epsilon_{xt}, \\ z_t &= g_{zt}\eta_{zt}, \qquad \eta_{zt} = 0.2 + 0.5\eta_{z,t-1} + \epsilon_{zt}, \end{aligned}$$

where $\epsilon_{xt} = \varepsilon_{t-1}$ and $\epsilon_{zt} = \varepsilon_{t-2}$. A more complex regression noise $u_t = h_t\varepsilon_t$ is used. We suppose that ε_t follows GARCH(1,1) process:

$$\varepsilon_t = \sigma_t e_t, \quad \sigma_t^2 = 1 + 0.7\sigma_{t-1}^2 + 0.2\varepsilon_{t-1}^2, \quad e_t \sim i.i.d.\mathcal{N}(0,1). \tag{8.1}$$

In this setting, $\{\varepsilon_t\}$ and $\{\eta_{xt}, \eta_{zt}\}$ are mutually dependent processes.

The scale factor h_t is either deterministic or stochastic trend:

$$\text{Deterministic}: \quad h_t = 0.5(t/n) + 0.5,\ t = 1, ..., n, \tag{8.2}$$

$$\text{Stochastic}: \quad h_t = |n^{-\gamma}\sum_{i=1}^{t}\xi_i| + 0.2, \tag{8.3}$$

where ξ_i is an ARFIMA$(0, d, 0)$ process with parameter $d = 0.4$, see Giraitis, Koul and Surgailis (2012), Chapter 7.

The time-varying intercept β_{1t} is a sine function:

$$\beta_{1t} = 0.5\sin(\pi t/n) + 1.$$

We centre on two types of time-varying parameter β_t:

$$\text{Deterministic}: \quad \beta_{2t} = 0.5\sin(2\pi t/n) + 1, \quad t = 1, ..., n, \tag{8.4}$$

$$\text{Stochastic :} \quad \beta_{2t} = |n^{-\gamma} \sum_{i=1}^{t} e_i| + 0.2, \tag{8.5}$$

where e_i is an ARFIMA$(0, d, 0)$ process with parameter $d = 0.4$.

We consider two data generating models, Model 8.1 and Model 8.2, that include deterministic and stochastic scale factors g_{xt}, g_{zt}, $t = 1, ..., n$:

$$\text{Deterministic : } g_{xt} = \frac{1}{2}\sin\left(\frac{2\pi t}{n}\right) + \frac{0.3t}{n} + 1,\ \ g_{zt} = \frac{1}{2}\cos\left(\frac{2\pi t}{n}\right) + \frac{0.4t}{n} + 1, \tag{8.6}$$

$$\text{Stochastic : } g_{xt} = |n^{-\gamma} \sum_{i=1}^{t} \upsilon_{xi}| + 0.2, \quad g_{zt} = |n^{-\gamma} \sum_{i=1}^{t} \upsilon_{zi}| + 0.2, \tag{8.7}$$

where $\{\upsilon_{xi}\}, \{\upsilon_{zi}\}$ are stationary ARFIMA$(0, d, 0)$ processes with memory parameter $d = 0.4$.

Model 8.1. *$y_t, t = 1, \cdots, n$ follows (34) with deterministic scale factors $g_{x,t}, g_{zt}$ as in (8.6), h_t as (8.2) and parameter β_{2t} as (8.4).*

Model 8.2. *$y_t, t = 1, \cdots, n$ follows (34) with stochastic scale factors $g_{x,t}, g_{zt}$ as in (8.7), h_t as (8.3) and parameter β_{2t} as (8.5).*

Tables 1 and 2 report estimation results for fixed parameter α in Model 8.1 and 8.2. They confirm good coverage rate for bandwidth $H = n^h$, $h = 0.6, 0.7$.

Table 1: Estimation of α in Model 8.1.

h	Bias	RMSE	CP	SD
0.4	0.0239	0.0326	79.2	0.0222
0.5	0.0121	0.0250	90.7	0.0219
0.6	0.0053	0.0224	94.6	0.0218
0.7	-0.0019	0.0223	94.7	0.0222

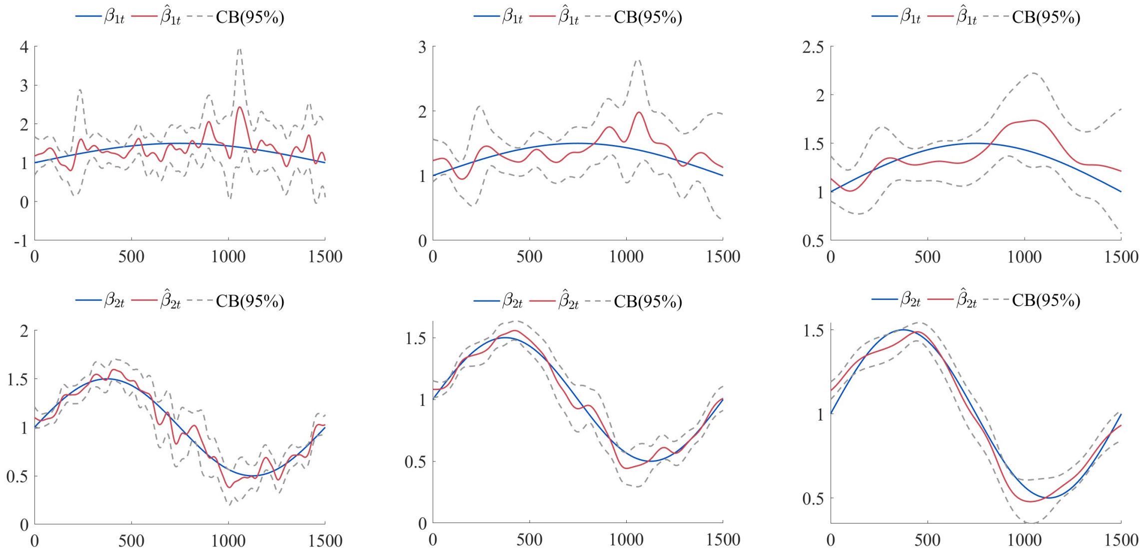
Table 2: Estimation of α in Model 8.2.

h	Bias	RMSE	CP	SD
0.4	0.0558	0.0807	83.4	0.0582
0.5	0.0277	0.0641	93.1	0.0578
0.6	0.0135	0.0590	94.6	0.0574
0.7	0.0060	0.0576	94.2	0.0572

Next we proceed to estimation results for the time-varying parameter β_t with pre-selected bandwidth $H = n^{0.6}$ and $H_z = n^h$, $h = 0.4, 0.5, 0.6$. Figure 1 displays estimation results for a single simulation in Model 8.1. It shows that the estimator $\widehat{\beta}_t$ tracks the path of the true parameter β_t and the true parameter is well covered across the time t by 95% confidence intervals. Empirical coverage rates, shown in Figure 2, are close to the nominal 95% which confirms the good finite-sample performance of the normal approximation established for PTVR estimator for components of β_t.

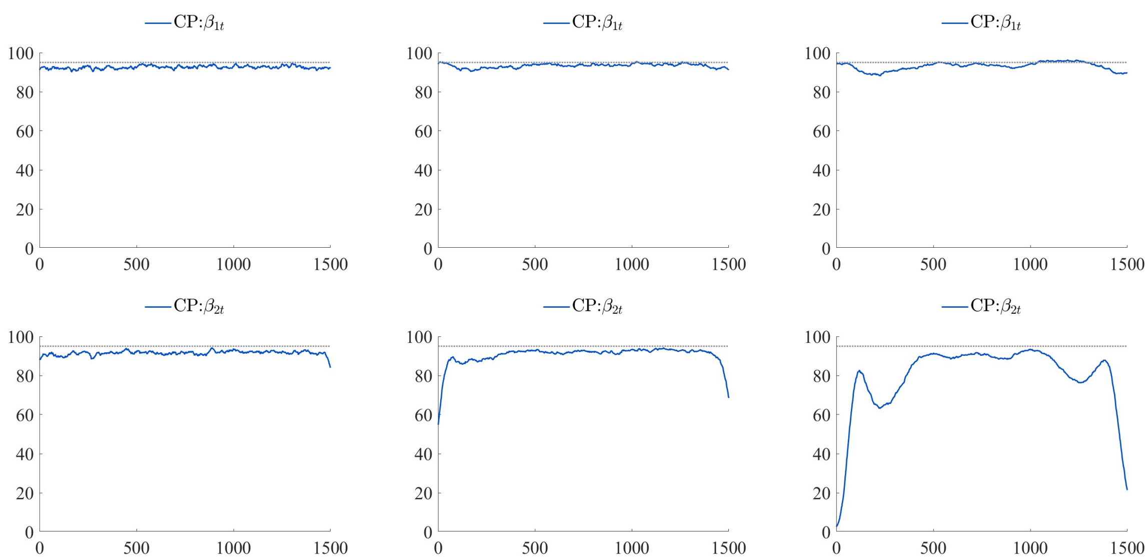
Figure 3 shows that overall the bias in estimation of β_t is small and increases as the bandwidth H_z increases. Figure 4 reveals that the RMSE is declining when the bandwidth H_z increases, but it can rise when there is a lot of variability in the time-varying parameter β_t, see panel (b).

The estimation results of Model 8.2 are similar to those for Model 8.1, confirming the applicability of PTVR estimation procedures to our complex regression setting. By addressing both deterministic and stochastic scale factors and parameters, our results reaffirm the theoretical and practical strengths of the PTVR estimation in dealing with complex data structure.



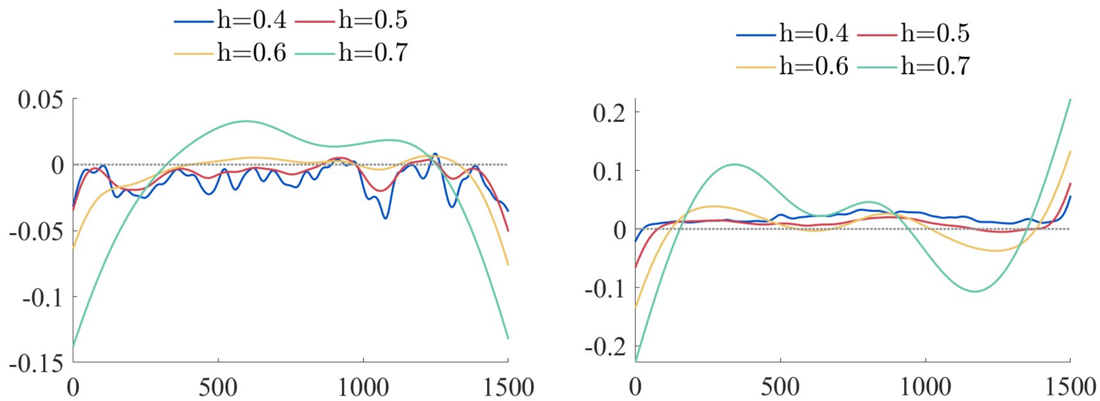
(a) $H = n^{0.6}, H_z = n^{0.4}$ (b) $H = n^{0.6}, H_z = n^{0.5}$ (c) $H = n^{0.6}, H_z = n^{0.6}$

Figure 1: PTVR estimates of parameters β_t and their 95% confidence bands for one simulation of Model 8.1. Sample size $n = 1500$.



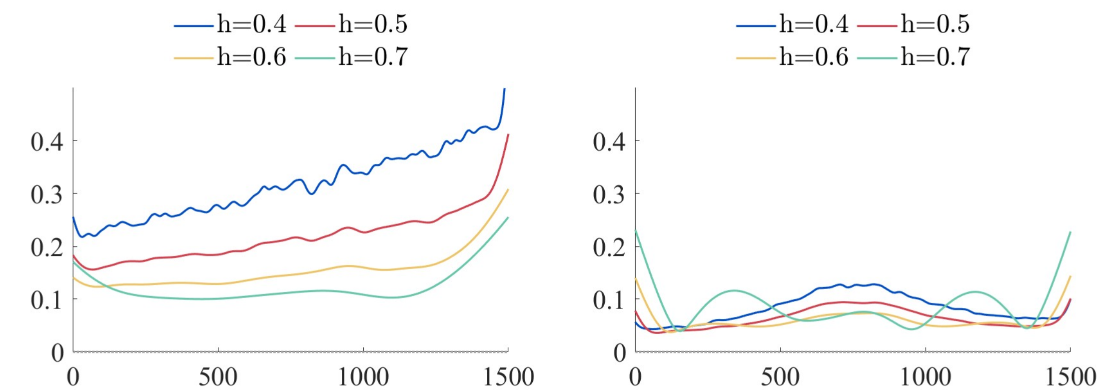
(a) $H = n^{0.6}, H_z = n^{0.4}$ (b) $H = n^{0.6}, H_z = n^{0.5}$ (c) $H = n^{0.6}, H_z = n^{0.6}$

Figure 2: Empirical coverage probability of 95% confidence intervals for β_t in Model 8.1. Sample size $n = 1500$.



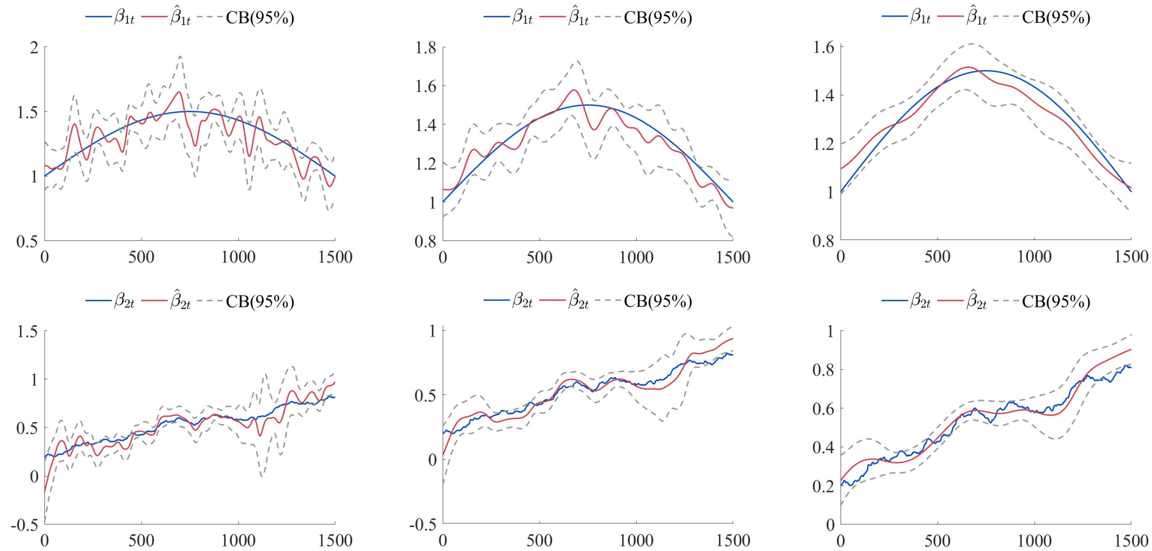
(a) β_{1t} (b) β_{2t}

Figure 3: Bias of β_t in Model 8.1. Sample size $n = 1500$. Bandwidth parameters $H = n^{0.6}, H_z = n^h$, $h = 0.4, 0.5, 0.6, 0.7$.



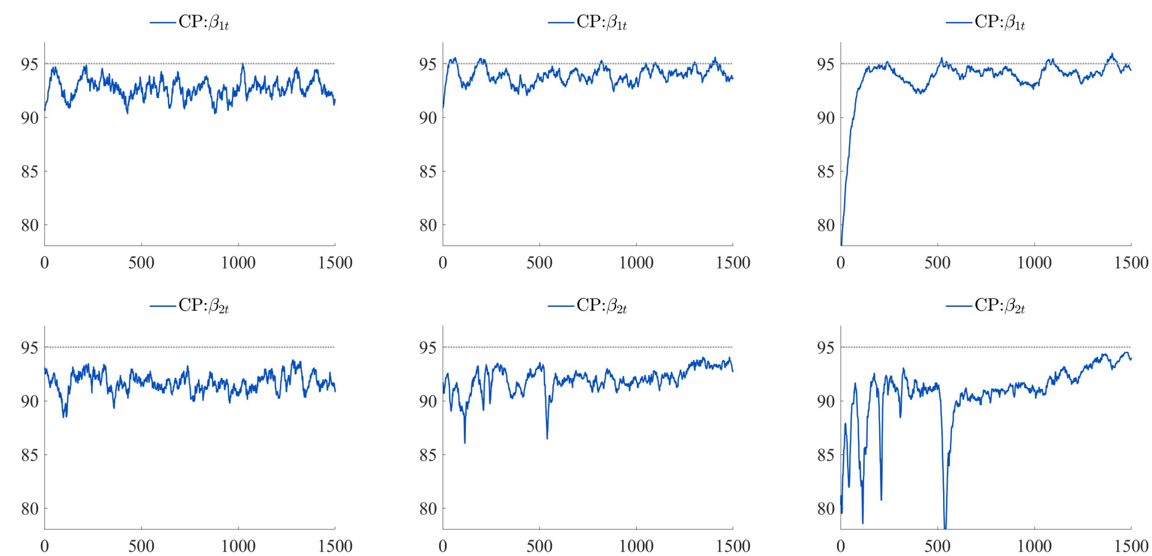
(a) β_{1t} (b) β_{2t}

Figure 4: RMSE of β_t in Model 8.1. Sample size $n = 1500$. Bandwidth parameters $H = n^{0.6}, H_z = n^h$, $h = 0.4, 0.5, 0.6, 0.7$.



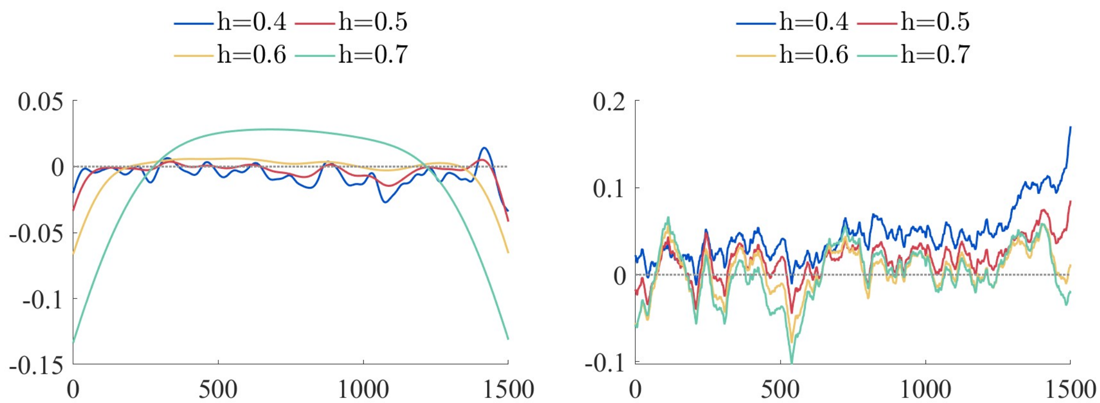
(a) $H = n^{0.6}, H_z = n^{0.4}$ (b) $H = n^{0.6}, H_z = n^{0.5}$ (c) $H = n^{0.6}, H_z = n^{0.6}$

Figure 5: PTVR estimates of parameters β_t and their 95% confidence bands for one simulation of Model 8.2. Sample size $n = 1500$.



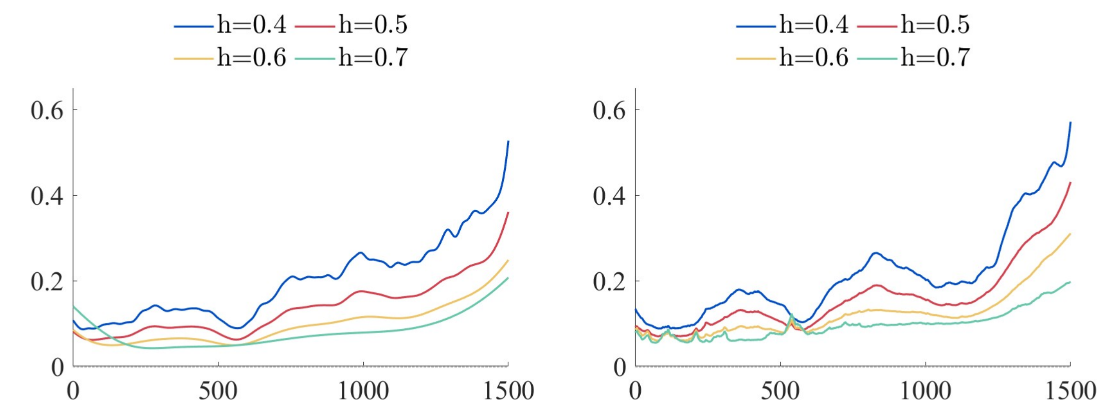
(a) $H = n^{0.6}, H_z = n^{0.4}$ (b) $H = n^{0.6}, H_z = n^{0.5}$ (c) $H = n^{0.6}, H_z = n^{0.6}$

Figure 6: Empirical coverage probability of 95% confidence intervals for β_t in Model 8.2. Sample size $n = 1500$.



(a) β_{1t} (b) β_{2t}

Figure 7: Bias of β_t in Model 8.2. Sample size $n = 1500$. Bandwidth parameters $H = n^{0.6}, H_z = n^h$, $h = 0.4, 0.5, 0.6, 0.7$.



(a) β_{1t} (b) β_{2t}

Figure 8: RMSE of β_t in Model 8.2. Sample size $n = 1500$. Bandwidth parameters $H = n^{0.6}, H_z = n^h$, $h = 0.4, 0.5, 0.6, 0.7$.

8.2 Robust standard errors vs. classical standard errors

In the main paper, for estimation of PTVR model we apply the robust standard errors, (21) for fixed parameter, and (31) for time-varying parameter. We use the term *classical standard errors* to refer to standard errors commonly used in the literature. In this section, we make a comparison between the performance of the PTVR estimation based on robust standard errors, as in (21) and (31), and classical standard errors, as in (23) and (33).

We focus on a simplified version of Model 8.1 of Section 8.1. We set all scale factors $g_{x,t}, g_{zt}$ and h_t equal to 1 and generate arrays of samples y_t, $t = 1, \cdots, n$ of the following model:

$$y_t = \alpha x_t + \beta_t' z_t + u_t, \quad u_t = \varepsilon_t \tag{8.8}$$
$$= \beta_{1t} + \alpha x_t + \beta_{2t} z_{2t} + \varepsilon_t,$$

where ε_t is GARCH(1,1) process as in (8.1). With $g_{x,t}, g_{zt} = 1$, the regressors in (8.8) become

$$x_t = \eta_{xt}, \qquad \eta_{xt} = 0.2 + 0.5\eta_{x,t-1} + \epsilon_{xt},$$
$$z_t = \eta_{zt}, \qquad \eta_{zt} = 0.2 + 0.5\eta_{z,t-1} + \epsilon_{zt},$$

where $\epsilon_{xt} = \varepsilon_{t-1}$ and $\epsilon_{zt} = \varepsilon_{t-2}$. We set the fixed parameter $\alpha = 1$ and employ the time-varying parameter $\beta_t = (\beta_{1,t}, \beta_{2t})'$ as below:

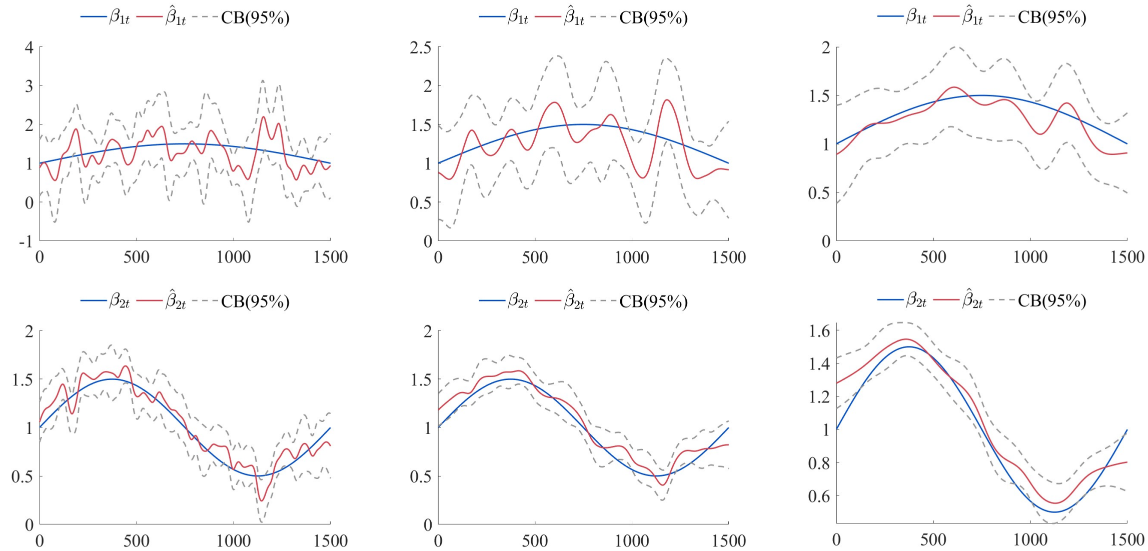
$$\beta_{1t} = 0.5 \sin(\pi t/n) + 1, \quad t = 1, ..., n,$$
$$\beta_{2t} = 0.5 \sin(2\pi t/n) + 1.$$

Table 3 reports the empirical coverage rate for the fixed parameter α based on robust standard errors (denoted by CP) and classical standard errors (denoted by CP_{st}). We observe that the robust standard errors produce coverage close to the nominal 95%, and, clearly, implementation of the classical standard errors leads to coverage distortions.

Table 3: Estimation of α in model (8.8)

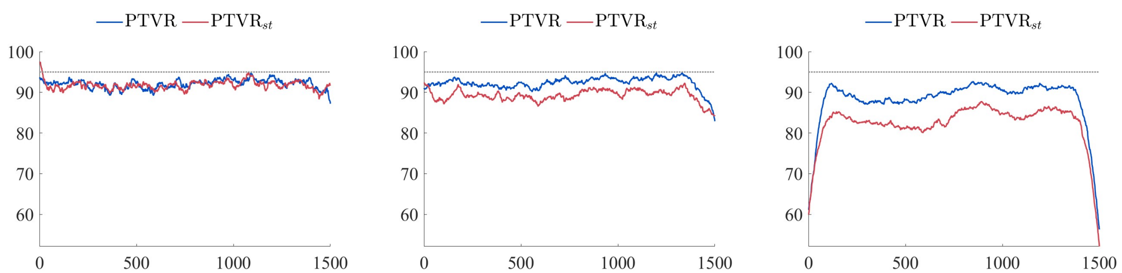
h	Bias	RMSE	SD	CP	CP_{st}
0.4	0.0395	0.0546	0.0377	79.2	62.2
0.5	0.0199	0.0423	0.0374	90.2	77.9
0.6	0.0098	0.0385	0.0372	92.4	82.3
0.7	0.0046	0.0374	0.0371	93.8	83.6

Figure 9 displays estimation results for time-varying parameter β_t for one single sample. We observe that the true value of β_t is well covered by the confidence intervals based on the robust standard errors. Figure 10 reports the empirical coverage (in %) of 95% confidence intervals for parameter β_{2t} based on robust standard errors (blue line) and classical standard errors (red line). The PTVR estimation based on robust standard errors achieves good coverage rate, while estimation with classical standard errors leads to size distortions due to the presence of a non i.i.d noise ε_t and the dependence between regressors and regression noise. Figure 11 reports the RMSE in estimation of β_{2t} with different bandwidth parameters $H_z = n^h$, $h = 0.4, 0.5, 0.6, 0.7$. It shows that RMSE decreases as h increases, but RMSE can rise because of rapid change in the time-varying parameter β_{2t}, see panel (b).



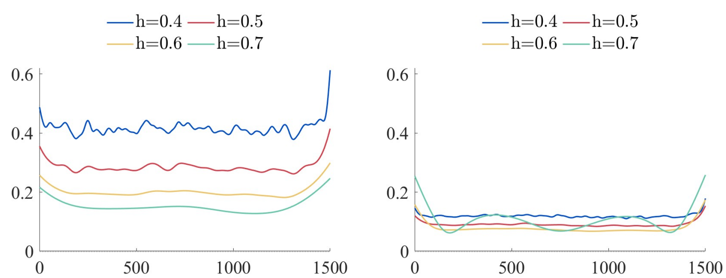
(a) $H = n^{0.6}, H_z = n^{0.4}$ (b) $H = n^{0.6}, H_z = n^{0.5}$ (c) $H = n^{0.6}, H_z = n^{0.6}$

Figure 9: PTVR estimates of parameters β_t and their 95% confidence bands for one simulation in (8.8). Sample size $n = 1500$.



(a) β_{2t}: $H = n^{0.6}, H_z = n^{0.4}$ (b) β_{2t}: $H = n^{0.6}, H_z = n^{0.5}$ (c) β_{2t}: $H = n^{0.6}, H_z = n^{0.6}$

Figure 10: Empirical coverage probability of 95% confidence intervals for β_t in (8.8). Sample size $n = 1500$.



(a) β_{1t} (b) β_{2t}

Figure 11: RMSE of β_t in (8.8). Sample size $n = 1500$. Bandwidth parameters $H = n^{0.6}, H_z = n^h$, $h = 0.4, 0.5, 0.6, 0.7$.

References

Giraitis, L., Koul, H.L., and D. Surgailis (2012): *Large Sample Inference for Long memory Processes*. Imperial College Press, London.

Giraitis, L., G. Kapetanios, and Y. Li (2024): "Regression modelling under general heterogeneity," *Preprint*, No. 983, Queen Mary University of London, https://www.qmul.ac.uk/sef/research/workingpapers/2024/items/983.html

School of Economics and Finance

School of Economics and Finance
Queen Mary University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 7356
Fax: +44 (0)20 8983 3580
Web: www.econ.qmul.ac.uk/research/workingpapers/